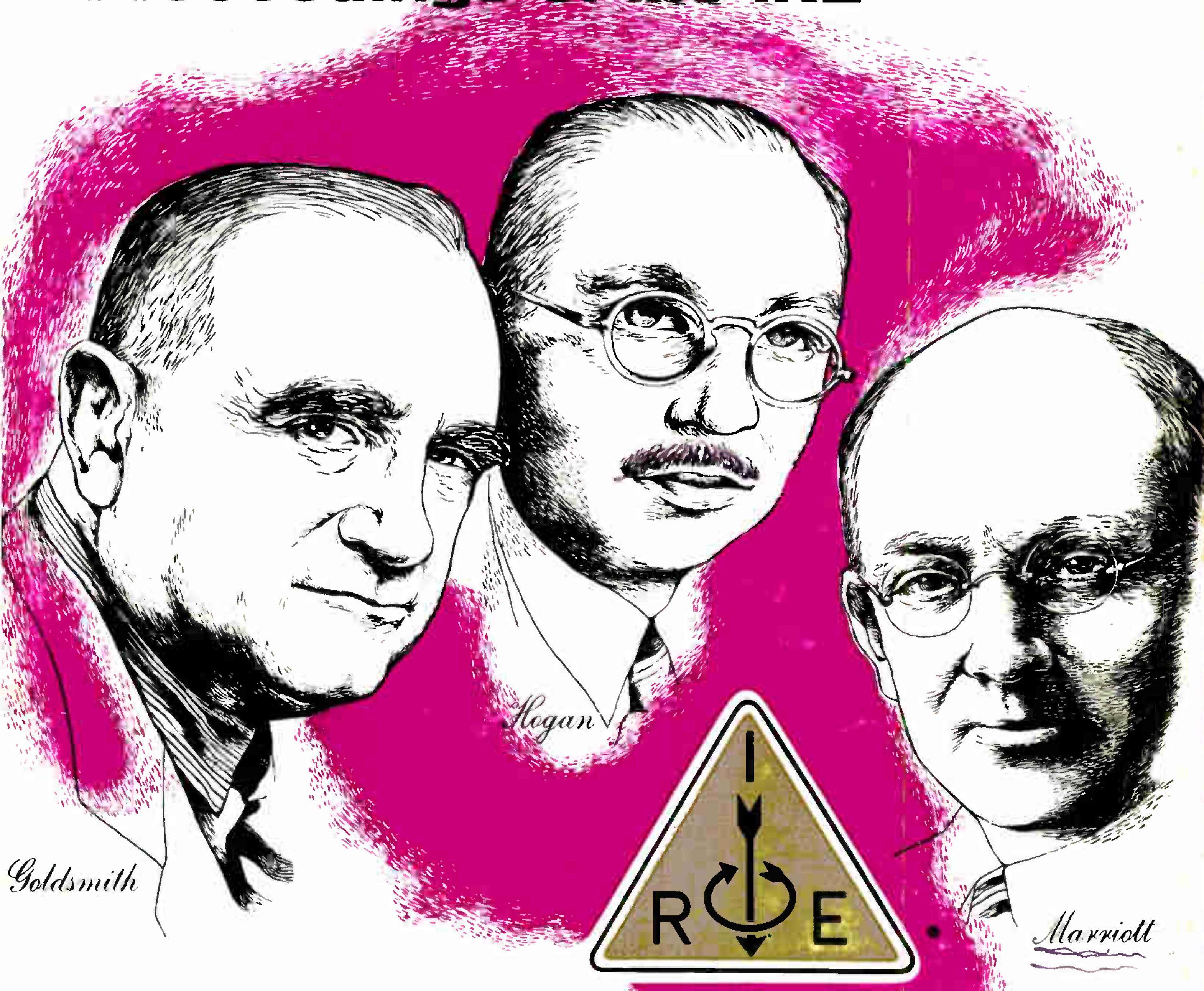
may 1962
the
institute
of
radio
engineers

Proceedings of the IRE



1912 Fiftieth Anniversary 1962

PIONEERS IN MINIATURIZATION

SPECIAL CUSTOM BUILT POWER TRANSFORMERS & REACTORS

Exceptional quality and reliability is provided in all UTC designs. Over 30 years of engineering knowledge and experience backed by complete environmental testing and life testing facilities assure the highest standard in the industry. Full analysis and evaluation of materials are conducted in UTC's Material and Chemical Laboratories. Rigid quality control measures coordinated with exhaustive statistical findings and latest production procedures results in the industry's highest degree of reliability. Range covered in Power Transformers is from milliwatts to 100 KVA.

New designs for both military and commercial types custom built to YOUR specifications for transistor and tube applications.

SOME TYPICAL APPLICATIONS:

- Current Limiting
- Filament
- Isolation
- Plate
- Transistor Inverter
- Transistor Supply

Write for catalog of over 1,000 STOCK ITEMS with UTC High Reliability IMMEDIATELY AVAILABLE from your local distributor.

SPECIAL CUSTOM BUILT IRON CORE COMPONENTS FOR EVERY PHASE OF THE ELECTRONICS ART

POWER TRANSFORMERS • AUDIO TRANSFORMERS
REACTORS • PULSE TRANSFORMERS • ELECTRIC WAVE FILTERS • LUMPED CONSTANT DELAY LINES
HIGH Q COILS • MAGNETIC AMPLIFIERS
SATURABLE REACTORS • REFERENCE UNITS

UNITED TRANSFORMER CORP.

150 VARICK STREET, NEW YORK 13, N.Y.

PACIFIC MFG. DIVISION: 3630 EASTHAM DRIVE, CULVER CITY, CALIF.

EXPORT DIVISION: 13 EAST 40th STREET, NEW YORK 16, N.Y. CABLES: "ARLAB"

VOLUME 50, NUMBER 5

May, 1962

published monthly by The Institute of Radio Engineers, Inc.

Proceedings of the IRE®

contents

COVER

The men who founded the IRE May 13, 1912.

MAY

published monthly by The Institute of Radio Engineers, Inc.

Proceedings of the IRE®

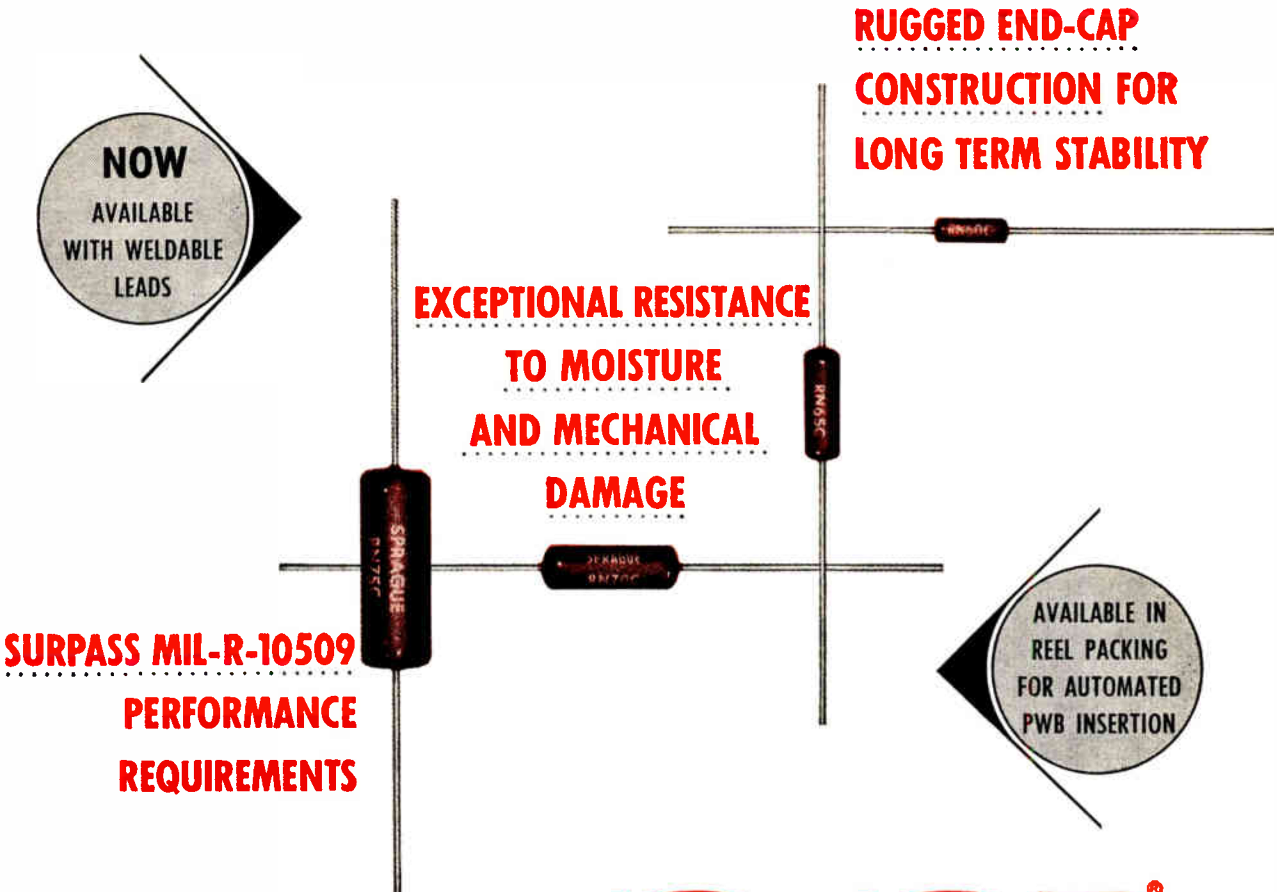
MAY

continued

continued

PROCEEDINGS OF THE IRE, published monthly by The Institute of Radio Engineers, Inc., at 1 East 79 Street, New York 21, N. Y. Manuscripts should be submitted in triplicate to the Editorial Department. Correspondence column items should not exceed four double-spaced pages (illustrations count as one-half page each). Responsibility for contents of papers published rests upon the authors, and not the IRE or its members. All republication rights, including translations, are reserved by the IRE and granted only on request. Abstracting is permitted with mention of source.

Thirty days notice is required for change of address. Yearly subscription price: members $9.00, one additional subscription $13.50; public libraries and colleges $13.50; non-members in the United States, Canada, and U. S. Possessions $18.00; abroad $19.00. Second-class postage paid at Menasha, Wisconsin, under the act of March 3, 1879. Acceptance for mailing at a special rate of postage is provided for in the act of February 28, 1925, embodied in Paragraph 4, Section 412, P. L. and R., authorized October 26, 1927. Printed in U.S.A.



FILMISTOR®

METAL FILM RESISTORS

OFFER 5 DISTINCT TEMPERATURE COEFFICIENTS TO MEET ALL CIRCUIT REQUIREMENTS

Providing close accuracy, reliability and stability with low controlled temperature coefficients, these molded case metal-film resistors outperform precision wirewound and carbon film resistors. Prime characteristics include minimum inherent noise level, negligible voltage coefficient of resistance and excellent long-time stability under rated load as well as under severe conditions of humidity.

Close tracking of resistance values of 2 or more resistors over a wide temperature range is another key performance characteristic of molded-case Filmistor Metal Film Resistors. This is especially important where they are used to make highly accurate ratio dividers.

Filmistor Metal Film Resistors, in ⅛, ¼, ½ and 1 watt ratings, surpass stringent performance requirements of MIL-R-10509D, Characteristics C and E. Write for Engineering Bulletin No. 7025 to: Technical Literature Section, Sprague Electric Co., 235 Marshall Street, North Adams, Mass.

For application engineering assistance write:
Resistor Division, Sprague Electric Co., Nashua, New Hampshire.

SPRAGUE COMPONENTS

RESISTORS
CAPACITORS
MAGNETIC COMPONENTS
TRANSISTORS
INTERFERENCE FILTERS
PULSE TRANSFORMERS
PIEZOELECTRIC CERAMICS
PULSE-FORMING NETWORKS
HIGH TEMPERATURE MAGNET WIRE
CERAMIC-BASE PRINTED NETWORKS
PACKAGED COMPONENT ASSEMBLIES
FUNCTIONAL DIGITAL CIRCUITS

SPRAGUE®
THE MARK OF RELIABILITY

On this fiftieth anniversary of the IRE, all of us look forward to the advances sure to be made in the next fifty years. But we thought it would be fun to look back fifty years and see what the world was like in the year that the IRE was founded.

Peg Salzberg, the wife of AIL's Chief Scientist, contributed the idea and most of the material for this ad. It is arranged in the format that THE NEW YORK TIMES *uses to summarize the news of the day.*

1912 NEWS SUMMARY

the year the IRE was born

International

U. S. Troops land in China. U. S. Marines land in Cuba, Honduras, and Nicaragua to maintain order and to protect American interests.

General Monteagudo, commander in chief of the Cuban Government troops, formally turns over the government of the province of Oriente to the civil authorities, declaring that the revolution is over.

Russian troops advance into Azerbaijan and northeast Khorosan.

The French General Gouraud defeats the Rogui of Sichtala, said to have been France's most troublesome enemy in Morocco.

War breaks out between Turkey and a confederation of Greece, Serbia, Bulgaria, and Montenegro.

Andrew Carnegie sends a letter to the heads of all leading nations, appealing for peace. "Get together often. Learn to trust each other."

Dr. Charles W. Eliot, former president of Harvard, returns from a trip around the world in the interests of universal peace. His conclusion: "International or national disarmament is not taken seriously by the leaders and thinking men of the more important peoples."

The Arbitration League (which includes such eminent Englishmen as Sir Arthur Conan Doyle, John Galsworthy, Sir William Osler, and Sir William Ramsay) issues a "memorial protesting against the use of aerial vessels in war" and appeals "to all Governments to foster an international understanding which shall prevent the world from what they think would add new hideousness to warfare. . . . To the argument that because men fight on earth and water they may just as well fight in the air, they reply that there has never yet been a moment when it was practically possible to ban war machines of the earth or the water, but there is a moment when it is practically possible to ban those of the air. That moment is now before us, before these machines have been proved and great vested interests have been formed."

The board of directors of the Nobel Institute announce that no Peace Prize will be awarded in 1912.

The Titanic, the world's largest passenger ship, strikes an iceberg on her maiden voyage and sinks—with a loss of 1500 lives. (Her wireless calls for help were heard by numerous ships, but the radio room of the California, only a few miles away from the sinking ship, had shut down for the night and did not receive the distress signal.)

Captain Robert Falcon Scott and four members of his British Antarctic Expedition reach the South Pole—only to find that Amundsen with a Norwegian party had arrived about a month earlier. (Scott's party suffered extremely severe weather, and on the way back their supplies ran short. Two men had already died when Scott made his last camp—only 11 miles from a supply depot. The death of his party was not known until 1913, when the rest of his expedition reached New Zealand. All subsequent expeditions were equipped with wireless.)

A survey of noted Englishmen revealed this list of "The Ten Greatest Living Men": Edison, Kipling, Roosevelt (Theodore), Marconi, Lister, Chamberlain, Taft, Roberts, William II, Metchnikoff. (Also rans: Wilbur Wright, Andrew Carnegie, J. Pierpont Morgan, John D. Rockefeller, Puccini.)

The population of the world is estimated at 1,623,300,000.

National

New Mexico and Arizona are admitted to the Union as the 47th and 48th states.

April floods in the Mississippi valley leave 30,000 homeless and cause damage of $50,000,000.

About 50,000 students will graduate from college this year; this represents about 66 percent more students than graduated ten years ago. . . . College tuition costs range from about $40 to $250; Harvard, Yale, and Princeton charge about $150.

A teacher sees the public school as "the most momentous failure in our American life. Can you imagine a more grossly stupid, a more genuinely asinine system tenaciously persisted in to the fearful detriment of over 17,000,000 children and at a cost to you of over $403,000,000 each year—a system that not only is absolutely ineffective in its results, but also actually harmful in that it throws every year 93 out of 100 children into the world of action absolutely unfitted for even the simplest tasks in life?"

There have been fewer lynchings in 1912 than in any previous year since 1884—64 (vs 71 in 1911). But suicides have increased from 12,242 in 1911 to 12,981 in 1912.

R. E. Olds, noted automobile designer, calls his Reo the Fifth his Farewell Car because he believes that it represents the ultimate in car design. The future, he says, will bring minor changes, folderol and fashion, but a 35-horsepower 4-cylinder car will doubtless remain the standard because greater power is unnecessary and expensive. In the future, just this one car will be made.

In six months, gasoline prices have increased from 6¢ per gallon to 16¢ per gallon. . . . "What do you know about mortgages?" "I know all about mortgages. The first mortgage is for the car; the second mortgage is for the upkeep."

Headline on front page of "The New York Times," December 28, 1912: "THE STOCK EXCHANGE IS TO REFORM ITSELF"

The 54-hour-a-week law becomes effective in New York State. . . . The Government sues the Erie Railroad for keeping firemen on duty more than 16 hours a day.

As a result of the showing of American moving pictures, American styles of hats, shoes, clothing, and haircuts have become very popular in Belgrade. The American ambassador suggests that all kinds of American products, including farm machinery, should be promoted in this fashion.

Commissions delegated to investigate vice conditions in many large cities have uniformly found the public dancehall next to the saloon in the potency of its influence for evil.

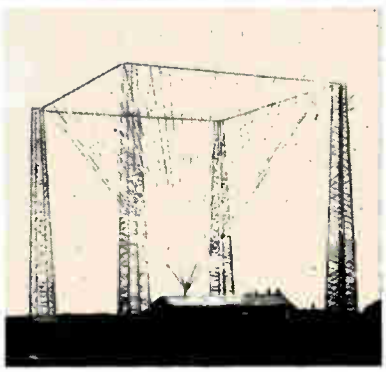
During 1912, Broadway showed 97 new plays, 36 new musical comedies, and numerous revivals. "Peg o' My Heart," "Little Women," and "Within the Law" were some of the plays: George M. Cohan, Weber and Fields, and Laurette Taylor were some of the performers. . . . The Metropolitan Opera stars include Tetrazzini, Mary Garden, Scotti and Caruso. Caruso earned about $90,000 in 1912 from "talking-machine records."

The Boston Red Sox win the World Series from the New York Giants. (Christy Mathewson was a Giant pitcher.)

The chestnut tree blight is spreading and affecting more and more trees. Scientists and tree experts have not been able to find any cure for it.

Wireless

In a new wireless telegraph system devised by Prof. Zehnder, no antenna is required.



Marconi Station at Glace Bay, Cape Breton Island

A law is passed (shortly after the Titanic disaster) requiring ships carrying more than 50 passengers to have wireless equipment capable of transmitting messages 100 miles or more, an auxiliary set with independent power supply, and at least two skilled operators.

A new law requires every station (operator) to obtain a license, to use special call letters, and to designate his normal sending and receiving wave lengths. (Fortescue, in a book written in 1912, describes the indiscriminate use of powerful transmitting apparatus by amateurs in the U.S.A. as resulting in chaos.)

Headline on front page of "The New York Times," February 1, 1912: "WIRELESS TO LINK ALL BRITISH EMPIRE" "Every part of the empire, on which it is the Briton's proud boast that the sun never sets, will be linked by an immaterial chain formed out of the encircled either. A new Puck will girdle the earth."

G. W. Pierce, a Harvard professor, demonstrates a "wireless telephone" with which he communicates between Gloucester and Cambridge, Mass.—a distance of about 35 miles.

The new wireless station of the Navy Department at Arlington, Va., which is one of the largest in the world, will send out time signals to all parts of the United States.

The Audion detector is coming into use, replacing crystals or the moving steel tape in a magnetic field which are often used on board ships.

Science and Industry

Headline in "The New York Times," February 4, 1912: "GREAT BIOLOGISTS UNITE TO SEEK A CANCER CURE" (Among them—Dr. Paul Ehrlich, Dr. A. von Wassermann, and Dr. Metchnikoff.)

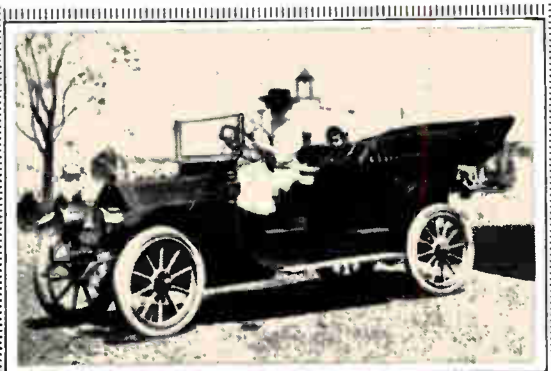
Headlines in "The New York Times," December 29, 1912: "DR. LEO LOEB MAKES BOLD EXPERIMENTS IN CANCER CASES. ST. LOUIS EXPERT ACHIEVES RESULTS FROM COLLOIDAL COPPER THAT ARE CONSIDERED REMARKABLE."

Dr. H. Glebeler reports "The discovery of radium, uranium and emanation radiations in the spectrum of Nova Geminorum."

Cutler-Hammer develops the "world's first practical, electrically-controlled, pushbutton-operated gear shift mechanism applicable to the sliding gear transmissions of the automobile." The picture shows Miss Claire Rochester in her Premier car, on which the C-H Magnetic Gear Shift is regular equipment. (Miss Rochester is the famous "Liberty Girl," who drove from coast to coast in 11 days, to procure funds to illuminate the Statue of Liberty.) Miss Rochester says she would not shift gears by hand any more.

To raise money for charity, Miss Rochester drove up the County Court House steps in San Antonio. Here is what she says about the C-H Magnetic Gear Shift: "For years I've shifted gears by hand. Now, I simply press a button on my Premier and electricity does the shifting for me. Today I wouldn't be bothered with the difficulty of shifting gears by hand. No other up-to-date motorist should and wouldn't if they saw the Magnetic Gear Shift of the Premier."

(In 1912, the 19-year-old Cutler-Hammer organization employed 250 people. The median age of the forty people who today make up AIL's senior staff was minus six years in 1912.)



Meetings with Exhibits

● As a service both to Members and the industry, we will endeavor to record in this column each month those meetings of IRE, its sections and professional groups, which include exhibits.

Δ

May 14-16, 1962

NAECON (National Aerospace Electronics Conference), Biltmore Hotel, Dayton, Ohio

Exhibits: Mr. Robert J. Stein, 19 Beverly Place, Dayton 19, Ohio

May 23-25, 1962

National Telemetering Conference, Sheraton Park Hotel, Washington, D.C.

Exhibits: Mr. L. H. King, Jansky & Bailey Co., Shirley Highway & Edsall Rd., Alexandria, Va.

May 24-26, 1962

IRE Seventh Region Conference on "Space Communications," Olympic Hotel, Seattle, Wash.

Exhibits: Mr. J. K. Schloss, 1520 N.E. 106th, Seattle 55, Wash.

June 12-14, 1962

Armed Forces Communications & Electronics Show, Sheraton Park Hotel, Washington, D.C.

Exhibits: Mr. William C. Copp, 72 W. 45th St., New York 36, N.Y.

June 18-19, 1962

Chicago Spring Conference on Broadcast and Television Receivers, O'Hare Inn, Chicago, Ill.

Exhibits: Mr. John H. Landeck, Admiral Corp., 3800 W. Cortland St., Chicago 47, Ill.

June 25-27, 1962

Sixth National Convention on Military Electronics, Shoreham Hotel, Washington, D.C.

Exhibits: Mr. L. D. Whitelock, 5614 Greentree Rd., Bethesda 14, Md.

June 25-27, 1962

Sixth National Conference on Product Engineering & Production, Los Angeles, Calif.

Exhibits: Mr. Lionel Crown, International Computer, 4301 Redwood Ave., Los Angeles 66, Calif.

August 21-24, 1962

Western Electronics Show and Conference (WESCON), Ambassador Hotel & Memorial Sports Arena, Los Angeles, Calif.

Exhibits: Mr. Don Larson, WESCON, 1435 La Cienega Blvd., Los Angeles, Calif.

(Continued on page 10A)

HYREL® ST Capacitors, developed and qualified for use in the Minuteman Missile, are NOW available to you in ALL RATINGS!

- Quality *100 times greater* than that of former high-reliability components! That's the ultra-high-reliability now demanded of electronic parts in the Minuteman missile's intricate guidance and control system.
- An unmatched test history of over 127 million unit-hours backs up the design of HYREL ST Capacitors to withstand the rigorous performance requirements specified for Minuteman components.
- The pioneer in solid tantalum capacitors, Sprague is one of 12 nationally-known manufacturers chosen to participate in the Air Force's Minuteman Component Development Program of Autonetics, a division of North American Aviation, Inc.
- All of the special processes and quality control procedures that make HYREL ST Capacitors the most reliable in the world can now help you in your military electronic circuitry. A tantalum capacitor engineer will be glad to discuss the application of these capacitors to your missile and space projects. Write to Mr. C. G. Killen, Vice-president, Industrial and Military Sales, Sprague Electric Company, 235 Marshall St., North Adams, Mass.

SPRAGUE COMPONENTS

CAPACITORS	INTERFERENCE FILTERS	HIGH TEMPERATURE MAGNET WIRE
RESISTORS	PULSE TRANSFORMERS	CERAMIC-BASE PRINTED NETWORKS
MAGNETIC COMPONENTS	PIEZOELECTRIC CERAMICS	PACKAGED COMPONENT ASSEMBLIES
TRANSISTORS	PULSE-FORMING NETWORKS	FUNCTIONAL DIGITAL CIRCUITS

SPRAGUE®
THE MARK OF RELIABILITY

Meetings with Exhibits

(Continued from page 8A)

August 29-September 1, 1962

Second International Conference on Information Processing, Munich, Germany

Exhibits: Dr. Heinz Billing, Max Planck Institute, Munich, Germany

September 3-7, 1962

National Advanced Technology Management Conference, Opera House, Seattle World's Fair Grounds, Seattle, Wash.

Exhibits: Mr. Hugh Fairclough, 1624 22nd Avenue East, Seattle 2, Wash.

October 1-3, 1962

Eighth National Communications Symposium, Hotel Utica & Utica Municipal Auditorium, Utica, N.Y.

Exhibits: Mr. Charles Glaviano, 45 Meadow Dr., Rome, N.Y.

October 2-4, 1962

Seventh National Symposium on Space Electronics & Telemetry, Fontainebleau Hotel, Miami, Fla.

Exhibits: Mr. Charles H. Doersam, Jr., Instruments for Industry, 101 New South Road, Hicksville, L.I., N.Y.

October 8-10, 1962

National Electronics Conference, McCormick Place, Chicago, Ill.

Exhibits: Mr. Rudy Napolitan, National Electronics Conference, 228 N. LaSalle St., Chicago, Ill.

October 15-18, 1962

Symposium on Space Phenomena & Measurement, Statler Hilton Hotel, Detroit, Mich.

Exhibits: Mr. J. B. Bullock, University of Michigan, Ann Arbor, Mich.

November 1-2, 1962

Sixth National Conference on Product Engineering and Production, Jack Tar Hotel, San Francisco, Calif.

Exhibits: Mr. W. Dale Fuller, Lockheed Missiles & Space Div., P.O. Box 504, Sunnyvale, Calif.

November 4-7, 1962

Fifteenth Annual Conference on Engineering in Biology and Medicine, Conrad Hilton Hotel, Chicago, Ill.

Exhibits: Professional Associates, Inc., 6520 Clayton Rd., Saint Louis, Mo.

Δ

Note on Professional Group Meetings: Some of the Professional Groups conduct meetings at which there are exhibits. Working committeemen on these groups are asked to send advance data to this column for publicity information. You may address these notices to the Advertising Department and of course listings are free to IRE Professional Groups.

HOW TO TORTURE A MICROWAVE COMPONENT

Metcom puts every one of the microwave components it manufactures through a very special kind of hell known technically as Dynamic Environmental Testing.

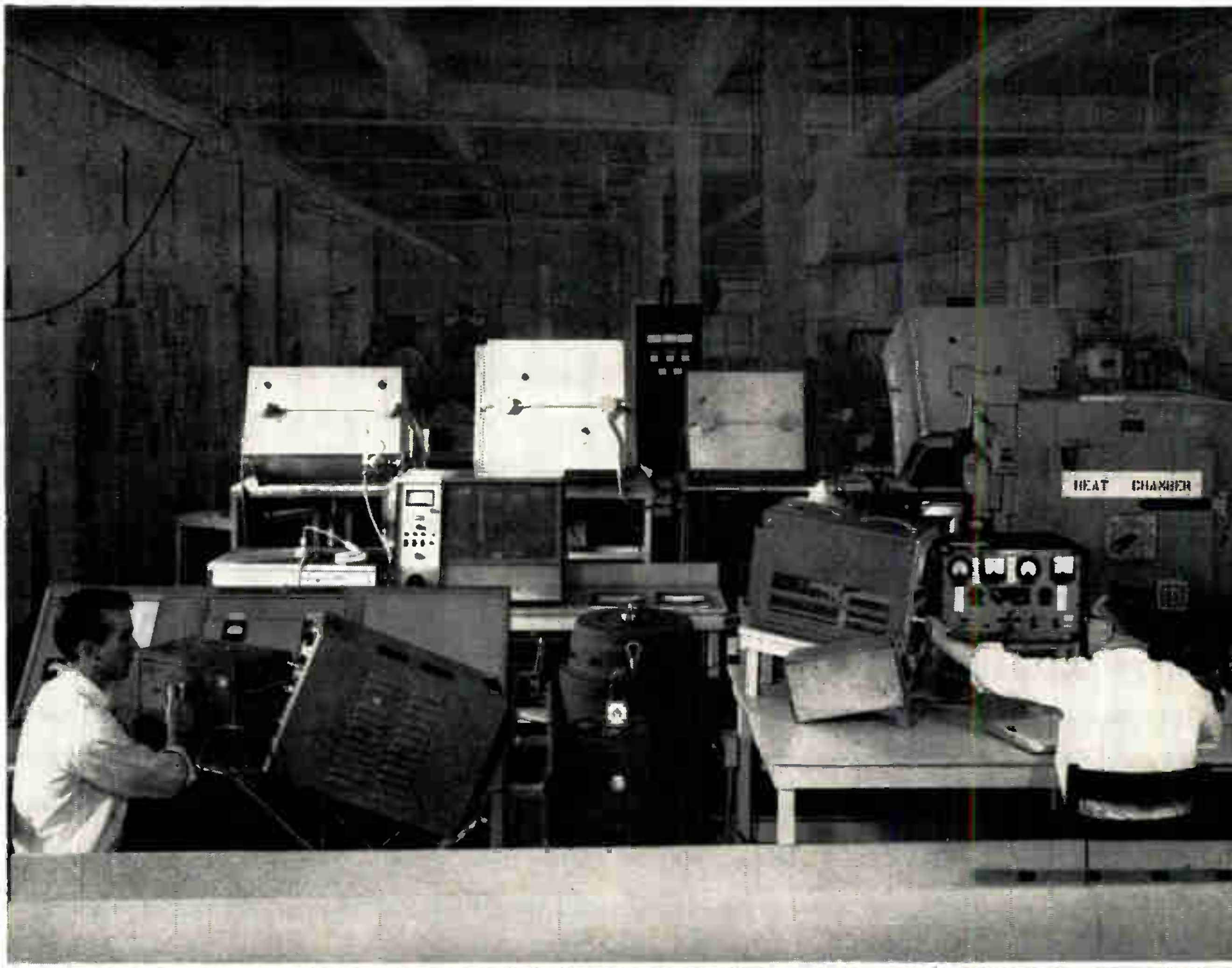
Here, for example, a beacon magnetron, hooked up to the barrel-shaped vibrator in the center of this picture, is about to receive, electronically, the shaking up of its life. When it has passed this particular test, it (and every other beacon magnetron with our name on it) will be placed in a vacuum, and alternately frozen and cooked at temperatures ranging from —54°C to 125°C. And that's not all. Some from each lot will be sterilized — some will be tested to ultimate destruction.

No part of our job is more important than subjecting our own products to this uniquely thorough dynamic environmental testing — to this very special kind of hell. Extra-rigorous testing is the best guarantee there is of extra-reliable performance.

Whatever your requirements in microwave components, *make sure they're by Metcom . . . because Metcom makes sure.*

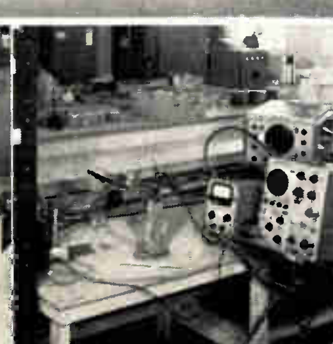
Write for our folder containing brief specifications on over 300 microwave tubes and devices.

for better microwave tubes and devices . . . . METCOM INC.
SALEM, MASSACHUSETTS



Anode assembly

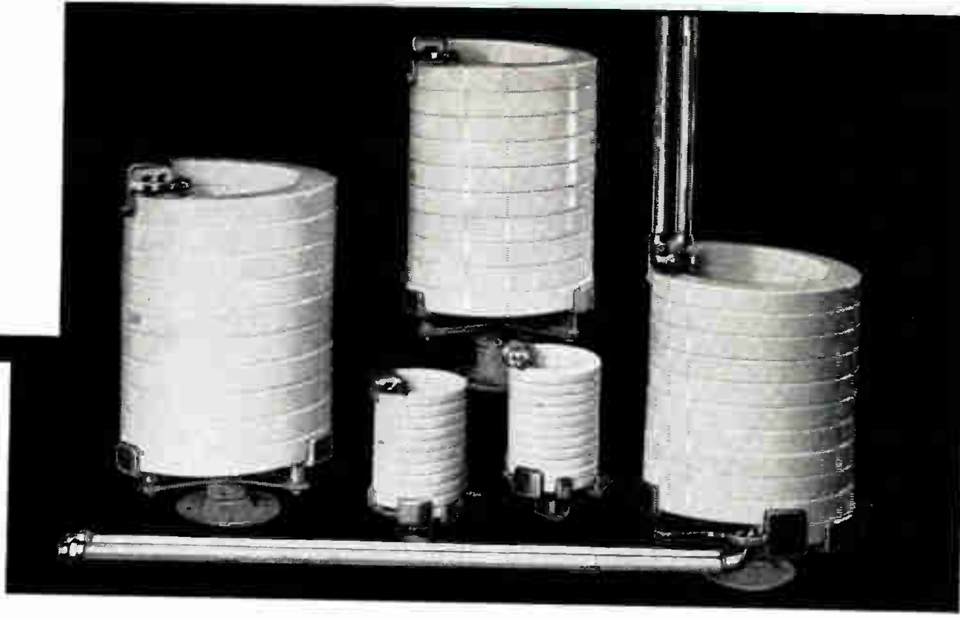
Magnetron aging



Magnetron life test

Magnetron acceptance test

Lapp

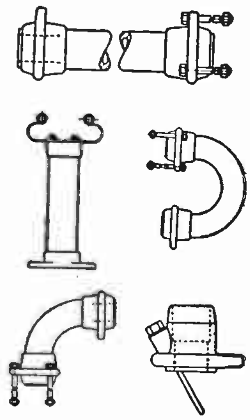


TUBE SUPPORTS
for air-cooled tubes

No longer is it necessary for equipment manufacturers to design supports for forced-air-cooled tubes. After producing many such special designs, Lapp has engineered and built a complete series of insulating supports for air-cooled tubes. Use of these supports facilitates design of transmitter and other high-power circuits. Simple, compact, efficient and attractive in appearance, in porcelain or steatite, they make for economical production, easy interchangeability, and availability of replacement units. Write for Bulletin 301, with complete description and specification data.

INSULATION FOR WATER-COOLED SYSTEMS

For carrying cooling water which must undergo a change in potential, use of Lapp porcelain eliminates troubles arising from water contamination and conductivity, sludging and electrolytic attack of fittings. Permanent cleanness and high resistance of cooling water is assured—for positive cooling and long tube life.



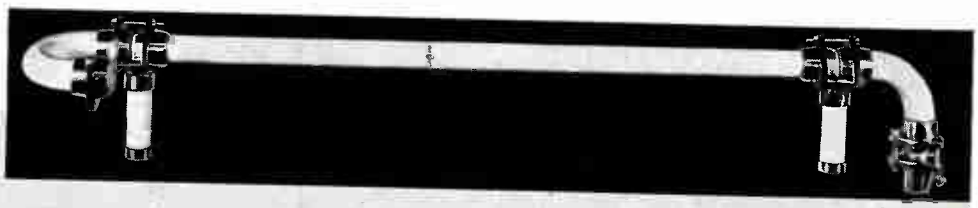
LAPP PORCELAIN WATER COILS

Twin-hole and single-hole models in sizes to provide flow of cooling water from 2 to 90 gallons per minute. Cast aluminum mounting bases, lead pipe or flexible metal hose for attachment.

LAPP PORCELAIN PIPE

Inside diameters ¾″ to 3″, in straight, 90° and 180° elbows, fittings. Swivel-type connections. Standoff insulators attach directly to fitting bolts.

Write for Bulletin 301 with specifications for Lapp Porcelain Water Coils and Porcelain Pipe.



LAPP GAS-FILLED CONDENSERS
Capacitance for high-voltage, high-current duty

The record of more than 18 years' service, in thousands of installations, proves the reliability of the Lapp Gas-Filled Condenser. For duty at high voltages and high currents, this unit offers the advantages of extreme compactness . . . low loss . . . high safety factor . . . constant capacitance under temperature variation . . . grounded tuning shaft . . . and rugged sturdiness. Units available in four tank diameters, 7″ to 18″ . . . in fixed or variable capacitance . . . capacitances to 30,000 mmf; current ratings to 390 amps at 1 mc; voltages to 85 Kv peak. Write for Bulletin 302.

ENTRANCE INSULATORS

To eliminate losses which occur in ordinary types of bushings at radio frequency, Lapp units use air as major insulation, porcelain path being lengthened by its "bowl" shape. In wide variety of types for most electrical and mechanical requirements.

STAND-OFF INSULATORS

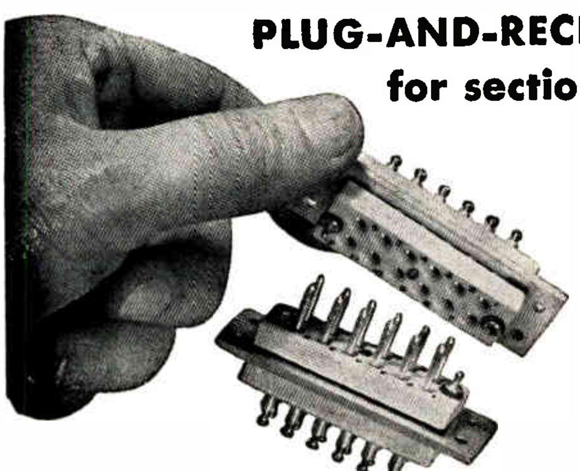
Dependable mechanical and electrical performance—and trim good looks. Hundreds of standard and special types available.

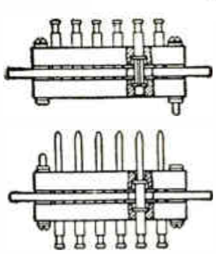
Write for Bulletin 301, listing entrance and stand-off insulators.

ANTENNA STRAIN INSULATORS

The largest of the porcelain rod insulators shown develops 12,000 lb. strength. Available, as specified, with rain shield and/or corona rings. Smaller units, in porcelain or steatite, for all strain or spreader use. Engineering and production facilities also available for special performance units See catalog 301.

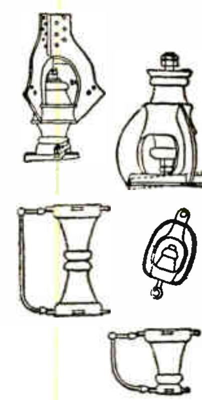
PLUG-AND-RECEPTACLE UNITS for sectionalizing circuits





For panel-rack or other sectionalized circuits, Lapp offers a variety of plug-and-receptacle units. Any number of contacts can be provided (in multiples of twelve). Male and female contacts are full floating, for easy alignment and positive contact. Contacts silver plated, terminals tinned for soldering. Insulation is steatite, the low-loss ceramic . . . non-carbonizing even under leakage flashover resulting from contamination, moisture or humidity. Write for complete electrical and mechanical specifications and recommendations.

ANTENNA TOWER INSULATORS



Most of the world's radio towers are supported by Lapp insulators, including the tallest guyed mast ever built. Through the Lapp "compression cone," immense loads can be carried by electrical porcelain. Write for data on Lapp footing insulators for self-supporting radiators, mast-base insulators, or guy insulators.

SPECIAL ASSEMBLIES

Lapp business is electrical porcelain, its application, design, manufacture, and assembly. Our skill is in our knowledge of the capabilities, and limitations, of ceramic insulation . . . in engineering ingenuity to meet specified requirements . . . and in efficient production. If you have need for insulating parts and associated sub-assemblies, we may be able to show you how they can be made most economically, to perform most efficiently.

Lapp

LAPP INSULATOR CO., INC.
RADIO SPECIALTIES DIVISION
150 SUMNER ST., LE ROY, N. Y.

Current IRE Statistics

(As of December 31, 1961)

Membership—98,084
Sections—111
Subsections—33
Professional Groups—29
Professional Group Chapters—295
Student Branches†—224

† See June, 1961, issue for a list.

Calendar of Coming Events and Author's Deadlines*

1962

May 1–3: Spring Joint Computer Conference, Fairmont Hotel, San Francisco, Calif.
May 3–4: 1st International Congress on Human Factors in Electronics, Lafayette Hotel, Long Beach, Calif.
May 4–5: Bay Area Symp. on Reliability and Quality Control, U. S. Naval Postgraduate School, Monterey, Calif.
May 8–10: Electronic Components Conf. Marriott Twin Bridges Hotel, Washington, D. C.
May 14–16: MAECON, Dayton, Ohio.
May 14–16: Symp. on Thermionic Power Conversion, Antlers Hotel, Colorado Springs, Colo.
May 14–18: 156th Annual Convention of the Medical Society of the State of New York.
May 18: Space Medicine Spectra Anno V of Space Exploration, Statler Hilton Hotel, New York, N. Y.
May 22–24: Nat'l Microwave Theory and Techniques Symp., Boulder, Colo.
May 22–24: Conf. on Self-Organizing Principles, Museum of Science and Industry, Chicago, Ill.
May 23–25: Nat'l. Telemetering Conf., Sheraton Park Hotel, Washington, D. C.
May 24–26: IRE 7th Region Conf. on "Space Communications," Olympic Hotel, Seattle, Wash.
June 18–19: Chicago Spring Conf. on Broadcast and Television Receivers, O'Hare Inn, Chicago, Ill.
June 25–27: 6th Nat'l Conf. on Product Engrg. and Production, Los Angeles, Calif.
June 25–27: 6th Nat'l Convention on Military Electronics, Shoreham Hotel, Washington, D. C.
June 25–30: Symp. on Electromagnetic Theory and Antennas, Tech. Univ. of Denmark, Copenhagen.
June 27–29: Joint Automatic Control Conf., New York Univ., New York, N. Y.
June 27–28: 9th Annual Symp. on Computers and Data Processings, Elkhorn Lodges, Estes Park, Colo.

* DL = Deadline for submitting abstracts.

(Continued on page 15A)

NBS Offers Radio Propagation Course

The National Bureau of Standards will present a three-week course in Radio Propagation from July 16–August 3, 1962, at the Boulder Laboratories, Boulder, Colo.

The course, sponsored by the NBS Central Radio Propagation Laboratory in association with the University of Colorado, is designed to give scientists and engineers from universities, industry, and government agencies access to the latest advances in radio propagation research and show how this knowledge can best be applied to the design and development of communication systems. It will consider communication via the entire range of useable radio frequencies in the atmosphere, space, underground, and underwater, and will extend into the modes of propagation which are being explored for the future.

The Central NBS Radio Propagation Laboratory in Boulder conducts continuous studies designed to develop new methods of communication and more efficient use of the methods which now exist. This Laboratory is the central information agency of the government on the propagation of radio waves at all frequencies along the surface of the earth, in the atmosphere, and in space. The course is being offered as part of the Laboratory's responsibility to help industry and government design the most efficient communication systems possible.

To facilitate transfer of academic credit for students working on graduate degrees in areas related to radio propagation, the course will be included in the Graduate program at the University of Colorado. Such students must plan an additional half day at the end of the course to write a final examination.

The 1962 course will employ a unified approach to both ionospheric and tropospheric propagation stressing their similarities and differences. Increased emphasis will be given to both fundamental physics and to systems applications. Lectures will be supplemented by regularly scheduled informal discussions conducted by all lecturers within a particular topic group.

Prerequisites for the course are a Bachelor's degree in electrical engineering, physics, or other suitable academic or practical experience. A tuition is $300 for the entire course running from July 16 through August 3. A small additional University registration fee will be required of students who also desire academic credit.

Registration will be limited and early application should be made to ensure consideration. To facilitate local arrangements, registration will be closed July 1, 1962. Further details of the course and registration forms will be available March 1, 1962, from: Edmund H. Brown, Education Director, Boulder Laboratories, National Bureau of Standards, Boulder, Colo.

PGVC Presents Chapter of the Year Award

The Institute of Radio Engineers Professional Group on Vehicular Communications presented its Chapter of the Year Award for 1960–1961 to the Chicago Chapter. The award was presented at the 12th National Conference of the Professional Group Vehicular Communications held in Minneapolis, Minn., on November 30 and December 1, 1961. J. E. Farley, presently Chairman of the Chicago Chapter, accepted the award for Chicago.

This Chapter organized an active professional program covering the fast moving developments in the mobile radio field last year. Included in their program were papers and meetings on Satellite Communications, Airlines Radio, and a field trip of the new communication facilities at O'Hare Airport, Chicago, Ill.

1960–1961 PGVC Chicago Chapter Officers: *Left to right:* G. P. Schleicher (Membership), Jack Germain (Chairman), J. E. Farley (Vice-Chairman), F. Hilton (Secretary-Treasurer), B. D. Wickline (Arrangements).

Captain Kenneth M. Gentry, USN, (*left*) receiving his decoration "Cavialeri Ufficiale" of the Order "Al Merito della Repubblica Italiana," from Admiral Pecori Geraldi, Chief of Italian Naval Staff.

Captain Kenneth M. Gentry Receives Decoration

Captain Kenneth M. Gentry, USN, (Retired) has been decorated by the Italian government for his work in establishing a Supreme Allied Commander Atlantic anti-submarine warfare research center in Italy. The decoration, called "Cavialeri Ufficiale" of the Order "Al Merito della Repubblica Italiana," was presented by Chief of Italian Naval Staff, Admiral Pecori Geraldi, in Rome.

Captain Gentry served as Subcommittee Chairman of the Naval Research Advisory Committee which visited NATO countries to investigate organizing the NATO research laboratory. The recommendations of this subcommittee and the subsequent negotiations between nine NATO nations resulted in setting up the center in LaSpezia, Italy. Captain Gentry served as its first Military Director, and for his work as director was awarded a Letter of Commendation from SACLANT, and a Letter of Appreciation from the Scientific Advisory Council of the nine NATO nations.

At present, Captain Gentry is Associate Director of Research and Development for Undersea Warfare and Electronic Countermeasures for the Motorola Military Electronics Division, Chicago, Ill.

IEE-IRE Cooperate In London TV Conference

The International Television Conference will be held in London, England, May 31 to June 7, 1962. The Conference is organized by the Electronics and Communications Section of The Institution of Electrical Engineers in association with the Television Society, the British Kinematograph Society and the IRE. The association of the IRE in the Conference will provide an important example of Anglo-American cooperation in the professional electronic field, which, it is hoped, will set a pattern for the future.

The Conference will overlap the last three days of the International Instruments, Electronics and Automation Exhibition at Olympia, London, and arrangements are being made for those attending the Conference to attend the exhibition. The technical program will also include visits to various laboratories and industrial organizations.

The Chairman of the Conference Organizing Committee is Robert C. G. Williams who is also Chairman of the IRE Advisory Committee in the United Kingdom. Information and registration forms may be obtained by writing the Secretary, The Institution of Electrical Engineers, Savoy Place, London, W.C.2, England.

The provisional sessions of the Conference are as follows:

Thursday, May 31—Opening Session.

Friday, June 1 (morning)—Systems Standards; (afternoon)—Signal Pick-Up Tubes.

Monday, June 4 (morning)—Frequency-Assignment Problems, with a parallel session on Television Recording—Programme and Commercial; (afternoon)—Standards Conversion, with a parallel session on Medical Uses.

Tuesday, June 5 (morning)—Point-to-Point Links, with a parallel session on Outside-Broadcast Equipment; (afternoon)—Studio Design, with a parallel session on both Scientific Applications and Data Transmission.

Wednesday, June 6 (morning)—Wire Broadcasting for Domestic Television, with parallel sessions on Studio Design and Lighting, and Industrial Applications; (afternoon)—Broadcast Transmitter Equipment, with a parallel session on Low-Light-Level Applications.

Thursday, June 7 (morning)—Domestic Receiver Equipment, with a parallel session on Space Applications; (afternoon)—Color Television.

Calendar of Coming Events and Author's Deadlines*

(*Continued from page 14A*)

June 28–29: 4th Nat'l Symp. on Radio Frequency Interference, Town House Hotel, San Francisco, Calif.

July 17–18: Data Acquisition and Processing in Medicine and Biology, Whipple Aud., Strong Mem. Hosp., Rochester, N. Y.

Aug. 14–16: Internat'l Conf. on Precision Electromagnetic Measurements (former title: Conf. on Standards and Electronic Measurements), Nat'l. Bur. Standards, Boulder, Colo.

Aug. 21–24: WSCON (Western Electronics Show and Conf.) Los Angeles, Calif.

Aug. 27–Sept. 1: 2nd Internat'l Federation of Information Processing Societies Congress, Munich, Germany.

Sept. 3–7: Internat'l Symp. on Information Theory, Brussels, Belgium

Sept. 3–7: Nat'l Advanced Technology, Management Conf., Seattle, Wash.

Sept. 3–7: 4th Internat'l Congress on Microwave Tubes, Kurhaus Hotel, Scheveningen, Netherlands

Sept. 4–7: ACM Nat'l Conf., Hotel Syracuse, Syracuse, N. Y.

Sept. 13–14: 10th Annual Engineering Management Conf., Hotel Roosevelt, New Orleans, La.

Sept. 13–14: Nat'l Symp. on Engineering Writing and Speech, Mayflower Hotel, Washington, D. C.

Sept. 19–20: 11th Annual Industrial Electronics Symp., Hotel Sheraton, Chicago, Ill.

Sept. 28–29: 12th Annual Broadcast Symp., Willard Hotel, Washington, D. C.

Oct. 1–3: 8th Nat'l Communications Symposium, Utica, N. Y. (DL*: June 1, 1962, G. Baldwin, Paris Rd.-R.D. 2, Clinton, N. Y.)

Oct. 7–12: AIEE 1962 Fall General Meeting—3rd Ann. Symp. on Switching, Circuit Theory and Logical Design, Chicago, Ill. (DL*: Apr. 1, 1962, W. Semon, Sperry Rand Res. Ctr., North Rd., Sudbury, Mass.)

Oct. 15–18: Symposium on Space Phenomena and Measurement, Statler Hilton Hotel, Detroit, Mich.

Oct. 22–24: ECCANE, Emerson Hotel, Baltimore, Md. (DL*: June 4, 1962, W. C. Vergara, Dept. 466-2, Bendix Radio, Towson 4, Md.)

Oct. 30–31: Spaceborne Computer Engineering Conf., Disneyland Hotel, Anaheim, Calif. (DL*: June 15, 1962).

Nov. 1–2: Sixth National Conf. on Product Engineering and Production, Jack Tar Hotel, San Francisco, Calif. (DL*: June 11, 1962, G. F. Reyling, Varian Associates, 611 Hansen Way, Palo Alto, Calif.)

Nov. 4–7: 15th Annual Conference on Engineering in Biology and Medicine, Conrad Hilton Hotel, Chicago, Ill. (DL*: June 1, 1962, D. A. Holaday, Univ. of Chicago, 950 E. 59 St., Chicago, Ill.)

Nov. 5–7: Northeast Res. and Engineering Meeting, Boston, Mass.

Nov. 12–14: Radio Fall Meeting, King Edward Hotel, Toronto, Ont., Canada.

* DL = Deadline for submitting abstracts.

Tudor R. Finch (*right*), chairman of the 1961 International Solid-State Circuits Conference, presenting one of the annual conference awards to Hewitt D. Crane (*left*) for his outstanding paper on the Neuristor delivered during the 1961 meeting.

Solid-State Circuits Conference Digests Offered

Digests of Technical Papers for the 1962 International Solid-State Circuits Conference held at the University of Pennsylvania and the Sheraton Hotel, Philadelphia, Pa., February 14, are being offered. The digest (112 pages) includes illustrated condensations of all of the conference papers.

Copies are now available at $5.00 each and can be obtained from: H. G. Sparks, The Moore School of Electrical Engineering, University of Pennsylvania, 200 S. 33rd Street, Philadelphia, Pa. Remittance (payable in U. S. currency) should be made payable to the order of: Solid-State Circuits Conference.

University of California Presents Summer Course

The seventh annual summer Reliability and Statistical Methods in Industry Course will be presented at the University of California, Los Angeles, Calif., July 30, through August 24, 1962. The course will be jointly sponsored by the Dept. of Engineering and University Extension in cooperation with the Los Angeles Sections of the American Society for Quality Control, American Society of Mechanical Engineers, American Institute of Industrial Engineers, American Institute of Electrical Engineers, the Small Business Administration and the Society for Non-Destructive Testing.

The program will be divided into two sessions, each consisting of four courses of two weeks each. The courses are planned for key people whose responsibilities are concerned with improved production including inspectors, managers, engineers and others involved in research and development.

Staff for this program is specially selected from all parts of the United States and includes recognized experts in their respective fields.

For a brochure of this program giving full details of the above courses and other general lectures and special features, contact: Edward P. Coleman, Professor of Engineering, University of California, Los Angeles 24, Calif.

New Geoscience PG Announces Officers

At the first organizational meeting of IRE's newest Professional Group, No. G-29 on Geoscience Electronics, nine members of an anticipated 15-member administrative committee were elected. The meeting was held in Dallas, Tex., on February 15, 1962, under the guidance of Chairman pro-tem Robert W. Olson, Texas Instruments, Dallas. The committee members are, in addition to Chairman Olson (F): M. A. Arthur (SM), Humble Oil, Houston, Tex.; W. T. Born (F), Geophysical Research, Tulsa, Okla.; R. A. Broding (SM), Century Geophysical, Tulsa, Okla.; Sidney Kaufman (SM), Shell Development, Houston, Tex.; B. H. List (M), Texas Instruments, Dallas, Tex.; W. Harry Mayne (SM), Petty Geophysical, San Antonio, Tex.; F. C. Smith, Jr. (SM), Dannemiller-Smith, Houston, Tex.; H. W. Smith (SM), University of Texas, Austin, Tex.

The remaining six positions on the administrative committee will be filled at its next meeting to be held during the 1962 SWIRECO Convention, April 11–13, in Houston, Tex. It is planned for these to represent other geographical areas and different branches of the geosciences from those listed above. Interested members of the IRE are invited to attend.

Space Symposium Call for Papers

A Symposium on Space Phenomena and Measurement will be held in Detroit, Mich., October 15–18, 1962. Co-sponsors for the Symposium are the IRE Professional Group on Nuclear Science, the Atomic Energy Commission, and the National Aeronautics and Space Administration.

Space environments and measurements will be emphasized during the three-day meeting but, in addition, a parallel session will be presented on selected topics on space nuclear propulsion, energy conversion, nuclear instrumentation and radiation effects. Typical of the space exploration topics to be included are: Measurement in the Radiation Belts, Solar and Galactic Cosmic Ray Behavior, Electromagnetic Fields in the Solar System, and X-Ray and Gamma Ray Astronomy.

Papers are being accepted for both sessions of this meeting. One hundred word abstracts should be submitted before July 1, 1962 to: Mr. Michael Ihnat, AVCO Corp., 201 Lowell St., Wilmington, Mass. A rough draft of the paper should be submitted to the same address prior to September 1, 1962.

Further details relating to the Symposium can be obtained by writing: Dr. Harold E. DeBolt, AVCO Corp., 201 Lowell St., Wilmington, Mass.

Industrial Engineering Seminars Held at Cornell

Cornell University, Ithaca, N. Y., will hold Industrial Engineering Seminars from June 12 through June 15, 1962. The Seminars are being sponsored by the Dept. of Industrial and Engineering Administration, Cornell University.

Participants enroll in one of the following eleven groups and also attend general sessions: Industrial Management, Engineering Administration, Capital Investment Planning (Theory and Practice), Operations Management of the Smaller Company, Work Measurement, Applied Operations Research, Systems Simulation Using Digital Computers, Techniques of Mathematical Programming, Queuing and Inventory Theory, Statistical Decision Making (Theory and Practice), and Statistical Reliability Analysis (Theory and Practice).

The discussion leaders and speakers will be specialists from both industry and the staff of Cornell University.

The Seminars are offered to operating management personnel in line supervision and staff positions in industrial engineering, production engineering, engineering administration, operations research, research and development, quality control, production and inventory control, cost control, materials management, warehousing, purchasing, data processing, and the controller's function.

For additional information write: J. W. Gavett, Seminars Coordinator, Dept. of Industrial and Engineering Administration, Upson Hall, Cornell University, Ithaca, N. Y.

PANORAMIC "the pioneer is the leader"

dependable CERTIFIED SPECIFICATIONS for accurate data

SUMMARY OF SPECIFICATIONS

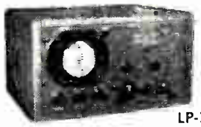
SPECTRUM ANALYZERS

½ cps to 44 kmc

BROADBAND SPECTRUM ANALYZERS 0.5 cps to 25 mc. Easy-to-operate direct reading analyzers display plot of signal amplitude vs. frequency on 5" CRT or chart. Features include: "quick look" wideband scans plus highly selective narrow scans. Center frequency, sweep width, and selectivity (resolution) are adjustable. Calibrated linear and 40-db log scales

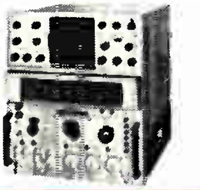
All these broadband spectrum analyzers are available with accessories and in systems (data on SY-systems available) which materially improve specifications and application versatility.

*This specification improved or adjustable with optional accessories.



LP-1a

Frequency Range	Model No.	Sweep Width	Resolution (I-F Bandwidth)	Sensitivity Range (Full Scale Linear)	Sweep Rate	Harmonic Products (Dynamic Range)
0.5 cps to 2500 cps	LF-2b (incl. chart recorder. 3" CRT optional)	Preset 2-500 cps on 2 ranges	0.1-20 cps	10 mv to 100 v	30 sec- 16 min/scan	—40 db
20 cps* to 22.5 kc.	LP-1a (5" CRT)	40 cps—20 kc log scan. Linear* 200, 1000,5000 cps	Automatic optimum. Best resolution 25 cps*	500 µv to 500 v	*1 cps	—60 db
200 cps to 300 kc	SB-7bZ (5" CRT)	0-200 kc	100 cps—2 kc	250 µv to 25 v	*6.7 cps.	—50 db
100 cps to 600 kc	SB-15a (5" CRT)	1 kc—200 kc	100 cps—4 kc + automatic optimum	200 µv to 200 v	*1-60 cps	—60 db
1 kc to 15 mc	SPA-3 (5" CRT)	0-3 mc	200 cps—30 kc	20 µv to 2 v	1-60 cps	—46 db
1 kc to 25 mc	SPA-3/25 (5" CRT)	0-3 mc	200 cps—30 kc	20 µv to 2 v	1-60 cps	—46 db



SPA-4a

MICROWAVE SPECTRUM ANALYZERS 10 mc to 44 kmc. Feature unsurpassed usable sensitivity and freedom from spurious effects for pulsed AM, FM and CW signals. Direct reading frequency dial ± 1%. They have adjustable dispersion, I-F bandwidth and adjustable sweep rate; 1-60 cps, free running or synchronized; 3 calibrated amplitude scales, linear, 40 db log, and power; calibrated frequency markers. Model SPA-1 has 2 interchangeable tuning heads; SPA-4a covers band to 44 kmc with 1 head.

Frequency Range	Model No.	Sweep Width (Dispersion)	Resolution (I-F Bandwidth)	Sensitivity (Signal = 2x Noise)	Other Features
10 mc to 44,000 mc 1 Tuning Head	SPA-4a (5" CRT)	2 swept oscillators 0 to 70 mc 0 to 5 mc	1 kc to 80 kc	L band —100 to —110 dbm.; S band — 90 to —100 dbm.; X band — 90 to —100 dbm.; K band — 70 to — 90 dbm.	All mixers built in • Dynamic range to —60 db. • Low residual FM • Synchroscope Output • Marker Modulation
50 mc to 4000 mc; (10 mc to 4000 mc. optional)	SPA-1 (5" CRT)	0 to 10 mc	9 kc to 80 kc	(With opt. internal preamp.) RF-2 head, 50-250 mc —100 dbm RF-3 head, 220-4000 mc 500 mc —100 dbm 1500 mc — 94 dbm 3000 mc — 76 dbm	Extremely low cost • Exceptionally stable • Available in 4 styles • Highly sensitive I-F preamp, optional.

FREQUENCY RESPONSE PLOTTERS

½ cps to 15 mc

COMPANION SWEEP GENERATORS FOR BROADBAND SPECTRUM ANALYZERS: Response to fundamental frequency only; discriminate against noise and hum; virtually unlimited dynamic range; single line plots. Accessories available for comparison testing, triangular scanning, and manual tuning.

Frequency Range	Sweeper Model No.	Used With Analyzer Model	Output Voltage (& Impedance)	Output Attenuator	Overall Flatness
0.5 cps to 2500 cps	G-5	LF-2a	0-5V (3000Ω)	0-60 db	±0.5 db (1-2500Ω)
20 cps to 22,500 cps	G-2a	LP-1a	50 mv-5v (1000,600, 3KΩ)	0-100 db	±1 db
200 cps to 300 kc	G-3a	SB-7bZ	250 µv to 2.5 v (600Ω)	0-80 db	±0.5 db (1-300 kc)
100 cps to 600 kc	G-15a	SB-15a	250 µv to 2.5 v (600Ω)	0-80 db	±1 db (0.2-600 kc)
1 kc to 15 mc	G-6	SPA-3 SPA-3/25	200 µv to 0.2 v (72Ω)	0-60 db	±2 db (1 kc-13.5 mc)

G-3a

SWEEP GENERATOR 20 cps-200 kc used with ext. scope (CRT scale furnished). Model SG1. Pre-set broad-band, log and adjustable linear scans. Built-in markers. SG-1R same as SG-1, plus synch pulse for testing tape recorder frequency response.

Frequency Range	Model	Scan Modes	Output Voltage (& Impedance)	Flatness	Sweep Rate
20 cps-200 kc	SG-1 SG-1R	40 cps to 20 kc log scan 400 cps to 200 kc log scan 0—200 kc lin. sweep widths	0-1 volt (600Ω)	±½ db	1 cps. Adjustable with TW-1a

COMMUNICATIONS SYSTEMS ANALYZERS

SSB, AM, FM

PANADAPTORS. For use with communications receiver with I-F = Panadaptor center frequency. Panadaptor response shaped to match receiver. Specify receiver model no. and I-F.

SUMMARY OF SERIES CAPABILITY

For Receivers With Intermediate Frequency of	Model Series	Sweep Width	Resolution	Sensitivity (At I-F)	Sweep Rate	Amplitude Scales	CRT Size
450 kc-30 mc	SA-8b (4 types)	0-100 kc to 0-10 mc	50 cps to 80 kc	150-2000 mv full scale	1-60 cps	Linear, 40 db log Power	5" (camera mount style optional)
450 kc-30 mc	SA-3 (10 types)	0-50 kc to 0-6 mc	2 kc to 50 kc	10 µv- 10 mv ¼" deflection	30 cps	linear (nominal)	3 inch

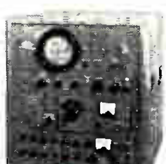
PANALYZORS scan through an adjustable sweep width about their center frequencies. An external VFO* is used for conversion of signals in Panalyzor input mixer. Mixer range up to 1000 mc.

*eg. Model RF-7a Panoramic VFO, 2-40 mc.

SUMMARY OF SERIES CAPABILITY

Model Series (Input Center Frequency)	Sweep Width	Resolution	Signal Sensitivity (VFO 0.1 volt rms)	Sweep Rate	Amplitude Scales
(†) SB-12b, 500 kc (or 455 kc)	0 to 100 kc + preset modes	10 cps-2.5 kc (see also SSB-3b)	2 mv full scale log	0.1-30 cps	linear 40 db log
SB-8b 3 types (500 kc-30 mc)	0 to 200 kc to 0 to 10 mc	50 cps-80 kc	10-100 mv full scale log	1-60 cps	linear 40 db log power
SB-3 5 types (500 kc-30 mc)	0 to 50 kc to 0 to 6 mc	2 kc-40 kc	1-10 mv, ¼" deflection	30 cps	linear (nominal)

(†) SB-12b also part of SSB-3b, (below)



SSB-3b

SINGLE SIDEBAND SPECTRUM ANALYZER—Model SSB-3b is a complete, compact unit for communications systems analysis. Advantages are 60 db dynamic range (at least 65 db upon request), selectivity for hum sidebands resolved down 60 db, and simple operation with preset modes and self-checking facilities.

SSB-3b Frequency Range	SSB-3B INCLUDES THESE COMPONENT INSTRUMENTS. Panalyzor Model SB-12b(s)	Model RF-7a Tuning Head	2 tone AF Gen. Model TTG-2	Resolution (I-F Bandwidth)	Sensitivity
2 mc to 40 mc (100 cps to 40 mc with Model REC-1 optional)	see SB-12b specs. above	2-40 mc Fast search + precise vernier tuning 1% accuracy	100 cps-10 kc (±2%) 0-4 volts rms metered outputs	(Same as SB-12b Minimum—60 db bandwidths = 50 cps)	2 mv full scale log, Virtually uniform from 100 cps-40 mc. 0-65 db input attenuator

TELEMETRY TEST INSTRUMENTS

(for IRIG FM/FM Systems)

TMI-1b

Model No.	Name	Features	Uses
TMI-1b (TMI-1b/120)	Telemetering Indicator	Analyzes 350 cps-85 kc band (to 120 kc optional) • Both Log & Linear Scans • Automatic optimum resolution • Internal markers	Subcarrier spectrum analysis • pre-emphasis • distortion measurements to —55 db • spillover • noise
TMC-1a	3 Point Calibration and Subcarrier Deviation Indicator	Xtal controlled ± 0.02%; + 7.5%, 0, —7.5%; also A to E • Special deviations available • Provides markers for uni-channel scans for TMI-1b analyzer (not used with TMI-1b/120)	Discriminator calibration ±0.02% With TMI-1b, monitors individual channels or VCO deviation linearity
TMC-411E	Simultaneous 11 Pt. Calibrator (also 3 pt. and 5 pt. optional)	Accuracy ± 0.002% • All ± 7½% channels + A to E • 18 channel outputs simultaneously or individually • auto./manual sequencing • special provisions available • all electronic, 7" high • distortion —40 db	Complete discriminator checkout in seconds! Multiple frequency reference source for data reduction and system calibration. With TMI-1b for VCO checks.

Specification subject to change without notice

Write for detailed specifications and catalog.

PANORAMIC ELECTRONICS, Inc.

FORMERLY PANORAMIC RADIO PRODUCTS, Inc.

522 So. Fulton Avenue, Mount Vernon, N. Y. • Phone: Area 914 OWens 9-4600

TWX: MT-V-NY-5229 • Cables: Panoramic, Mount Vernon, N. Y. State

Panoramic electronics, inc.

Joint Conference On Automatic Control

The University of Texas, Austin, Tex. has been selected as the location of the fourth Joint Automatic Control Conference, June 19–21, 1963. The first invitation of papers will be made in February 1962, with abstracts due at the end of September and manuscripts by November 1. Such timing, an innovation with the 1962 JACC, will permit the mailing of complete preprints to pre-registrants as much as a month before the conference.

Since the 1963 congress of the International Federation on Automatic Control follows the JACC by only three months and should overlap it only slightly in attendance, authors of IFAC papers are being urged to submit them also for JACC presentation. Such dual submission is approved by the American Automatic Control Council, and is in fact encouraged for purposes of practice and criticism. Because of IFAC copyright restrictions, such papers will be preprinted in abstract only.

Of the five member societies of the sponsoring council, the American Institute of Chemical Engineers has prime responsibility for the 1963 JACC. Dr. Theodore J. Williams, of Monsanto Chemical Co., St. Louis, Mo., is General Chairman; Dr. Otis L. Updike, University of Virginia, Charlottesville, Va., is Technical Program Chairman; and Dr. William A. Cunningham, University of Texas, Austin, Tex., is Local Arrangements Chairman. Papers on control theory, applications, or components will be solicited through the sponsoring societies.

Pacific Energy Conversion Conference Set for August

The Pacific Energy Conversion Conference will be held at the Fairmont Hotel, San Francisco, Calif., August 13–16, 1962.

The Conference is being sponsored by the American Institute of Electrical Engineers. Its purpose is to bridge the gap between scientific investigation and practical application. Topics to be covered are: Energy Storage, Thermoelectric, Magnetohydrodynamics, Thermionic, Photovoltaic, and Fuel Cells.

Air Force MARS Announces Schedule

The schedule of broadcasts of the Air Force MARS Eastern Technical Net, operating Sundays from 2 to 4 P.M. EDST on 3295, 7540, and 15,715 kc, has been announced for May and early June:

May 6—"Rectifiers for Amateur Radio Power Supplies." A panel of rectifier engineers will discuss the theory, operation and application of germanium and silicon rectifiers. J. Mengenast, moderator.

May 13—"Presentation of the 1961 Edison Radio Amateur Award." Rebroadcast of ceremonies attendant to the presentation of the 1961 Edison Award to William G. Welsh, WISAD/6. Major General J. B. Bestic, Director of Telecommunications, U. S. Air Force, speaker.

May 20—"Novel Equipment for the MARS and Amateur Radio Station," R. Gunderson, Braille Technical Press.

May 27—"Applications of Drift Transistors to Radio Receivers," J. Englund, RCA Research Laboratories.

June 3—"Transistor Multivibrator and Logic Design Considerations," B. Lohman, RCA Semiconductor Products Division.

Group Travel Plan To IFIP Congress-62

IRE members planning to attend the IFIP Congress-62 may take advantage of the new reduced-fare group travel plan. This major international conference on information processing, which is organized by the International Federation for Information Processing, will be held in Munich, Germany, August 27–September 1, 1962. Papers and exhibits on computing and data processing will be contributed by experts from many countries.

The reduced fare applies to a group of 25 or more who travel together, round-trip, on the same regularly scheduled flights. For example, the fare from New York to Munich and back is $353.00 for each member of such a group, compared with the $568.80 individual fare on the same economy-class jet flights. This substantial saving is available to paid-up IRE members of any grade and accompanying members of their immediate family. Groups with various flight dates and routings will be arranged as interest warrants. For further information please write to: W. Buchholz, IBM Development Lab., P.O. Box 390, Poughkeepsie, N. Y.

Reservations should be made soon because seasonally heavy travel at the time of the Congress may limit the number of groups that can be organized.

National Advanced-Technology Management Conference Set

The National Advanced-Technology Management Conference will be held at the Opera House on the Seattle World's Fair Grounds, Seattle, Wash., September 3–7, 1962. The sponsors for the conference will be: the American Institute of Chemical Engineers, the American Institute of Electrical Engineers, the American Society of Civil Engineers, the American Society of Mechanical Engineers, the IRE, the Society of American Military Engineers, the Society for Advancement of Management, the Institute of Management Sciences, the National Society of Professional Engineers, and the University of Washington.

The theme of the conference is: An Exposition of Modern Systems Management . . . Its Development and Practical Application During an Era of Accelerating Technology. The conference is designed to create an atmosphere suitable for the exchange of ideas, both theoretical and practical, on this subject for advanced-technology program managers.

Panel subjects under consideration are: 1) Information Management in Both Civil and Military Science and Technology, 2) How to Organize our Forces to Compete with the Russians in the Military Area, 3) A Summary of the Conference—to allow the exchange of ideas about what the panel members have found most significant in what has been presented.

Workshop subjects under consideration are: 1) Planning the overall program, 2) Controlling the configuration of the end product, 3) Quality assurance, 4) Final system evaluation, 5) Allocation and control of resources, 6) "Exploding" technology.

The Proceedings of the conference can be purchased for $8.00 by writing to the: National Advanced-Technology Management Conference, 512 First Avenue North, Seattle 9, Wash.

Professional Group News

At its meeting held March 8, 1962, the IRE Executive Committee approved the following new chapters: Joint PG on **Audio and Broadcasting**—Montreal Chapter; PG on **Broadcasting** and **Broadcast and Television Receivers** to be combined to form one Joint Chapter—Philadelphia Chapter; PG on **Communications Systems** and **Vehicular Communications** to be combined to form one Joint Chapter—Philadelphia Chapter; PG on **Antennas and Propagation** and **Microwave Theory and Techniques** to be combined to form one Joint Chapter—Philadelphia Chapter.

Military Electronics Convention Proceedings

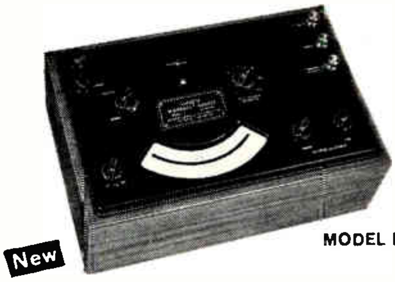
Convention Proceedings for the Third National Winter Convention on Military Electronics, held in Los Angeles, Calif., February 7–9, 1962, may be obtained from the IRE Business Office. Mark requests attention: Ray Banks, 1435 South La Cienega Boulevard, Los Angeles 35, Calif.

Schedule Changed for MAECON

The dates for the 1962 Mid-American Electronics Conference scheduled for November 13 and 14, at the Hotel Muehlebach, Kansas City, Mo., have been changed. The conference will be held instead November 19 and 20 at the Continental Hotel, Kansas City, Mo.

The conference's theme will be "Electronics and Automation." Technical Papers Chairman is Dr. J. N. Warfield, University of Kansas, Lawrence, Kan.

For information regarding booth space write: Charles E. Day, c/o R. W. Farris Co., Inc., 1715 Baltimore, Kansas City, Mo.



MODEL LE

10 NEW INSTRUMENTS

Lindeck Microvolt Source for use as:
• A comparator in the calibration of volt ratio boxes, saturated standard cells and similar instrumentation.
• A direct measuring instrument in the microvolt range.

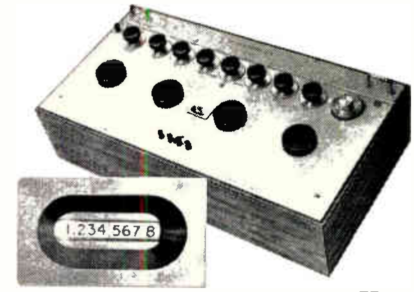
Ranges: 0-1/2/5/10/20/50/100/200/500 μv; 1/2/5/200 mv; 2 v.

Accuracy: ±.5% of full scale

Readout: SRIC Model C. 100 division, 6.3" hand-drawn scale. Accuracy .25%. Diamond pivoted!

Application is described in NBS Paper RP1419, "Testing and Performance of Volt Ratio Boxes". Copies available upon request.

Model PCM "Pocket-Pot." A new multirange, high sensitivity, miniature potentiometer, with self-contained galvanometer and battery operated standardization circuit. In-line readout. Continuous resolution on slidewire with 1 mv. divisions. Infinite resistance at null. Total measuring range 0-5.100 v. Plug in unit available to increase measuring range to 500 v. and 1 amp. (Model PC-S). Accuracy ±.05% of reading; ±.5 mv. on x1 range; ±.1 mv. on x.1 and x.01 ranges.



SINGLE WINDOW READOUT

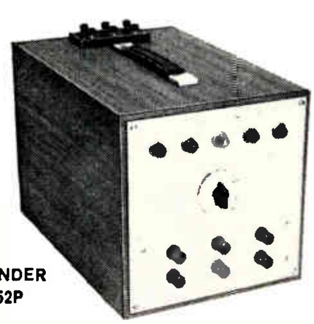
COMMANDER TYPE 9120

New **A 4 dial, 7 figure precision potentiometer** with single window readout and a resolution of 1 part in 20 million. Total measuring range of 2.099,999,9 volts in steps of .1 μv. is achieved without the necessity of switching ranges. **Accuracy ±(.0015% of reading +.1 μv.)**

Facilities are incorporated to enable the user to completely "Self Check" the potentiometer within its guaranteed accuracy. All positions on the measuring dials are calibrated steps (no slidewire). Thermal emf's are less than .1 μv.

SENSITIVE RESEARCH

ELECTRICAL MEASURING INSTRUMENTS SINCE 1927



COMMANDER TYPE 9152P

New **A portable constant temperature air bath** designed to eliminate the hazards involved in transporting saturated standard cells. It is an ideal transfer standard when direct intercomparison of saturated standard cells is desired between a calibration laboratory and NBS or between variously located calibration facilities of the same organization. Enclosure accommodates three saturated cadmium cells constructed with a porous partition to greatly improve portability. Cells can be certified to ±.0001%.

Weight (30 lbs.) and box size (18" L x 10" W x 10" H) are convenient for hand carrying. Provision is made for operation from a portable battery pack, a 110 volt AC line or a car cigarette lighter outlet. Enclosure is maintained at a nominal 32°C and is guaranteed not to vary more than ±.01C° even after continuous exposure to ambient temperatures down to 0°F (−17.4°C) in still air.

MODEL PC-R

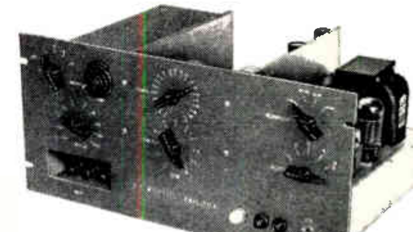
"Plug-in" Wheatstone Bridge for use with PocketPot to make DC resistance measurements from .1 Ω to 9.999 megohms. Accuracy is ±1% from .1 Ω to 10 Ω; ±.2% above 10 Ω to 1 megohm; ±2% above 1 megohm to 9.999 megohms.

New **COMMANDER Type 9770:** Constant current source for standardizing a DC potentiometer or supplying a stable current to any circuit in the range of 10 ma. to 100 ma. Current stability is ±(1 ppm + stability of reference cell).

New **COMMANDER Type 9180:** A 3 dial, 5 figure version of the Dauphinee Potentiometer. Accuracy ±.002% of reading. Range: −10μv. to +2.10100 v. in steps of 10 μv. on X1. Additional ranges of X.1, X.01 and X.001 with resolution to .01 μv. Thermal emf's less than .1 μv. "Self Checking".

New **COMMANDER Type 9174:** A dual range, 2 dial, 4 figure microvolt potentiometer. Thermal emf's less than .02 μv. Accuracy ±.01% of reading. Range: −10 μv. to + 1010 μv.

New **Model ESX:** AC/DC, 0-150 v., electrostatic voltmeter with DC scale expansion enabling practical readings down to 2 v. Input resistances from 1 x 10^{14} Ω.



COMMANDER TYPE 9790

New **A DC amplifier** for the measurement of low level voltages in the fractional microvolt range. Features are:

1. .03 μv. lowest range.
2. Immediate recovery from overloads of 1,000,000%.
3. .003 μv. stability over a 24 hour period.
4. Facility for compensation of stray thermal emf's.
5. Easy accessibility for maintenance purposes.

Output is 1 ma. into 1500 Ω, or 1 mv., 10 mv. or 100 mv. F.S. Input resistance of 50 Ω to 1000 Ω depending on input transformer selected. Continuously variable gain control. Calibration signals to the input (accuracy ±5%) are .03/.1/.3/1/3/10/30/100/300 μv. Response time is 2 to 5 seconds to 90% F.S.

SENSITIVE RESEARCH INSTRUMENT CORPORATION
New Rochelle, N. Y.

COMMANDER instruments are manufactured for Sensitive Research by Guildline Instruments, Ltd., Smiths Falls, Canada.

Conference on Self-Organizing Systems

Museum of Science and Industry, Chicago, Ill., May 22–24, 1962

Tuesday Morning, May 22

"The Organization of Organization," O. Selfridge, Lincoln Lab., Massachusetts Institute of Technology, Lexington, Mass.

"Self-Organizational Systems," M. D. Mesarovic, Case Institute of Technology, Cleveland, Ohio.

"Self-Organization in the Time Domain," D. M. MacKay, University College of North Staffordshire, Dept. of Communication, Keele, Staffordshire, Eng.

"Integrative Processes," J. D. Cowan, Warren S. McCulloch, Research Lab. of Electronics, Massachusetts Institute of Technology, Cambridge, Mass.

Tuesday Afternoon

"Information Input Overload," J. G. Miller, Mental Health Research Institute, University of Michigan, Ann Arbor, Mich.

"Learning Signal Detection," M. Kac, Rockefeller Institute, New York, N. Y.

"Optimization Through Evolution with Sexual Recombinations," H. J. Bremermann, University of California, Dept. of Mathematics, Berkeley, Calif.

The Automatic Formation of a Program which Represents a Theory," S. Amarel, Radio Corporation of America, Princeton, N. J.

Wednesday Morning, May 23

"Natural and Artificial Synapses," L. D. Harmon, Bell Telephone Labs., Murray Hill, N. J.

"Logical Aspects of Neuristor Systems," H. Crane, Stanford Research Inst., Menlo Park, Calif.

"Some Probabilistic Aspects of Automata with a Pushdown Memory," M. P. Schultzenberger, Harvard University Medical School, Boston, Mass.

"Efficient Adaptive Systems and their Realization," J. H. Holland, University of Mich., Ann Arbor, Mich.

"Empirical Laws and Physical Theories; Discussion of the Respective Roles of Information and Imagination," L. Brillouin, Columbia University, New York, N. Y.

Wednesday Afternoon

"Majority Logic," S. Muroga, International Business Machines Corp., Yorktown Heights, N. Y.

"Interaction Between a Group of Subjects and an Adaptive Automaton to Produce a Self-Organizing System for Decision Making," G. Pask, Systems Research Ltd., Richmond, Surrey, Eng.

"Cybernetic Ontology and the Transjunctional Operator," G. Gunther, University of Illinois, Urbana, Ill.

"Some Problems of Basic Organization in Problem-Solving Programs," A. Newell, Carnegie Institute of Technology, Pittsburgh, Pa.

"Training Sequences for Mechanized Induction," R. J. Solomonoff, Zator Co., Cambridge, Mass.

Thursday Morning, May 24

"Adaptive 'Neuron' Memory System," B. Widrow, Stanford University, Stanford, Calif.

"A Comparison of Several Perceptron Models," Frank Rosenblatt, Cornell University, Ithaca, N. Y.

"A New Class of Multi-Layer Series in Coupled Perceptrons," A. G. Konheim, International Business Machines Corp., Yorktown Heights, N. Y.

"Simple Tests for Linear Separability as Applied to Self-Organizing Machines," R. C. Singleton, Stanford Research Inst., Menlo Park, Calif.

Thursday Afternoon

"Remarks on the Algebra of Functors," K. Menger, Illinois Institute of Technology, Chicago, Ill.

"A Feedback Coding Theory of Learning and Cognition," H. H. Kantner, Armour Research Foundation, Illinois Institute of Technology, Chicago, Ill.

"Some Similarities Between the Behavior of a Neural Network Model and Electrophysiological Experiments," P. G. Farley, Lincoln Lab., Massachusetts Institute of Technology, Lexington, Mass.

"The Representation of Information by Neural Net Models," P. H. Greene, University of Chicago, Chicago, Ill.

1962 National Microwave Theory and Techniques Symposium

Boulder, Colo., May 22–24, 1962

Tuesday Morning, May 22

Session I—Microwave Frontiers

Chairman: *Dr. D. D. King, Electronic Communications, Inc., Timonium, Md.*

"Millimeter Wavelength Resonant Structures," *R. W. Zimmerer, M. V. Anderson, G. L. Strine, and Y. Beers, National Bureau of Standards, Boulder, Colo.*

"Superconducting Delay Line," *R. J. Allen, A. J. Cummings, and F. M. Kudo, Martin Marietta Corp., Baltimore, Md.*

"Ferroelectric Phase Shifters for VHF and UHF," *M. Cohn and A. F. Eikenberg, Electronic Communications, Inc., Timonium, Md.*

Tuesday Afternoon

Session II—Theory, Waveguide Properties

Chairman: *Prof. A. A. Oliner, Microwave Research Institute, Polytechnic Institute of Brooklyn, Brooklyn, N. Y.*

"Microwave Propagation in an Overdense Bounded Magnetoplasma," *B. Wieder, National Bureau of Standards, Boulder, Colo.*

"Measurements of Field Strength on Resonator Boundaries by Perturbation of Radiation Field," *I. Hefni, Lincoln Lab., M.I.T., Lexington, Mass.*

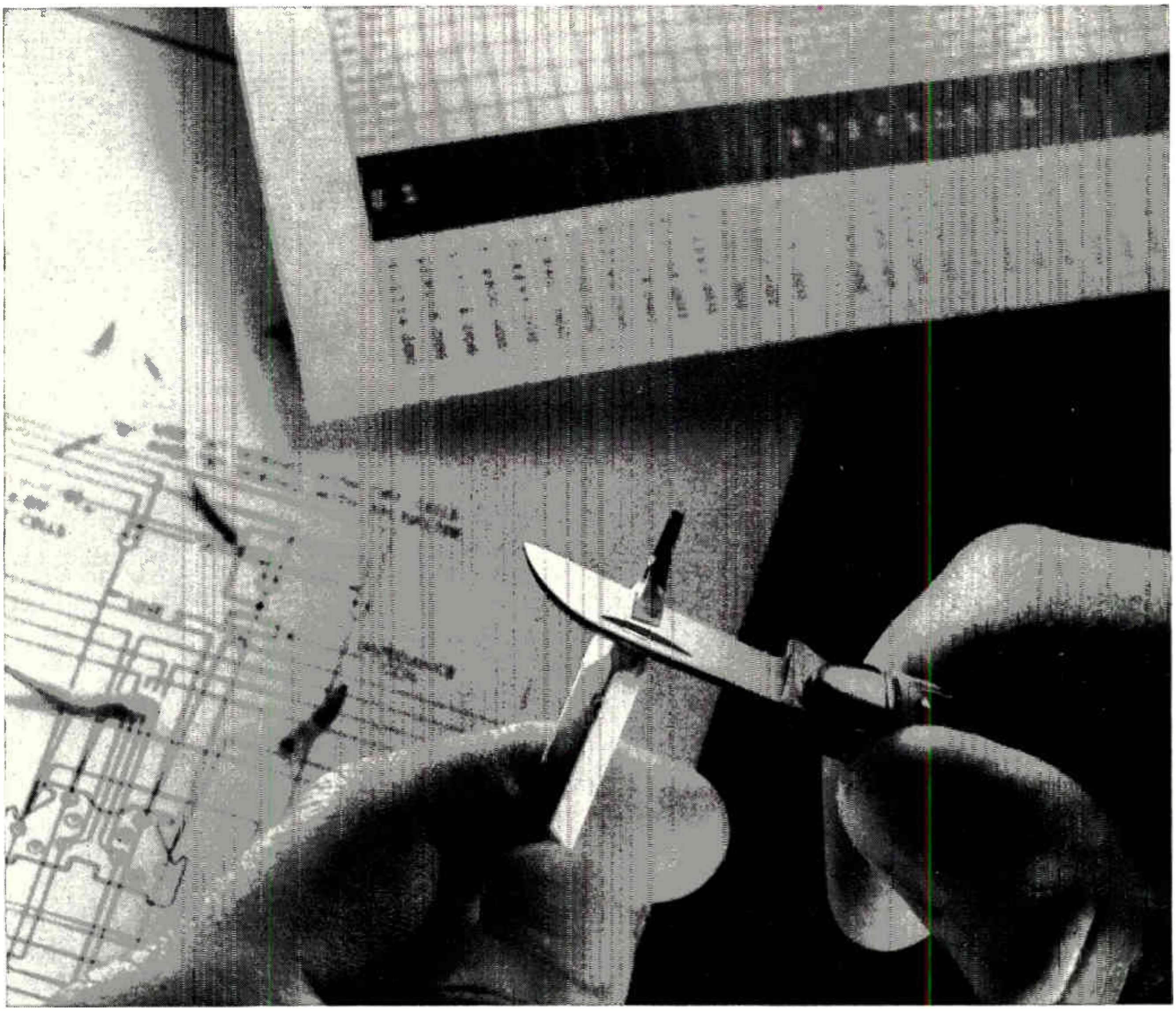
"General Power Loss Method for Attenuation of Cavities and Waveguides," *J. J. Gustincic and R. E. Collin, Case Institute of Technology, Cleveland, Ohio.*

Session III—Theory, Waveguide Discontinuities

Chairman: *Prof. A. A. Oliner, Microwave Research Institute, Polytechnic Institute of Brooklyn, Brooklyn, N. Y.*

"Perturbation Theorems for Waveguide Junctions, with Applications," *D. M. Kerns and W. T. Grandy, Jr., National Bureau of Standards, Boulder, Colo.*

"Investigations of the Reflection from a Junction of an Ideal Rectangular Waveguide with One Having Rounded Inside Corners," *W. J. Anson, R. W. Beatty, D. M. Kerns, and W. T. Grandy, Jr., National Bureau of Standards, Boulder, Colo.*



Sorry, but your resistor specs just went out of date

Now Corning supplies total ΔR less than 3% with no derating

Start a design with known limits of resistance deviation as tight as that and you can specify other components with more certainty and more freedom.

Drop an amplifier stage? Use broader tolerance, cheaper tubes or transistors?

That's what our new *Corning Design Tolerances* are all about. They give you a percent deviation from nominal that includes the purchase tolerance, maximum ΔR due to TC, and maximum load-life drift. They're based on extended performance at full power and 70°C. ambient for over 30,000 hours.

Model		Resistance (ohms)	Corning Design Tolerance
NF (Meets Mil-R-10509D)	60 65	100 to 100K 100 to 348K	3%
N (Meets Mil-R-10509D)	60 65 70	10 to 133K 10 to 499K 10 to 1 meg.	3%
C (Meets Mil-R-22684)	20 32 42S	51 to 150K 51 to 470K 10 to 1.3 meg.	5% (plus purchase tolerance of either 2% or 5%)

We've assigned them to resistors that cover the 10 ohm—1 meg. range: the fusion-sealed NF, the precision N, and the low-cost, high-performance general purpose C. They're available fast, and at factory prices, from your local Corning distributor.

A new folder, "Design Tolerances for Tin Oxide Resistors," gives you full information. Write for a copy to Corning Glass Works, 542 High St., Bradford, Pa. . . . and sharpen your pencil.

CORNING
Electronic Components

"Electromagnetic Diffraction by a Planar Array of Circular Disks," *W. H. Eggimann and R. E. Collin, Case Institute of Technology, Cleveland, Ohio.*

"The Representation of Waveguides Containing Small Ferrimagnetic Ellipsoids," *L. K. Anderson, Bell Telephone Labs., Murray Hill, N. J., and H. J. Shaw, Stanford University, Stanford, Calif.*

Wednesday Morning, May 23
Session IV—Filters

Chairman: *Dr. S. B. Cohn, Rantec Corp., Calabasas, Calif.*

"Interdigital, Band-Pass Filters," *G. L. Matthaei, Stanford Research Institute, Menlo Park, Calif.*

"Microwave Bandstop Filters with Narrow Stop Bands," *L. Young, G. L. Matthaei, Stanford Research Institute, Menlo Park, Calif., and E. M. T. Jones, TRG-West, Menlo Park, Calif.*

"The Design of Branch-Guide Couplers, with Applications to the Suppression of Spurious Frequencies," *L. Young, Stanford Research Institute, Menlo Park, Calif.*

"An Electronically Turnable Band-Pass Microwave Filter," *I. Kaufman, Space Technology Labs., Inc., Canoga Park, Calif., and W. H. Steier, University of Illinois, Urbana, Ill.*

Session V—High Power

Chairman: *Dr. S. B. Cohn, Rantec Corp., Calabasas, Calif.*

"A New Technique for Multimode Power Measurement," *J. J. Taub and J. Goldberg, Airborne Instruments Lab., Deer Park, L. I., N. Y.*

"A Novel High-Power Harmonic Suppressor," *E. Wantuch and R. M. Maines, Airtron, Inc., Morris Plains, N. J.*

"Layered Media as High Power Microwave Absorbers," *S. R. Steele and R. J. Briggs, Raytheon Co., Burlington, Mass.*

"Fine Grain Spectrum Analysis of Pulsed Microwave Amplifiers," *R. F. Koontz, Radio Corp. of America, Moorestown, N. J.*

Wednesday Afternoon
Session VI—Measurements

Chairman: *R. W. Beatty, National Bureau of Standards, Boulder, Col.*

"Subtle Differences in System Noise Measurements and Calibration of noise Standards," *T. Mukaihata, B. Walsh, M. F. Bottjer, and E. B. Roberts, Hughes Aircraft Co., Culver City, Calif.*

"An Accurate Millimeter Wave Loss and Delay Measurement Set," *M. B. Chasek, Bell Telephone Labs., Murray Hill, N. J.*

"Microwave Phase Comparator," *J. A. Kaiser, H. B. Smith, Jr., W. H. Pepper, and J. H. Little, Diamond Ordnance Fuze Labs., Washington, D. C.*

"Waveguide Perturbation Techniques in Microwave Semiconductor Diagnostics," *K. S. Champlin and D. B. Armstrong, University of Minnesota, Minneapolis, Minn.*

Session VII—Components

Chairman: *R. W. Beatty, National Bureau of Standards, Boulder, Colo.*

"Further Developments in Dielectric Waveguide Devices for Millimeter Wavelengths," *C. E. Barnes, Bell Telephone Labs., Murray Hill, N. J.*

"Phase Shift Characteristics of Dielectric Loaded Waveguide," *G. F. Bland and A. G. Franco, International Business Machines, Yorktown Heights, N. Y.*

"Ferrite Switches in Coaxial or Strip Transmission Line," *C. E. Fay, Bell Telephone Labs., Murray Hill, N. J.*

"A 6 KW Peak Power Varactor Duplexer," *J. C. Hoover, Sperry Microwave Electronics Co., Clearwater, Fla.*

Thursday Morning, May 24
Session VIII—Limiters

Chairman: *M. T. Lebenbaum, Airborne Instruments Lab., Deer Park, L. I., N. Y.*

"Bandwidth of Tem Diode Limiters," *R. V. Garver and J. A. Rosado, Diamond Ordnance Fuze Labs., Washington, D. C.*

"Frequency-Selective Limiting," *K. L. Kotzebue, Watkins-Johnson Co., Palo Alto, Calif.*

"Operation of a Microwave Garnet Limiter," *R. L. Comstock and L. J. Varnerin, Bell Telephone Labs., Murray Hill, N. J.*

"A Coincidence Region Power Limiter Using Monocrystal Lithium Ferrite at 6500 Mc," *S. Okwit, Airborne Instruments Lab., Deer Park, L. I., N. Y.*

Session IX—Parametric Amplifiers

Chairman: *M. T. Lebenbaum, Airborne Instruments Lab., Deer Park, L. I., N. Y.*

"Multiple Pumped Parametric Amplifiers," *G. Spacek and R. Brewster, General Motors Corporation Defense Systems, Santa Barbara, Calif.*

"Broadband Parametric Amplifiers," *R. Pettai, Micro State Electronics Corp., Murray Hill, N. J., and B. B. Bossard, Radio Corp. of America, N. Y., N. Y.*

"Large Signal Properties of Non-Degenerate Varactor Parametric Amplifiers," *D. Jackson, Boeing Co., Seattle, Wash.*

"Design of a Wideband Tunnel Diode Preamplifier for Phased Array Radars," *D. W. MacGlashan, Bendix Corp., Baltimore, Md.*

Thursday Afternoon
Session X—(Invited Session) Future of Microwave and Solid State

Chairman: *Dr. K. Tomiyasu, General Electric Co., Schenectady, N. Y.*

"Masers and Millimeter Waves," *Prof. F. S. Barnes, University of Colorado, Boulder, Colo.*

"Microwaves and Solid State," *Dr. B. Lax, Lincoln Lab., M.I.T.., Lexington, Mass.*

"Coherent Wave Optics, New Frontier of the Microwave Art," *A. G. Fox, Bell Telephone Labs., Holmdel, N. J.*

IRE Seventh Region Conference

OLYMPIC HOTEL, SEATTLE, WASH., MAY 24–26, 1962

"Space Communications" is the theme of the 1962 IRE Seventh Region Conference which will stress the impact of electronics on the cultural and social life of the future.

All hotel reservations are being handled only through the Seventh Region Conference Housing Bureau. The bureau will hold hotel space until April 23. Your request must be accompanied by a deposite of $15.00 a room, payable to "IRE Seventh Region Conference." Acknowledgment will be made by the hotel. Cancellations must be received by May 8 for any refund.

The Board of Directors of the IRE and of WESCON are also meeting during the period of the Seventh Region Conference.

The program will be as follows:

Thursday Morning, May 24
Session 2A—Antennas

Moderator: *Dr. N. Yaru, Hughes Aircraft Co.*

"Giant Aperture Antennas for Space Applications," *C. J. Sletten and P. Blacksmith, Jr., Electromagnetic Radiation Lab. AFCRL.*

"Antennas for the Naby Space Surveillance System," C. A. Bartholomew, Space Surveillance Group NRL.

"Cassegrain Design Procedure," P. A. Jensen, Hughes Aircraft Co.

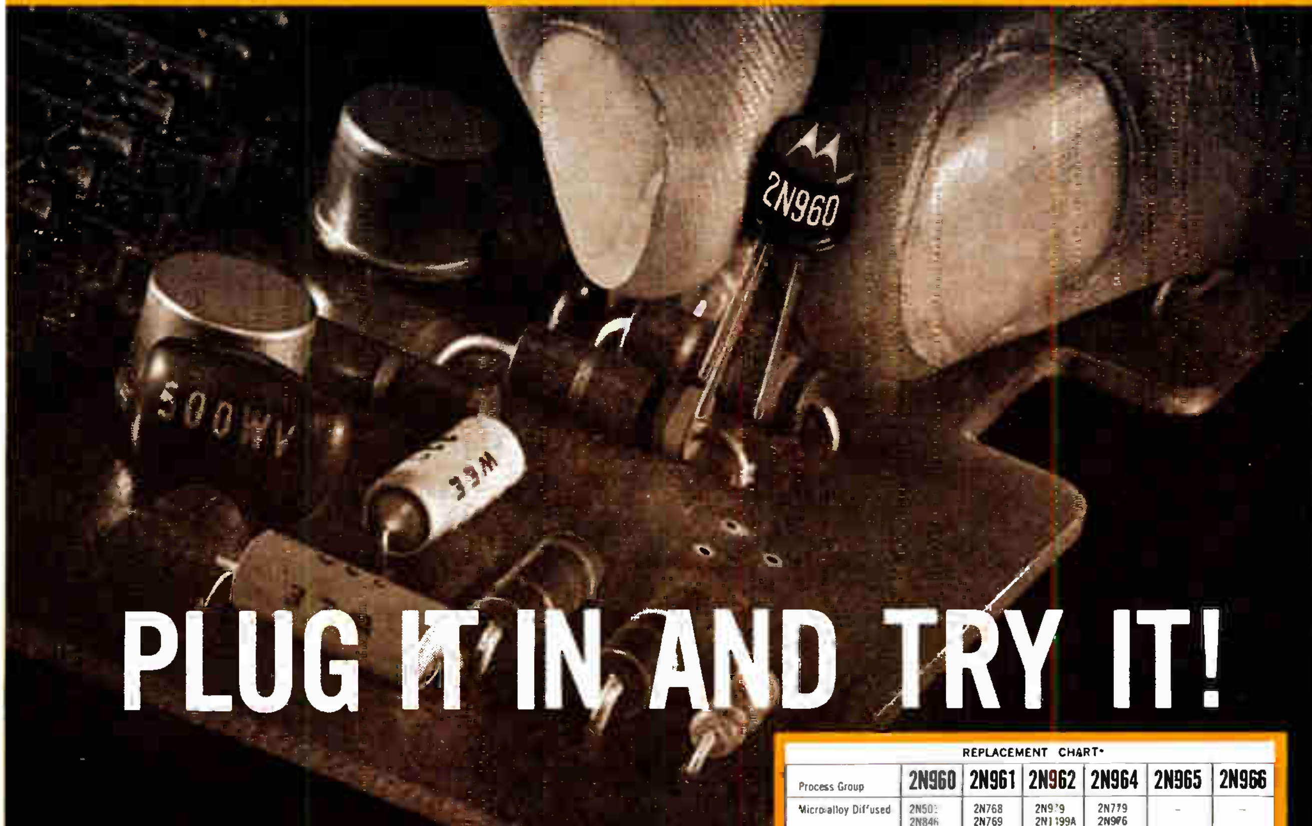
"High Power Dual Polarized Monopulse Feed," E. D. Burnett, Hughes Aircraft Co.

Session 3A—Electronics in Education

Moderator: *T. L. Martin, Dean, College of Engineering, University of Arizona.*

Panel members: *Prof. W. W. Harman, Electrical Engineering Dept., Stanford University.*

How much better is MOTOROLA'S 2N960 Series Epitaxial Mesa?



REPLACEMENT CHART*						
Process Group	2N960	2N961	2N962	2N964	2N965	2N966
Micro-alloy Diffused	2N502 2N846 2N984	2N768 2N769	2N979 2N1199A 2N1200	2N779 2N976 2N982 2N983	–	–
Mesa	2N705 2N710 2N711B 2N781	2N711 2N711A	2N782	2N795 2N934 2N994 2N1301	2N796 2N1683	2N1309 2N794
Micro-alloy	2N1122A	2N1122	2N1411 2N1427	–	–	2N393
Alloy	2N583	–	–	2N582 2N584	–	–
Drift	–	2N643 2N644 2N645	2N602 2N1450	–	2N609	2N603

*Interchangeability of types shown is on the basis of performance in most switching circuit applications.

Is it true that Motorola's germanium, epitaxial, 2N960 switching series will supplant nearly all other germanium micro-alloy, drift, mesa, and other transistor types for high-speed switching applications? And is it true that the Motorola series is even faster and performs better than advertised? The best way to find out is to try it!

Use the adjacent Replacement Chart and Specifications as a guide, and try any of these six remarkable new Motorola devices in your circuits. Judge for yourself what the advantages are. Samples are available from your local Motorola District Office. Ask also about Motorola's low prices – in many cases, they are considerably lower than those for old type devices.

Some of the Published Advantages of the Motorola 2N960 Series:

- faster switching time
- guaranteed minimum Beta over a wide current range . . . specified at 10, 50, and 100 mA
- low saturation even at 100 mA
- rugged Motorola mesa construction
- the most comprehensive *and conservative* published specifications of any similar switching transistors
- proven reliability from the world's largest manufacturer of germanium epitaxial transistors

MOTOROLA GERMANIUM EPITAXIAL SWITCHING TRANSISTORS							
	2N960	2N961	2N962	2N964	2N965	2N966	UNITS
h_{FE} (MIN) 10, 50, 100 mA	20	20	20	40	40	40	—
V_{CE} (SAT) MAX @ 10 mA	.20	.20	.20	.18	.18	.18	Volts
@ 50 mA	.40	.40	.40	.35	.35	.35	Volts
@ 100 mA	.70	.70	.70	.60	.60	.60	Volts
f_T (MIN) I_E=20 mAdc V_{CB}=1.0 Vdc	300 mc all types						
Q_T (MAX) I_C=10 mAdc I_B=1 mAdc	80	80	90	80	80	90	pc
I_C=100 mAdc I_B=5 mAdc	125	125	150	125	125	150	pc
τ_{RE}	0.6 nsec typical all types						
τ_{FE}	0.5 nsec typical all types						
All types have 150 mW dissipation in free air. 300 mW at 25°C case temperature							

For more information on this important new mesa series, contact your Motorola District Office, or call or write: Motorola Semiconductor Products Inc., Technical Information Department, 5005 East McDowell Road, Phoenix 8, Arizona.

MOTOROLA DISTRICT OFFICES:
Belmont, Mass. / Burlingame, Calif. / Chicago / Cleveland / Clifton, N. J. / Dallas / Dayton / Detroit / Garden City, L. I. / Glenside, Pa. Hollywood / Minneapolis / Orlando, Fla. / Phoenix / Silver Spring, Md. Syracuse / Toronto, Canada.

MOTOROLA
Semiconductor Products Inc.
A SUBSIDIARY OF MOTOROLA, INC. 1882
5005 EAST McDOWELL ROAD • PHOENIX 8, ARIZONA

Prof. D. K. Weaver, Director of Electronics Research Lab., Montana State University.

Prof. A. V. Eastman, Chairman, Electrical Engineering Dept., University of Washington.

Prof. D. O. Peterson, Director of Electronics Research Lab., University of California, Berkeley, Calif.

Thursday Afternoon

Session 1A—Propagation and Ionospheric Physics

Moderator: *R. A. Helliwell, Radio Science Lab., Stanford University.*

"Topside Sounder," *R. W. Knecht, National Bureau of Standards.*

"Measurement of Ion Composition Above the F_2 Peak," *T. W. Flowerday and D. D. McKibbin, Lockheed Missile and Space Co.*

"Interpretation of VLF Propagation-Measurement on the LOFTI Satellite," *R. L. Smith, Radio Science Lab., Stanford University; L. H. Rorden, Stanford Research Institute; R. A. Helliwell, Radio Science Lab., Stanford University.*

Session 1B—Satellite Communication

Moderator: *Dr. G. H. Keitel, Philco Corp., Palo Alto, Calif.*

(Title to be announced), *S. H. Reiger,* The Rand Corp.

"A Review of Defense Communications Satellite Program," *R. L. Clark, Department of Defense, Washington, D. C.*

"The Impact of Communication Satellites on Civilian Communications," *R. T. Haviland, General Electric, Philadelphia, Pa.*

Session 2C—Advanced Circuit Packing Techniques

Moderator: *Lt. Col. J. K. Schloss, USAF (Ret.)*

"Principles, Practices and Problems of Protective Packaging of Semi-conductor Devices," *Dr. S. S. Flaschen, Semiconductor Product Div., Motorola Inc., Phoenix, Ariz.*

"Protection and Interconnection of Space Quality Microcircuits," *G. J. Selvin, Microelectronics Lab., Sylvania Electric Products.*

"Packaging of Miniaturized Electronics," *E. C. Singletary, Texas Instruments Inc., Dallas, Tex.*

"System Packaging with Molecular Electronics," *J. McKinley, Air Arm Division, Westinghouse Electric Corp.*

Friday Morning, May 25

Session 2B—Devices

Moderator: *W. G. Shepherd, Head of Electrical Engineering Dept., University of Minnesota.*

"High Power Linear Electron Beam Tube for Space Communication." *C. M. Veronda, Sperry Electronic Tube Div.*

"The Solid-State Traveling-Wave Maser," *Dr. R. W. DeGrasse, Microwave Electronics Corp.*

"MM Wavelength Vacuum Tube Development," *G. Convert, Compagnie generale de telegraphie Sans Fil, Paris, France,*

"The Design of Optical Masers," *Dr. J. C. Axtell, The Boeing Co.*

Session 3C—Electronics in Business

Moderator: *H. W. Haines, Vice President Finance, The Boeing Co.*

"Impact of Electronics on Suppliers of Equipment," *W. Hume, President of I.B.M. Data Processing Div.*

"Impact of Electronics on Systems and Methods of Applying Electronics in Business," *N. Ream, Director of System and Procedures, Lockheed Co.*

"Impact of Electronics on Business in European Market," *H. Reinoud, Director in Chief, Netherlands Postal and Telecommunications Services.*

"Impact of Electronics on Finance, Banking and Stock Market." *J. H. Moller, Director Merrill Lynch, Pierce, Finner and Smith.*

Session 3E—Student Papers

Moderator: *Prof. I. J. Sandorf, University of Nevada.*

"Impact of Electronics on People Business and Management in Century 21," *J. Diebold, President John Diebold and Associates.*

Friday Afternoon

Session 1D—Space Experiments and Communication

Moderator: *S. H. Reiger, The Rand Corp.*

"Ranger III Lunar Impact Television Experiment," *R. C. Heyser, Jet Propulsion Labs.*

"JPL Venus Radar Experiment," *M. H. Brockman, Jet Propulsion Labs.*

"Influence of the Sun on Space Communication," *D. W. Swayze, Philco Corp., Palo Alto.*

"An Approach to Space Craft Data Handling and Experiment Control," *R. W. Waller, Chief Engineer, Western Div., Computer Control Co.*

"A Microwave Radiometer for a Venus Fly-By Probe," *D. Jones, Jet Propulsion Labs.*

Session 2D—Secondary Power

Moderator: *W. C. Scott, Chief, Space Power Technology Program, NASA.*

"Solar Thermo—Mechanical Power," *J. A. Rudy, New Devices Labs., Thompson Ramo-Wooldridge.*

"Solar Photovoltaic Devices," *A. Klammer, Jr., Power Sources Unit, The Boeing Co.*

"Biological Fuel Cell," *Commander F. W. Anders, Bureau of Ships, U.S.N.*

"Nuclear Radioisotope Power," *Capt. R. T. Carpenter, Auxiliary Power Branch, U. S. Atomic Energy Commission.*

Session 3D—Electronics in 2012 AD

Moderator: *Dr. L. Fields, Manager, Microwave Tube Div., Hughes Aircraft Co.*

"Automatic Handbook," *N. Rochester, I.B.M.*

"Electronic Nirvana," *D. E. Noble, Executive Vice President, Motorola, Phoenix.*

"Electronic Mastering of Ship," *Rear Adm. Bennett, Vice President Director of Engineering, Sangamo Electric Co.*

(Title to be announced), *W. O. McGuigan, Assistant General Manager Engineering, Stanford, Research Institute.*

"Electronics and Health Care," *V. K. Zworykim, Vice President and Technical Consultant, David Sarnoff Research Center RCA Research Labs., Princeton, N. J.*

Saturday Morning, May 26

Session 1C—Radio Astronomy

Moderator: *R. S. Lawrence, National Bureau of Standards.*

"Jupiter Radiation," *Dr. J. D. Warwick, University of Colorado.*

"University of Illinois Radio Telescope," *Dr. G. W. Swenson, University of Illinois.*

"Radar Reflections from the Solar Corona," *Dr. J. M. Chisolm, Lincoln Labs.*

"The 300 ft. Transit Radio Telescope at the National Radio Astronomy Observatory," *Dr. J. W. Finley, National Radio Astronomy Observatory.*

Session 2E—Radiation Effects on Electronic Equipment

Moderator: *W. L. Brown, Bell Telephone Lab.*

"Radiation Environment for Space Electronics," *Prof. S. F. Singer, University of Maryland, JPL.*

"Radiation Effects on Solar Cells," *Dr. J. M. Denny, Space Technology Labs.*

"The Implication of Nuclear-Powered Space Systems to the Radiation Effects Problem," *Dr. J. C. Lee, Lockheed Missiles and Space Div.*

"Prediction of Transient Radiation Effects on Electronic Equipment," *Dr. G. L. Keister, The Boeing Co.*

Chicago Spring Conference

O'HARE INN, DES PLAINES, ILL., JUNE 18–19, 1962

The Chicago Spring Conference will be held at the O'Hare Inn, Des Plaines, Ill., June 18 and 19, 1962. R. W. Galvin of Motorola will be the guest luncheon speaker.

The tentative program for the conference follows:

Monday Morning, June 18
Session I

"Cross-Modulation and Modulation Distortion of RF Transistors," *G. E. Theriault, Radio Corp. of America.*

"A New UHF TV Local Oscillator," *L. R. Maguire, Sylvania Electric Products, Inc.*

"Transistors in UHF Television Tuners," *S. M. Weaver, Texas Instruments Incorporated.*

"Gamma Distortion in the TV System," *G. W. Fyler, Zenith Radio Corp.*

Monday Afternoon
Session II

"A High-Performance Silicon Transistor Television Receiver," *J. A. MacIntosh, S. A. Schwartz, P. J. Bénéteau, Fairchild Semiconductor, Division of Fairchild Camera and Instrument Corp.*

"Recent Developments in Deflection Circuits," *K. W. McGlashan, Advance Ross Electronics Corp.*

"Transistor Failure Modes in High Power Switching Operation," *J. W. Mathews Philco Corp.*

"Highlights of Stabilized TV Deflection," *J. Sennik, Dominion Electrohome Industries Ltd.*

"A New Silicon High Performance Video Output Transistor," *R. Gudis and C. Kile, Lansdale Division, Philco Corp.*

Tuesday Morning, June 19
Session III

"A Four-Transistor Line Operated Receiver," *S. H. Kuehler and K. G. Cherry, Texas Instruments Incorporated.*

"Ferrite Antennas for AM Broadcast Receivers," *H. J. Laurent and C. A. B. Carvalho, Bendix Radio Division, Bendix Corp.*

"Television Stereo System," *R. B. Dome, General Electric Co.*

"A Mechanical Ultrasonic Signal Source for Remote Control Systems," *L. J. Sienkiewicz and H. T. Goldstein, Westinghouse Electric Corp.*

Tuesday Afternoon
Session IV

"Practical Design Considerations for Low-Cost FM Stereophonic Receivers," *R. J. Nelson and O. P. Hart, Radio Corp. of America.*

"A Two-Transistor FM Broadcast Band Tuner for a 12-Volt Automobile System," *W. F. King, Amperex Electronic Corp.*

"New Capacitor Products for Use in Radio and Television Receivers," *W. M. Robinson, Cornell-Dubilier Electronics Division, Federal Pacific Electric Co.*

"The Design of a Low Cost, High Sensitivity FM Tuner," *D. Ruby, Zenith Radio Corp.*

"A Self-Protecting Transistor Power Amplifier," *G. Randolph and N. Kramer, Knight Electronics Corp.*

Professional Groups*

Aerospace and Navigational Electronics (G-11)—G. M. Kirkpatrick, Electronics Equipment and Systems Lab., GE Co., Syracuse, N. Y.; H. R. Mimno, Cruft Lab., Harvard Univ., Cambridge 38, Mass.

Antennas and Propagation (G-3)—Dr. H. Fine, Applied Propagation Branch, Technical Research Div., FCC, Washington, D. C.; S. A. Bowhill, Pennsylvania State Univ., University Park, Pa.

Audio (G-1)—C. M. Harris, Electronics Res. Labs., Columbia Univ., New York 27, N. Y.; M. Camras, Armour Res. Foundation, Tech. Ctr., Chicago 16, Ill.

Automatic Control (G-23)—J. M. Salzer, Ramo-Wooldridge, 5500 E. Segunda, Hawthorne, Calif; G. S. Axelby, Westinghouse Air Arm Div., Friendship Airport, Baltimore 3, Md.

Bio-Medical Electronics (G-18)—G. N. Webb, Dept. of Medicine, Biophysical Div., Johns Hopkins Hospital, Baltimore 5, Md.; L. B. Lusted, Dept. of Radiology, Univ. of Rochester, Rochester 20, N. Y.

Broadcast and Television Receivers (G-8)—J. F. Bell, Zenith Radio Corp., 6001 W. Dickins Ave., Chicago 39, Ill.; C. W. Sall, RCA, Princeton, N. J.

Broadcasting (G-2)—R. F. Guy, 264 Franklin St., Haworth, N. J.; W. L. Hughes, School of Elec. Engrg., Oklahoma State University, Stillwater, Okla.

Circuit Theory (G-4). Dr. J. H. Mulligan, Jr., College of Engrg., New York Univ., University Heights, New York 53, N. Y.; M. E. Van Valkenburg, Dept. of E.E. Univ. of Illinois, Urbana, Ill.

Communications Systems (G-19)—R. L. Marks, Rome Air Dev. Ctr., Griffiss AFB, N. Y.; E. J. Baghdady, Elec. Engrg. Dept., M.I.T., Cambridge 39, Mass.

Components Parts (G-21)—F. E. Wenger, Headquarters, ARDC, Andrews AFB, Washington 25, D. C.; G. Shapiro, Engineering Electronics Sec., Div. 1.6, NBS, Connecticut Ave. and Van Ness St., Washington 25, D. C.

Education (G-25)—G. E. Moore, Westinghouse Elec. Corp., East Pittsburgh, Pa.; W. R. LePage, Dept. of E.E., Syracuse Univ., Syracuse 10, N. Y.

Electron Devices (G-15)—W. A. Adcock, Texas Instruments, 6000 Lemmon Ave., Dallas 9, Tex.; G. Wade, Spencer Lab., Raytheon Co., Burlington, Mass.

Electronic Computers (G-16)—W. L. Anderson, 1408 Lexington Rd., Falls Church, Va.; Prof. N. R. Scott, Dept. of Elec. Engrg., University of Michigan, Ann Arbor, Mich.

Engineering Management (G-14)—T. W. Jarmie, Engineered Electronics Co., 1441 E. Chestnut Ave., Santa Ana, Calif.; A. H. Rubenstein, Dept. of Indus. Engrg., Northwestern Univ., Evanston, Ill.

Engineering Writing and Speech (G-16)—J. M. Kinn, Jr., *IBM Journal*, 17th Floor, 545 Madison Ave., New York, N. Y.; H. B. Michaelson, IBM Res. Ctr., Box 218, Yorktown Heights, N. Y.

Geoscience Electronics (G-29)—R. W. Olson (Acting Chairman), Texas Instruments, Inc., Dallas, Tex.; Editor to be advised.

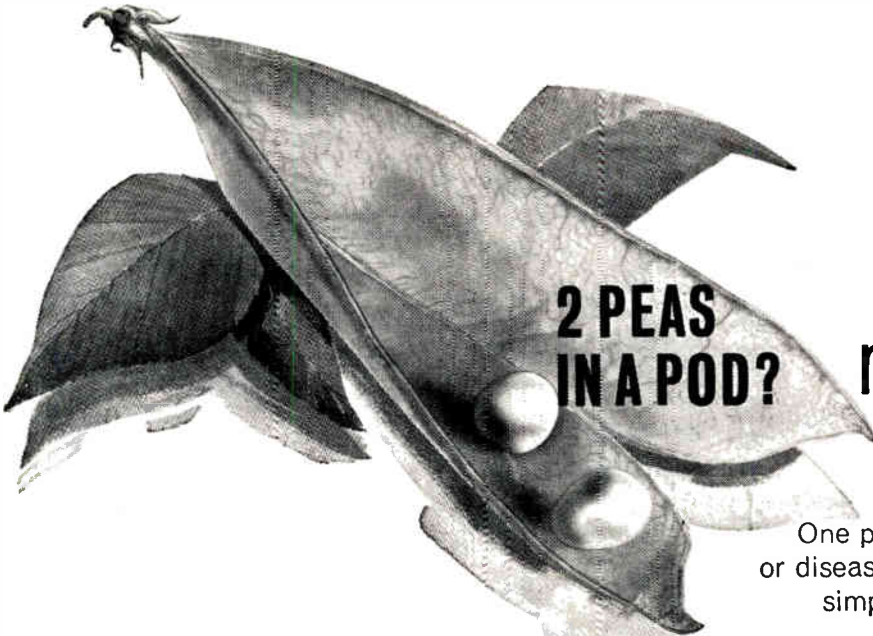
Human Factors in Electronics (G-28)—R. R. Riesz, Bell Telephone Labs., Murray Hill, N. J.; J. I. Elkind, Bolt Beranek and Newman, Inc., 50 Moulton St., Cam-Bridge, Mass.

Industrial Electronics (G-13)—J. E. Eiselein, RCA Victor Div., Camden, N. J.; R. W. Bull, Coleman Instruments, Inc., 42 Madison St., Maywood, Ill.

Information Theory (G-2)—G. L. Turin, E.E. Dept., Univ. of California, Berkeley, Calif.; A. Kohlenberg, Melpar Inc., 11 Galen St., Watertown, Mass.

Instrumentation (G-9)—H. W. Lance, Nat'l.

* Names listed are Group Chairman and TRANSACTIONS Editors.



2 PEAS IN A POD?

no guarantee of similarity!

One pea may be damaged by insects or disease . . . or may be quite different simply through a caprice of nature.

THERE IS A DIFFERENCE IN DEFLECTION YOKES

TYPE Y58
Highest Resolution Stator Yoke

TYPE Y25
Compact Rotating Yoke

TYPE Y16
Low Capacity Push-Pull Yoke for Speed

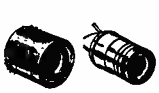
TYPE F10
Precision Electro-magnetic Focus Coil

Specialists in yokes, Syntronic's team of experts knows more about yoke design, engineering and quality control than anyone else. A solid record of leadership for over 10 years—acknowledged throughout the industry.

This specialized experience means more to you than you may know. "Little" things make a big quality difference . . . like fast spot recovery time . . . smallest, roundest dots . . . geometrical precision . . . rapid settling time . . . least spot growth from center to edge . . . smallest residual magnetism. Mighty important, actually. Syntronic cares enough to make them right. You can depend on it.

Syntronic claims—and delivers—the most precise yokes that can be made. Every time. Year after year. To your exact requirement. And Syntronic's team of experts can help solve your yoke problems.

Since there is a difference, benefit from it. Get the most from your display.



TYPE Y15
Precision Low Current Yoke

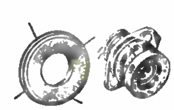
TYPE Y57
1" I.D. Push-Pull Yoke

TYPE Y59
7/8" I.D. Transistor Yoke

TYPE F40
Dynamic-static Focus Coil

syntronic INSTRUMENTS, INC.

100 Industrial Road, Addison, Illinois
Phone: Kingswood 3-6444



Bureau of Standards, Boulder, Colo.; G. B. Hoadley, Dept. of E.E., North Carolina State College, Raleigh, N. C.

Microwave Theory and Techniques (G-17)—T. N. Anderson, EMT Corp., Syosset Industrial Park, 210 Michael Dr., Syosset, L. I., N. Y.; D. D. King, Electronic Commun., Inc., 1830 York Rd., Timonium, Md.

Military Electronics (G-24)—W. L. Doxey, EC Dept., USARDL, Fort Monmouth, N. J.; D. R. Rhodes, Radiation Inc., Box 37, Melbourne, Fla.

Nuclear Science (G-5)—L. Costrell, Nucleonic Inst. Sec., NBS, Washington 25, D. C.; R. F. Shea, Dig Power Plant Engrg., Knolls Atomic Power Lab., General Electric Company, Schenectady, N. Y.

Product Engineering and Production (G-22)—A. R. Gray, Astronics of Florida, Inc., 1110 S. Woods St., Box 5606, Orlando, Fla., D. B. Ehrenpreis, 325 Spring St., New York, N. Y.

Radio Frequency Interference (G-27)—H. E. Dinger, Naval Research Lab., Washington 25, D. C.; O. P. Schreiber, Technical Wire Prods., Inc., 129 Dermondy St., Cranford, N. J.

Reliability and Quality Control (G-7)—L. J., Paddison, Sandia Corp., Sandia Base, Albuquerque, N. Mex.; W. X. Lamb, Jr., 22124 Dumetz Rd., Woodland Hills, Calif.

Space Electronics and Telemetry (G-10)—K. M. Uglow, Jr., 2228 Oriole Dr., Sarasota, Fla.; F. T. Sinnott, Mail Zone 549-30, Convair-Astronautics, San Diego 12, Calif.

Ultrasonics Engineering (G-20)—V. Salmon, Stanford Research Inst., Menlo Park, Calif.; O. E. Mattiat, Res. Div., Acoustica Associates Inc., 415 E. Montecito St., Santa Barbara, Calif.

Vehicular Communications (G-6)—R. P. Gifford, General Electric Co., Rm. 206, Mountain View Rd., Lynchburg, Va.; W. G. Chaney, AT&T Co., 195 Broadway, New York 7, N. Y.

Sections*

Akron (4)—C. M. Kelly, 2915 Silver Lake Circle, Cuyahoga Falls, Ohio; D. Grant, 1206 Thorndale Dr., Akron 20, Ohio.

Alamogordo-Holloman (7)—W. S. Stotts, 1507 Michigan Ave., Alamogordo, N. Mex., H. T. Castillo, 2611 Pontiac Ave., Alamogordo, N. Mex.

Albuquerque-Los Alamos (7)—R. K. Moore, 608 Vassar Dr., N.E., Albuquerque, N. Mex.; F. H. Jean, 1017 Alcazar, N.E., Albuquerque, N. Mex.

Anchorage (7)—H. D. DeVoe, Box 644, Anchorage, Alaska; J. T. Little, Star Route B, Box 3453, Spenard, Alaska.

Atlanta (3)—B. E. Montgomery, 6520 Bridgewood Valley, Atlanta 5, Ga.; R. S. Duggan, Jr., 1112 Mason Woods. Dr., N.E., Atlanta 6, Ga.

Baltimore (3)—J. T. Wilner, 2610 N. Charles St., Station WBAL, Baltimore 18, Md.; E. B. Cole, Jr., 103 Glenwood Rd., Bel Air, Md.

Bay of Quinte (8)—M. B. Broughton, Elec. Engrg. Dept., Royal Military College, Kingston, Ont., Canada; E. E. Pascal, 45 Benjamin St., STE-32, Belleville, Ont., Canada.

Beaumont-Port Arthur (6)—H. W. Bartlett, 3790 Ironton, Beaumont, Tex.; R. Cunningham, Rte. 1, Box 181, Kirbyville, Tex.

Benelux—Herre Rinia, Parklaan 24, Eindhoven, Netherlands; G. J. Siezen, 3 Konig Willem III Laan, Blaricum, Holland.

Binghamton (1)—F. W. Schaaf, R.D. 1, Apalachin, N. Y.; D. W. Deno, 3149 Briar Cliff Ave., R.D. 1, Vestal, N. Y.

Boston (1)—G. P. McCouch, Alcott Rd., Concord, Mass.; R. E. Scott, 114-H Northeastern Univ., Boston 15, Mass.

Buenos Aires—J. M. Barcala, 71-385 La Plata, FCNGR, Republica, Argentina; E. J. Kirkscether, Arenales 3598-A, Buenos Aires, Argentina.

Buffalo-Niagara (1)—D. P. Sante, 4530 Greenbriar Rd., Williamsville 21, N. Y.; F. A. Timberlake, 169 Greenhaven Terrace, Tonawanda, N. Y.

Canaveral (3)—C. E. Mattox, 209 Beverly Rd., Cocoa, Fla.; W. J. Haberhern, Jr., 118 Atlantic Blvd., Eau Gallie, Fla.

Cedar Rapids (5)—S. M. Morrison, 2034 Fourth Ave., S.E., Cedar Rapids, Iowa; A. E. Anderson, 1054 34 St., N.E., Cedar Rapids, Iowa.

Central Pennsylvania (4)—W. J. Ross, 105 Elec. Engrg. Dept., University Park, Pa.; J. L. Brown, Jr., Ordnance Research Lab., P.O. Box, 30 State College, Pa.

Chicago (5)—W. B. McClelland, 1601 Cherry St., Park Ridge, Ill.; P. D. Huston, 233 S. Villa Ave., Elmhurst, Ill.

Chile—J. del Rio, Corp. de Radio de Chile, Casilla 1407, Santiago, Chile; J. de la Quintana, IBM World Trade Corp., Cassila 3630, Santiago, Chile.

China Lake (7)—D. W. Kermode, USNOTS, P.O. Box 577, China Lake, Calif.; D. P. Kasper, 1803-B Young Circle, China Lake, Calif.

Cincinnati (4)—A. J. Bissonette, 680 Tyler, Milford, Ohio; J. W. Murray, 5603 Green Acres Court, Cincinnati 11, Ohio.

Cleveland (4)—R. Kazarian, 2031 Lakeview Ave., Rocky River, Cleveland 16, Ohio; J. I. Sheetz, Ohio Bell Telephone Co., 700 Prospect Ave., Rm. 704, Cleveland 15, Ohio.

Colombia—W. Westphal, Carrera Fa, #40-62, Bogota D. E., Colombia, S. A.; B. Alzate, Calle 48, #23-26, Apt. 401, Bogota, Colombia, S. A.

Columbus (4)—F. A. Wolf, 354 Chatham Rd., Columbus 14, Ohio; E. N. Wyler, 109 W. Stafford, Worthington, Ohio.

Connecticut (1)—S. C. Wadhams, 13 Derwen St., Glenbrook, Conn.; P. N. Johnstone, Twin Lakes Rd., North Branford, Conn.

Dallas (6)—T. Prickett, Jr., 805 Clearwater Dr., Richardson, Tex.; C. K. Hager, 2709 Greenbrook Dr., Garland, Tex.

Dayton (4)—R. J. Framme, 316 Claranna, Dayton 9, Ohio; S. Weissman, 805 Eppington Dr., N., Trotwood, Ohio.

Denver (6)—H. S. Johnson, 1241 Monaco Pkwy., Denver 20, Colo; F. P. Venditti. Denver Res. Inst., Univ. of Denver, Denver 10, Colo.

Detroit (4)—D. V. Stocker, Dept. of Elec. Engrg., Wayne Univ., Detroit 2, Mich.; L. E. Barber, 402 N. Troy St., Royal Oak, Mich.

Egypt—H. M. Mahmoud, 24 Falaki St., Cairo, Egypt; El-Garhi I. El-Kashlan, 46 Suez-Kanal St., Engineering City, Imbaba, Cairo, Egypt.

Elmira-Corning (1)—T. G. Keeton, 611 McDowell Place Elmira, N. Y.; R. K. Ainey, Westinghouse Elec. Corp., Box 284, Elmira, N. Y.

El Paso (6)—A. M. Rojas II., 2326 Copper St., El Paso, Tex.; C. W. Brown, 325 Encino Drive, El Paso, Tex.

Emporium (4)—H. W. Herbert, Sylvania Electric Products, Inc., Emporium, Pa.; R. F. Bergdahl, R.F.D. 2, Emporium, Pa.

Erie (1)—R. S. Page, 1224 Idaho Ave., Erie, Pa.; K. I. Challstrom, Buffalo Rd., R.D. 1, Harborcreek, Pa.

Evansville-Owensboro (5)—K. G. Miles, 2105 East Gum St., Evansville 14, Ind.; F. R. Snyder, 2521 Iroquois Dr., Owensboro, Ky.

Florida West Coast (3)—R. E. Henning, Sperry Microwave Electronics Co., Box 1828, Clearwater, Fla.; C. W. Benfield. 605 First St., S.E., Largo, Fla.

Fort Huachuca (7)—J. W. Virden, 102 Crandel St., Fort Huachuca, Ariz.; L. J. Knight, 135 Jeffords St., Fort Huachuca, Ariz.

Fort Wayne (5)—C. F. Mason, 1519½

* Numerals in parentheses following sections designate region number. First name designates Chairman; second name Secretary.

FABRICATING MICROMODULES TO MOBILE ROOMS

Magnetic Shields Custom Fabricated to Any Size or Shape

The industry's widest range of production facilities . . . in 3 factories:

PRESS FORMING AND DRAWING • HYDROFORMING • SPINNING • HAND FABRICATION OF PROTOTYPES

Already Tooled for Samples or Production

- Simple designs to elaborate complexes.
- Single or multiple laminae constructions.
- Micromodule size to a completely prefabricated mobile shielded room weighing 5 tons which can be transported anywhere — an achievement unique with Magnetic Shield Division.*

ALSO AVAILABLE

- Processed stock for fabrication in your own plant.
- Co-operative design facilities.

Use this SINGLE CONVENIENT SOURCE for all magnetic shielding requirements. Saves you countless design hours . . . helps speed your project . . . lowers your costs.

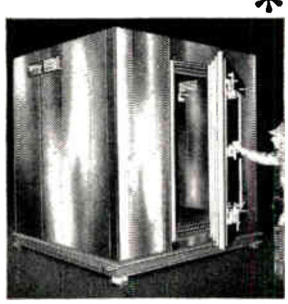
We recommend NETIC and CO-NETIC magnetic shielding materials because they are non-shock sensitive, non-retentive, do not require periodic annealing and provide completely effective shielding for optimum results.

NETIC and CO-NETIC are widely specified for satellite, missile, protecting recording tapes, data processing and for innumerable other military, scientific and laboratory applications as well as for commercial applications.

*

PHONE YOUR NEAREST SALES OFFICE TODAY:

BALTIMORE, MARYLAND HOpkins 7-3766
UNION CITY, NEW JERSEY UNion 4-9577
MERIDEN, CONNECTICUT BEverly 7-9232
MIAMI, FLORIDA HIghland 4-1118
WESTMOUNT, MONTREAL, QUEBEC WEllington 7-1167
DALLAS, TEXAS FLeetwood 1-1615
LOS ANGELES, CALIFORNIA WEbster 1-1041
PALO ALTO, CALIFORNIA DAvenport 1-5064
SAN DIEGO, CALIFORNIA ACademy 4-1717
SEATTLE, WASHINGTON EAst 3-8545
PHOENIX, ARIZONA AMhurst 4-4934
HOUSTON, TEXAS HOmestead 5-7780

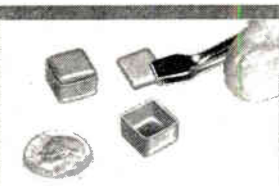


MAGNETIC SHIELD DIVISION

Perfection Mica Company

1322 N. ELSTON AVENUE, CHICAGO 22, ILLINOIS

ORIGINATORS OF PERMANENTLY EFFECTIVE NETIC CO-NETIC MAGNETIC SHIELDING

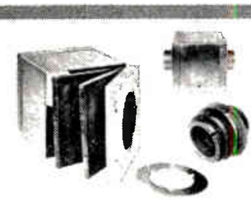


Micro miniature shield and cover, punch press operation.

Sequence of shield cans, punch press or spinning.

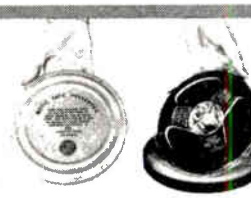
Complex configuration multi-lamina shield, hydroformed.

CRT shield illustrating combination of hand fabrication, spinning and sizing.

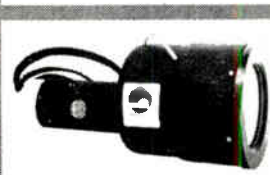


Backward wave tube shield assembly design, involving hand fabrication and hydroform or spinning.

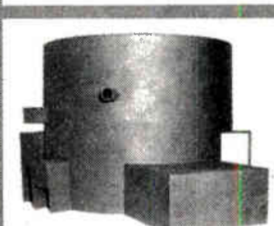
Special purpose shield, hand fabrication (levitated gyro).



Tape preserver can be spun, hydroformed or punch press fabricated.



Data storage tube shield, hydroform or spinning, plus hand fabrication.



Large fabricated special structure (shaker table shield), approx. 60" dia. and 57" high.

Composite photo demonstrating that magnetic shielding qualities of NETIC alloy material are not affected by vibration, shock, (including dropping), etc.

Crescent Ave., Fort Wayne, Ind.; D. W. Von Berg, 4015 Wedgewood Dr., Fort Wayne, Ind.

Fort Worth (6)—S. T. Lanham, Sr., 5017 Vicki, Fort Worth 17, Tex.; M. F. Shepherd, 2324 Clearview, Fort Worth, Tex.

France—J. R. Pernice 15 Rue de la Faisanderie, Paris 16, France; J. D. Lebel, Lab. de Physique, Ecole Normal Supereure, 24 Rue Lhomond, Paris 5, France.

Gainesville (3)—A. D. Sutherland, 1626 N.W. 14 Ave., Gainesville, Fla.; A. T. Raczynski, Sperry Elec. Tube Div., Gainesville, Fla.

Geneva—G. C. Gross, I.T.U., Palais Wilson, Geneva, Switzerland; Secretary to be advised.

Hamilton (8)—F. H. Edwards, United-Carr Fastener Company of Canada, Ltd., 231 Gage Ave., N., Hamilton, Ont., Canada; C. C. Minchel, Canada Westinghouse Co., Elec. Tube Div., Plant 2, Hamilton, Ont., Canada.

Hawaii (7)—E. A. Piety, 2030 Home Rule St., Honolulu 17, Hawaii; G. H. Stagner, 525 Lunalilo Home Rd., Honolulu 16, Hawaii.

Houston (6)—L. A. Mullan, Radio Station KTRH, Rice Hotel, Houston 2, Tex.; J. A. Beyer, 5221 Lincrest Lane, Houston 27, Tex.

Huntsville (3)—D. E. Rowe, 8602 Camille Dr., S.E., Huntsville, Ala.; J. A. Power, 4003 Marie St., Apt. 1, Huntsville, Ala.

India—Officers to be advised.

Indianapolis (5)—S. M. Stuhlbarg, 4904 Staughton Dr., Indianapolis, Ind.; R. C. Olsen, 4040 Adams Court, N., Indianapolis, Ind.

Israel—Chairman to be advised; H. Fischler, Weizmann Inst. of Science, Dept. of Electronics, Box 26, Rehovoth, Israel.

Italy—A. Marino, Via Guido d'Arezzo 14, Rome, Italy; G. P. Tarchini, Laboratorio Ricerche Elettr., Via de Parlamento N. 33, Borgolombardo, Milan, Italy.

Ithaca (1)—H. L. Heydt, GE Advanced Elec. Ctr., Tomkins County Airport, Ithaca, N. Y.; M. E. Hollar, Advanced Electronics Ctr., G.E. Co., Ithaca, N. Y.

Kansas City (6)—F. A. Spies, Bendix Aviation Corp., Box 1159, Kansas City 41, Mo.; K. D. Benedict, 1032 East Alton, Independence, Mo.

Kitchener-Waterloo (8)—B. S. McConachie, Raytheon Canada, Ltd., 61 Laurel St., E., Waterloo, Ont., Canada; W. J. Grey, Raytheon Canada, Ltd., Box 8, 400 Phillips St., Waterloo, Ont., Canada.

Las Vegas (7)—J. H. Jurmain, Edgerton, Germeshausen, Grier, 300 Wall St., Box 1912, Las Vegas, Nev.; L. S. Kreyer, 617 Twin Lakes Dr., Las Vegas, Nev.

Little Rock (6)—R. V. Anders, 5801 Windamere Dr., Little Rock, Ark.; W. C. Allsopp, Jr., Ark. Industrial Electric Co., 5020 Club Rd., Little Rock, Ark.

London (8)—T. W. W. Stewart, 1 Susan Ave., London, Ont., Canada; R. W. Johnson, 9 Ingleside Place, London, Ont., Canada.

Long Island (2)—J. W. Kearney, Airborne Instruments Lab., Walt Whitman Rd., Melville, L. I., N. Y.; C. H. Doersam, Jr., Instruments for Industry, 101 New South Rd., Hicksville, L. I., N. Y.

Los Angeles (7)—J. J. Guarrera, 17160 Gresham St., Northridge, Calif.; E. E. Ingebretsen, 15435 Tupper St., Sepulveda, Calif.

Louisville (5)—H. T. Smith, Kentucky Rural Elec. Co-op Corp., 4515 Bishop Lane, Louisville 18, Ky.; E. L. Marven, Sr., 811 Camden Ave., Louisville 15, Ky.

Lubbock (6)—H. A. Spuhler, Texas Tech. College, Dept. of E.E., Lubbock, Tex.; E. F. Kilcrease, 2114 38 St., Lubbock, Tex.

Miami (3)—G. A. Liedholz, 6060 S.W. 114 St., Miami 56, Fla.; J. R. Haeger, 1240 S. Alhambra Circle, Coral Gables 46, Fla.

Milwaukee (5)—T. H. Houle, 1000 S. 56 St., West Allis 14, Wis.; K. J. Schlager, 1519 N. 118 St., Wauwatosa 13, Wis.

Mobile (3)—C. R. Hicks, 4128 N. Marietta Dr., Mobile, Ala.; E. G. Schone, Box 4102, Mobile, Ala.

Montreal (8)—H. H. Schwartz, 5212 King Edward Ave., Montreal 29, Que., Canada; A. Breton, Ecole Polytechnique, 2500 Marie Guyard Ave., Montreal 26, Que., Canada.

Newfoundland (8)—Chairman to be advised; F. A. Mulloy, 55 Calver Ave., St., John's Newf., Canada.

New Orleans (6)—G. Allen, 4239 St. Charles Ave., Apt. C, New Orleans, 15, La.; N. R. Landry, 6079 Louisville St., New Orleans 24, La.

New York (2)—H. W. Pollack, 300 Lennox Road, Brooklyn, N. Y.; E. Schutzman, College of Engrg., New York University, University Heights, New York 53, N. Y.

North Carolina (3)—W. E. Lanford, 3260 Nottingham Rd., Winston, Salem, N. C.; J. I. Barron, Southern Bell T and T Co., Box 240, Charlotte, N. C.

Northern Alberta (9)—L. N. Donovan, 24 Birch Dr., Box 136, St. Alberta, Alta., Canada.; R. A. McLean, 9516 69 A St., Edmonton, Alberta, Canada.

Northern New Jersey (2)—H. S. Wertz, 111 Parsonage Rd., Short Hills, N. J.; M. E. Levine, Tung-Sol Electric, Inc., 545 N. Arlington Ave., East Orange, N. J.

Northwest Florida (3)—F. J. Sites, 29-B Boatner Dr., Eglin Village, Eglin AFB, Fla.; H. G. Wise, Box 878, Fort Walton Beach, Fla.

Oklahoma City (6)—E. Black, 1701 N.W. 35 St., Oklahoma City 18, Okla.; J. M. Ashworth, 3934 N.W. 34 St., Oklahoma City 12, Okla.

Omaha-Lincoln (5)—J. A. Rogers, 3815 Marcy St., Omaha 5, Neb.; C. Hyde, 312 Ferguson Hall, Univ. of Neb., Lincoln 8, Neb.

Orlando (3)—J. R. Matzinger, 1825 Whitehall Dr., Winter Park, Fla.; C. E. McGinnis, 1720 Shawnee Trail, Maitland, Fla.

Ottawa (8)—W. R. Conway, 2075 Woodcrest Rd., Ottawa 1, Ont., Canada; S. G. Jones, 1910 Haig Dr., Ottawa 1, Ont., Canada.

Philadelphia (3)—R. M. Showers, Moore School of Elec. Engrg., 200 S. 33 St., Philadelphia 4, Pa.; P. A. Lathrop, 19 Sweetgum Rd., Levittown, Pa.

Phoenix (7)—H. R. Hyder, III., 2523 N. 57 St., Phoenix, Ariz.; R. K. Peterson, 2034 W. Rancho Dr., Phoenix 15, Ariz.

Pittsburgh (4)—R. I. Van Nice, 112 Delaware Court, Glenshaw, Pa.; R. P. O'Shea, Rear 125 Woodvale Ave., Johnstown, Pa.

Portland (7)—M. L. Morgan, 12907 S.W. 62 Ave., Portland 19, Ore.; C. H. Moulton, 1625 S.W. 87 Ave., Portland 1, Ore.

Princeton (2)—R. D. Lohman, RCA, Somerville, N. J.; O. E. Dow, RCA Labs., Princeton, N. J.

Quebec (8)—R. M. Vaillancourt, 638 Ave. Mon Repos, Ste. Foy, Quebec 10, Que., Canada; K. A. Laurie, 775 Calixa LaVallee, Apt. 15, Quebec 6, Que., Canada.

Regina (8)—S. K. Smith, c/o Saskatchewan Govt. Tel., 2350 Albert, Regina, Sask., Canada; G. H. Beuker, 2901 Athol St., Regina, Sask., Canada.

Rio de Janeiro—J. A. Wiltgen, Caixa Postal 450, Rio de Janeiro, DF, Brazil; C. J. Chapin, c/o Riggs National Bank, DuPont Circle Branch, 1913 Massachusetts Ave., N.W., Washington 6, D. C.

Rochester (1)—J. L. Wheeler, 535 Rondo Lane, Webster, N. Y.; D. W. Healy, Jr., Univ. of Rochester, River Campus Station, Rochester 20, N. Y.

Rome-Utica (1)—R. A. Zachary, Jr., 11 Arbor Dr., New Hartford, N. Y.; C. J. Civin, 4 Colonial Dr., New Hartford, N. Y.

Sacramento (7)—A. O. Rohde, 3160 Adams Rd., Sacramento 25, Calif.; J. C. Bissett, 8717 Jonnie Way, Fair Oaks, Calif.

St. Louis (6)—G. E. Barnard, 639 N. 69 St., East St. Louis, Ill.; E. A. Kuhlman, 9020 Cardinal Ter., Brentwood 17, Mo.

Salt Lake City (7)—A. W. Vodak, 2587 Sherwood Dr., Salt Lake City, Utah.; H. J. Redd, 1771 Severn Dr., Salt Lake City 17, Utah.

San Antonio-Austin (6)—W. L. Donaldson, 129 El Cerrito Circle, San Antonio 12, Tex.; C. R. Graf, 207 Addax Dr., San Antonio 1, Tex.

San Diego (7)—E. W. Carlson, 3154 Bremerton Pl., La Jolla, Calif.; A. C. Greeley, 2615 Kim Pl., San Diego 11, Calif.

San Francisco (7)—S. F. Kaisel, Microwave Electronics Corp., 4061 Transport St., Palo Alto, Calif.; Acting Secretary: A. T. Waterman, Jr., Electronics Research Lab., Stanford University, Stanford, Calif.

Schenectady (1)—T. G. Mihran, G.E. Research Lab., Box 1088, Schenectady, N. Y.; F. J. Ellert, G.E. Co., Bldg. 37, Rm. 578, 1 River Rd., Schenectady, N. Y.

Seattle (7)—W. J. Siddons, 6539 39 St., N.E., Seattle 15, Wash.; W. L. Green, 7202 N. Mercer Way, Mercer Island, Wash.

Shreveport (6)—E. J. Culling, 3252 Sarah St., Bossier City, La.; E. C. Strickland, 2914 Bolch St., Shreveport, La.

South Bend-Mishawaka (5)—H. W. Vogtmann, Bendix Mishawaka Div., 400 S. Beiger St., Mishawaka, Ind.; N. O. Kindt, 50635 Dresden Dr., South Bend 17, Ind.

South Carolina (3)—J. Taylor, 6417 Westshore Rd., Columbia, S. C.; J. Bouvy, 4700 Oakwood Rd., Columbia, S. C.

Southern Alberta (8)—R. E. Smith, 1507–20 A St., N.W., Calgary, Alta., Canada; G. E. Parkinson, 735 35 St., N.W., Calgary, Alta., Canada.

Syracuse (1)—G. F. Platts, 101 Iroquois Lane, Liverpool, N. Y.; G. M. Kirkpatrick, 202 David Dr., N. Syracuse 12, N. Y.

Tokyo—Miyaji Tomota, Yokogawa Elec. Works, Ltd., 3000 Kichijoji, Musashino-Shi, Tokyo, Japan; Fumio Minozuma, 16 Ohara-Machi, Meguro-Ku, Tokyo, Japan.

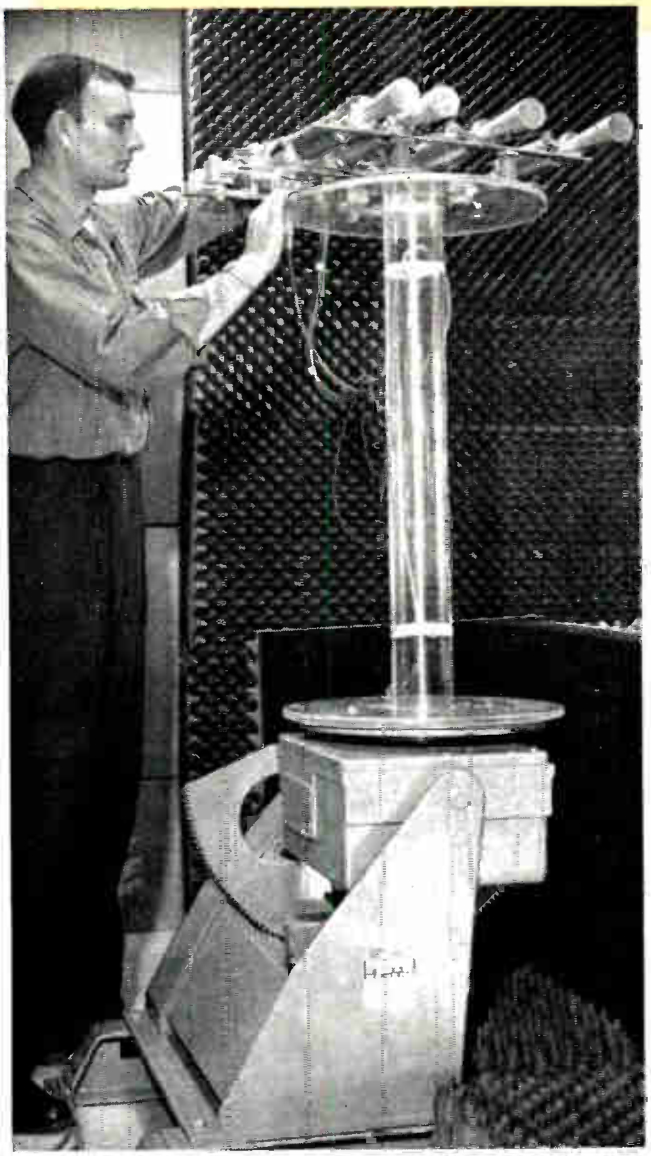
Toledo (4)—R. B. Williams, Jr., 5945 Summit St., Sylvania, Ohio; H. R. Holmes, 3557 149 St., Toledo, Ohio.

Toronto (8)—G. T. Quigley, Philips Indus-

AT TEXAS INSTRUMENTS:

New capabilities in antenna research

WITH THE AID OF SCIENTIFIC-ATLANTA EQUIPMENT



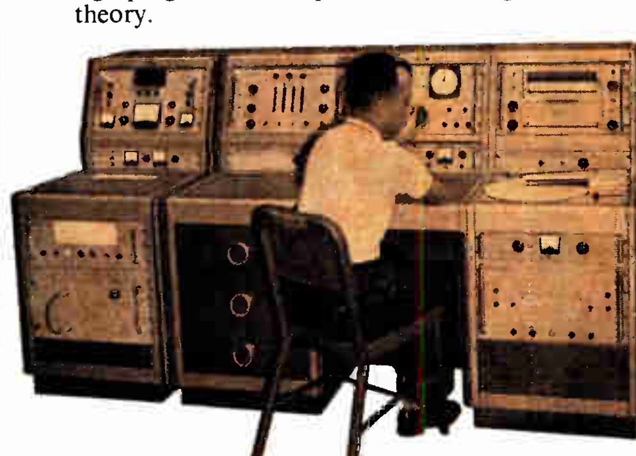
Antenna positioner installed in Texas Instruments' anechoic chamber (microwave darkroom) is one part of the Scientific-Atlanta instrumented facility.

Recently the Apparatus Division of Texas Instruments Incorporated built a new antenna research laboratory to allow more advanced work on a wide range of applications. These include radar equipment and systems for airborne early warning, submarine detection, and airport and airborne surveillance. Scientific-Atlanta supplied the antenna pattern instrumentation consoles and positioning equipment which includes the following major items:

a wide range, low noise receiving system (model 402A); an antenna pattern recording system (model APR 20/30); an antenna pattern integrator (model API-1); and an azimuth over elevation antenna positioner (model PAE 5).

This instrumentation, along with signal sources from S through K_u bands, a complete line of standard test equipment, an anechoic chamber, an inside test range, and 150 acres of outside test ranges, gives Texas Instruments one of the most sophisticated antenna test facilities in the country.

Texas Instruments engineers report that their Scientific-Atlanta equipment enables them to make quick, accurate and reliable one-way pattern/gain measurements, antenna backscatter experiments, and studies of advanced methods of electronic scanning, all contributing to missile antenna design programs and experiments relating to antenna theory.



Texas Instruments' installation uses a completely equipped Scientific-Atlanta console.

For more details on the Texas Instruments facility and how Scientific-Atlanta can design, construct and install an antenna test facility that suits *your* needs, please write to:

SCIENTIFIC-ATLANTA, INC.

2162 Piedmont Road, N.E. Atlanta 9, Georgia Phone: 875-7291

Challenging opportunities for electronic engineers with five to ten years' circuit design experience. Please send complete resume to Personnel Manager. An equal opportunity employer.

16

IRE People

Univac Military Operations, St. Paul, has hired two new engineers to supervise the electronic design of advanced micro-miniature aerospace computers.

W. C. Timm

T. M. Sammis

They are: **Theodore M. Sammis,** (SM'61), supervising development engineer, and **Walter C. Timm** (S'53–A'55–M'60) principal development engineer, both former senior group engineers at McDonnell Aircraft Corp., St. Louis, Mo.

Mr. Sammis supervised the design of miniature aerospace digital equipment at McDonnell. He previously was a supervising engineer at Hazeltine Electronics Corp., Little Neck, N. Y.

Before joining McDonnell, Mr. Timm was on the technical staff of Bell Telephone Laboratories, Whippany, N. J. He is a member of the American Institute of Electrical Engineers.

❖

Joseph D. Schantz (S'33–A'35–SM'44) has been appointed manager of the ZMAR program at the eastern operation of Sylvania Electronic Systems, Waltham, Mass., a major division of Sylvania Electric Products, Inc.

J. D. Schantz

Since July 1960, Mr. Schantz has served as Director of Engineering at the division's Waltham Laboratories, which recently combined with the Data Systems Operations at Needham, Mass., to form the eastern operation. Previously, he served successively as Chief Electronics Engineer, Assistant Manager of the avionics division, and finally as Director of Plans and Programs for Bell Aircraft Corp., Buffalo, N. Y. Mr. Schantz also served as a research engineer with RCA and as manager of the military projects division of Farnsworth Television and Radio Corporation.

He received his Bachelor's degree in electrical engineering from Gettysburg College and his Master's degree from the University of Michigan.

John D. Smith (M'59) has been appointed head of the Mechanical Engineering Department at the Data Systems Division of Litton Systems, Canoga Park, Calif.

Mr. Smith was graduated from the University of Denver in 1941 with a B.S. degree in electrical engineering, and from the University of Southern California in 1955 with an M.S. degree in electrical engineering.

In the newly-created post Mr. Smith will be responsible for mechanical engineering for the advanced command and control systems developed and manufactured by the division. Before joining Litton systems he was manager of the displays laboratory at the Hughes Ground Systems Division, Fullerton.

Mr. Smith is a member of the AIEE.

❖

It was recently announced that **Dr. Thomas E. Tice** (S'46–A'50–SM'55–F'61) has been appointed to the position of Chief Engineer of the Antenna and Microwave Group, at Motorola Inc., Military Electronics Division, Scottsdale, Ariz.

T. E. Tice

Dr. Tice, a native of Florence, Ala. was prior to this appointment the director of Ohio State University's Antenna Laboratory, Columbus, Ohio, and was responsible for research in microwave circuits, wave propagation, antennas, and electromagnetic field theory. Concurrently, Dr. Tice held the rank of Professor, OSU Department of Electrical Engineering.

He was a former communications officer with the U. S. Army Signal Corps and was responsible for the installation and maintenance of Army airways communications equipment and air navigational aids in North Africa and Arabia.

Dr. Tice served as Editor-in-Chief for McGraw-Hill's "Techniques for Airborne Radome Design" in 1957.

Dr. Tice is a member of Sigma Xi, Tau Beta Pi, Eta Kappa Nu, and Chi Beta Phi. He is also a member of the Air Force Advisory Group on Passive Satellite Communications, the IRE Seventh Region Student Paper Competition Committee, and the IRE Phoenix Section Awards Committee.

❖

Eugene F. Shell (M'51) of Endicott, N. Y., has been promoted to Assistant Systems Manager at IBM's Space Guidance

(Continued on page 36A)

A SIGNIFICANT BREAK-THROUGH IN TRANSISTOR TECHNOLOGY...

ECDC*

The Best High-Power, High-Speed Switching Transistors Ever Developed Provide the Optimum Combination of Voltage, Power, and Speed

The Sprague ECDC Transistor is the first to combine the optimum features of the electro-chemical precision-etch techniques and diffused collector techniques in one highly-mechanized process.

The ECDC Transistor meets these 7 conditions for an "ideal" transistor:

1. Very low collector-to-case thermal resistance through the utilization of high thermal conductivity material as the collector body, resulting in high power dissipation.
2. Thin base width for high radiation resistance and lower storage time.
3. Precision-etched emitter pit permits placement of emitter junction at proper resistivity for optimum breakdown voltage and frequency response.
4. High conductivity surface surrounds emitter pit and close emitter-to-base spacing results in extremely low base resistance.
5. Precision-etched collector provides optimum control of collector series resistance with attendant low saturation voltage, low storage time, and high breakdown voltage.
6. Low collector series resistance as a result of the use of high conductivity material for the mass of the collector area.
7. The structure and manufacturing processes are suited for automated production equipment with immediate in-process feedback.

*Trademark of Sprague Electric Co.

TYPES 2N2099 & 2N2100

These P-N-P Germanium Electro-Chemical, Diffused-Collector Transistors are especially designed for core and film memory driver applications. They feature excellent beta linearity from less than 1 ma to over 400 ma, high frequency response, and low saturation resistance. The low-height TO-9 case is ideally suited to meet equipment designers' needs.

TYPES 2N2096 & 2N2097

Types 2N2096 and 2N2097 are electrically identical to Types 2N2099 and 2N2100, respectively, except for their TO-31 Case, with its threaded stud mounting.

For complete information on ECDC Transistors, write Product Marketing Section, Transistor Division, Sprague Electric Company, Concord, New Hampshire.

SPRAGUE COMPONENTS

TRANSISTORS	INTERFERENCE FILTERS	HIGH TEMPERATURE MAGNET WIRE
CAPACITORS	PULSE TRANSFORMERS	CERAMIC-BASE PRINTED NETWORKS
MAGNETIC COMPONENTS	PIEZOELECTRIC CERAMICS	PACKAGED COMPONENT ASSEMBLIES
RESISTORS	PULSE-FORMING NETWORKS	FUNCTIONAL DIGITAL CIRCUITS
MICROCIRCUITS	TOROIDAL INDUCTORS	ELECTRIC WAVE FILTERS

SPRAGUE®
THE MARK OF RELIABILITY

IRE People

(Continued from page 34A)

Center in Owego.

E. F. Shell

Mr. Shell joined IBM in May, 1952 after having served as an engineer with the Air Force in a civilian capacity at Wright Patterson Air Force Base, Dayton, Ohio.

His IBM assignments have been in design coordination and development engineering in the Owego bomb navigation programs. He was one of the original group assigned to manage the advanced system study in October, 1955 that led to development of the B-70 ASQ-28 Bombing Navigational Missile Guidance System.

In May, 1958 he was named Manager of the ASQ-28 Project Office and in April 1960 was appointed Functional Manager of the ASQ-38 Project Office.

❖

S. Paul Shackleton (M'31–SM'43), engineering consultant, has joined the staff of ITT Communication Systems, Inc., Paramus, N. J.

S. P. Shackleton

Since his retirement from Bell Telephone Laboratories, Inc., he has served as consultant to NSA and the President's Committee on Scientists and Engineers in Washington and to RCA in New York. He has also been employed as assistant to the Director of Research at Maxson Electronics and with Lehigh Design Co., Inc. of Arcs Industries. He is presently working on worldwide communication systems under DOD contracts at ITT Communications Systems, Inc.

❖

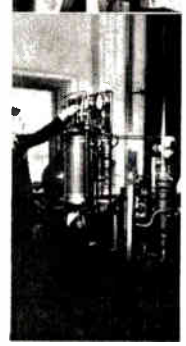
Tom Tracy (M'55) has been named Marketing Manager of Datamec Corporation, Sunnyvale, Calif.

T. Tracy

Mr. Tracy formerly was Northwest District Manager of Ampex Instrumentation Products, with headquarters in Palo Alto. Before joining Ampex in 1959, he was associated with the instrument, transistor and data handling divisions of Minneapolis-Honeywell for 10 years as a

(Continued on page 40A)

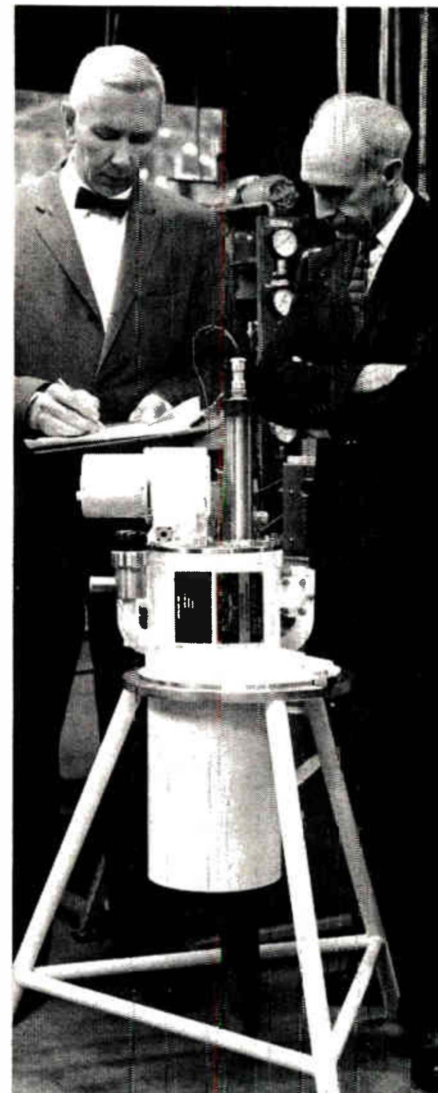




THE CRYODYNE® HELIUM REFRIGERATOR

a new milestone

now from ADL



Albert Hatch (left) and Dr. Howard McMahon discuss the results of a test run. Dr. McMahon and William E. Gifford, now Professor of Mechanical Engineering at The University of Syracuse, were co-inventors of the refrigeration cycle used in the CRYODYNE Refrigerator. Albert Hatch is in charge of engineering development of the Refrigerator at ADL.

Closed-cycle refrigeration at liquid helium temperatures is now available in proven hardware. Arthur D. Little, Inc. — under the initial sponsorship of International Business Machines Corporation and continuing sponsorship by the U.S. Army Ordnance Corps through Bell Telephone Laboratories, Inc. — has extended the development of its patented Gifford-McMahon cycle to produce the ADL CRYODYNE HELIUM REFRIGERATOR. The Refrigerator will provide up to 250 milliwatts of refrigeration at liquid helium temperatures and will function normally regardless of physical orientation. A number of units were completed in 1961 and by December 31st had accumulated, in aggregate, more than 10,000 operating hours. A CRYODYNE Refrigerator on one endurance test run exceeded 1500 hours of continuous operation without maintenance or control manipulation.

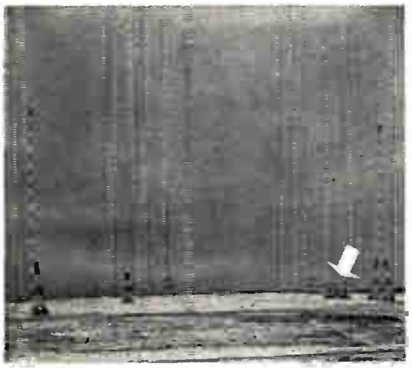
Because of its demonstrated reliability, its compact size and simplicity of design, the CRYODYNE Refrigerator lends itself to a wide range of application in cooling superconductive, quantum electronic, and IR devices. For complete technical data and price information, write: CRYOGENIC EQUIPMENT, Arthur D. Little, Inc., 20 Acorn Park, Cambridge 40, Massachusetts.

APPLIED SCIENCE • RESEARCH EQUIPMENT
ENGINEERING • MANAGEMENT CONSULTING

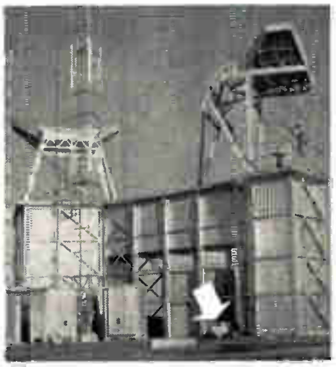
RESEARCH FOR INDUSTRY

Arthur D. Little, Inc.

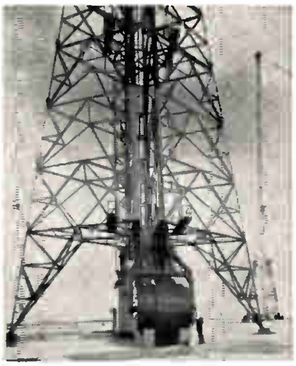
NOW OPERATIONAL...ONE YEAR AHEAD OF SCHEDULE...WOR



2800 acres . . . an area greater than two dozen Pentagon Buildings . . . two identical antenna arrays . . . center towers nearly as high as the Empire State Building support the gigantic spider web of steel towering a thousand feet up and embracing two square miles . . . nearly an entire peninsula at Cutler, Maine. (Arrow indicates comparative size of Helix House to tower.)

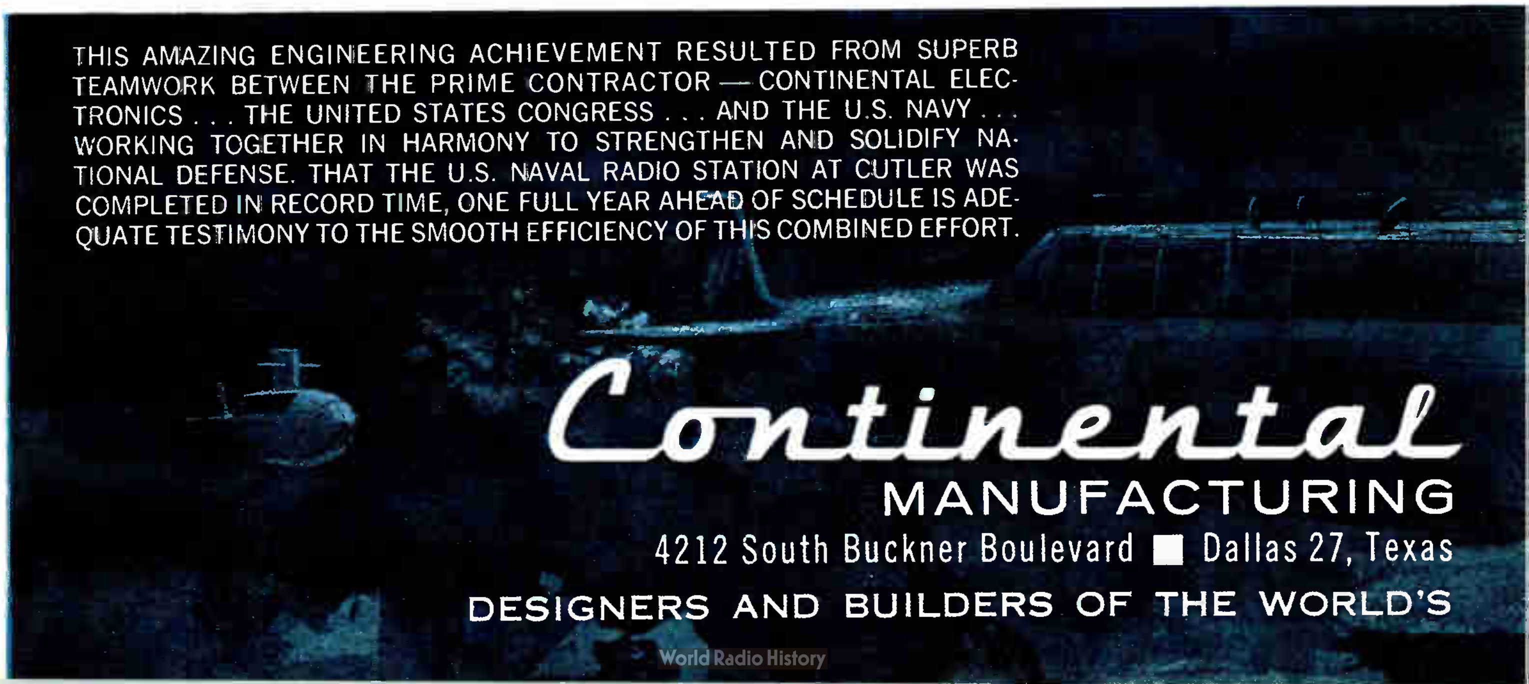


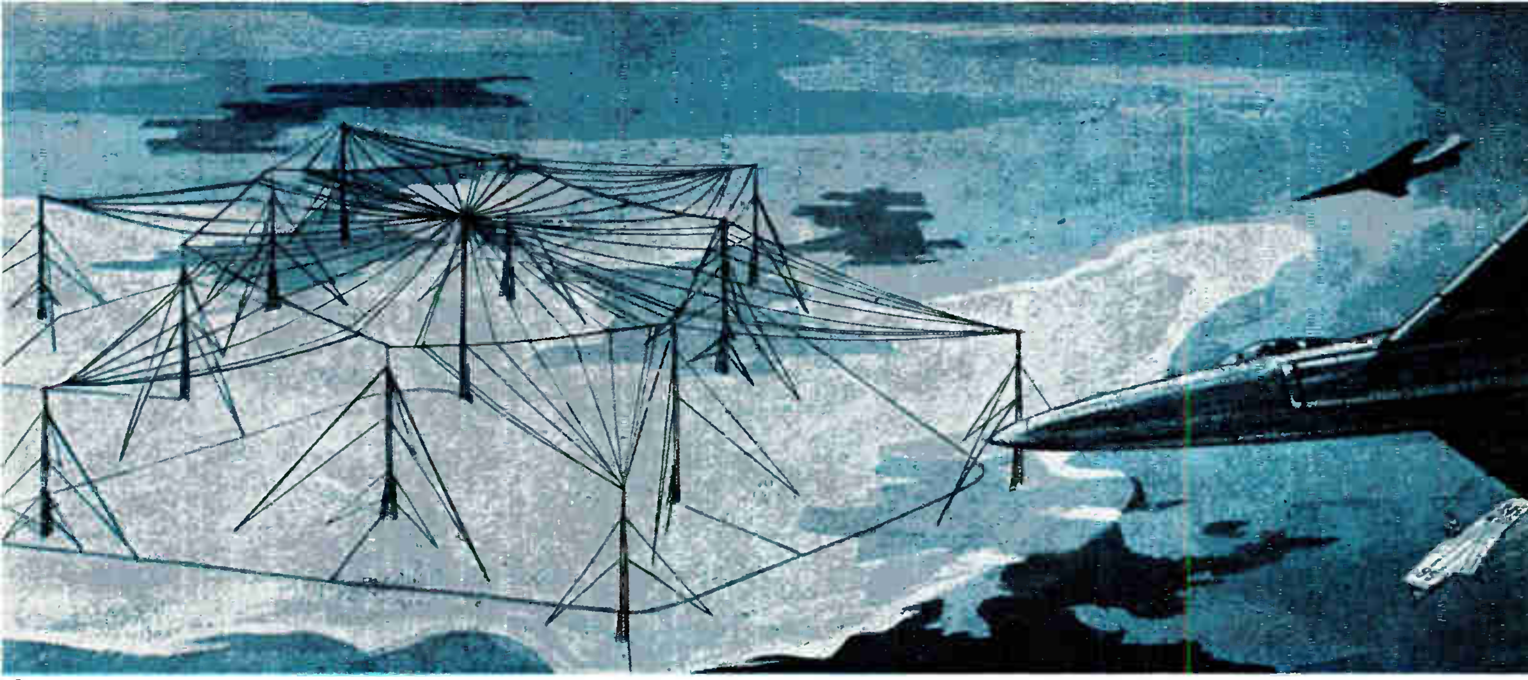
(Arrow points to truck. Compare Helix House size in first photo.) 8-story Helix House contains antenna coupling and automatic de-icing equipment to rid the immense antenna system of ice. Buried beneath the ground: another 11 million feet of copper wire in the radiating system terminating in the sea water itself.



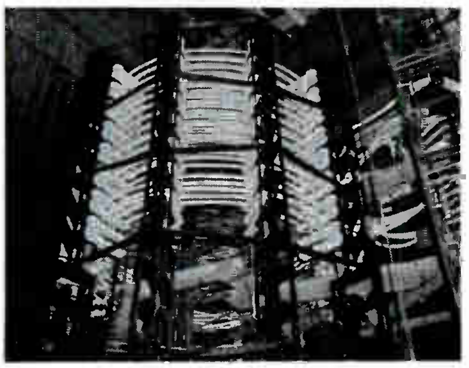
42 counter-weight towers—36 of them like this—carrying tremendous counter balances of 202 tons each to maintain and correct antenna tension and strain from winds up to 150 knots or ice forming on the 64 miles of bronze antenna.

Enormous variometer coil for inductance to tune the antenna system through a range of 14 to 30 KC . . . very low frequency. These VLF radio waves penetrate the depths of the sea to submerged submarines.





'S MOST POWERFUL TRANSMITTER...2,000,000 WATTS VLF



Huge Helix coil 20 feet in diameter and 40 feet tall is wound with 3½ inch Litz Wire . . . just one of the scores of huge components that combine to give this new communication station maximum power . . . range . . . reliability . . . and the special penetration possibilities VLF possesses that no normal high frequency radio provides.

Control console and portion of the unique CEMC Type-125 2,000,000 watt VLF Transmitter that propagates along the curvature of the earth instead of bouncing off the IONOSPHERE: thus eliminating dead communication areas or skip distances to give this Naval voice of command greater range and improved reliability.

In one instant 2,000,000 watts will blast the U. S. Navy's voice of command around the world. At the control console, during operation, push-button simplicity provides a new and highly reliable major element of command to the U. S. Navy . . . another element of that might by which the nation promotes the keeping of the peace.

IRE People

(Continued from page 36A)

sales engineer and sales supervisor.

Mr. Tracy was graduated from the California Institute of Technology with a B.S. degree in Applied Physics.

❖

The Polytechnic Institute of Brooklyn announced the appointment of **Dr. John G. Truxal** (S'47–A'48–SM'54–F'59) to a vice-presidential position. Dr. Truxal, the former head of the electrical engineering department since 1957, was named vice-president for educational development.

Dr. Truxal, a native of Lancaster, Pa., received his B.A. degree from Dartmouth in 1943, a B.S. in electrical engineering from Massachusetts Institute of Technology in 1947, and a Doctor of Science degree from M.I.T. in 1950.

He joined the Polytechnic faculty in 1954 as an associate professor of electrical engineering after serving with the faculties at M.I.T. and Purdue University.

Dr. Truxal is the author of "Controls Systems Engineering," and the co-author of "Principles of Electronic Instrumentation." He is a member of the IRE Editorial Board, the American Institute of Electrical Engineers, Phi Beta Kappa, and Tau Beta Pi.

❖

(Continued on page 118A)

Here is MEASURED RELIABILITY!

DUR-MICA CAPACITORS TYPE M2DM

Ten thousand EL-MENCO high reliability dipped mica capacitors were put on life test at 85°C with 225% of the rated DC voltage applied–After 26,500,000 actual test unit-hours no failures of any type occurred.

The accumulated 26.5 x 10^6 test unit-hours without any failures can be used to calculate many different failure rates depending upon the confidence level desired. However, we shall explore the meaning of the results at a 90% confidence level.

Assuming no acceleration factor for either temperature or voltage, we have verified a failure rate of less than 0.01% per 1000 hours. (Actually, there is a temperature effect and it has been found that, with the DC voltage stress remaining constant, the life decreases approximately 50% for every 10°C rise in temperature. There is also a voltage effect such that, with the temperature stress remaining constant, the life is inversely proportional to the 8th power of the applied DC voltage.)

Assuming no temperature acceleration factor and assuming the voltage acceleration exponent is such as to yield an acceleration factor as low as 100, we have nevertheless verified a failure rate of less than 0.0001% per 1000 hours.

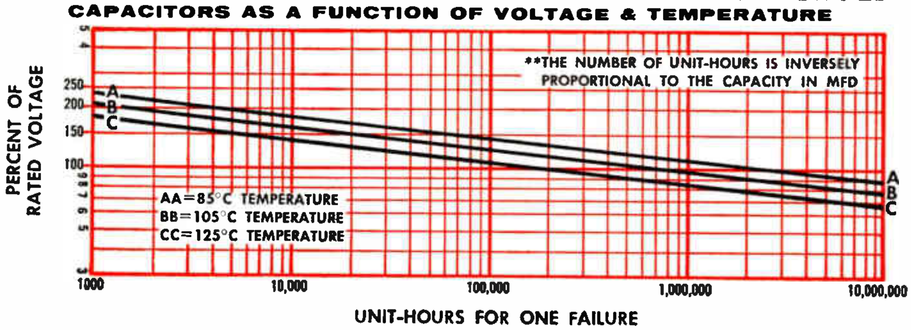
Assuming no temperature acceleration factor and assuming the voltage acceleration factor is on the order of 250 (test results are available to confirm this) we have accumulated sufficient unit-hours to verify a failure rate of less than 0.00004% per 1000 hours!

Note that all the above failure rates are calculated at a 90% confidence level!

Only 1 Failure in 14,336,000 Unit-Hours for 0.1 MFD Capacitors

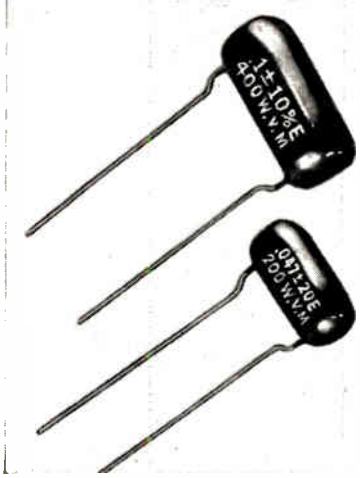
Life tests have proved that El-Menco Mylar-Paper Dipped Capacitors — tested at 105°C with rated voltage applied—have yielded a failure rate of only 1 per 1,433,600 unit-hours for 1.0 MFD. Since the number of unit-hours of these capacitors is inversely proportional to the capacitance, 0.1 MFD El-Menco Mylar-Paper Dipped Capacitors will yield ONLY 1 FAILURE IN 14,336,000 UNIT-HOURS.

MINIMUM LIFE EXPECTANCY FOR 1.0 MFD MYLAR-PAPER* DIPPED CAPACITORS AS A FUNCTION OF VOLTAGE & TEMPERATURE**



* Registered Trade Mark of DuPont Co.

MYLAR-PAPER DIPPED CAPACITORS TYPE MPD



Write for Reliability Study and technical brochures.

THE ELECTRO MOTIVE MFG. CO., INC.

WILLIMANTIC, CONNECTICUT

MANUFACTURERS OF El-Menco Capacitors

Dipped Mica • Molded Mica • Silvered Mica Films • Mica Trimmers & Padders
Mylar-Paper Dipped • Paper Dipped • Mylar Dipped • Tubular Paper

ARCO ELECTRONICS, INC., Community Drive, Great Neck, L. I., New York
Exclusive Supplier to Jobbers and Distributors in the U. S. and Canada

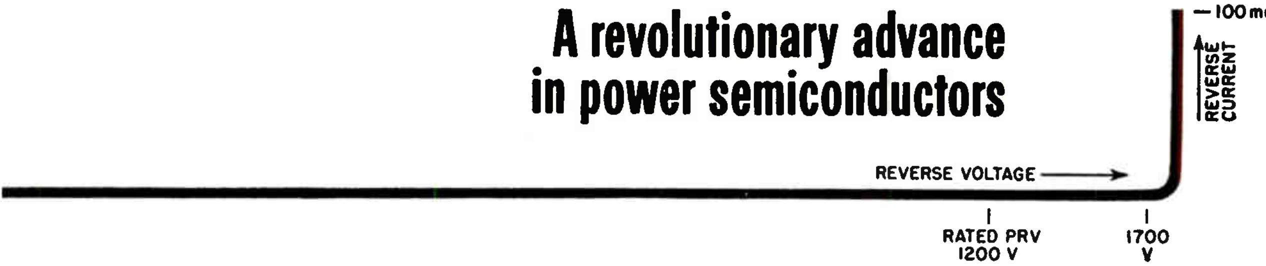
West Coast Manufacturers Contact:
COLLINS & HYDE CO., 535 Middlefield Road, Palo Alto, California
1250 E. Artesia Avenue, Long Beach, California

THE CONTROLLED RECTIFIER

For complete technical information, call your G-E Semiconductor Products District Sales Manager. Or write Rectifier Components Department, Section 23E62, General Electric Company, Auburn, New York. In Canada: Canadian General Electric, 189 Dufferin Street, Toronto, Ontario. Export: International General Electric, 159 Madison Avenue, New York 16, N. Y.

GE ELECTRONICS
PROGRESS IN RECTIFIER COMPONENTS

A revolutionary advance in power semiconductors



AVALANCHE

A NEW 12 AMPERE RECTIFIER WITH PEAK REVERSE RATING OF 1200 VOLTS, AND A SELF-PROTECTIVE REVERSE "AVALANCHE" POWER DISSIPATION RATING UP TO 3900 WATTS

The most important step forward in rectifier technology since General Electric introduced the SCR, the ZJ218 Controlled Avalanche Rectifier is the first of a new generation of G-E power semiconductors which eliminate a *fundamental* silicon rectifier limitation. Carefully controlled *non-destructive* internal avalanche breakdown *across the entire junction area* protects the junction surface, eliminates destructive local surface heating that permanently impairs or destroys the conventional rectifier's reverse blocking ability. In effect, ZJ218 has built-in "zener" diode protection, even well beyond 1200 volts. 600, 800, 1000 and 1200 PRV types are now available . . . *all* with these outstanding features:

SELF-PROTECTION AGAINST VOLTAGE TRANSIENTS

PRV	AVALANCHE VOLTAGE @ 25°C	
	Min.	Max.
1200	1500	1930
1000	1250	1550
800	1000	1290
600	750	1030

- *self-protection against voltage transients . . . dissipates up to 3900 watts peak power in reverse direction*
- *new high reliability standards at PRV's up to 1200 volts (as well as lower voltages)*
- *protection of other circuit components (including transformers) from overvoltage through rigidly spec'd max./min. avalanche characteristics*
- *simplified rectifier series operation in high voltage applications . . . eliminates need for shunt resistors*
- *permits continuous operation in avalanche breakdown region at high voltage . . . unharmed by hi-pot and megger tests*

GENERAL ELECTRIC

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

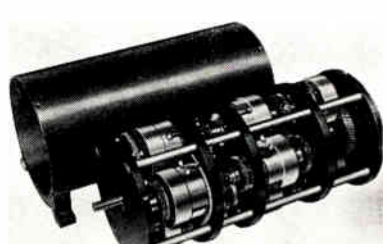
Burnell Dedicates Guillemin Laboratory

Science and industry joined in the dedication (Oct. 26, 1961) of **Burnell & Co., Inc.'s,** new Guillemin Research Laboratory in Cambridge, Mass. Honoring Dr. Ernst A. Guillemin, eminent M.I.T. scientist who is also vice president in charge of research of Burnell & Co., the lab is believed to be the first facility of its kind devoted exclusively to research in electronic filters and networks.

Above, Dr. Guillemin (center) is shown in a discussion with Dr. Lan J. Chu (right), professor of electrical engineering at M.I.T., and Lewis G. Burnell, executive vice president and director of engineering of Burnell & Co., who were among the guests at the dedication. The firm's main plant is located at 10 Pelham Parkway, Pelham, N. Y.

Multi-Speed Reversible Transmission

With continuous running input, this transmission, designed by **Autotronics, Inc., Dept. T.,** P. O. Box 208, Florissant, Mo., delivers eight binary ratio output speeds in both directions plus an electromagnetically actuated instant brake. Output torque is in excess of 45 lb. in. at each speed.



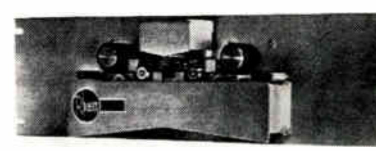
With 1800 rpm input, the output speeds are 360, 180, 90, 45, 22.5, 11.25, 5.625, and 2.8 rpm in both clockwise and counterclockwise directions. Speed and/or direction change takes place within 25 milliseconds.

Size is 3.6″ diameter by 7.3″ long excluding shaft extensions. Transmission is shown in picture with cover removed.

Price dependent upon specifications and quantity ordered. Contact: Art Lee, Sales Manager, for additional information.

Photocell Punched Tape Reader

A high speed unit is the newest addition to the line of Photocell Punched Tape Readers, according to an announcement of **Rheem Manufacturing Co., Electronics Div.,** 5200 W. 104th St., Los Angeles, Calif. The new model, the RR-1000, is nominally a 1,000 character per second unit, and features completely transistorized circuits, photovoltaic sensing cells and rugged two speed motor drives.



This new high speed unit is believed to be the lowest priced high speed tape reader currently being marketed. The uni-directional model is priced at **$1,450.00** and the bi-directional model at **$1,590.00.**

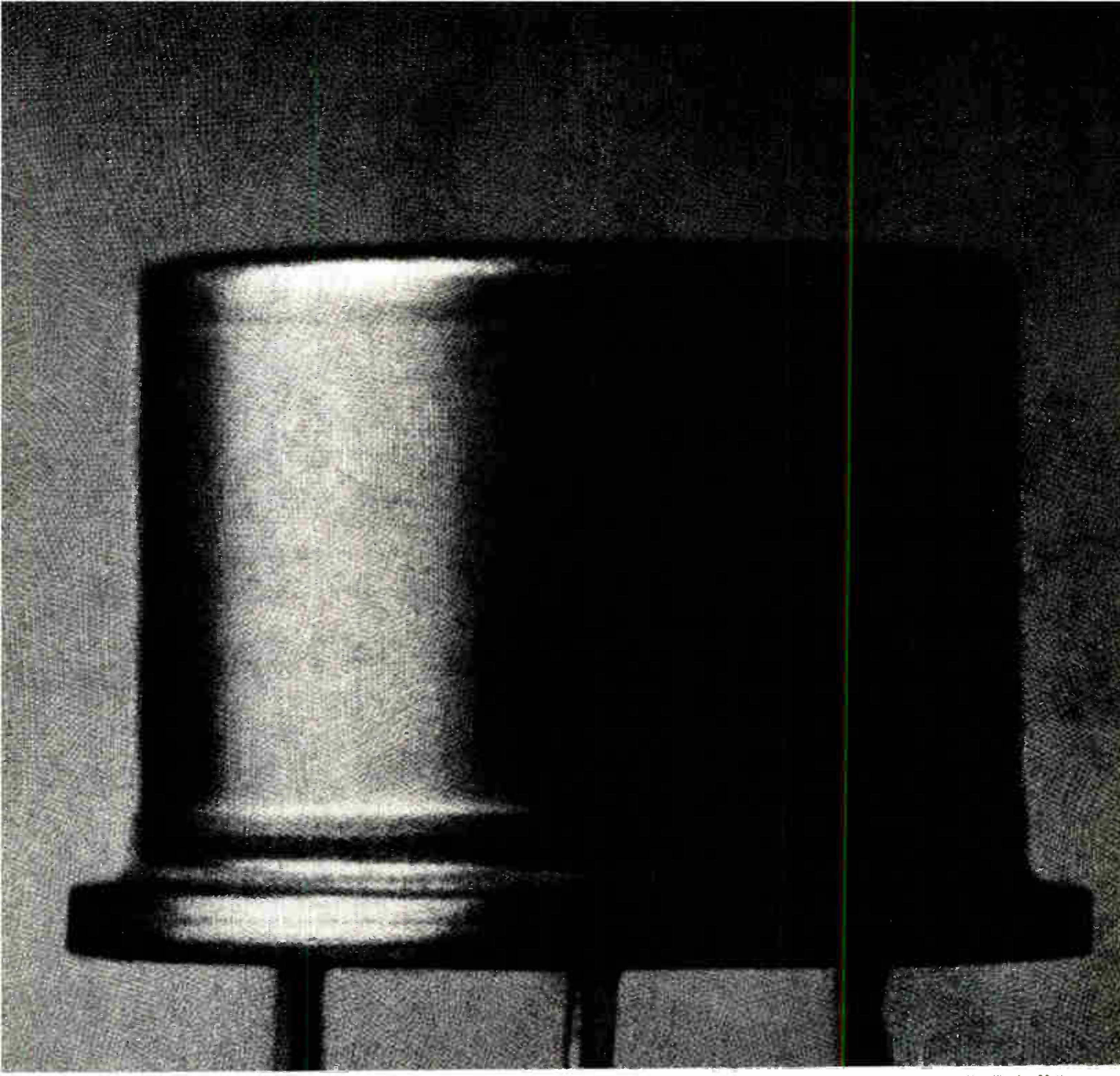
The standard model includes 2:1 speed ratio selectable with low level input, and 5, 7, and 8 level tape selection. Outputs can be selected for positive or negative logic and positive or negative voltage level. The reader is $5\frac{1}{4}$ inches high and 19 inches wide for rack mounting.

For details contact Dr. Wm. Leone at the firm.

Mechanical Filter

A new low cost mechanical filter for commercial and amateur communications equipment is available from **Collins Radio Co., Components Div.,** P. O. Box 1891, 19700 San Joaquin Rd., Newport Beach, Calif.

(Continued on page 46A)



Magnification: 20 times

Old hat

Six months ago this Fairchild transistor was the newest thing on the market. Now there's a better one. Made by Fairchild. Meeting the challenge of your own products is a criterion of leadership in this fast-moving, fast-changing industry. That's why the search to make it (1) work better, (2) do more and (3) cost less—goes on 24-hours a day at Fairchild.

FAIRCHILD SEMICONDUCTOR

FAIRCHILD SEMICONDUCTOR / 545 WHISMAN RD., MOUNTAIN VIEW, CALIF. / YORKSHIRE 8-8161 / TWX: MN VW CAL 853 / A DIVISION OF FAIRCHILD CAMERA AND INSTRUMENT CORPORATION

NEWS New Products

Based on innovations in design and manufacturing, the new 455 kc mechanical filter will be available at cost reductions of as much as 35 per cent under the standard version of the same filter.

The new filter has steep skirted selectivity with a 6 db bandwidth of 2.1 kc and a 60 db bandwidth of 5.3 kc—a shape factor of just over 2.5 to 1. One of the most widespread applications of the new filter is expected to be in single sideband transmitters and receivers.

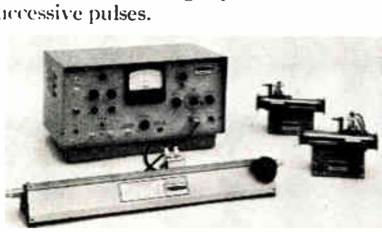
The filter is packaged in a phenolic case and is suited for circuit board manufacturing techniques involving dip soldering. The filter is 2½ inches long, and slightly more than ½ inch wide and ½ inch high, not including mounting studs and terminals.

Production quantities will be available with 120-day delivery after May, 1962. Specific application, price and delivery information can be obtained from the firm.

Phase Meter For Pulsed Microwaves

Wiltron Co., 717 Loma Verde Ave., Palo Alto, Calif., has developed a phase meter for pulsed microwave signals.

This new phase measuring capability is of importance to microwave tube manufacturers, radar and communication system designers, transmitter designers and other users of microwave pulsed signals. It means that for the first time there is an instrument for measuring the RF phase characteristics during a pulse and between successive pulses.



The meter takes a μs sample of the pulsed signal being measured. After calibration, the average phase during this sample is presented on a direct reading meter in degrees. The point of sampling can be progressively moved through a pulse to achieve a plot of phase for each portion of the pulse. In addition, the entire pulse phase pattern can be read out on an oscilloscope. This permits the observation of dynamic phase changes, with a response time of about 0.2 microsecond.

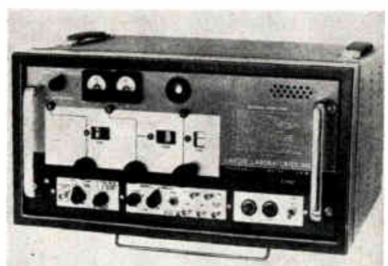
The Model 305 sells for $2,450.00 and must be used with one of Wiltron's Phase Detectors which are in the range of $950.00 and above. Delivery is 7 weeks.

(Continued from page 44A)

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

VHF Frequency Meter

The LA-70B VHF Frequency Meter, designed by **Lavoie Laboratories, Inc.,** Morganville, N. J., can measure from 10 kc to 3,000 mc with an accuracy of 0.0001% between 20 mc to 3000 mc. In addition, it is capable of generating frequencies from zero to 3000 mc. Using an oven-mounted 1 mc quartz crystal for frequency control, the instrument has a stability of 0.001% and resetability of 0.000025%.



It is a direct reading meter, no temperature correction curves or calibration book is required, and the instrument operates in an ambient temperature range of 4°F to 130°F.

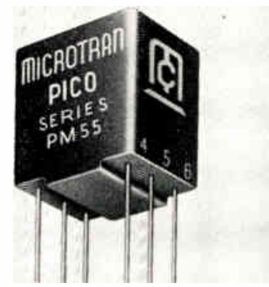
Useful for any frequency measurement task, and particularly mobile radio service, the instrument is useful in input, oscillator and intermediate or harmonic frequency measurements and any frequency encountered in VHF receivers. Since any desired single frequency over the entire range of zero to 3,000 mc is available as an active output, alignments of VHF receivers and other passive elements can also be performed.

The LA-70B also features: output frequencies of zero to 3000 mc with output levels in the order of a few tenths of a volt; internal impedance (as a generator) 3000 ohms; internal impedance (as a frequency meter) 50 ohms; with power requirements of 115 v, ±10%, 60 cps, 130 watts maximum during warmup to 105 watts during operation.

Transformers and Inductors

Microtran Company, Inc., Valley Stream, N. Y., has produced several new lines of products.

Among these are a new line of PICO miniaturized transformers for miniaturized printed-circuit board applications, and a new stock line of toroid inductors.



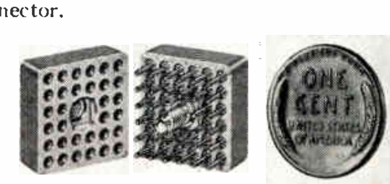
The new PICO series with leads on 0.1″ grids is 5/16″×13/32″×15/32″ high and weighs 0.1 ounce. Gold-plated, high-strength nickel alloy leads permit both reliable soldered joints and high-density welded packaging. Epoxy molded construction per MIL-T-27, Grade 5, Class R, 10,000-hour life. The 20 different items in the PICO series are available in impedance ratings from 3.2 ohms to 200,000 ohms.

The line of toroid inductors is available in a wide choice of frequency ranges and inductances. Types of construction available are: open frame, plastic dipped, molded with lugs. The inductors are hermetically sealed with a standard accuracy of inductance of ±2%.

Detailed information, specifications, comprehensive catalogs, pricing, and list of distributors may be obtained from the firm.

High Density Contact Connector

Winchester Electronics, Inc., 19 Willard Rd., Norwalk, Conn., has announced a new subminiature high density connector.



Named the SRD series, the connector is designed for use with instrumentation employing small modular packages. Extremely suitable for logic units in computors, and allied applications, the connector, 0.6″ square, has center jackscrew for engagement and disengagement and is readily adapted for potting.

The component has 32 contacts; solder dup and dip solder terminals; pin contact brass, socket phosphor bronze—gold plated over silver. Polarization is by pins in place of contacts. Standard molding is diallyl phthalate. Contact spacing is 0.094″.

(Continued on page 48A)

NEWS New Products

(Continued from page 46A)

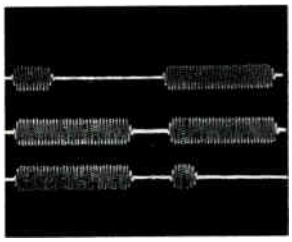
High Power Pulsed Oscillator

Oscillator, PG-650C, an improved version of the original model, manufactured by **Arenberg Ultrasonic Laboratory, Inc.**, 94 Green St., Jamaica Plain 30, Mass., is used as a variable frequency pulse modulated RF source for applications requiring high power as well as extreme stability. Its chief application has been in all phases of ultrasonic testing, nuclear magnetic resonance and transient analysis.

The standard units provide 0 to 300 v peak to peak pulses into 93 ohms from 5 mc to 90 mc. The pulse length is continuously adjustable from 1½ to 20 μsec and provision is made for external or internal triggering from 50 to several thousand cps. Three calibrated delay ranges cover 20 to 120, 100 to 1100, and 1000 to 11,000 μsec. In addition, fine controls giving a 1% change of the coarse controls on pulse length, prf and delay are provided.

The pulse length can be reduced to a minimum of ½ μsec and on special units extended to 50 or 100 μsec with a small increase in the minimum continuously variable range.

The rise time is 0.4 μsec or 3 cps, whichever is greater, on standard units and the fall time is shorter. Provision can be made for increasing the rise or shaping the envelope. The RF pulse is phase stable with respect to the trigger and jitter can be held to 0.005 μsec. Droop or overshoot is less than 5%. Because of the push-pull gated oscillator action, the total harmonic content is less than 10% consisting mainly of the third harmonic and no noise or background is generated between pulses. RF leakage is detectable only with great difficulty.



Variable Pulse Output

The unit is contained in a shielded aluminum box with a rack mounting front panel. The output may be taken either from a balanced or unbalanced shielded U.H.F. receptacle.

The RF tuning is easily adjusted via a wide band spread vernier dial and read on a straight line chart. The range can be extended from 300 kc to 150 mc on the standard unit with extra coils, or from 12 kc to 210 mc on a modified unit.

Other features of the standard unit now are:

(1) Gated output for blanking or intensifying purposes available during the RF pulse.

(2) Provision for two pulses of the same frequency and amplitude separated by a variable delay initiated by the same trigger. A modification will allow these to have different widths.

(3) External modulation of the RF pulses from a gate to allow any sequence of variable pulse widths and spacings. The application is mostly for nuclear magnetic resonance work but external modulation will allow pulsing up to a 110 kc repetition rate for accurate measurements of sound velocity.

(Continued on page 52A)

Patents and the Electronics Company

The basis for Patent Law in the Constitution of the United States reads: "The Congress shall have power . . . to promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries."

Patents Are Needed!

In the early days of the patent law, individual inventors obtained patents on their inventions and benefited in proportion to their commercial successes. Today the individual has much less opportunity to work out successful ventures alone and more often combines forces with others. At the other extreme is the large corporation with an elaborate system for exploiting the inventions of its engineers even to the point of license pools some of which have been knocked down by application of Anti-Trust Laws. With serious difficulties at both ends of the scale, the small company has come into its own in patent matters. The small company with proper handling can reap great rewards from the skillful use of the Patent Laws of this and foreign countries. Some of these rewards are set forth below:

Worth Patenting?

1. The growth and stability of a small electronics business depends on many factors such as marketing, management, financing, but fully as important is its technical ability. Technical ability means, among other things, a certain amount of development. This development requires good engineers and good ideas. Many of these good ideas are worth patenting.

Stimulant to Engineers

2. The patenting of new ideas is one of the best stimulants to more and better ideas. This is really the heart and core of the intent of the Constitutional provision for patents and it really works. New products require new ideas. New ideas flow readily from creative engineers in the right environment. One of the ways to help create this environment is to give the engineer the recognition that the proper handling and patenting of his ideas provides. A well integrated and competent patent department can provide much of this stimulating environment for the engineer. In a small company in the electronics field a competent patent engineer/attorney can provide this patent department at a nominal and predictable cost.

Patents Stabilize

3. Patents held by small electronics companies provide a strong stabilizing influence on the company's operations and its standing and recognition in the field. Patents help establish a field of priority. This established field of priority may be most valuable in its implicit warning to competition. This is especially true of competition from other companies who recognize the value of patents and ask and give respect for the fields of priority which they establish. This valuable aspect of patents grows with the growth of the patent coverage obtained, but requires the obtaining of good patents. All this usually takes place, when properly handled, without patent suits and only with the warnings which are implicit with the issuance of carefully drawn and apparently valid patents.

Patents Are Assets

4. Patents are property and have asset value. There is a striking contrast between the company which after years of operation has no patents and the one which has gathered many. Many businesses are sold or obtain financing on asset values greatly augmented by their patents. Sometimes the patents are themselves sold at considerable profit. A company which might otherwise be an empty shell, when provided with a patent folio can show substantial values. Patents are an equity accumulated over the years which might otherwise become almost completely lost in the normal course of business. This asset value alone makes patents an extremely good business practice. There are other important ways in which the asset value of patents may be used. The licensing of patents especially for noncompetitive uses may be a source of considerable revenue. Patents which are not being worked may be sold. The patent department should be aware of these possibilities and bring them to the attention of management for decision.

Patents Do Protect

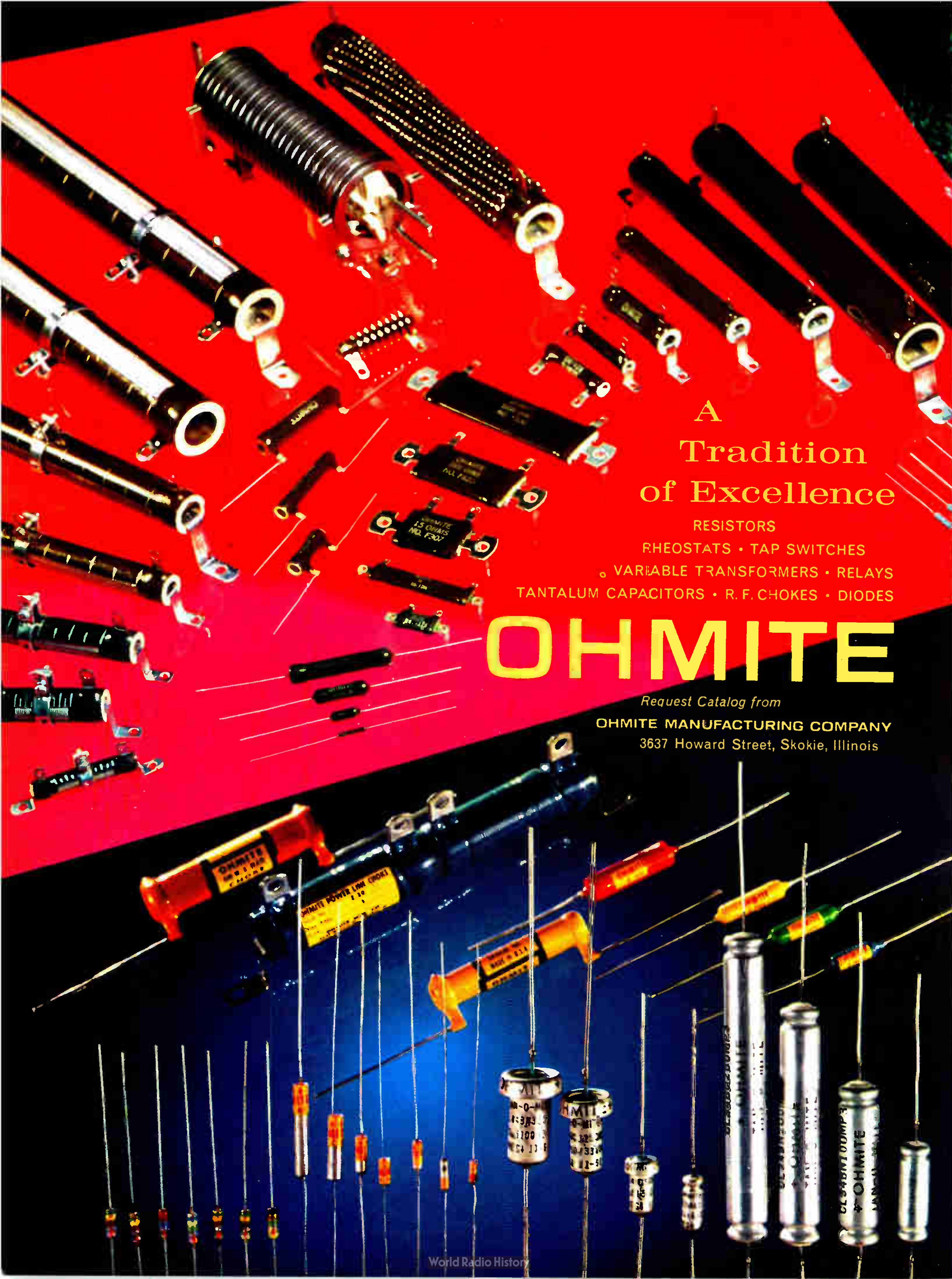
5. Finally, and this section is made the last for good reason, patents may be used to sue others. The small company is admittedly at a disadvantage here with respect to the large corporation since patent suits are expensive. However, with careful management and the proper attitude with regard to patents, suits are a last resort. This is a subject with too many implications to be discussed fully here. It is merely suggested that patents may be dealt with and many important and real values realized without resorting to patent suits. The statement that a patent is only a right to sue, glosses over the many advantages and values set forth above, which may be had without even considering suing others.

A Patent Is Insurance

A patent program may be regarded as an insurance program; insurance that full value will be received for contributions to advancement of the electronics art.

Printed through the courtesy of:

Alfred W. Barber Laboratories
32-44 Francis Lewis Blvd.
Flushing 58, N.Y.
Tel. 212-IN 3-3306


A
Tradition
of Excellence
RESISTORS
RHEOSTATS • TAP SWITCHES
VARIABLE TRANSFORMERS • RELAYS
TANTALUM CAPACITORS • R. F. CHOKES • DIODES
OHMITE
Request Catalog from
OHMITE MANUFACTURING COMPANY
3637 Howard Street, Skokie, Illinois

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

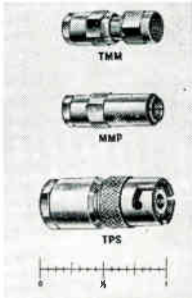
(Continued from page 48A)

(4) Provision for use as a pulsed amplifier in order to use an external CW source of greater frequency stability than is possible with a gated single stage oscillator. This mode can be extended to several gated stages of amplification.

(5) Operation as a CW oscillator of 10 to 20 watts output in the standard frequency range.

Connectors

Automatic Metal Products Corp., 314-323 Berry St., Brooklyn 11, N. Y., announces the development of five new series of micro-miniature connectors. These new series are as follows: TMM (Threaded) 50 ohm; TMM (Threaded) 75 ohm; MMP (Push On) 50 ohm; MMP (Push On) 75 ohm and, TPS (Bayonet) 50 ohm.



In style and configuration, these connectors closely parallel the present sub-miniature connectors as shown in the firm's brochure #sub min-1-660-5M. They have improved electrical and environmental characteristics, and are recommended for the most severe applications at frequencies up to "X" band.

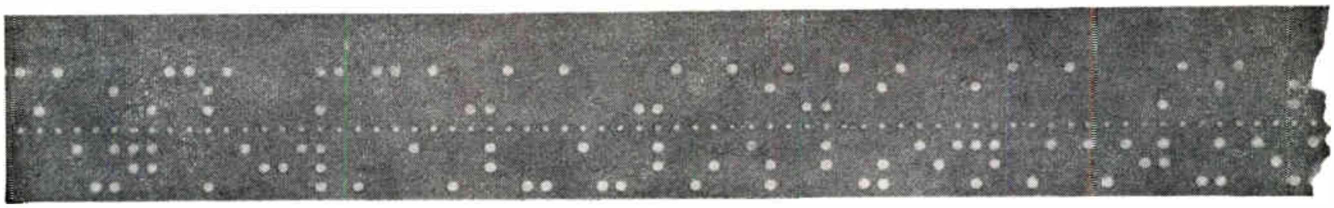
Standard cable clamping in all of these series is Wedge-Lock/Captive Contact. Optional cable clampings are Wedge-Crimp/Captive Contact, or an external crimp type clamp. These cable clamping techniques are said to cut assembly time as much as 70% over previous methods. As an optional extra, neoprene or silastic strain relief boots in either black or standard RETMA colors may be supplied.

All connectors have Teflon insulation. They may however, be supplied with alternate insulation materials for special applications such as radiation resistance, etc.

In the Threaded and Push-on series, plug members contain the female contact and jack members contain the male contact. In the Bayonet series, contacts are reversed. Optional polarization of all of these connectors may be obtained by reversing the male and female contacts. Additional polarization may be obtained by reversing both the contacts and insulators. (By utilizing the various polarization methods, three connectors may be placed side by side on a panel without the chance of accidental mismating.)

(Continued on page 54A)

RFI ANALYSIS & PREDICTION BY COMPUTER!



Any equipments or components you are producing for missile weapons systems can be analyzed for Radio Frequency Interference by Capehart's INTERDICT* Group. Under our analysis procedures, RFI is computer-predicted on finished products, prototypes, or preliminary designs. Every item of electronic equipment which produces radiated or conducted RF signals is considered a transmitter, and every electronic equipment which is subject to malfunction as a result of exposure to conducted or radiated spurious signals is considered a receiver. We will make specific recommendations and provide design changes to bring your equipment within the RFI specifications or inform you if it is already within standards. Capehart's INTERDICT Group performs this same service for any other type of installation, site or system where RFI may be a potential problem. Our approach in RFI diagnosis of over-all operational requirements utilizing computer programming has been proven on a number of electronic equipments, missile systems and military sites including Cape Canaveral and Vandenberg AFB.

** Interference Detection and Interdiction by Countermeasures Team*

COMPLIANCE WITH CURRENT MIL SPECS: MIL-I-26600 • MIL-I-6051 • MIL-I-6181D

Engineers interested in joining Capehart's INTERDICT Group, or in other opportunities at Capehart, are invited to contact our Professional Recruiter, at the address below.

Capehart® CORPORATION

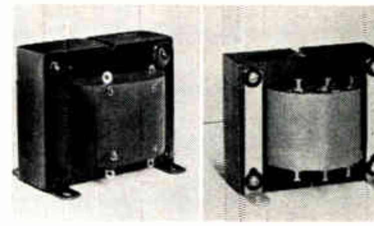
INTERDICT GROUP, Dept. I, CAPEHART CORPORATION
87-46 123rd Street, Richmond Hill 18, New York • HIckory 1-4400

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 52A)

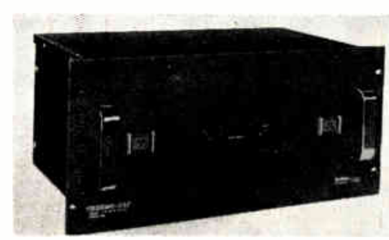
Transformers and Chokes



A new stock line of high current rectifier transformers and chokes with ratings from 9 to 56 volts at 1 to 200 amperes is now available from **Signal Transformer Co.**, 1661 McDonald Ave., Brooklyn 30, N. Y. Primarily designed for low voltage, high current applications, these items find extensive use in power supplies for transistor circuits, computers, magnets, battery charging and electro-plating. Units can be obtained for 50 or 60 cps operation in both open frame and encapsulated types. Available for immediate shipment from the firm.

Solid State Temperature Reference

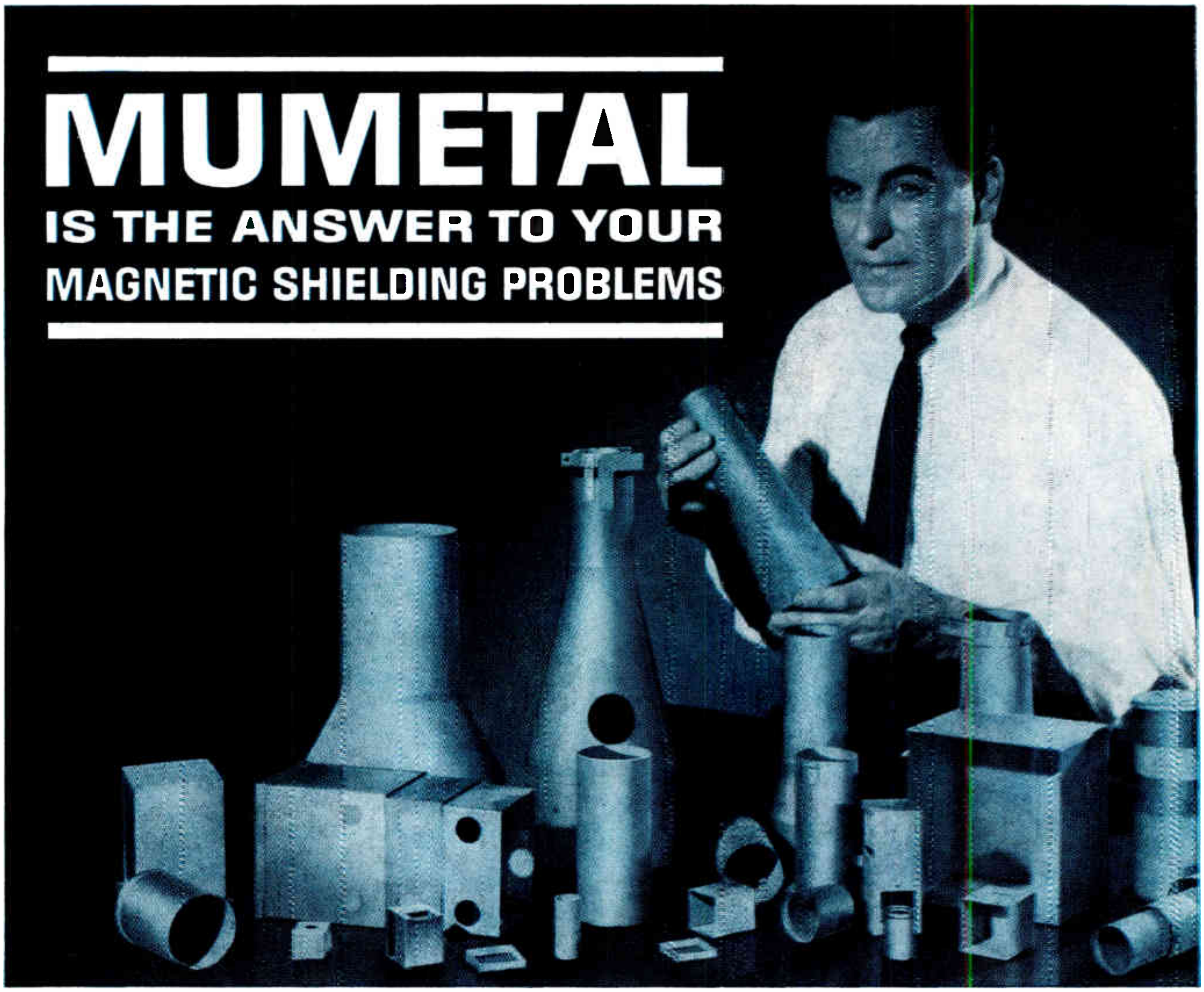
A rack-mounted version of the Thermo-Ref solid-state temperature reference is now available from **Genistron, Inc.**, 2301 Federal Ave., Los Angeles 64, Calif., from all applications requiring close temperature measurement tolerances and field ruggedness with laboratory accuracy.



Two 40-channel configuration of the Genistron unit provides a choice of junction materials—chromel-alumel in the Thermo-Ref Model 4010 and copper-constantan in the Model 4020, and other materials as specified.

Reference temperature is 212°F (100°C) with a special scale, visual calibrator, indicating meter for voltage check point. Junction uniformity and temperature stability are both ±.2°F over the standard operating range of −20°F to +150°F which allows laboratory accuracy in typical applications. Power requirements are 10 watts maximum at 115 V, 60 cps with nominal 15-minute warmup. TC channel to channel isolation is in excess of 50 megohms. Custom modifications can be supplied to meet specific requirements. The Thermo-Ref meets applicable military environmental specifications. Designed for

(Continued on page 56A)



Instant relief to interference caused by extraneous magnetic fields is the net result of shields made of Allegheny Ludlum's Mumetal. These shields protect components against stray external fields or prevent neighboring parts from being affected by a field-generating component inside the shield. In electronics, Mumetal and shielding are practically synonymous terms.

To develop its optimum shielding properties, Mumetal must be properly annealed in a pure, dry, high temperature hydrogen atmosphere after fabrication. When properly annealed, Mumetal has extremely high permeability and is capable of attenuating stray fields to negligible proportions.

In general, high permeability, shielding excellence and strain sensitivity go hand in hand. In the optimum condition, Mumetal is relatively soft. Shields in this condition should be handled with care in order to preserve optimum shielding efficiency.

In many applications, fabricating or field conditions are encountered which make it impossible to avoid straining the material after the high temperature hydrogen anneal. Even when strained, Mumetal shields remain extremely effective.

The inherent ductility of Mumetal offers fabricating advantages in forming, drawing, and spinning operations.

For all your shielding requirements, insist on Allegheny Ludlum Mumetal. And for more information, ask for a copy of EM12, a 20 page technical Blue Sheet describing Mumetal, its properties, annealing details, etc. Write ***Allegheny Ludlum Steel Corporation, Oliver Bldg., Pittsburgh 22, Pa., Address Dept. IRE-5***

ALLEGHENY LUDLUM

PIONEERING on the Horizons of Steel

4106

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 54A)

standard 19-inch rack mounting, the unit is 10 inches deep.

Typical usage includes rack-mounted Thermo-Ref models as an integral component of temperature monitor systems in missile test stand operations, and other applications when temperature accuracy is important.

For additional information, contact the firm.

Octave Amplifiers

A new series of packaged octave amplifiers is now available from **LEL, Inc.**, 75 Akron St., Copiague, N. Y. These low noise, broadband RF amplifiers cover a range of frequencies from 40 mc to 640 mc.

The Model OMP-5-160 has a gain of 30 db minimum, ±0.5 db ripple and a linear output of greater than 1 volt over the range of 160 to 320 mc. It is available as a self-contained rack-mounted unit as shown or as an amplifier only, Model OMX-5-160.* The rack mounted units employ a solid state regulated power supply. Price and delivery upon request.

* Average noise figure is 6 db.

Miniature Preamplifier

Quanta Dyne, P.O. Box 252, Menlo Park, Calif., has introduced its Model Q-63 miniature amplifier. This ruggedized, ultrasensitive, low-noise, solid-state preamplifier can be used for original equipments or test bench. Included among its many uses are microwave crystal boosting and video preamplification.

The Q-63 provides nominally 50 db gain over the bandwidth of 30 cps to 100 kc and 40 db gain over the bandwidth of 10 cps to 600 kc, measured with typical input impedance of 7 kohms, output impedance of 4 kohms, source impedance of 1 kohm, and load impedance of 1

(Continued on page 58A)

MICRO-MESH SIEVES DETERMINE PARTICLE SIZE DISTRIBUTION-Accurately

Set of Buckbee Mears Co. Micro-mesh sieves, standard sizes from 20 to 90 microns. Smaller sizes to order. Background: Enlarged evaporation mask for Mesa transistor.

Buckbee Mears research in subminiature mesh has solved many problems for the industry. A few current items are—

EVAPORATION MASKS—Mesa Transistors

SHADOW MASKS—Color TV

STORAGE MESH—Radar Tubes

CALIBRATED DIALS—Test Equipment

Anything that can be drawn can be reproduced. Send us your problems

SETS FROM 20 TO 90 MICRONS

The accurate measurement of particle size distribution of cracking catalyst has become practical. Buckbee Mears has developed screen sets which have been thoroughly tested as reported on by Ketjen N. V. Amsterdam*, and ASTME*. One comment "micro-mesh sieve method produces data with great accuracy and reproducibility." Another "the sieves made by Buckbee Mears Company appear especially attractive."

According to these papers, previous methods had disadvantages and results by different methods varied considerably. Now BMC sieves are recommended as a primary standard for testing size distribution of dry materials (cracking catalyst).

**Copies of these papers will be sent on request also complete sieve information.*

buckbee mears

245 EAST SIXTH STREET • ST. PAUL 1, MINNESOTA

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 56A)

megohm (high input impedance module to work in conjunction with the amplifier is also available).

The noise of the Q-63 is nominally 3 microvolts, measured with a source impedance of 1 kohm and a load impedance of approximately 6 K.

Mechanical specifications list the size at $\frac{3}{4}'' \times \frac{3}{4}'' \times 2''$ less the BNC 1094/U connectors, and the weight at 2 oz.

The external power supply requirements of 9 volts at approximately 3 ma may be met by a Quanta Dyne miniature battery pack, Model Q-630.

Ferrite Isolator

A new ferrite isolator is now being manufactured by **Instrument Systems Corp., Microwave Div.,** 129-07 18th Ave., College Point 56, N. Y.

Model 3005-I is an X Band, medium power waveguide isolator, operating on the Faraday rotation principle, having substantial isolation characteristics coupled with minimum insertion loss.

In normal applications, it is placed between the microwave source and its transmission line to improve the stability and reliability of the generator. The nonreciprocal characteristic prevents impedance variations, due to long line effects and load mismatches, from appearing at the source, thus providing broad band operation in addition to the elimination of frequency pulling.

Constructed of aluminum, the isolator operates into a 3:1 VSWR mismatched load. The device has been qualified under applicable government specifications for Naval and Air Force environments. Cost per unit is $210.

More information on isolators and microwave ferrite components of all kinds may be obtained by writing to Manager, Microwave Division, at the firm.

AC Power Supply

A portable ac power supply that provides an adjustable output voltage from 1 to 140 volts has been developed by **Industrial Timer Corp.,** 1407 McCarter Highway, Newark 4, N. J., for a wide variety

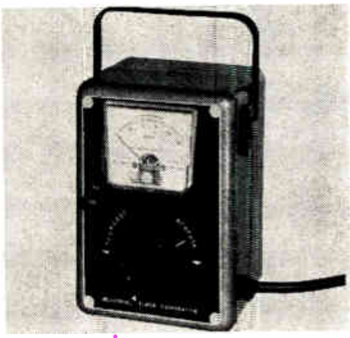
(Continued on page 60A)

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(*Continued from page 58A*)

of applications in both laboratory and production installations.



Called Model VVS, the unit is basically a package comprising a variable transformer capable of controlling loads up to 5 amperes, voltmeter, input line and output receptacle. Model VVS features a large voltmeter showing the output voltage, which is adjusted with a knob on the front of the unit.

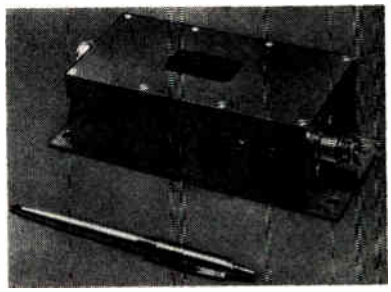
The portable unit is self-contained, with a carrying handle, heavy duty line cord and plug, and a convenient two-prong build-in receptacle for output voltage.

Applications for Model VVS are numerous, including laboratory testing and production control.

List price of the Model VVS is $55.00, subject to Industrial Timer's standard quantity discount schedule.

Solid State Frequency Multiplier

A high efficiency, passive solid state frequency multiplier was exhibited at the IRE show by **Microwave Associates, Inc.**, Burlington, Mass.



The MA-8028 multiplier can be driven with 3 watts CW power at about 36 to 40 mc and will deliver 200 mw at the 32nd harmonic in the L-band range. The bandwidth is 1.25%.

Three stages, featuring a compact lumped-circuit design, are incorporated in this multiplier, two of which are quadruplers and the third stage a doubler. A total of 5 varactors are used. No external bias is required.

(*Continued on page 64A*)

29 Fields of Special Interest-

The 29 Professional Groups are listed below, together with a brief definition of each, the name of

Aerospace and Navigational Electronics

Annual fee: $3.

The application of electronics to operation and traffic control of aircraft and to navigation of all craft.

Mr. George M. Kirkpatrick, Chairman, General Electric Co., Syracuse, N.Y.

42 Transactions, *9, Vol. 1, No. 3; Vol. 2, No. 1-3; Vol. 4, No. 1, 2, 3; Vol. 5, No. 2, 3, 4; Vol. 6, No. 1, 2, 4; Vol. 7, No. 1, 2, 3, 4; Vol. 8, No. 1, 2, 3, 4

Antennas and Propagation

Annual fee: $6.

Technical advances in antennas and wave propagation theory and the utilization of techniques or products of this field.

Dr. Harry Fine, Chairman, Applied Propagation Branch FCC, Washington, D.C.

47 Transactions, *Vol. AP-2, No. 2; AP-4, No. 4; AP-5, No. 1-4; AP-6, No. 1, 2, 3, 4; AP-7, No. 1, 2, 3, 4; AP-8, No. 1, 2, 3, 4, 5, 6; AP-9, No. 1, 2, 3, 4, 5, 6.

Audio

Annual fee: $2.

Technology of communication at audio frequencies and of the audio portion of radio frequency systems, including acoustic terminations, recording and reproduction.

Dr. Robert W. Benson, Chairman, Elec. Eng. Dept., Vanderbilt University, Nashville 5, Tenn.

61 Transactions, *Vol. AU-1, No. 6; *Vol. AU-2, No. 4; Vol. AU-3, No. 1, 3, 5; Vol. AU-4, No. 1, 5-6; Vol. AU-5, No. 1, 2, 3, 4, 5, 6; AU-6, No. 1, 2, 3, 4, 5, 6; AU-7, No. 1, 2, 3, 4, 5, 6; AU-8, No. 1, 2, 3, 4, 5, 6; AU-9, No. 1, 2, 3, 4, 5.

Automatic Control

Annual fee: $3.

The theory and application of automatic control techniques including feedback control systems.

Mr. John M. Salzer, Chairman, 909 Berkeley St., Santa Monica, Calif.

16 Transactions, PGAC-3-4-5-6, AC-4, No. 1, 2, 3; AC-5, No. 1, 2, 4; AC-6, No. 1, 2, 3.

Bio-Medical Electronics

Annual fee: $3.

The use of electronic theory and techniques in problems of medicine and biology.

Mr. George N. Webb, Chairman, Dept. of Med. Biophysical Div., Johns Hopkins Hospital, Baltimore 5, Md.

23 Transactions, 8, 9, 11; ME-6, No. 1, 3, 4; ME-7, No. 2, 3, 4; BME-8, No. 2, 3, 4.

Broadcast & Television Receivers

Annual fee: $4.

The design and manufacture of broadcast and television receivers and components and activities related thereto.

Mr. John F. Bell, Chairman, Zenith Radio Corp., 6001 W. Dickens Ave., Chicago 39, Ill.

30 Transactions, *7, 8; BTR-1, No. 1-4, BTR-2, No. 1-2-3; BTR-3, No. 1-2; BTR-4, No. 2, 3-4; BTR-5, No. 1, 2; BTR-6, No. 1, 2, 3; BTR-7, No. 1, 2, 3.

Broadcasting

Annual fee: $2.

Broadcast transmission systems engineering, including the design and utilization of broadcast equipment.

Mr. Raymond F. Guy, Chairman, 264 Franklin St., Haworth, N.J.

21 Transactions, No. 10, 11, 12, 13, 14; BC-6, No. 1, 2, 3; BC-7, No. 1, 2, 3, 4

Circuit Theory

Annual fee: $4.

Design and theory of operation of circuits for use in radio and electronic equipment.

Dr. James H. Mulligan, Jr., Chairman, College of Eng., New York University, New York 53, N.Y.

35 Transactions, CT-4, No. 3-4; CT-5, No. 1, 2, 3; CT-6, No. 1, 2, 3, 4; CT-7, No. 2, 3, 4; CT-8, No. 2, 3

Communications Systems

Annual fee: $2.

Radio and wire telephone, telegraph and facsimile in marine, aeronautical, radio-relay, coaxial cable and fixed station services.

Mr. Ralph L. Marks, Chairman, Rome Air Dev. Center, Griffiss AFB, N.Y.

23 Transactions, CS-5, No. 2, 3; CS-6, No. 1, 2; CS-7, No. 1, 3, 4; CS-8, No. 1, 2, 3, 4; CS-9, No. 1, 2, 3.

Component Parts

Annual fee: $3.

The characteristics, limitation, applications, development, performance and reliability of component parts.

Mr. Floyd E. Wenger, Chairman, Headquarters ARDC, Andrews AFB, Washington 25, D.C.

25 Transactions, CP-4, No 1, 2, 3, 4; CP-5, No. 2, 3, 4; CP-6, No. 1, 2, 3, 4; CP-7, No. 1, 2, 3, 4; CP-8, No. 1, 2, 3

Education

Annual fee: $3.

To foster improved relations between the electronic and affiliated industries and schools, colleges, and universities.

Mr. George E. Moore, Chairman, Westinghouse Electric Corp., East Pittsburgh, Pa.

15 Transactions, Vol. E-1, No. 3, 4; E-2, No. 1, 2, 3, 4; E-3, No. 1, 2, 3, 4; E-4, No. 1, 2, 3.

Electron Devices

Annual fee: $3.

Electron devices, including particularly electron tubes and solid state devices.

Mr. Willis A. Adcock, Chairman, Texas Instruments Co., Dallas 9, Tex.

36 Transactions, *Vol. ED-1, No. 3, 4; ED-3, No. 2, ED-4, No. 2, 3, 4; ED-5, No. 2, 3, 4; ED-6, No. 1, 3; ED-7, No. 2, 3, 4; ED-8, No. 1, 2, 3, 4, 5, 6.

Electronic Computers

Annual fee: $4.

Design and operation of electronic computers.

Dr. A. A. Cohen, Chairman, Remington-Rand Univac, St. Paul 16, Minn.

40 Transactions, EC-6, No. 2, 3; EC-7, No. 1, 2, 3, 4; EC-8, No. 1, 2, 3, 4; EC-9, No. 1, 2, 3, 4; EC-10, No. 1, 2, 3, 4

Engineering Management

Annual fee: $3.

Engineering management and administration as applied to technical, industrial and educational activities in the field of electronics.

Mr. T. W. Jarmie, Chairman, Engineered Electronics, 1441 East Chestnut Ave., Santa Ana, Calif.

23 Transactions, EM-4, No. 1, 3, 4; EM-5, No. 1-4; EM-6, No. 1, 2, 3; EM-7, No. 1, 2, 3; EM-8, No. 1, 2, 3.

Engineering Writing and Speech

Annual fee: $3.

The promotion, study, development, and improvement of the techniques of preparation, organization, processing, editing, and delivery of any form of information in the electronic-engineering and related fields by and to individuals and groups by means of direct or derived methods of communication.

John M. Kinn, Jr., Chairman, IBM Journal, 545 Madison Ave., New York, N.Y.

10 Transactions, Vol. EWS-1, No. 2; EWS-2, No. 1, 2, 3; EWS-3, No. 1, 2; EWS-4, No. 2, 3.

THE INSTITUTE OF RADIO

IRE's 29 Professional Groups

the group chairman, and publications to date.

* Indicates publications still available

Geoscience Electronics

Annual fee: None

Research and development in electronic instrumentation for geophysics and geochemistry, especially gravity measurements, seismic measurements, magnetics, well-logging, space exploration, meteorology, oceanography, etc.

Human Factors in Electronics

Annual fee: $2.

Development and application of human factors and knowledge germane to the design of electronic equipment.

Mr. Robert R. Riesz, Chairman, Bell Tel. Labs, Murray Hill, N.J.

4 Transactions, HFE-1, No. 1, 2; HFE-2, No. 1, 2.

Industrial Electronics

Annual fee: $3.

Electronics pertaining to control, treatment and measurement, specifically, in industrial processes.

Mr. J. E. Eiselein, Chairman, RCA Victor Div., Camden, N. J.

16 Transactions, *PGIE 1, 3, 5, 6, 7, 8, 9, 10, 11; IE-7. No. 1, 2, 3; IE-8, No. 1, 2.

Information Theory

Annual fee: $4.

The theoretical and experimental aspects of information transmission, processing and utilization.

Dr. George L. Turin, Chairman, Dept. of E.E., Univ. of California, Berkeley 4, Calif.

31 Transactions, PGIT-4, IT-1, No. 3; IT-2, No. 3; IT-3, No. 1, 2, 3, 4; IT-4, No. 1, 2, 3, 4; IT-5, No. 1, 2, 3, 4; IT-6, No. 1, 3, 4, 5; IT-7, No. 1, 2, 3, 4.

Instrumentation

Annual fee: $3.

Measurements and instrumentation utilizing electronic techniques.

Mr. Harvey W. Lance, Chairman, Natl. Bureau of Standards, Boulder, Colo.

21 Transactions. PGI-4, Vol. 1-6, No. 2, 3, 4; Vol. 1-7, No. 1, 2; Vol. 1-8, No. 1, 2; Vol. 1-9, No. 1, 2, 3; Vol. 1-10, No. 1, 2.

Microwave Theory and Techniques

Annual fee: $3.

Microwave theory, microwave circuitry and techniques, microwave measurements and the generation and amplification of microwaves.

Mr. Tore N. Anderson, Chairman, EMT Corp., Syosett, Long Island, N.Y.

40 Transactions, MTT-4, No. 3; MTT-5, No. 3, 4; MTT-6, No. 1, 2, 3, 4; MTT-7, No. 2, 3; MTT-8, No. 1, 2, 3, 4, 5, 6; MTT-9, No. 1, 3, 4, 5, 6.

Military Electronics

Annual fee: $2.

The electronics sciences, systems, activities and services germane to the requirements of the military. Aids other Professional Groups in liaison with the military.

Mr. Willie L. Doxey, Chairman, EC Dept., USASRDL, Fort Monmouth, N.J.

14 Transactions, MIL-1, No. 1; MIL-2, No. 1; MIL-3, No. 2, 3, 4; MIL-4, No. 2-3, 4; MIL-5, No. 1, 2, 3, 4.

Nuclear Science

Annual fee: $3.

Application of electronic techniques and devices to the nuclear field.

Mr. Louis Costrell, Chairman, N.B.S., Washington 25, D.C.

22 Transactions, NS-1, No. 1; NS-4, No. 2; NS-5, No. 1, 2, 3; NS-6, No. 1, 2, 3, 4; NS-7, No. 1, 2-3, 4; NS-8. No. 1, 2, 3, 4.

Product Engineering & Production

Annual fee: $2.

New advances and materials applications for the improvement of production techniques, including automation techniques.

Mr. Alfred R. Gray, Chairman, Astrodyne Reliability Inc. Orlando, Fla.

10 Transactions, No. 3, 4, 5, 6; PEP-5, No. 1, 2, 3, 4.

Radio Frequency Interference

Annual fee: $2.

Origin, effect, control and measurement of radio frequency interference.

Mr. Harold E. Dinger, Chairman, Naval Research Lab., Washington 25, D.C.

3 Transactions, RFI-1, No. 1, RFI-2, No. 1, RFI-3 No. 1.

Reliability and Quality Control

Annual fee: $3.

Principles and practices used in reliability and quality engineering of electronic products.

Mr L. J. Paddison, Chairman, Sandia Corp., Sandia Base, Albuquerque, N.M.

22 Transactions, *3, 5, 10, 11, 12, 13, 14, 15, RQC-9, No. 1. 2, 3; RQC-10. No. 1, 2, 3.

Space Electronics and Telemetry

Annual fee: $3.

The control of devices and the measurement and recording of data from a remote point by radio.

Mr. Kenneth V. Uglow, Chairman, Electro-Mechanical Research Inc., Sarasota, Fla.

20 Transactions, TRC-1, No. 2-3; TRC-2, No. 1; TRC-3, No. 2, 3; TRC-4, No. 1; SET-5, No. 1, 2, 3, 4; SET-6, No. 1, 2, 3-4; SET-7, No. 1, 2, 3, 4.

Ultrasonics Engineering

Annual fee: $2.

Ultrasonic measurements and communications, including underwater sound, ultrasonic delay lines, and various chemical and industrial ultrasonic devices.

Dr. Vincent Salmon, Chairman, Stanford Research Inst., Menlo Park, Calif.

10 Transactions, PGUE, 5, 6, 7; UE-7, No. 1, 2; UE-8, No. 1.

Vehicular Communications

Annual fee: $3.

Communications problems in the field of land and mobile radio services, such as public safety, public utilities, railroads, commercial and transportation, etc.

Mr. Richard P. Gifford, Chairman, General Electric Co., Lynchburg, Va.

18 Transactions, 5, 8, 9, 10, 11, 12, 13; Vol. VC-9, No. 1, 2, 3; Vol. 10, No. 1, 2.

USE THIS COUPON

Miss Emily Sirjane PG-5-62
IRE—1 East 79th St., New York 21, N.Y.

Please enroll me for these IRE Professional Groups

.... $ ....

.... $ ....

Name ....

Address ....

Place ....

Please enclose remittance with this order.

Professional group membership is limited to active IRE members.

ENGINEERS
1 East 79th Street, New York 21, N.Y.

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 60A)

The body, of aluminum construction, is $5\frac{3}{8}''$ long, $2\frac{3}{4}''$ wide and 2" deep excluding mounting pads and connectors. Total weight is less than two lbs. The nominal input impedance is 50 ohms with type "N" input and output connectors.

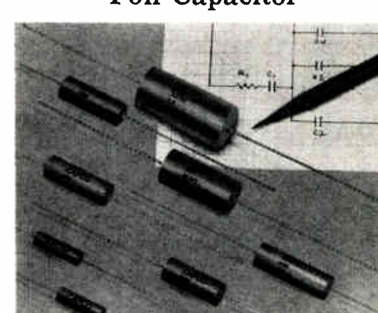
The multipliers are particularly useful in telemetering or beacon transmitters, antenna alignment, signal generation, radar or fixed-frequency communications driver stages, or in local oscillator service. A major advantage is that precise frequency control can be achieved in microwave channels using crystal-controlled drivers at frequencies under 100 mc.

These units are now in production and are available within six weeks after receipt of order. Prototypes are priced at about $1600 with substantial price reduction on quantity orders.

Bouldin New V. P. For Westronics, Inc.

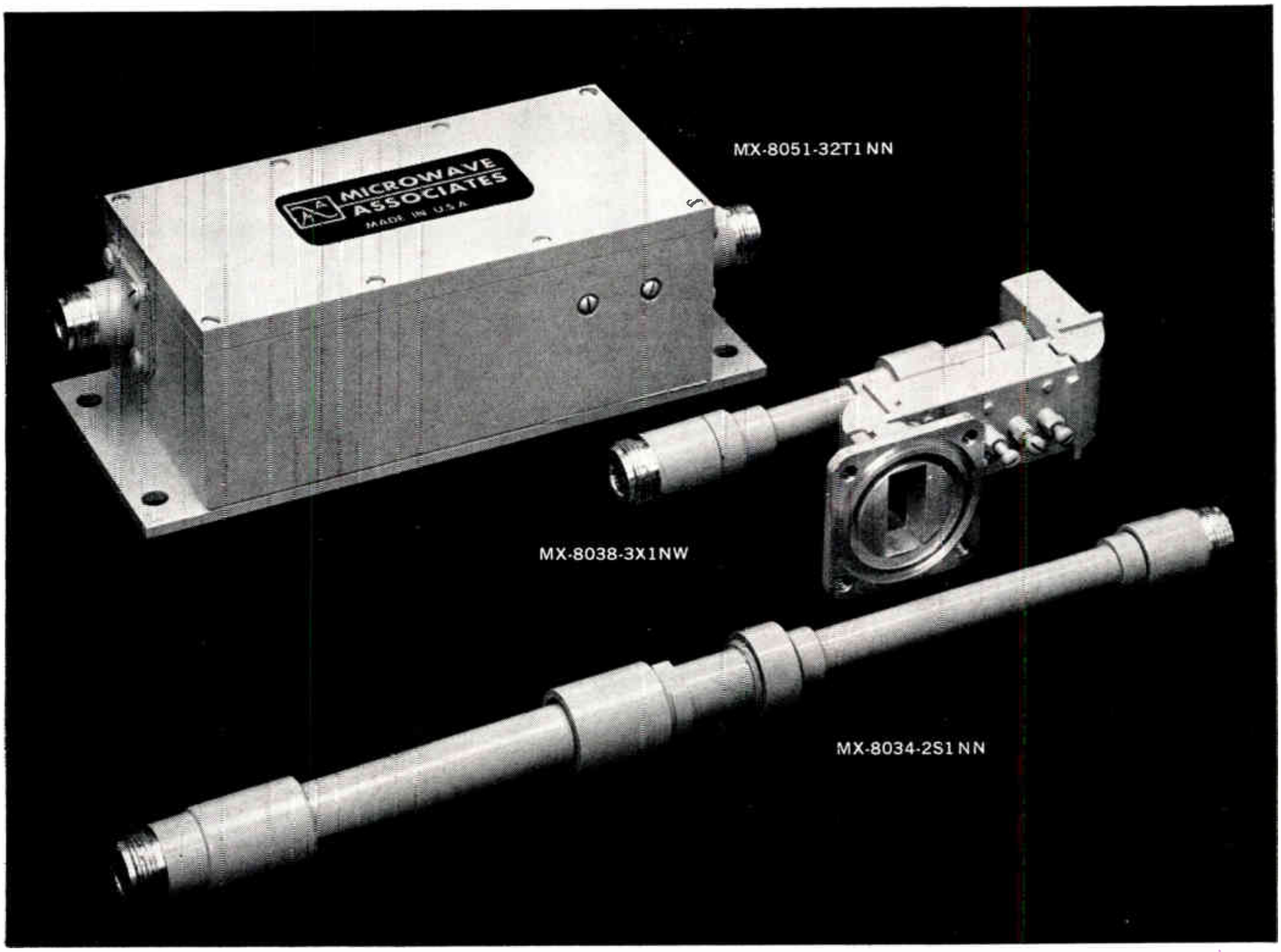
Wood W. Bouldin has been elected vice president in charge of sales at **Westronics, Inc.**, 3605 McCart, Ft. Worth, Texas. He joined the firm as a sales engineer in March 1959, having previously served three years in the USAF as a B-36 pilot and two years at Chance Vought as a flight test engineer. He is a graduate of John Reagan High School, Houston, and Texas A&M where he earned his B.S. degree in 1954 in Industrial Technology. Westronics, Inc. specializes in its own proprietory line of industrial electronic recording and indicating instruments for both domestic and foreign markets.

Mylar and Foil Capacitor



A Mylar and foil capacitor, Type MW, has been announced by the **Capco Div., Texas Research and Electronic Corp.**, 6612 Denton Dr., Dallas 35, Texas. These capacitors are specifically designed for low drift requirements. Fully stabilized, they

(Continued on page 66A)



Harmonic Generators for Microwave Power

AVAILABLE models of varactor harmonic generators can now provide outputs over most of the microwave spectrum. Microwave Associates' several years of experience in designing and building a wide range of varactor harmonic generators make possible a standard line of which several typical units are described here.

These solid-state sources of microwave power provide precise frequency multiplication, lightweight, small size, and have low power requirements.

We would like to discuss with you how these units can serve your requirements for highly reliable, fixed-frequency RF power sources. If you have special environmental or space requirements, modifications of units in our standard line will most probably save you considerable time.

Model	Application	Input Frequency (Mc)	Output Frequency (Mc)	Input Power (Nominal) Watts	Output Power (Nominal) Watts
MX-8052-16T1 NN	VHF — L	60	960	10	2
MX-8052-24T1 NN	VHF — L	40-43	960-1025	12	1
MX-8051-32T1 NN	VHF — L	40	1280	4	0.2
MX-8034-2S1 NN	L — S	1125	2250	2	1
MX-8057-128X1MW	VHF — X	64	8192	10	0.175
MX-8045-2X1WW	C — X	4500	9000	1.2	0.5
MX-8038-3X1NW	S — X	2800-3100	8400-9300	0.25	0.010

The above units are representative of our current line of over 60 standard harmonic generators. Write for our new four page short form Harmonic Generator Catalog.

MICROWAVE ASSOCIATES, INC.

United Kingdom Sales:
Microwave and Semiconductor Devices, Ltd.
Skimpot Trading Estate, Luton, Bedfordshire, England
Other Export Sales: Microwave International Corp.
36 W. 44th Street, N.Y.C., N.Y., U.S.A. Cable: Microken

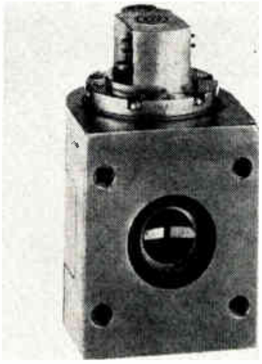
BURLINGTON, MASSACHUSETTS • WESTERN UNION FAX
TWX: BURLINGTON, MASS. 942 • BROWNING 2-3000

NEWS New Products

(Continued from page 64A)

are available from stock in 1% capacity tolerance. The high moisture resistance of the epoxy-impregnated wrap eliminates the need for encapsulation. These capacitors have a positive temperature coefficient with dissipation factors of less than 0.45% for values under 1.0 mfd and less than 0.75% for higher values at 25°C. Available in standard ratings from 0.0005 to 2.00 mfd at 200, 400 and 600 volts dc, they are priced from $.30 to $1.20 in 100 lots.

Tunable Magnetron



Metcom, Inc., 76 Lafayette St., Salem, Massachusetts, offers for the first time a waveguide output 2 KW X Band tunable magnetron. The MXM-28 has been designed to withstand missile type environmental conditions in addition to working under pure vacuum atmosphere. In addition, the MXM-28 has been designed to undergo sterilization under customer's specification for space environs.

The MXM-28 offers 2 KW peak power at 3 KV and 3 amperes input under extreme shock, vibration, and temperature environs.

The MXM-28 is available 60–90 days.

Magnetic Shields For Dewar Flasks

Magnetic Shield Div., Perfection Mica Co., 1322 No. Elston Ave., Chicago 22, Ill., offers a new line of diversionary Netic and Co-Netic magnetic shields which minimize effects of the earth's magnetic field as well as all other low level fields on samples being tested under cryogenic conditions in any size dewar flask. Two, three or four concentric cylindrical shields can be used, one inside

the other. The shield illustrated is 36″ high and 15″ OD. It consists of three inner shields of high permeability Co-Netic alloy and one outer shield of heavy gauge

(Continued on page 68A)

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 66A)

Netic alloy. Shield sections are symmetrically separated by rectangular aluminum spacers positioned around the shield's periphery parallel to its axis. This multi-sectional air gapped construction was selected because attenuations of 1,000 times or better were required. Conduction of external magnetic fields to the inner low level area is minimized by progressively shortening the inner Co-Netic shields and extending the outer Netic shield beyond the inner shields at the open end.

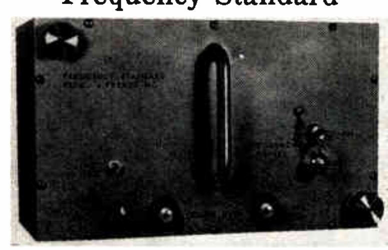
A multi-section viewing port is provided. This consists of four sliding panels, one in each shield section, which can be manually opened or closed sequentially by inserting a hand hook in the panel holes. (See illustration.) The entire shield can be parted at the middle of its length on a butt joint junction seam with overlapping flange.

In addition to dewar flask application, the shield's basic design permits general laboratory usage for a wide variety of other magnetic shielding applications. Although designed specifically for dc magnetic fields, the shield will prove even more effective in medium or low level ac fields. Electrostatic shielding is also provided by simply grounding the outer shield.

Low retentivity and low shock sensitivity are inherently displayed by the Netic and Co-Netic alloys used. All shields have been annealed and require no further heat treating.

Delivery is four to six weeks. Prices are subject to design complexities and size, and range from $150.00 to $1500.00.

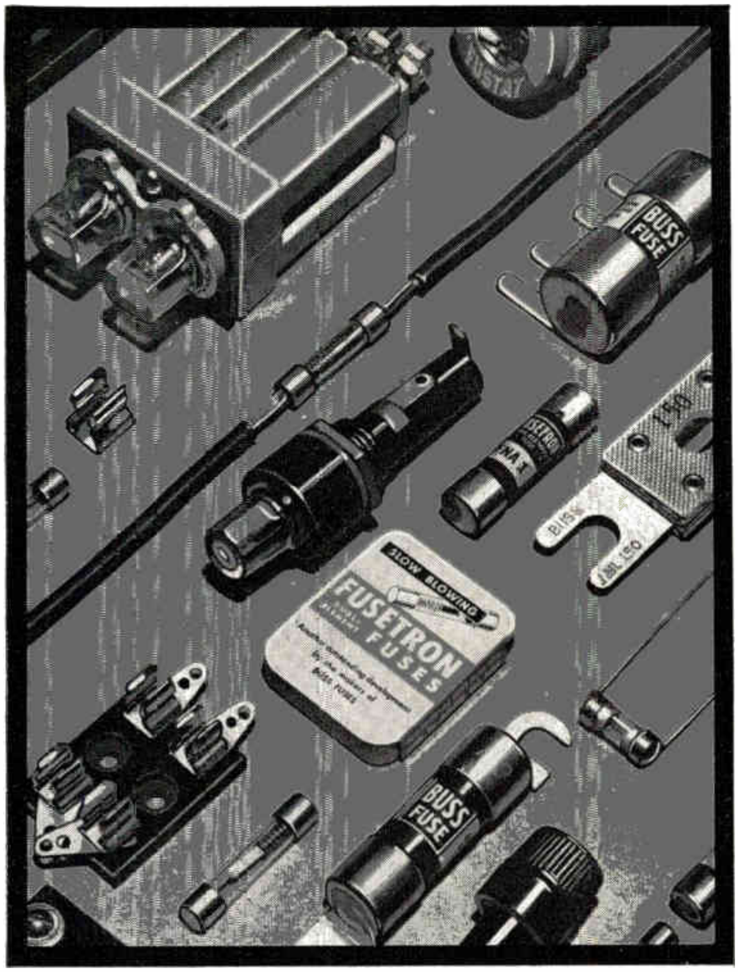
Crystal-Controlled Frequency Standard

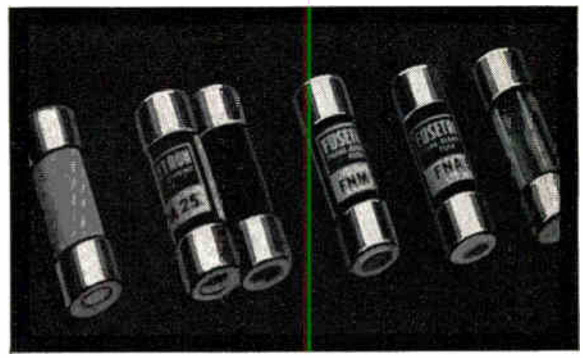


The model S1455, a 4.999600-mc crystal-controlled frequency standard, was designed by **Reeves-Hoffman Div., Dynamics Corp. of America,** Cherry and North Sts., Carlisle, Pa., as an aid to precise navigation and for use as a master oscillator of "TRANSIT" operational test equipment. Operating ambient-temperature range is from 0 to 60°C. Frequency is 5 mc or frequencies nearby; aging, 1 part 10^8 per week; short-term stability 5 parts 10^{11} per second. Output is 55 millivolts, sine wave; power required, 2.7 watts at 12 volts dc regulated to 1% or better, or 33 watts at 28 volts dc regulated to 5%. Output impedance is 50 ohms.

The oscillator measures 5 by 5 by 8½ in., including hardware, and weighs 8 lb. Mounting is by two captive ¼-inch screws.

(Continued on page 81A)





In the *complete* BUSS line...

you can quickly find the right fuse and fuseholder to meet every demand!

Dual-element "slow-blowing", single-element "quick-acting" and signal or visual indicating type fuses . . . plus a companion line of fuse clips, blocks and holders . . . are available from one source — BUSS. You'll save time and trouble by turning first to BUSS when you need fuses and fuseholders.

To safeguard against 'kicks' or complaints, every BUSS fuse is tested in a sensitive electronic device. Any fuse not correctly calibrated, properly constructed and right in all physical dimensions is automatically rejected to assure dependable protection under all service conditions.

Save engineering time on special problems in electrical protection. At your request, the BUSS fuse engineers are at your service to help you determine the fuse or fuse mounting best suited to your needs.

In many cases it is possible to find, in the complete BUSS line, a fuse and fuse mounting already available in local wholesalers' stocks, so that your device can be easily serviced.

For more information on BUSS and FUSETRON small dimension fuses and fuseholders . . . Write for bulletin SFB.

562

BUSSMANN MFG. DIVISION,
McGraw-Edison Co.
University at Jefferson, St. Louis 7, Mo.

BUSS fuses are made to protect - not to blow, needlessly.
BUSS makes a complete line of fuses for home, farm, commercial, electronic, electrical, automotive and industrial use.

FUSETRON · TRUSTWORTHY NAMES IN ELECTRICAL PROTECTION · BUSS · McGRAW EDISON

Industrial Engineering Notes*

Association Activities

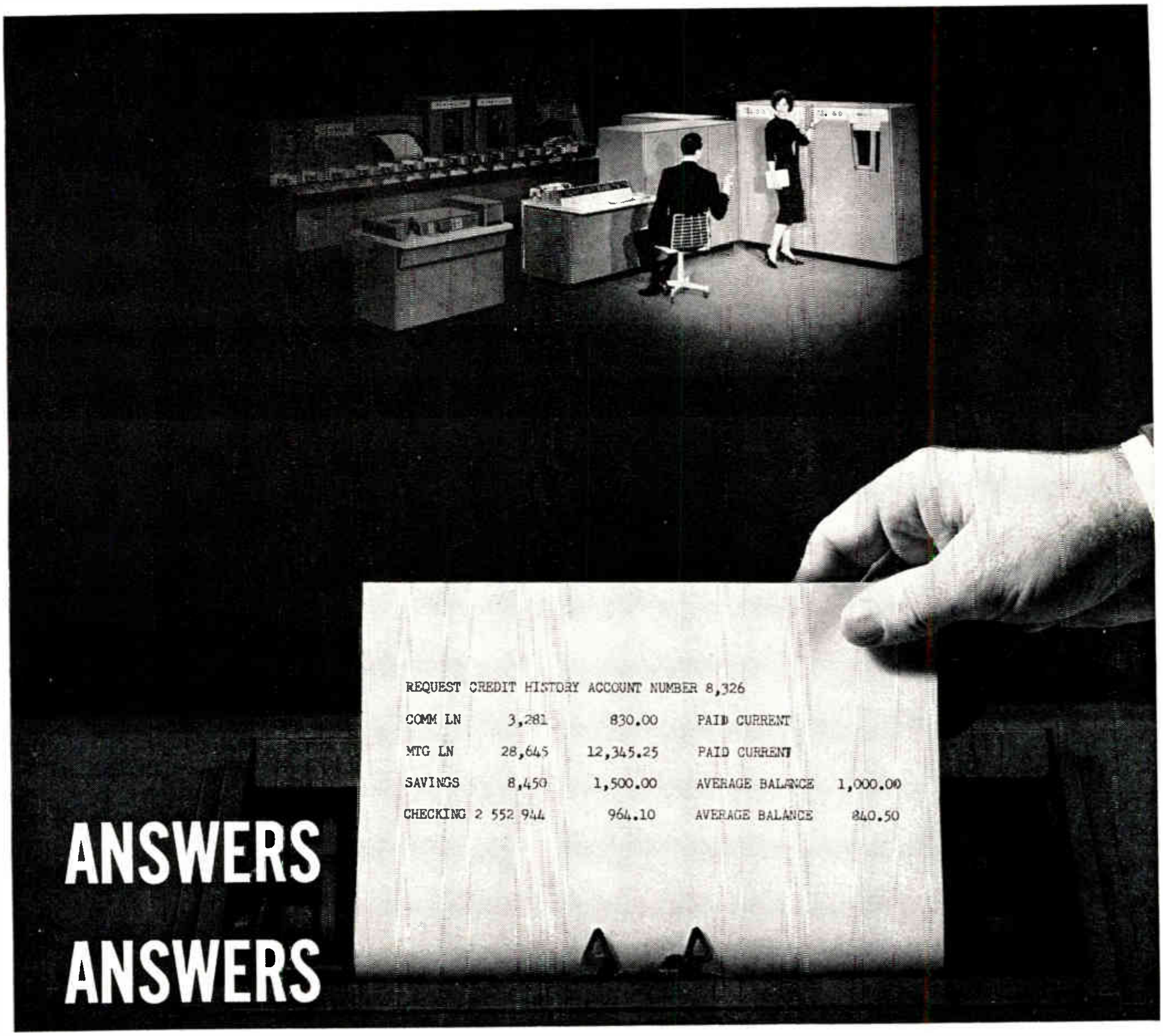
The Military Relations Department has been asked to gauge the interest of EIA member companies in participating in a program of exchanging abstracts of electronic data processing programs aimed at solving EDP problems. Abstracts would be limited to computer techniques, excluding electronic accounting machine procedures. Each abstract would consist of a one-page written explanation of the purpose, scope, and frequency of the system being abstracted, together with a brief description of the system. This would be accompanied by a skeleton schematic diagram of the system. The purpose of the program is to exchange useful information and assist management in effectively applying EDP techniques. Any participant could contact other participants directly if additional information were desired, but there would be complete discretion in each company as to how much it revealed. The abstracts would not be made available to EDP manufacturers for propaganda purposes. To insure mutual benefit, all participants would be required to submit at least two abstracts before receiving any from other companies. Firms interested in participating may contact Manager, Military Relations Department, EIA Headquarters. . . . **The Federal Communications Commission should freeze assignments of television channel 14 until engineering standards can be established to forestall interference to land mobile radio frequencies, EIA told the FCC last week.** EIA pointed out that UHF channel 14 (470-476 Mc) lies immediately adjacent to the land mobile allocation (450-470 Mc). "If the Commission were to permit unrestricted allocation of channel 14 in these areas, serious harm could be caused to land mobile users by the TV side bands," the Association declared. "Furthermore, it is conceivable that interference could be caused by mobile users to TV reception. This interference would be existent even with the use of vestigial side band filters, which (the Commission) proposes to eliminate." The comments on channel 14 were among a number by EIA in a filing in FCC Docket 14229, "Fostering Expanded Use of the UHF Television Channels." Comments on individual proposals in the docket reflected the views of several industry groups, represented in the Association by the TV Broadcast Equipment Section and the Land Mobile Communications Section, both of the Industrial Electronics Division; and the Consumer Products Division. EIA declared it supported the move by the Commission to expand use of UHF television channels. Specifically, it made the following points: 1) UHF TV transmitting equipment can be designed and built for remote control operation, and requirements proposed by the FCC appear adequate for remote control and monitoring. 2) Elimination of restrictions on antenna directivity for UHF stations, as proposed by the FCC, will permit greater flexibility in station location and the possibility of meeting market coverage requirements with lower transmitter power. 3) On the FCC's recommendation to eliminate the vestigial side band attenuation for all UHF receivers, EIA restated an earlier comment that attenuation should be a function of transmitter power and recommended a formula for attenuation to replace one proposed by the Commission. The Association also said the recommended elimination of the vestigial side band filter might result in degraded performance in some existing receivers having significant response in the lower side band region. 4) The proposed reduction of the ratio of visual-to-aural power should not be made, EIA declared, because reduction in sound power results in deterioration of receiver performance and a reduction in service coverage.

Governmental and Legislative

The President last week established the long-expected Director of Telecommunications Management to administer Government radio frequencies. The new post was placed in the Office of Emergency Planning (OEP), which took over the telecommunications functions of the Office of Civil and Defense Mobilization when that agency's responsibilities were split between the Defense Department and OEP last year. Concurrent with his Executive Order establishing the position, President Kennedy disclosed the impending appointment of Dr. Irvin Stewart as Director of Telecommunications Management. The 63-year-old Texan is a former member of the Federal Communications Commission and a past President of West Virginia University. Dr. Stewart will coordinate telecommunications activities of the Executive Branch of the Government and will be responsible for formulating its telecommunications policies and standards. He will also have authority to assign radio frequencies to Government agencies. The White House emphasized the move will not affect the authority of the FCC to regulate non-Federal use of radio or other communications. From 1934 to

(Continued on page 72A)

* The data on which these Notes are based were selected by permission from *Weekly Report*, issues of February 26, March 5 and 12, 1962, published by the Electronic Industries Association, whose helpfulness is gratefully acknowledged.



ANSWERS...with CRAM and Remote Inquiry

provided by the NCR 315 Computer System

The NCR 315 CRAM (Card Random Access Memory) Computer System is more than just another back-office electronic accounting machine . . . it is a vital electronic tool which contributes to a higher-quality of management and personnel effectiveness.

For example: in a bank, Inquiry Units placed in the various departments enable authorized personnel to interrogate the computer files at will . . . to promptly answer requests for balances . . . to quickly obtain valuable credit information . . . to obtain current, up-to-the-minute reports on investments, trusts, loans, and other essential data.

In industry, Inquiry Units can be located at dozens of remote locations, enabling people to communicate with the computer files . . . even from hundreds of miles away. With the NCR 315 you will be able to keep a "current-finger" on the pulse of your business . . . to get immediate answers to questions about inventories, production, sales . . . and a host of other timely facts people must have to effectively manage . . . and to act while the "iron is hot."

The NCR 315 CRAM System provides many exclusive advantages for Remote Inquiry, Data Processing, On-Line Accounting, Operations Research, Engineering and Scientific Applications. For more information, call your nearby NCR representative or write to Data Processing Systems and Sales, Dayton 9, Ohio.

NCR PROVIDES TOTAL SYSTEMS—FROM ORIGINAL ENTRY TO FINAL REPORT—
THROUGH ACCOUNTING MACHINES, CASH REGISTERS OR ADDING MACHINES, AND DATA PROCESSING
The National Cash Register Co.•1,039 offices in 121 countries•78 years of helping business save money

NCR

Industrial Engineering Notes

(Continued from page 70A)

1937, Dr. Stewart was a member of the Federal Communications Commission and Chairman of the FCC Telegraph Division. He was Director of the Commission on Scientific Research from 1937 to 1946 and in the latter year was elected President of West Virginia University, a position he held until 1958. Dr. Stewart joined the State Department in 1930 as head of electrical communications for the Treaty Division. In this post, he was a member of various American delegations to international radio conferences in Washington, Copenhagen, Madrid, and Mexico City.

Industry Marketing Data

Electronics output and employment will reach another all-time high in 1962, the Business and Defense Services Administration of the Department of Commerce reported last week. Because of increasing defense procurement, accelerating space exploration programs, and greater industrial demand, total output of systems and equipment is expected to exceed $7.2 billion in 1962, about 7 per cent more than in 1961, BDSA said. Total output of components should exceed $3.7 billion. The projections do not include the value of electronic research, development, evaluation, and test expenditures or distribution, service, installation, and operating revenues, the agency pointed out. Manufacturers' shipments of consumer products are expected to increase modestly in value, mainly because of new product promotion of color television receivers, stereophonic sound equipment, electronic toys and kits, and other products. Output of commercial and industrial electronic equipment will be stimulated by increased capital goods expenditures, including plant modernization programs and rapidly expanding space activities. Substantial increases in defense expenditures for advanced weapons systems of high electronic content may be partially offset by cutbacks and cancellations of new programs. Because most electronic components are used in the manufacture of new equipment, output of components will increase proportionately. Both domestic and foreign competition will be more intense; thus profit levels may not keep pace with rising output, BDSA said. Electronic research, development, test, and evaluation activities—a major part of total electronic effort—will continue to increase rapidly and may exceed $2.8 billion in 1962, according to BDSA. The major impact of advanced weapons and space programs will be felt in the research and development area. Electronics output in 1961 generally followed expectations, it was reported. Factory value of electronics production reached an all-time high of $6.7 billion—about 5 per cent above the 1960 level of $6.4 billion. Output of consumer electronic products declined in

(Continued on page 74A)

Industrial Engineering Notes

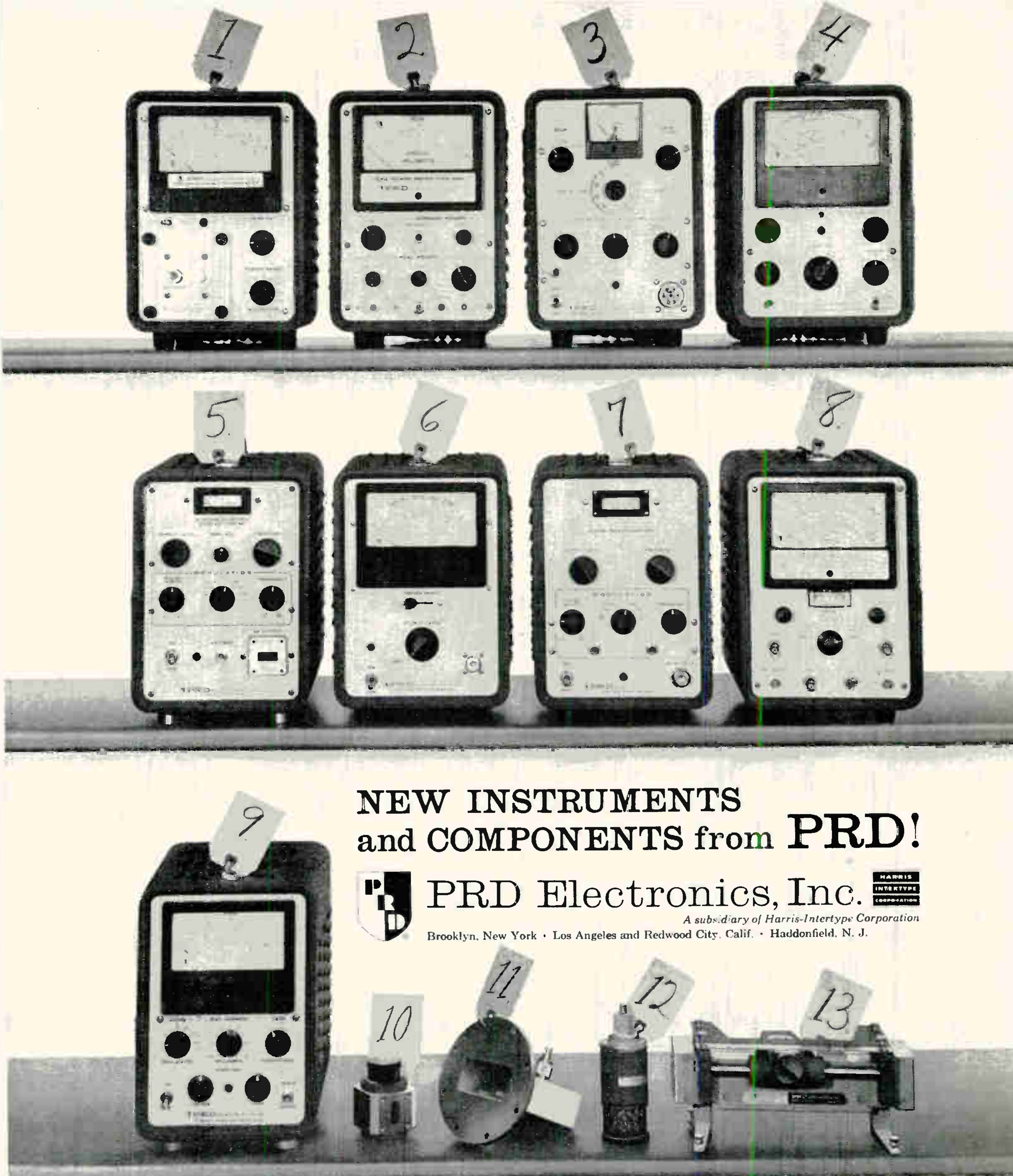
(Continued from page 72A)

the early months of 1961 but increased rapidly in the latter part of the year. The volume of production of most consumer electronic products in 1961 exceeded that in 1960, but the total value of shipments declined almost 4 per cent. This was attributed to a general shift in consumer buying to less expensive models and a decline in phonograph sales. Production and sales of consumer electronic products are now at high levels, and output is expected by BDSA to increase about 7 per cent in 1962 . . . **Japan's exports of electronic products to the United States during the first nine months of 1961 totaled $78.4 million, an increase of 24 per cent over the same period in 1960, according to Japanese statistics, the Electronics Division of the Commerce Department's Business and Defense Services Administration disclosed last week.** Radio receivers and chassis together with sound recorders and reproducers accounted for 76 per cent of the total value, compared with 81 per cent of the total during January-September 1960, it was reported. BDSA said products registering substantial gains were sound recorders and reproducers, up $6 million; tube-type radios, up $5.7 million; and radios with less than three transistors, up $3.6 million. Exports to the United States of radios with three or more transistors declined $7.3 million, while shipments of this type radio to the rest of the world increased. During January-September 1961, Japanese exports of television receivers to the United States totaled nearly a million dollars. Exports of transistors doubled in dollar volume—from $0.8 million to $1.6 million—while tripling in quantity over exports for January-September 1960 . . . **Factory sales of transistors totaling more than 18 million a month during the last quarter of 1961 boosted the year's total to 190,916,354, a whopping 62,987,768 more than the 127,928,586 sold the previous year, according to the EIA Marketing Data Department's year-end compilation.** But while sales volume was up by an impressive margin, sales value dipped from $301,432,285 in 1960 to $299,538,760 last year, a drop of $1,893,525. December transistor sales at the factory remained relatively stable compared to the previous month. In December, 18,166,839 units worth $24,883,538 were sold, against 18,342,285 valued at $24,034,703 sold the month before.

Engineering

Two new recommended technical standards have been completed by EIA engineering committees and published by the Engineering Department. They are: 1) RS-255—*Simulated Life Test Circuit for Semiconductor Rectifier Diodes* (25 cents). Defines a simulated life

(Continued on page 76A)



NEW INSTRUMENTS and COMPONENTS from PRD!

PRD Electronics, Inc.

A subsidiary of Harris-Intertype Corporation

Brooklyn, New York • Los Angeles and Redwood City, Calif. • Haddonfield, N. J.

1. PRD N680 Calorimetric Power Meter 2. PRD 668 Peak Power Meter 3. PRD 809-A Klystron Power Supply 4. PRD 277-B Standing Wave Amplifier 5. PRD X712 Signal Source 6. PRD 904-A Noise Generator 7. PRD S712 Signal Source 8. PRD 279 Ratiometer 9. PRD 650-C Microwave Power Meter 10. PRD 4000 Series Waveguide Switches 11. PRD 6608 Bolometer Mounts 12. PRD 3302 Calibrated Susceptances (For PRD 219) 13. PRD 232 & 233 Slotted Line and Carriage ...and many, many more! Send for data! PRD Electronics, Inc., 202 Tillary St., Brooklyn 1, N. Y.

Industrial Engineering Notes

(Continued from page 74A)

test circuit for use when actual life tests would require the use of large quantities of power due to the size of the rectifier diode being tested or the number of units under test. 2) RS-256—*Deflecting Yokes for Cathode Ray Tubes* (25 cents). Defines the mechanical and electrical parameters of 70°, 90°, 110°, and 114° deflecting yokes.

Copies of the standards have been mailed to each EIA member company. Additional copies are available from the EIA Engineering Department, 11 West 42nd Street, New York 36, N.Y. Minimum order is $1 unless EIA Standard Coupons are used. Books of coupons may be ordered from the Department for $25 each.

Section Meetings

ALBUQUERQUE-LOS ALAMOS

"Optical Masers," G. Dacey, Sandia Corp.; "Presentation on Photography," R. Cudney, Le Gaunt Studio; 2/21/62.

ATLANTA

"Georgia's Future in Peaceful Uses of Atomic Energy," W. B. Harrison, Georgia Inst. of Technology; 2/19/62.

BALTIMORE

Discussion of the provisions of the proposed IRE-AIEE Merger, P. E. Haggerty, Pres. of IRE; 2/12/62.

BEAUMONT-PORT ARTHUR

"Management of Polaris Program," W. F. Raborn, Jr., USN; 2/19/62.

CANAVERAL

"Errors in Sampled Data Systems," L. Gardenhire, Radiation, Inc.; D. McRae, Research Engineer; 2/15/62.

CEDAR RAPIDS

Installation of New Officers; 1/26/62.

"The ABC's of Wave Propagation in Plasma," M. Z. von Krzywoblocki, Michigan State Univ. 2/14/62.

COLUMBUS

"Economic Need For an Ohio Research Center," W. E. Chope, Industrial Nucleonics Corp.; 2/13/62.

ERIE

"Electrical Breakdown in Gases," A. C. J. Leyn, Duquesne University; 12/28/61.

"Piezoelectric Ceramics and Their Applications," L. Shoot, Erie Technical Ceramics, State College; 2/28/62.

FLORIDA WEST COAST

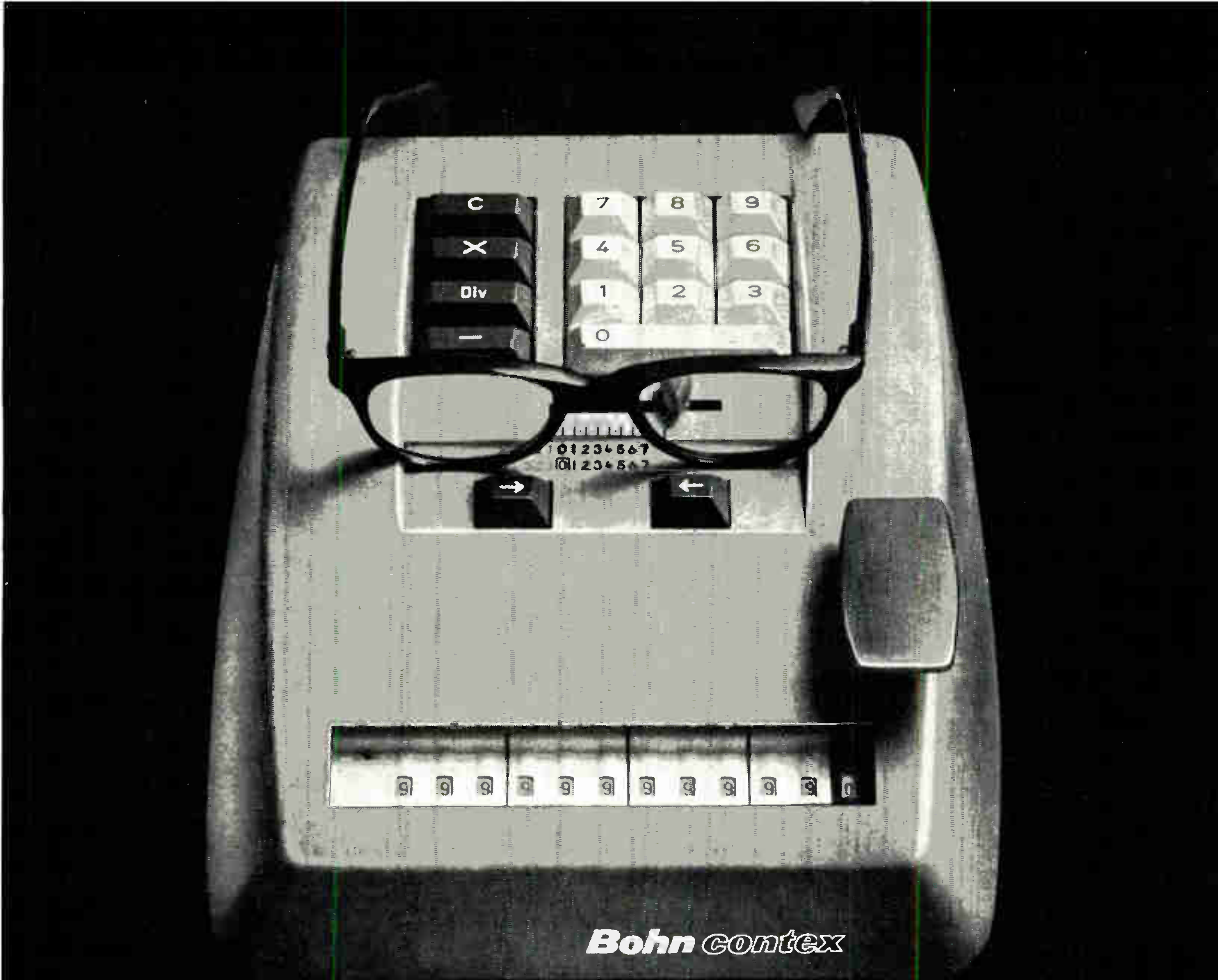
"Space Engineering," B. G. MacNabb, General Dynamics; 2/22/62.

FORT WAYNE

"Electro-Luminescence," M. Wasserman, General Telephone Labs.; Joint with AIEE; 1/11/62.

"George Washington—President and Engineer," J. Thimlar, Bowmar Instrument; "Plastics In Construction," E. Zieglar; 2/22/62.

(Continued on page 78A)



BRAINCHILD (it figures!) **ONLY $125** (it doesn't figure!)

This smart little machine weighs only six pounds, yet does just about everything the big, expensive machines do...and a couple of things they can't.

The big machines add, subtract, multiply, divide.

So does the Bohn Contex.

Here's what the expensive calculators *can't* do.

They can't double as high-speed 10-key adding-subtracting machines.

The Bohn Contex can.

They can't be operated by anyone after just a few minutes instruction.

The Bohn Contex can.

Most important of all, the Bohn Contex is the first compact, 10-key machine that's truly portable. Goes to work anywhere.

Businessmen, salesmen, engineers, architects, students, housewives find the Bohn Contex the easiest, fastest way to figure bills, invoices, percentages, extensions, estimates, income tax..........................All this for **$125***

For hour-after-hour use, the Bohn Contex *Electric*. Same size, same features, but greater speed and ease because the motor does the work.....................Only **$235***

To a businessman who pays $500 to $1,000 for an office calculator, the price of our Brainchild seems too good to be true. Our suggestion: *borrow a Bohn Contex and see how it figures to save you money*. Any dealer listed below will lend you a hand (or electric) model. If the dealer in your city is not listed, or you're not in a borrowing mood, just mail the coupon.

*PLUS F. E. T. CARRYING CASE $14.95

Bohn Business Machines Inc., Dept. BY-10
444 Park Avenue South, New York 16, N. Y.
Send detailed "specs" on your Brainchild ☐ *manual* ☐ *electric.* ☐ *Send name of nearest dealer.*

NAME____________________

FIRM____________________

ADDRESS____________________

CITY____________STATE____________

BOHN CONTEX CALCULATOR

Atlanta, DeKalb Office Equip.; Tidwell Co. **Bakersfield**, Valley Office Supply **Baltimore**, Lucas Bros. **Billings**, Western Office Equip. **Boise**, Kalbus Office Supply **Boston**, Thorp & Martin **Bridgeport**, Frank McGuire **Buffalo**, Office Automation **Charleston, S.C.**, Fred Johnson Inc. **Chicago**, Horder's; Benbow **Cincinnati**, Pounsford's **Columbus**, Dorsey's General Office Equip. **Dallas**, S. L. Ewing **Denver**, Kistler's **Des Moines**, Berry-Myron **Detroit**, Macauley's **Eugene**, Clary Bus. Mach **Fresno**, Stanton Office Mach **Houston**, Maverick-Clarke **Indianapolis**, Rex. Bus. Mach. **Jacksonville**, Frank Coar Agency **Kansas City**, Kansas City Loose Leaf **Lancaster, Pa**., Weaver Book Store **Lansing**, Hasselbring **Los Angeles**, Office Machines Inc.; Long Beach, California Calculator; Sherman Oaks, Valley Stationers **Louisville**, Fetter's **Memphis**, S. C. Toof **Miami**, Accurate Bus. Mach. **Milwaukee**, Clark Bus. Mach. **Minneapolis**, Farnham's **Newark**, New Jersey Office Supply **New Orleans**, Amann Bus. Mach. **New York**, Mineola, L. I., D. Waldner **Ogden**, Weber Office Supply **Palo Alto**, Delmer Israel **Pasadena**, Anderson Typewriter **Philadelphia**, Pomerantz **Phoenix**, Heinze, Bowen & Harrington **Pittsburgh**, Milton Wiener **Portland**, Oregon Typewriter **Raleigh**, Business Machines Inc. **Rochester**, Dunphrey Duplicating **Sacramento**, David's Office Equip. **St. Louis**, Von Soosten & Co. **Salem**, Needham's **Salinas**, Peninsula Typewriter **Salt Lake City**, Salt Lake Desk **San Bernardino**, Stockwell & Binney **San Diego**, Nelson Bus. Mach. **San Francisco**, Quality Office Mach. **San Jose**, Kennedy Bus. Mach. **Santa Barbara**, Lund Office Equip. **Savannah**, Chatham Typewriter **Seattle**, J. K. Gill & Lowman-Hanford **Springfield, Mo.**, Carroll Typewriter **Tucson**, Tucson Office Supply **Washington, D.C.**, Ginn's, Stockett-Fiske **Wichita**, Wilbur E. Walker **Wilmington**, Hilyard's **Winston-Salem**, Hinkle's **In Canada**, Scandinavian Bus. Mach., Montreal

Section Meetings

(Continued from page 76A)

HUNTSVILLE

"Fundamentals of Masers," A. J. McKinnon, Western Electric Co.; 11/17/61.

"Plasma Physics Fundamentals," R. C. Watson, Jr., Brown Engineering Co.; 12/21/61.

"Saturn Telemetry System," J. E. Rorex, NASA—MSFC, Astrionics Lab.; 1/26/62.

"Radio and Light Waves, Their Similarities and Differences," G. F. Hull, Sylvania Engineering Systems Lab.; Demonstrations; 2/23/62.

INDIANAPOLIS

Tour of Western Electric Company Plant; 2/15/62.

KANSAS CITY

"Simulation of the Fractional Operators on the Analog Computer," C. Halijak, University of Kansas; 2/13/62.

KITCHENER-WATERLOO

"Automatic Programming of a Radio Station," G. A. Robitaille, CFPL Radio; 2/19/62. (Canada)

LONDON

"A Precision Frequency Standard for X-Band," A. Marriage, Univ. of Western Ontario; "High Frequency Ionospheric Reflection," R. Turnbull, Univ. of Western Ontario; 3/5/62. (Canada)

LOS ANGELES SECTION

"Image Storage in Information Retrieval," L. D. Stevens, IBM Corp.; Joint with Orange Belt Subsection; 2/13/62.

LOUISVILLE

Installation of Officers for year 1961–1962; 9/28/61.

"Transistor Theory and Application," (Film), E. W. Szecewa & G. Reiling, RCA; 10/19/61.

Executive Council Meeting; 12/27/61.

"Keeping Current in Plastics," N. James, E. I. DuPont Co.; Joint with AIEE; 2/9/62.

MIAMI

"A Synchronous Cardiac Pacemaker," W. Keller & D. Nathan, Cordis Corp.; 2/21/62.

MONTREAL

Visit to Ecole Polytechnique, J. C. Bernier; 2/14/62.

NEW ORLEANS

"Educational Television Program Production Systems and Studio Techniques," W. S. Hart, WYES-TV; Tour of facilities; 2/23/62.

NEW YORK

"Fiber Optics,"—Panel Discussion, R. J. Potter, IBM Corp.; J. W. Hicks, Mosaic Fabrications; L. Curtiss, American Cystoscope Makers, Inc.; E. Snitzer, American Optical Co.; 3/7/62.

NORTH CAROLINA

"Sound Recording," F. L. Hopper, Bell Telephone Labs.; 2/16/62.

OMAHA-LINCOLN

"Contemporary Trends In Engineering Education," R. W. Schmelzer, Rensselaer Polytechnic Institute; NSPE—Engineers Week Exhibition; 2/22/62.

PHILADELPHIA

Discussion of Proposed Consolidation of IRE and AIEE; R. M. Showers; 11/17/62.

"Crossroads for Engineers," P. Haggerty, President of IRE; 2/3/61.

SAN DIEGO

"Applications of Pattern Recognition Techniques to Medical Data Processing," P. Broome, General Dynamics/Astronautics; 2/7/62.

SOUTH CAROLINA

"Muscles to Missiles," J. L. Powell, Dept. of Defense; 2/20/62.

SOUTHERN ALBERTA

"The Cause and Prevention of Signal Loss In Parabolic Microwave Antennas," J. C. Annett, Andrew Antenna Corp.; 2/21/62.

TORONTO

"Solid State Power Inverters and Converters," B. C. Hansen, Univ. of Toronto; 2/22/62.

TULSA

"The Two-Million Watt VLF Navy Transmitter at Cutler, Maine," Mark W. Bullock, Continental Electronics; 2/15/62.

VIRGINIA

"Active Satellites for World Wide Communication," L. Pollack, IT&T Co.; 2/16/62.

WESTERN MASSACHUSETTS

Talk on The Proposed IRE-AIEE Merger, R. L. McFarlan, Datamatic Corp. and Raytheon Mfg. Co.; 3/5/62.

WINNIPEG

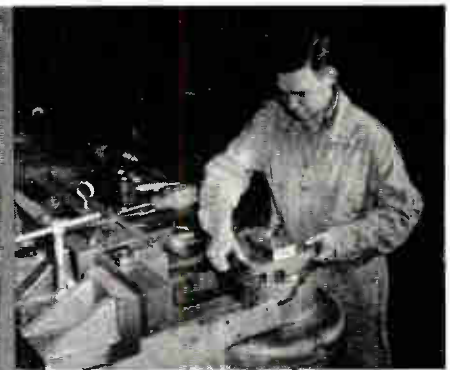
"Microwave Antennas," J. Annett, Andrew's Antenna Co.; 2/19/62.

(Continued on page 80A)

here's how a four-bend waveguide is processed at specialty automatic

Waveguide, taken from stock . . . is first cut to proper length . . . and then annealed for bending.

Special laminated shims must now be inserted into the waveguide to prevent its collapse during the bending process.



The proper die is then placed in position in the automatic bending machine preparatory to making the first bend.

Here the waveguide is shown still in the bending machine just after the first of the four bends has been made.

The shims are now removed in preparation for the next step in the bending process.

After each bend the waveguide is sized and the angle is checked to assure that all tolerances are correctly maintained.

Prior to each of its subsequent bends the waveguide must be properly positioned to give accurate register of the next bend (illustration left). The second, third and fourth bends are shown here as each has been processed on the automatic bending machine.

After the final bend has been made, sized and checked, rigid inspection takes place, assuring accuracy of the completed waveguide.

Whatever your needs — bends, twists, offsets, coils, transitions, fabrications, complete assemblies — SPECIALTY AUTOMATIC is ready to precision manfacture to your prints.

Specialty AUTOMATIC MACHINE CORPORATION

80 Cambridge Street, Burlington, Massachusetts • BRowning 2-0800

Section Meetings

(Continued from page 78A)

SUBSECTIONS

BUENAVENTURA

"Project 'Oscar,'" D. Stoner, Stoner Electronics Sales; 2/14/62.

CATSKILL

"Economic Growth Through Engineering," J. Medaris, Lionel Corp.; Joint with Technical Societies; 2/21/62.

EASTERN NORTH CAROLINA

"Thermal Considerations in Transistor Circuit Design," B. Rittman, Minneapolis-Honeywell; 2/9/62.

MID-HUDSON

"Economic Growth Through Engineering," J. B. Medaris, Lionel Corp.; Joint with Technical Societies; 2/21/62.

NORTHERN VERMONT

"Communications Satellites," J. D. Tebo, Bell Tel. Labs.; Joint with Technical Societies; 2/19/62.

PALM BEACH

"The Electronic Heart," Dr. Nathan, Dr. Center & Mr. W. Keller, Cordis Corp.; 12/19/61.

PANAMA CITY

"Microminiaturization Techniques," L. Farabee, RCA Corp.; 2/27/62.

SANTA ANA

"Fallout Defense," W. F. Libby, UCLA; 2/13/62.

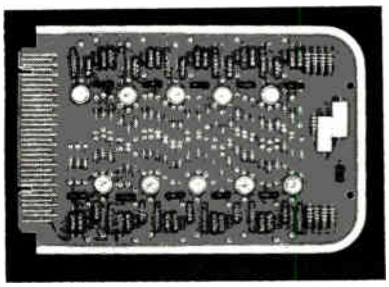
NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 68A)

Octal/Decimal Decoder

Computer Control Co., 983 Concord St., Framingham, Mass., announces the recent addition of the Model OD-30 Octal /Decimal Decoder to its compatible series of dc to 1 and 5 megacycle S-PAC digital modules.



The OD-30 contains a prewired binary-to-octal decoder plus additional circuitry to expand the matrix for BCD-to-decimal decoding. Three additional inputs are provided to permit the matrix to be expanded

(Continued on page 82A)

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 81A)

to 16, 32, or 64 outputs by connecting additional decoders. One or more of these additional input lines may be used for strobing or sampling the matrix.

The BCD-to-decimal decoder uses the octal matrix for the "zero" through "seven" output lines, and two additional independent gates, included on the PAC, for output lines "eight" and "nine." The two independent gates are standard NAND gates and may be used as such if BCD-to-decimal decoding is not required. One of these gates has six inputs, the other has two. Both gates have modes for expanding the fan-in to a maximum of ten. Each output will drive the specified number of standard S-PAC loads plus stray capacitance.

Packaging features include a $4\frac{1}{2}'' \times 7''$ glass-impregnated epoxy card, etched wiring, an anodized aluminum frame, and a 34-terminal hermaphroditic type connector designed for insertion into a standard, prewired 19- or 28-connector S-BLOC housing.

APD Data Acquisition System

Sampling rate of 200 channels per second is now available in a new high speed version of an analog to pulse duration (APD) data acquisition system for process control applications developed by Genisco, Inc., 2233 Federal Ave., Los Angeles 64, Calif.

The new unit adds the advantages of high speed sampling to the accuracy and stability of the original APD. The system directly converts a low level electrical input signal to a pulse linearly related in duration to the input signal amplitude. It is applicable with DC-voltage, low impedance sensing instrument sources such as thermocouples, resistance thermometers, strain gage transducers.

The pulse duration output is digitally measured for display, recording, or for further processing for input to digital computing equipment. The system retains the permanent "sample and hold" feature of the original APD and can be designed to operate either in sequential or parallel data sampling modes. (Output values can be in engineering units.) Modular design permits systems of several hundred channels to be built.

Operational characteristics include 5-

(Continued on page 84A)

GIVE YOUR YOUNG ENGINEERS A HEAD START

Send today for this professional development kit prepared by successful engineers

The first five years after graduation are the most important years of an engineer's career. That's why this carefully planned program of professional development can help your engineers—just as it is now helping so many others.

This kit, sponsored by eight of the foremost engineering societies provides a framework for planning, focuses attention on six vital areas:

1. Starting your career
2. Education after college
3. Professional identification
4. Responsible citizenship
5. Selected reading
6. Personal appraisal

The complete kit contains a 48-page reference manual, a statement of the responsibilities and ethical principles of the engineering profession, a selected reading list, a personal appraisal booklet, and a brochure outlining the complete six point program.

Start now to help your engineers plan their careers with professional help . . . send in the coupon today.

COMPLETE KIT $2.00—*quantity discounts on request*

ENGINEERS COUNCIL FOR PROFESSIONAL DEVELOPMENT
345 East 47th Street, New York 17, N.Y.

Gentlemen: Please send me ______ *First Five Years* Kit (s) for which $_____ is enclosed.

Name ________________

Position ________________

Firm ________________

Address ________________

City, Zone, State ________________

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 82A)

millisecond sampling time for each channel, resolution of 0.1% (1 part in 1000), repeatability of ±0.1% ± one output count per 1000; linearity of ±0.1% of full scale input, sensitivity of 5 millivolts minimum full scale input. Source impedance is a nominal 100 ohms for thermocouples. Converter output is 10-volt pulse with less than 2 microsecond rise and fall times, output impedance of 1800 ohms. Power requirements: Approximately 1 watt /channel, 110 volts, 60 cps. Special sampling rates from 5 to 500 samples/second are available on order. For additional information, contact the firm.

Dual Preset Counter-Controller

A new all transistorized 100 kc dual preset counter-controller has been announced by **Computer Measurements Co.**, 12970 Bradley Ave., San Fernando, Calif.

The 780 Series counters are designed to solve control problems, such as motor overspeed, shearing to length, high speed coil winding, weighing, sorting, batching, func-

(Continued on page 85A)

tion programming and timing, and variable pulse generation. CMC offers this instrument with a two-year free service warranty.

The counter-controller is a direct-reading solid state electronic counter providing output information in the form of pulses and relay closures at any two preselected counts within the capacity of the unit. Capable of operating at rates up to 100,000 counts per second, the instrument accepts input information from any standard transducer such as a photocell, tachometer, flow and pressure pickups, switch closures, and so forth. Selection of the preset numbers is accomplished by thumb wheels adjacent to each decade. Input sensitivity is 0.1 volt.

The series is available in two-through six-decade models. Power consumption is 40 watts. The unit measures $5\frac{1}{4}''$ H $\times 17''$ W $\times 12''$ D and weighs 21 pounds. It is available in either rack or cabinet mounting. Prices range from \$975.00 to \$1350.00. Delivery 60 days.

For complete technical information, please address Bob Steele, sales manager at the firm.

(Continued on page 86A)

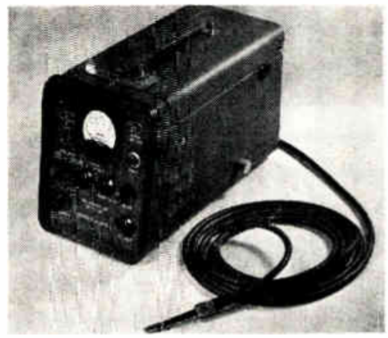
NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 85A)

Halogen-Sensitive Leak Detector

General Electric of Schenectady, N. Y. has introduced a new halogen-sensitive, electronic leak detector specifically designed for use in halogen-contaminated atmospheres. Designated the H-5P and featuring a new proportioning probe, this detector is sensitive enough to detect halogen leaks as small as 1/100 ounces Refrigerant-12 per year even when contamination levels reach 1000 parts R-12 per million parts air.



The H-5P leak detector can be used to test pressurized enclosures by manually probing seams, joints and welds. When a leak is detected, an audible alarm sounds.

In operation, the H-5P draws air from the atmosphere and from a pure air filter in the control unit. The proportioning probe rations the amount of air coming from both sources. By adjusting the probe's proportioning valve, the operator can reduce the amount of contaminated air entering the detector and increase the amount of pure air coming from the filter. As a result, the effects of contamination are neutralized and the H-5P operates at full sensitivity.

Because of the new proportioning probe, the detector's sensitive element is protected against exposure to large amounts of contamination. General Electric engineers say that this will result in minimum maintenance of the unit and longer sensitive element life.

High Power TWT Supply

A TWT supply designed specifically for high power TWT's is available as a standard unit. Developed by **The Narda Microwave Corp.**, Plainview, L. I., N. Y. and designated Model 15101, the unit is available in eight weeks.

The beam supply is continuously variable from −2 to −12 kv by means of both coarse and fine controls on the front panel. The DC bias supply is also continuously variable from 0 to 250 V negative with respect to the cathode, and it floats at the beam voltage. The DC filament supply can be varied between 0 and 10 V and is connected to float at the cathode voltage. The grid pulser, which provides for complete control of the pulse width, amplitude, and repetition rate, is referenced to the bias supply and, therefore, also floats at approximately the beam voltage. In all cases, each of these voltages is completely controllable from ground level at the front panel.

Meters are provided to monitor both voltage and current of the filament and bias supplies, and the high voltage power supply. The filament and bias meters are located behind a shielded viewing panel. An output is also provided for viewing the performance of the grid pulser.

Other features included are a vac-ion power supply, and a provision for external synchronization at any repetition rate compatible with the 0.005 maximum duty cycle and pulse width range.

The price is $9,925. For additional information and more complete specifications, write directly to Narda.

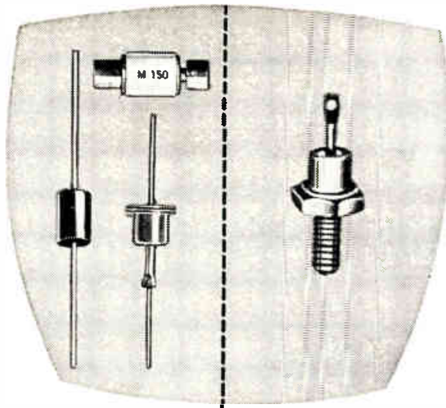
Magnetic Tape Adapter

A magnetic tape adapter unit which provides IBM 1401 computer users with read/write capability from GE/ERMA

(Continued on page 88A)

Low Current Silicon Rectifiers

22 types, with ratings from 0.15 amps to 1.50 amps; 100 to 2800 piv.

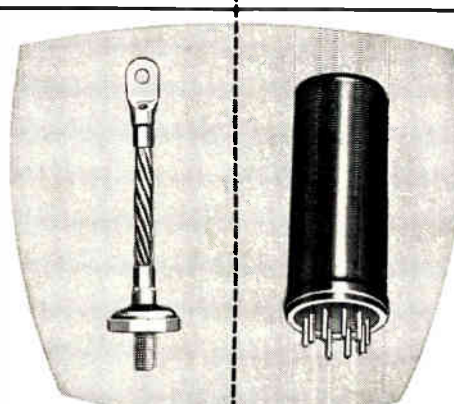


Medium Current Silicon Rectifiers

36 types, with ratings from 2 amps to 35 amps; 100 to 600 piv; many with choice of positive or negative base polarity.

High Current Silicon Rectifiers

36 types, with ratings from 50 amps to 1000 amps; 100 to 600 piv; most with choice of base polarity.

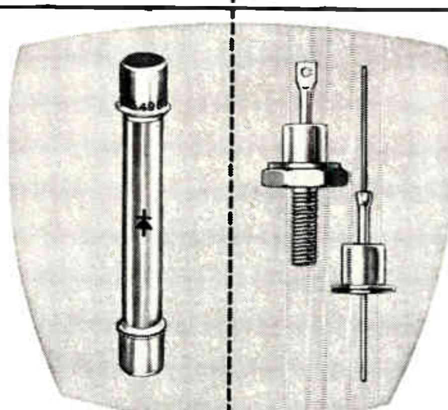


Tube Replacement Silicon Rectifiers

Long life, cool-operating, compact units replacing 95% of all popular vacuum tube rectifiers. PIV ratings from 1600 to 10,400; dc output current ratings, 250 to 750 ma.

High Voltage Cartridge Silicon Rectifiers

Ferrule mounted and axial lead series, each in 18 different types; 600 to 16,000 piv.

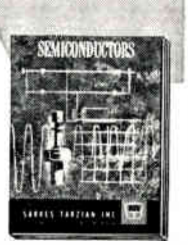


Silicon Voltage Regulators

Regular series: 93 types in ¼, 1, and 10 watt classifications; 5.6 to 100 volts breakdown. Special series: 17 types, 1 watt, 6 to 105 volts breakdown; excellent performance at low cost.

7 invitations

to invention in circuit design... seven lines of Tarzian semiconductor products offering dependable performance at realistic prices, plus interested and informed application engineering service and production and development facilities to help you solve special problems or meet special requirements. Tell us your need or send for our newest catalogs, or both.



Klipvolt Surge Suppressors

136 types, polarized, single phase and three-phase non polarized; maximum discharge currents from 0.25 to 33.0 amps. New application data catalog.

SARKES TARZIAN, Inc.

World's Leading Manufacturers of TV and FM Tuners • Closed Circuit TV Systems • Broadcast Equipment • Air Trimmers • FM Radios • Magnetic Recording Tape • Semiconductor Devices

SEMICONDUCTOR DIVISION • BLOOMINGTON, INDIANA

In Canada: 700 Weston Rd., Toronto 9 • Export: Ad Auriema, Inc., New York

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 86A)

or GE 210 to IBM format and vice versa, is now available from **Electronic Engineering Co. of California,** Box 58, Santa Ana, Calif.

The 754 magnetic tape adapter operates as a program control input/output unit of IBM 1401 and ties to it in place of one IBM 729 Tape Unit. Data may be edited or updated during the conversions as programmed by the IBM 1401. The complete ERMA and IBM error detection circuitry are provided in the system. All data translated to and from the EECO 754 is in IBM coding and format. All normal capabilities of the IBM 1401 are available to process data.

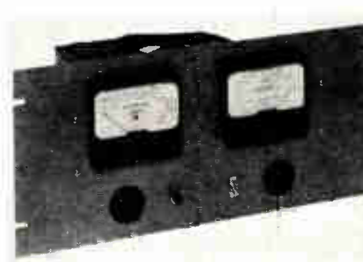
The 754 is housed in two equipment cabinets 69" high, with a total width of 48", a depth of 24", and weighs 800 lbs. All active components in the system are solid state construction.

Price is $85,000 and delivery is 120 days after receipt of order.

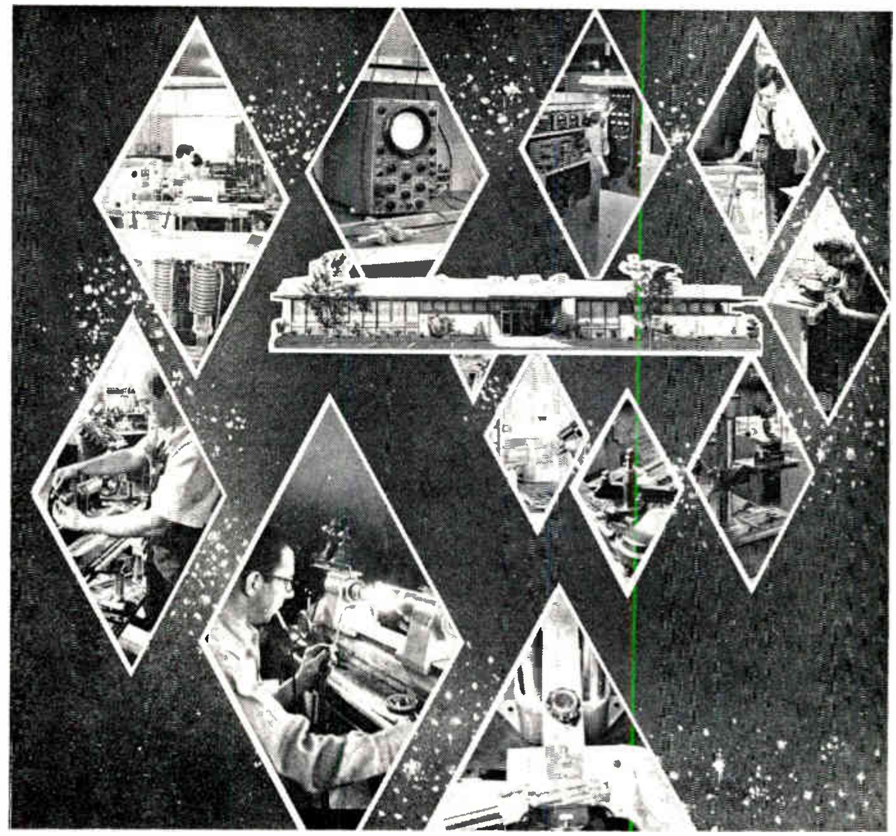
For additional information write Data Processing Section, Sales Department, of the firm.

Ionization Gage

A new, low-cost, cold cathode ionization gage (referred to as a Philips or Pennings gage), designed for industrial applications and capable of vacuum measurements from 10^{-2} to 10^{-7} mm Hg, has been announced by **The Fredericks Co.,** Bethayres, Pa. Designated the Televac Model 7A Cold Cathode Ionization Gage, this instrument is of particular value in such applications as vacuum metallurgy, electron beam welders, vacuum welding, vac-



(Continued on page 90A)



NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 88A)

uum brazing, vacuum evaporating, vacuum sintering, vacuum annealing, vacuum heat treating, particle accelerators, vacuum jacketing for liquid oxygen and hydrogen, and missile launch instrumentation.

Sensing tubes are metal, and include a removable anode assembly for ease of cleaning and inspection of inside of sensing tube. Tubulation is of one-inch ID to permit rapid response to pressure changes. The Televac Model 7A permits operation in either single, dual or three range scales from 10^{-2} to 10^{-7} mm Hg.

By adding a thermocouple vacuum gage as an integral part of the Model 7A, complete and independent coverage of the range from 1 to 10^{-3} mm Hg can be provided.

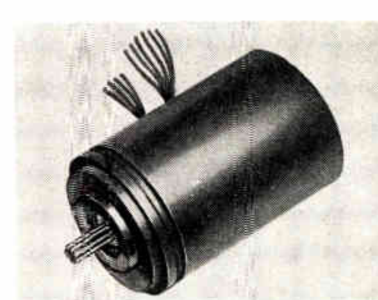
The Model 7A has no measurable drift or meter oscillation and can be coupled with an automatic controller for indication, remote alarm or external control. Terminals can also be provided for driving a 10 millivolt, full scale, potentiometer-type recorder.

Several explosion-proof versions for use in liquid oxygen and hydrogen vacuum dewar-type tanks are available.

The instrument is produced in cabinet panel, chassis and rack mounted configurations. Dimensions of the cabinet mounted unit are: Height, 7⅜″; width, 14¼″; and depth, 10″. The rack mounted unit measures 7″ high (other heights available upon request), 19″ wide and 11¾″ deep. For details write to the firm.

Motor Generator

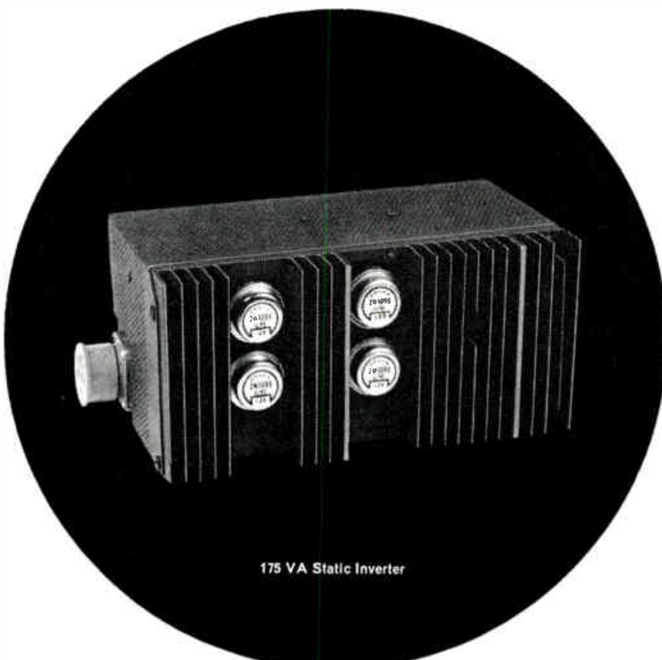
Thirty-five per cent more space and a reduction of 23 per cent in weight, is made available through the use of the motor generator designed and developed by **The Bendix Corp., Eclipse-Pioneer Div.,** Teterboro, N. J. Designated as the FV-6000-1-A1, the new unit is 1.375 inches long in a size 10 frame, as compared to the conventional length of 2.1 inches. The unit is comparable in operating characteristics to the Eclipse-Pioneer FV-100 series motor generators, and is being made available in the same variety of shaft configurations.

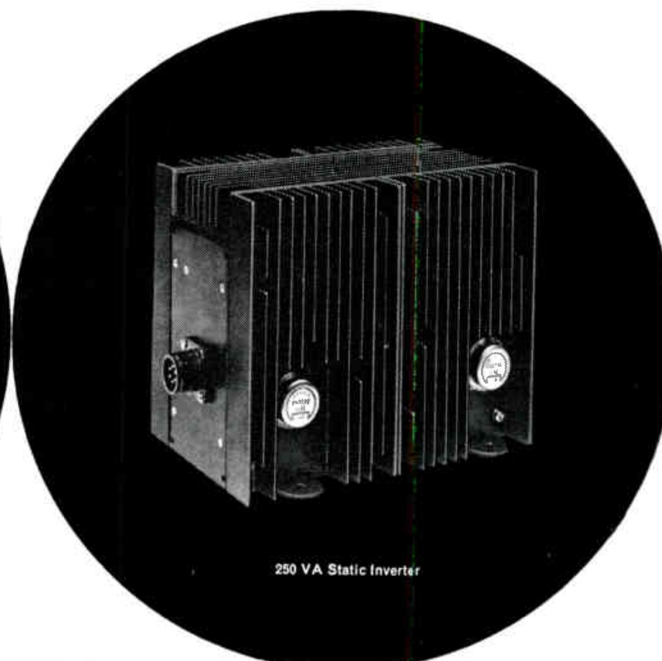


(Continued on page 92A)

PRECISION WITH SIMPLICITY
FROM DELCO RADIO

That's the big feature in Delco Radio's new 175 VA and 250 VA static inverter power supplies. These *all-transistor* units offer increased reliability through simplified circuits. Both static inverters are designed for either airborne or ground applications and will withstand overload and output short circuit conditions indefinitely, delivering at least 110% of rated output before going into overload protection. Units automatically recover to full output upon removal of overload and short circuit. Units are designed to meet the environmental requirements of MIL-E-5272C. For further information on military electronics write Delco Radio's Military Sales Department.


175 VA Static Inverter


250 VA Static Inverter

ELECTRICAL SPECIFICATIONS

175 VA STATIC INVERTER

Input

Voltage: 27.5 VDC ± 10% per MIL-STD-704

Output

Power:	175 VA single phase 0.5 lag to 1.0 power factor
Voltage:	115 V adjustable from 110 to 120 volts
Regulation:	1-volt change for any variation of load between zero and 110% of full load, and input voltage between 25 VDC and 30 VDC
Frequency:	400 ± 1 cps. Frequency changes less than 1.0 cps. for all environment, load and input voltage variation
Distortion:	Less than 5% total harmonic
Efficiency:	80% at full load

250 VA STATIC INVERTER

Input

Voltage: 27.5 VDC ± 10% per MIL-STD-704

Output

Power:	250 VA single phase 0.6 lag to 1.0 power factor
Voltage:	115 V adjustable from 110 to 120 volts
Regulation:	0.7 volt for any variation of load between zero and 110% of full load, and input voltage between 25 VDC and 30 VDC
Frequency:	400 ± .5 cps. Frequency changes less than 1.0 cps. for all environment, load and input voltage variation
Distortion:	Less than 5% total harmonic
Efficiency:	80% at full load

DELCO RADIO

Division of General Motors • Kokomo, Indiana

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

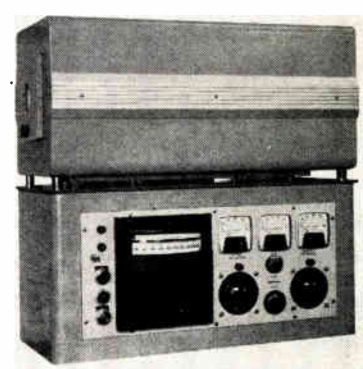
(Continued from page 90A)

In existing installations, the FV-6000-1-A1 can replace conventional units with comparable operating characteristics, resulting in additional space for expansion. In designing new installations, the unit will provide for an overall reduction of the space envelope.

The unit consists of a two-phase, four-pole, 400-cps induction motor and a single-phase generator, the rotors of both elements being mounted on a common shaft in a single housing. One motor phase is center-tapped, while the other has a split winding for either parallel or series operation. The generator produces a single-phase output, proportional to speed, of 0.42 volt at 1000 rpm, with a linearity of ½ of 1% up to 4000 rpm. Stainless steel bearings and laminations assure resistance to corrosion. The component weighs 2.3 ounces, and has an operating temperature range of −55 to +125°C.

Bench Size Diffusion Furnace

A bench size precision diffusion furnace has been introduced by **Sutherland Macklem, Inc.** The 16″×16″×36″ furnace is mounted on a power and control console only 15 inches high. In one to four 2¼″ alumina muffles it provides thermal flats within ±0.8°C over 10 to 14 inches long, at any temperature between 500°C and 1300°C. At maximum operating temperature, the case temperature remains under 95°C.



The wound heating elements are operated in three sections, with single-point, three-mode, millivoltmeter temperature control. Eight-point profile adjustment and continuous differential end-section controls set the thermal flat uniformity over a wide range of operating conditions. Two, three, and four tube models have separate heating elements, with individual adjustment for each.

The control system is based on silicon-controlled-rectifier ac power modulators, and fail-safe magnetic Transphasers® with error-rate damping control and reset vernier. Digital set-point, temperature re-

(Continued on page 94A)

MODEL	6db BANDWIDTH
TL-2D5A	2 kc
TL-4D9A	4 kc
TL-6D12A	6 kc
TL-8D16A	8 kc
TL-10D18A	10 kc
TL-16D25A	16 kc
TL-20D32A	20 kc
TL-30D45A	30 kc
TL-40D55A	40 kc
TL-45D60A	45 kc
TL-32E48C	32 kc
TL-50D85C	50 kc

(suffix "A" denotes 455 kc center frequency; suffix "C" denotes 500 kc center frequency)

Clevite Ceramic Ladder Filters

Now in stock in 12 bandwidths... 80 db rejection in 0.1 cu. in.

Clevite ceramic ladder filters provide more selectivity for their size than any conventional i-f filter. They are fixed tuned and need no alignment—are non-magnetic and non-microphonic. Leading manufacturers now have Clevite ladder filters in their communications equipment. Improve your newest design with these unique filters. Write now for complete specifications—Bulletin 94012, or for selectivity curves available on each stock model. ■ Dimensions: $\frac{5}{16}$" diameter x 1½" long. ■ Selectivity: 60 db/6db shape factor from 1.3:1 to 2.6:1. ■ Center Frequency Stability: within 0.2% for 5 years, and within 0.2% from −40° to +85° C. ■ Impedance: 1200-1500 ohms. ■ Designed for military environment.

CLEVITE ELECTRONIC COMPONENTS
DIVISION OF CLEVITE 232 FORBES ROAD, BEDFORD, OHIO

Field Sales Offices:
Maplewood, N.J. / Chicago, Ill. / Denver, Colo. / Inglewood, Calif.

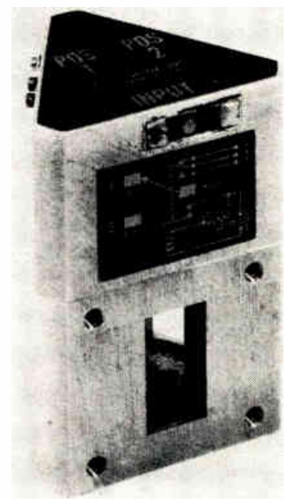
These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 92A)

cording, multi-point control, and other features, are available as modifications to the basic unit. The basic (single tube) unit, complete, is priced at $3500 (packing and shipping extra). Delivery 2 to 4 weeks, or from stock as available.

Microwave Switch

Production of a new group of microwave switches, which are now standard items, is announced by **Waveguide, Incorporated,** 851 W. 18th St., Costa Mesa, Calif.



These high speed switches are designed to meet the environmental conditions of MIL-E-5272. For reliability micro switches are used for both position indicating and solenoid control circuit. Precision instrument bearings are used at all friction points to insure maintenance of operating characteristics. They may be supplied for return to either outlet port upon failure of external circuit. They are of light-weight aluminum construction with maximum body width of 1.865" and are suitable for airborne, missile, space and shipboard application.

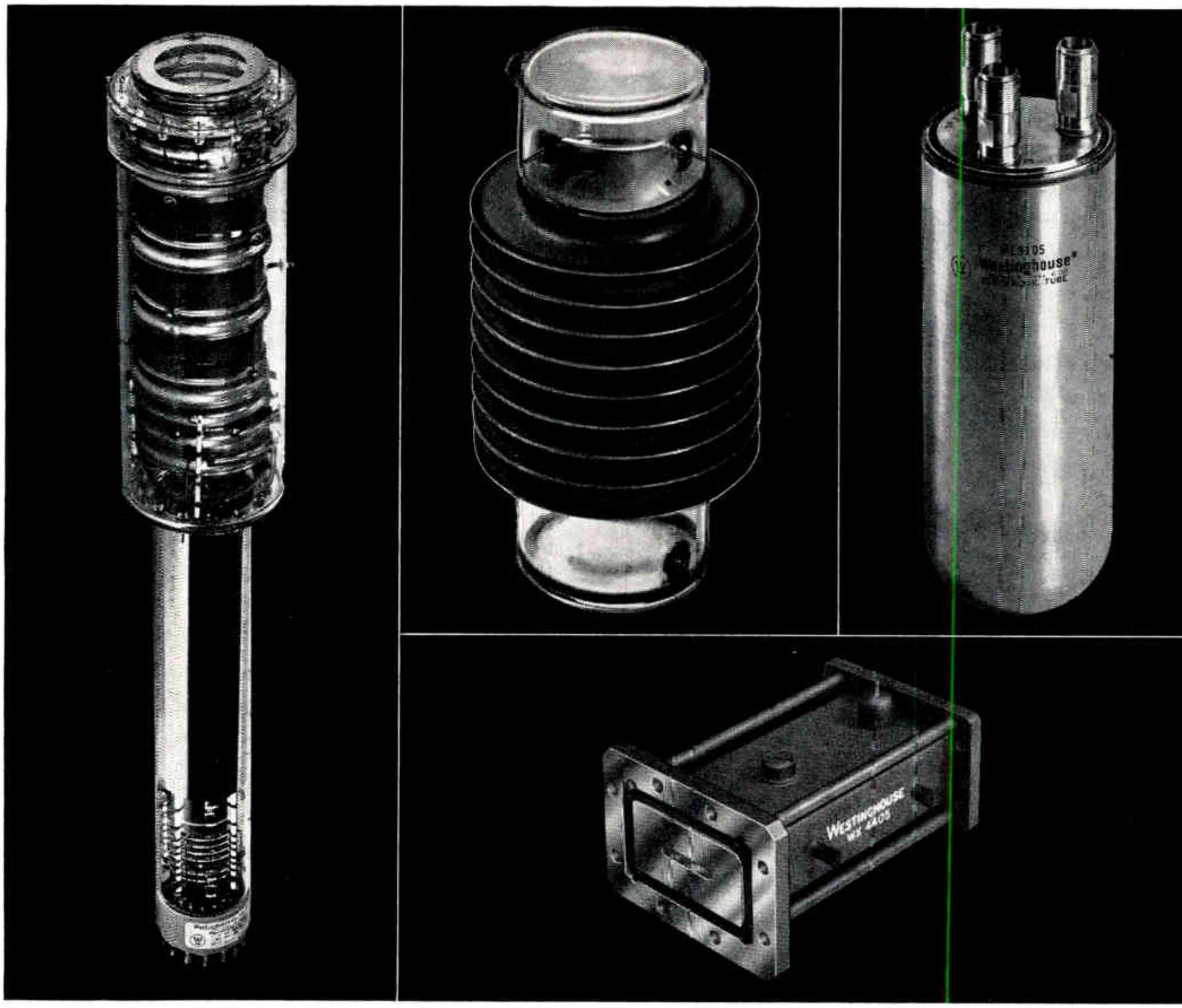
Additional specifications are as follows: VSWR, maximum 1.08:1; Insertion Loss, 0.2 db maximum; Isolation, 50 db minimum; Maximum Pull-in current, 900 M. Amperes ma at 28v dc; Maximum Hold-in current, 200 M. Amperes ma at 28v dc; Frequency Range of Line, 7.05—18.0 kmc; RF connectors: Mate with Standard Waveguide flanges; Life, Design-tested to greater than 200,000 cycles.

For further information write to the firm.

RF Power Amplifier

Vector Manufacturing Co., Inc., Southhampton, Pa. is marketing for use in aircraft, missiles and spacecraft, a pressurized

(Continued on page 96A)



Westinghouse electronic tube leadership

Westinghouse leadership in electronic tubes dates back to the first dry battery WD-11 tube made in 1920. In the great progress of tube advances that followed, Westinghouse engineers have contributed hundreds of improved tube types, and many entirely new ones. Today, Westinghouse continues to stay ahead of the field by creating new electronic tube capabilities that give design engineers dramatic increases in systems and equipment performance. The four tubes illustrated are only four examples of recent Westinghouse tube developments:

Microwave switching tube WX-4405 is a C-band, high-power, fast recovery time crystal protector featuring 0.1 to 0.2 microseconds recovery time.

Image intensifier orthicon for extremely low-light level viewing gives threshold pictures at 10^{-9} ft. candles illumination on the photocathode. Provides highest known sensitivity with good resolution.

Astracon—new 150,000X photon image amplifier tube—displays sufficient electron gain to display single photoelectron phenomenon.

High performance **neutron detectors** exemplified by the new ECONOLINE WL-8105 ionization chamber. In neutron detector production and development Westinghouse today has world-wide leadership—more Westinghouse neutron detectors are in use today than all other makes combined.

These tube developments are typical of Westinghouse leadership in electronic tubes for: atomic energy; commercial, industrial and military communications; radar; imaging systems; information storage and display; industrial processing; military aircraft and aerospace systems.

For the future, electronic tubes by Westinghouse have a planned place in systems developments. As micro-electronics, cryogenics, light frequencies, space explorations, new environments and technologies become common accomplishments, Westinghouse will provide electronic tube components to perform system functions that will advance the "state of the art." Whatever function your system has to perform, work with Westinghouse . . . for special system needs. Westinghouse Electric Corp., Electronic Tube Division, Elmira, N. Y. *You can be sure . . . if it's*

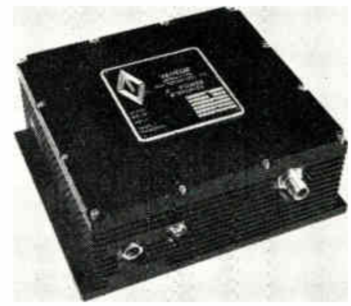
Westinghouse

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 94A)

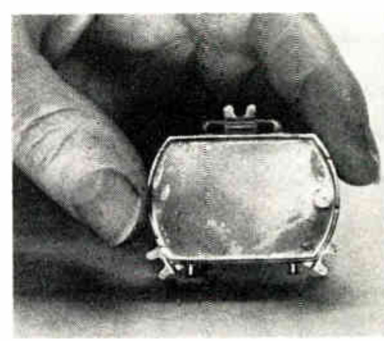
VHF telemetry power amplifier with an output of 50 watts. Using a unique innovation in circuitry, the anode is mechanically and electrically connected to the housing, thereby providing a superior thermal connection for vacuum tube cooling.



This 6 lb. amplifier of 115 cubic inches is designed for application in FM telemetry systems in the 215 to 260 mc range. The unit features compactness and excellent overall efficiency (typically 35%) with a power gain of 10 db, and operates from a standard 28 volt dc source.

RFI Filter

Efficient and economical filters to keep spurious incident RF energy from triggering sensitive U. S. Air Force cameras are now being produced by **Astron Corp.**, 255 Grant Ave., East Newark, N. J. The filters, designated AF-1637, are used to keep the line which controls the indexing of the film in airborne cameras free of accidental or intentional RF interference.



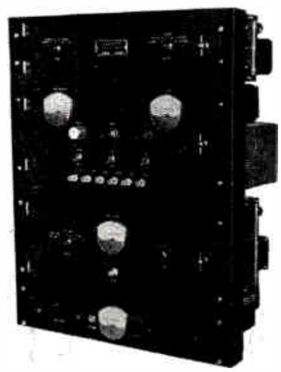
The AF-1637 filter operates conventionally, by presenting a very high impedance across the circuit but a low impedance to ground, thus trapping out signals at other than the operating frequency. The unit is a stabilized multiple-pi filter comprising four capacitors, two of which are single and two double, making six capacitive constants in all. It also contains two center-tapped inductors. The entire circuit was designed and fabricated to fit inside a housing less than $1\frac{3}{4}$ inches long, $1\frac{1}{4}$ inches wide and $\frac{1}{2}$ inch high including its standoff lugs.

NEWS New Products

The AF-1637 filter meets the environmental requirements of Department of Defense specification MIL-F-1573D. Electrical specifications are 50 ma and 28v dc.

Universal Transmitter

The 637 transmitter, a fixed frequency, crystal controlled unit, was designed by **Erco Radio Laboratories, Inc.**, 637 Stewart Ave., Garden City, N. Y., as a universal application transmitter. It can be used for any air-ground service in the 225–400 mc communications band.



Modular construction features permit plug-in multiplier strip and I.P.A. strip to form a complete 10 watt RF unit in an $8\frac{3}{4} \times 19''$ rack panel construction for low power applications. It requires the addition of a modulator panel unit which is completely solid state and requires $3\frac{1}{2}''$ of panel space. This unit features individual drawer-type modules with a low distortion compressor and low power drain.

The power amplifier unit, when added to the above combination together with its power supply, completes the 50 watt AM package. This unit can then operate in Class AB_1 mode as above

(Continued on page 98A)

(Continued from page 97A)

for AM, or can be switched to Class C condition for use as an amplifier for FM applications. It may also be used with a high level modulator for Class C AM operation at increased power specifications: Carrier Power is 50 Watts A3 in AB_1 mode. Frequency Range is 225 to 400 mc. Frequency Control is with an oven regulated CR32/U crystal for 0.002% stability. Output Impedance is 50 ohms unbalanced. Input Impedance is 600 ohms balanced. Frequency Response is 300 to 400 cps ±3 DB. Distortion is less than 6% at 1000 cps for 95% modulation. Carrier Noise is at least 40 db below full modulation. Control may be local or remote. On remote, up to 3000 ohms of loop resistance may be tolerated.

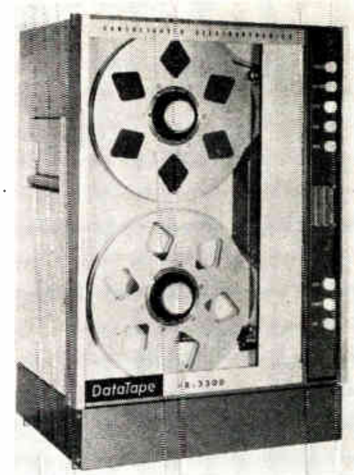
Power Source is 115 volts 50/60 cps single phase or 220 volts 50/60 cps can be provided on special order.

Panel Space Required is standard 19" relay rack; 28 inches is required for the 50 watt A3 model.

Technical Bulletin 637-T is available on request.

Portable Solid-State Recorder

A low cost precision portable, solid-state magnetic tape recorder/reproduce system has been introduced by the **Data Recorders Div., Consolidated Electrodynamics Corp.,** a subsidiary of Bell & Howell, 360 Sierra Madre Villa, Pasadena, California.



The Type PR-3300 unit can record and/or reproduce up to 14 channels of direct, FM or PDM data. Frequency response is 100 cps to 100 kc direct and 0 to 10 kc FM. Maximum cumulative flutter is 0.3% peak-to-peak up to 300 cps and 0.55% peak-to-peak up to 10 kc at 60 ips.

The aluminum carrying case is 22⅝ inches high by 18⅝ inches wide by 15⅛ inches deep.

Six standard speeds are available in pairs of 60, 30; 15, 7½; and 3¾, 1⅞ ips. Speed pairs are selected electrically by HIGH-LOW pushbuttons on the front panel. All control functions are relay op-

(Continued on page 99A)

(Continued from page 98A)

erated to permit remote operation.

The PR-3300 accepts either 1-mil or 1.5-mil mylar base tape or 1.5-mil acetate base tape. Reels up to 10½ inches in diameter with EIA-type hubs can be accommodated. CEC precision reels are supplied as standard equipment.

Operating power requirements are 115 volts ±10 v, 60 cps single phase. Power consumption is approximately 500 watts for a complete 14-channel system; 450 watts for a 7-channel system. Details are given in document 106-60.

Transistorized Pulse Generator

Pulse generator Model 521 is the latest addition to the line of digital equipment produced by **Digital Electronics Corp.**, 161 Sullivan Lane, Westbury, L. I., N. Y.

The generator is a low-cost, versatile, general-purpose fully transistorized unit designed to meet the requirements for research, development, and testing of low- and high-frequency pulse and digital equipment. A circuit in the Model 521 makes it possible to vary the frequency and width in a single multivibrator circuit as opposed to two in other types.

Supplying a pulse repetition rate from cps to 500 kc in five bands, the generator has a pulse width from 0.8 to 120,000 microseconds, a rise time of 50 nanoseconds, and a fall time of 200 nanoseconds. The amplitude is either positive or negative, 15 volts maximum, 5 volts into a 100-ohm load. For other pulse generators with dual polarities, a separate amplifier is needed for each polarity, with the Model 521, only one is needed. To produce the two polarities, a different voltage is applied to the amplifier with a switch on the panel.

Short-circuit protected, this unit operates with a nominal duty factor of 80%. With full amplitude at all loads and at all frequencies, the minimum duty factor is 60%. A synchronizing output pulse (−5 volts, 1 microsecond wide) is available at terminals on the front panel. To ensure synchronization with associated equipment, this signal is constant and is independent of the frequency and amplitude variations of the pulse output.

Weight is 2 pounds, dimensions are 8.25″ wide by 5.0″ high by 5.75″ deep. Power required is 115 volts, 60 cps, 5 watts.

Price is $89.50 f.o.b. Westbury, L. I., N. Y.

Availability is 30 days. For details contact E. Wiseman, at the firm.

(Continued on page 100A)

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 99A)

Miniature DC Amplifier

The new Model 611 Plug-In dc Amplifier developed by **Wilson Greatbatch, Inc.**, 10647 Main St., Dept. IR, Clarence, N. Y., features small size, $\frac{3}{8}'' \times \frac{3}{4}'' \times \frac{7}{8}''$ high; light weight, $\frac{1}{4}$ oz.; better than 0.02%/°F zero stability, 150 K input impedance, better than 0.3% linearity. Any dual power supply from ± 6 V to ± 28 V dc may be used. This amplifier is suited for bridge inputs. Over 500 amplifiers will fit on a square foot of chassis space.

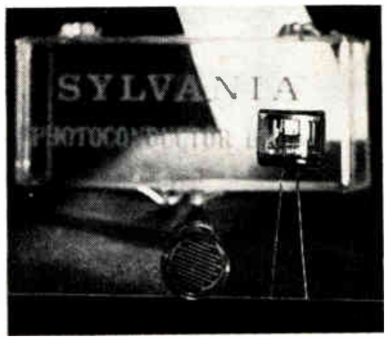
Gain may be set between 5 and 100 by means of an external resistor. Flat from dc to over 6 kc.

The Model 611 is suitable as a gain element for high-density physiological telemetering systems. One amplifier suffices for body temperature, GSR, respiration or pO_2 measurements. Two amplifiers cascaded suffice for EKG, EMG or Ocularmotor measurements. Three amplifiers cascaded suffice for EEG brain wave measurements.

Available from stock at $225 each.

Cadmium Sulfide Photoconductors

Two new cadmium sulfide photoconductor cells were displayed at the 1962 IRE Show by **Sylvania Electric Products Inc.**, a subsidiary of **General Telephone & Electronics Corp.**, 730 Third Ave., New York.



(Continued on page 102A)

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 100A)

The new units, Types 8142 and 8143, have an outside bulb diameter of 0.5 inch and are end aluminated. They feature high sensitivity and hermetically sealed-in-glass construction. The cells are gaseous back filled for a high dissipation safety factor and include a special confidence feature, a blue dot compound which turns pink if the cell envelope becomes damaged.

Ultra-Fast Rise Time Delay Line

Ad-Yu Electronics Lab., Inc., 249 Terhune Ave., Passaic, N. J., has introduced a new type of delay line with bandwidth over 30 mc. The ratio of rise time can be made less than 2%, yet the maximum attenuation is less than 3 db. Chiefly, this is due to utilization of a new principle of network for linear phase shift, and adoption of a new method to achieve correct amounts of second and third order of mutual coupling among sections. Two or more units of this delay line with identical impedance, may be connected in tandem. In this case, the total time delay will be the sum of each individual unit, and the rise time will equal approximately the square root of the sum of the squares of the rise time of each unit.

There are a number of features for this series: Phase linear beyond 30 mc; rise time less than 2% of the total delay; distortion less than 2% of the signal amplitude; light weight, small physical size, and rugged construction.

The specifications are: Attenuation per microsecond is less than 0.5 db minimum; 1 db maximum. Maximum input voltage, 500 volts for most types. The cutoff frequency is 0.35 **n/Td** and the 3 db bandwidth is at 0.175 **n/Td**, where **Td** is total delay and **n** is the number of sections. Temperature coefficient is 0.005%/°C. Accuracy of delay is ±2%; and accuracy of characteristic impedance is ±10% (better accuracy can be supplied on request). Delay can be made from 0.2 μs up to 18 μs. Impedance available in the following values: 50, 75, 93, 150, 220, 280, 320, 400, 500 and 700 ohms. Price: $49.00 and up. Delivery: 2–3 weeks.

(Continued on page 104A)

Amperex®

announces a significant step toward *instantaneous* electron emission . . .

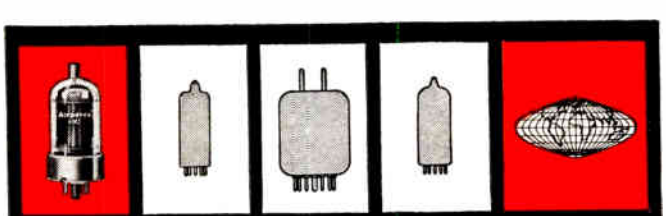
The

HARP CATHODE

Amperex sought the best means of producing the world's fastest heating electron tube; found it in the **Harp Cathode.** As developed by **Amperex, Harp Cathode** tubes permit warm-up in as little as 100 milliseconds. This major advantage, combined with low cathode inductance, low operating voltage, and extreme ruggedness—make Harp Cathode tubes the logical choice for transistorized mobile and airborne communications, and any other application where instantaneous response is a necessity.

Described simply, the Harp Cathode is a flat rectangle, strung harp-fashion with many superfine wires. Its unique advantages are largely a function of this physical arrangement. Thus, the minute wire size provides a high surface-to-volume ratio, resulting in instantaneous availability of thermal energy at the emissive surface. The quantity of electrically parallel, directly heated wires assures low cathode inductance. The low (1.6V) filament voltage affords the closest approach to the "unipotential" cathode. Moreover, in actual tests, tubes incorporating the Harp Cathode have given longer service life than tubes with conventional cathodes.

The **Amperex Harp Cathode Type 8042 Power Amplifier** (25 watts plate dissipation; 175mc ICAS) is available in pre-production quantities. It is the forerunner of **Amperex** Instant Heating Harp Cathode mobile transmitting tubes which will include versions of the types 6360, 6939, 6907 and 5894 push-pull tetrodes.



ask Amperex about the instant heating tube best suited to your particular application.

AMPEREX ELECTRONIC CORPORATION • 230 Duffy Avenue • Hicksville, Long Island, New York

IN CANADA: PHILIPS ELECTRON DEVICES LTD., 116 VANDERHOOF AVE., TORONTO 16, ONTARIO

NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(Continued from page 102A)

Counter-Timer Frequency Meter

Transistor Specialties, Inc., Engineers Hill, Plainview, L. I., N. Y., announces the availability of the Model 373 Counter/Timer, for the measurement of frequencies up to 10 megacycles, and time intervals (period measurements) from 1 microsecond to 10 seconds. By restricting the crystal "clock" circuitry to 1 mc, and employing only a single signal channel, very significant economies are effected, as are simplification of the panel controls and manipulations associated with each test setup.

The Model 373 is, according to the manufacturer, the most compact, lowest-priced all-solid-state 10 mc counter available. The level-sampling circuit, with 85 mv sensitivity, is used in the amplifier channel, and the readout is the superior 7-Nixie in-line type used on the most versatile and expensive counters.

The same high stability and accuracy (± 1 count, ± 3 parts in 10^7/week) are achieved in the Model 373, as in the most complex and advanced counters in its class. As in all TSI Counter/Timers, the instrument provides full self-testing, as well as frequency-standard output, consisting of every decade of frequency from 1 mc to 0.1 cps, derived from the crystal "clock."

Reliability is enhanced by the use of NOR-logic to reduce the semiconductor count, and generous derating guarantees a long, trouble-free life. Power consumption is only 35 watts.

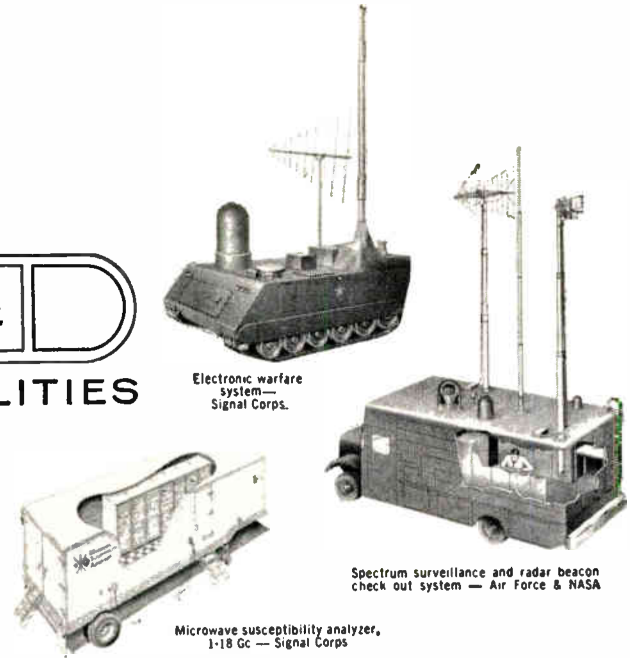
The 373 is illustrated in the attached photograph. The cabinet measures $8\frac{3}{4}'' \times 10\frac{1}{4}'' \times 12\frac{1}{2}''$. The unit weighs 19 lbs.

Additional information can be obtained from the manufacturer. Price is $1,880 (single units), delivery within 30 days.

Taut Band Measuring Instruments

A full new line of electrical measuring instruments featuring highly compact taut band suspension systems—interchangeable with conventional movements—has been

(Continued on page 106A)

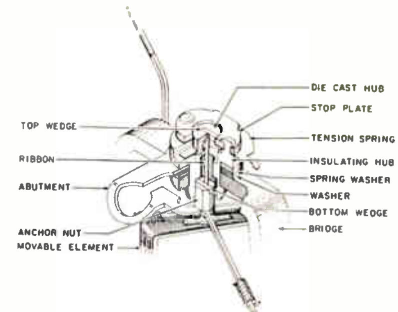


NEWS New Products

These manufacturers have invited PROCEEDINGS readers to write for literature and further technical information. Please mention your IRE affiliation.

(*Continued from page 104A*)

introduced by **Weston Instruments Div., Daystrom,** 614 Frelinghuysen Ave., Newark 4, N. J.



First instruments offered by Weston are 3.5, 4.5, 5.5, and 7.5-inch taut band models which, in external appearance, are exactly like the standard Weston rectangular 1900 series line in these sizes. Successful engineering design of taut band instrument movements, which are dimensionally interchangeable with conventional pivot-and-jewel movements, has made this possible.

In taut band instruments, the sensitive moving element is suspended by 2 slender metallic ribbons above and below the moving coil. When current is applied to the instrument, the rotation of the moving coil twists the suspension ribbons. When the current is removed, the ribbons return the movement to its original position.

The taut band method has the advantage of eliminating the friction encountered between pivots and jewels, and thereby increases repeatability. Sensitivity is also about five times greater than that of conventional instruments. The taut band suspension is extremely rugged and its stability is said to be superior to that of pivot-and-jewel instruments.

Professional Group Meetings

Aerospace and Navigational Electronics

Baltimore—January 16

"Exotic Methods in Space Communications," Dr. L. R. Bittman, Martin Marietta, Baltimore, Md.

Antennas and Propagation

Boston—February 27

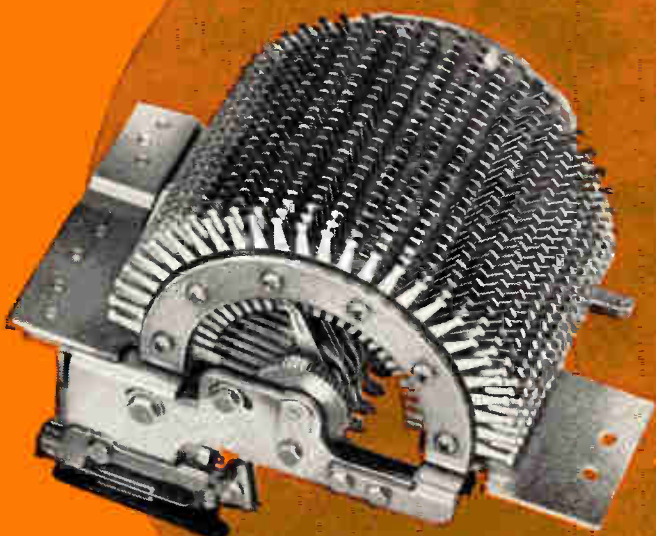
"Effects of Noise—Antennas—Transmission Losses," Dr. J. I. DeBettencourt, Consultant, Raytheon, Norwood, Mass.

(*Continued on page 108A*)

CLARE Stepping Switches

give designers—

- *more levels per switch*
- *more levels in less space*
- *more simplified circuitry*
- *NO synchronization problem*

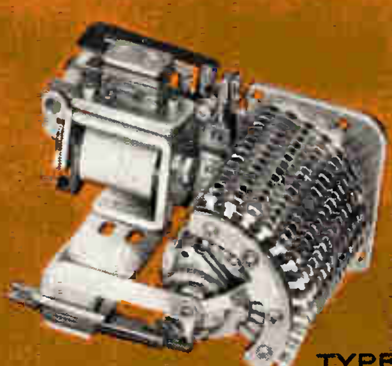


TYPE 26

Type for type, CLARE Stepping Switches provide more levels per switch . . . more levels per inch of height. The 12-level, 52-point, switch shown (CLARE Type 26), for instance, is 4$\frac{11}{16}$ in. high. It has four more 52-point levels than comparable 52-point switches, yet it is but $\frac{13}{16}$ in. higher than a comparable 8-level, 52-point switch. The smaller (Type 211), five-level, 33-point switch provides twice the levels of any comparable 33-point switch.

This greater working capacity per switch . . . and per inch . . . of CLARE Stepping Switches permits more simplified circuitry and avoids synchronization problems which arise when multiple stepping switches are necessary to do what is often a one-switch job with these high-capacity CLARE units.

CLARE Stepping Switches include a full line of spring-driven cam-operated or direct-drive switches with capacities from 10 to 52 points. CLARE engineering will cooperate with designers to develop special switches to meet unique requirements.



TYPE 211

For precise stepping switches famous for long life, high capacity and freedom from maintenance

IT HAS TO BE CLARE

Send for **DESIGN MANUAL 202A**

C. P. Clare & Co., 3101 W. Pratt Blvd., Chicago 45, Illinois. Cable Address: CLARELAY. In Canada: C. P. Clare Canada Ltd., 840 Caledonia Road, Toronto 19, Ontario. In Europe: Europelec, les Clayes-sous-Bois (S. et O.), France

C. P. CLARE & CO.

Relays and related control components

Professional Group Meetings

(*Continued from page 106A*)

"Granite Medium—Its Electrical Properties and Depth of Entry Required for Optimum Illumination," G. J. Harmon, Consultant, Raytheon, Norwood, Mass.

Chicago—February 9

"Antenna and Propagation Problems in Satellite Communications," Dr. R. F. Yang, Andrew Corp., Chicago, Ill.

Chicago—January 30

"Communications Antennas for Mercury Space Capsule," H. Walls, McDonnell Aircraft Corp., St. Louis, Mo.

Antennas and Propagation Microwave Theory and Techniques

Columbus—February 27

"VHF Ionospheric Scatter Communication," R. C. Kirby, Nat'l Bureau of Standards, Boulder, Colo.

Los Angeles—February 15

"Antennas and Microwave Research at Ohio State University," T. Tice, Motorola, Phoenix, Ariz.

Orange Belt-Santa Ana—December 7

"The Microwave Industry as seen by the Microwave Journal," W. Bazzy, Publisher, Microwave Journal.

"The Role of the Microwave Engineer in Today's Society," T. Saad, Editor in Chief, Microwave Journal.

Philadelphia—January 16

"Microwave Variable Attenuators & Modulators Using Pin Diodes," Kuhn, Hewlett-Packard, Palo Alto.

"Wide Band Tunnel Diodes," Koop, Philco Res. Div., Blue Bell, Pa.

San Diego—January 11

"Microwave Measurements Utilizing Swept Frequency Techniques," H. Poulter, Hewlett-Packard Co., Palo Alto, Calif.

Syracuse—January 18

"Propagation of Elastic Waves in Crystals at Microwave Frequencies," Dr. H. Hsu, General Electric Co.

Toronto—March 5

"New Developments in Collinear Arrays," W. V. Tilston, Sinclair Radio Labs., Ltd.

Audio

San Diego—January 26

"Recent Developments in Miniaturization and Transistorization of Professional Audio Components," H. Souther, Sonotec, Inc., Santa Ana.

(*Continued on page 111A*)


• coax and waveguide couplers • attenuators • mounts •
• bolometers • isolators • detectors • frequency meters •
• r. f. probes •
UHF to 90 Gc!
NARDA
can supply you with more than 600 PRECISION microwave instruments, components and hardware!
Write for your free catalog!
• klystron mounts •
• thermistors •
• tuneable shorts •
• ridged waveguides •
• ferrite circulators •
• waveguide adapters & accessories • stands • horns •
• hybrid junctions • instruments and modulators •
SEND FOR FREE CATALOG TODAY!
Address: Department PIRE-1-13
the narda microwave corporation
PLAINVIEW, L. I., NEW YORK • AREA CODE 516 GE 3-9000

Professional Group Meetings

(Continued from page 108A)

Automatic Control

Boston—February 20

"Attitude Control of Spacecraft," Hildebrandt, M.I.T. Instrumentation Lab.

"Attitude Control of Satellites," E. G. Ogletree, M.I.T. Instrumentation Lab.

Boston—January 17

"Some Guidance and Control Aspects of a Satellite Rendezvous Vehicle," E. Capen, C. King, and E. Wallner, Jr., RCA, Burlington, Mass.

Pittsburgh—February 12

"Phase Plane and Phase Space Techniques for Nonlinear Control," Dr. Z. H. Meiksin, University of Pittsburgh.

Pittsburgh—January 29

"Introduction to Nonlinear Control," Dr. J. E. Gibson, Purdue University.

Bio-Medical Electronics

Baltimore—February 13

"Some Similarities Between a Neural Network Model and Electrophysiological Exp.," Dr. B. G. Farley, M.I.T. Lincoln Lab., Lexington, Mass.

Chicago—February 9

"The Role of the Eng neer in the Hospital," Dr. D. A. Holaday, University of Chicago Medical School.

Memphis—March 1

"Hearing—How Hi the Fi?" D. F. Austin, M. D., J. J. Shea, Jr., M. D., and K. Kinchen.

Memphis—January 18

"Problems in Artificial Circulation," J. W. Pate, M.D., and W. Lee, M.D., University of Tennessee Medical Units, Memphis.

Broadcasting

Pittsburgh—February 5

"FM Stereo Broadcasting," R. M. Linz, General Electric Co., Utica, N. Y.

Circuit Theory

Albuquerque-Los Alamos—February 19

"Fundamentals of Non-Linear Systems, Part I," Dr. W. W. Koepsel, University of New Mexico.

Philadelphia—February 22

"Network Synthesis in the Time Domain," Dr. B. Liu, Bell Telephone Labs., Murray Hill, N. J.

(Continued on page 112A)

Professional Group Meetings

(Continued from page 111A)

Communications Systems

Northern New Jersey—February 20

"Advanced Industrial Communications in Venezuela," V. J. Nexon and G. Karger, Microwave Services International, Inc., Denville, N. J.

Component Parts

Baltimore—February 27

"The Glass Sealed Magnetic Switch (Reed Type)," L. Brown, Gordos Corp., Bloomfield, N. J.

Houston—January 23

"Magnetic Recording Media," J. Hanks, Minnesota Mining and Manufacturing Co.

Los Angeles—January 8

"Characterization of Radiation Sensitive Semiconductor Devices," J. Hodges, Texas Instruments.

Los Angeles—December 5

"Second Annual Symposium on Space Vehicles," Co-Sponsored by PGRQC, PGED and PGCP—all day meeting.

Los Angeles—November 29

"Integrated Circuits; Component Engineer's Responsibilities for the Future," D. Noble, Motorola.

Los Angeles—October 16

"Standards Laboratory Requirements, Accuracies and Precision Measurements," W. Hogan, Leeds-Northrop, E. Anderson, U. S. Navy, J. Orme, Aerojet.

Component Parts
Electron Devices

Omaha-Lincoln—February 15

"Testing Insulation," K. Loebl, Anaconda Wire and Cable, Chicago.

Electron Devices
Microwave Theory and Techniques

San Francisco—January 11

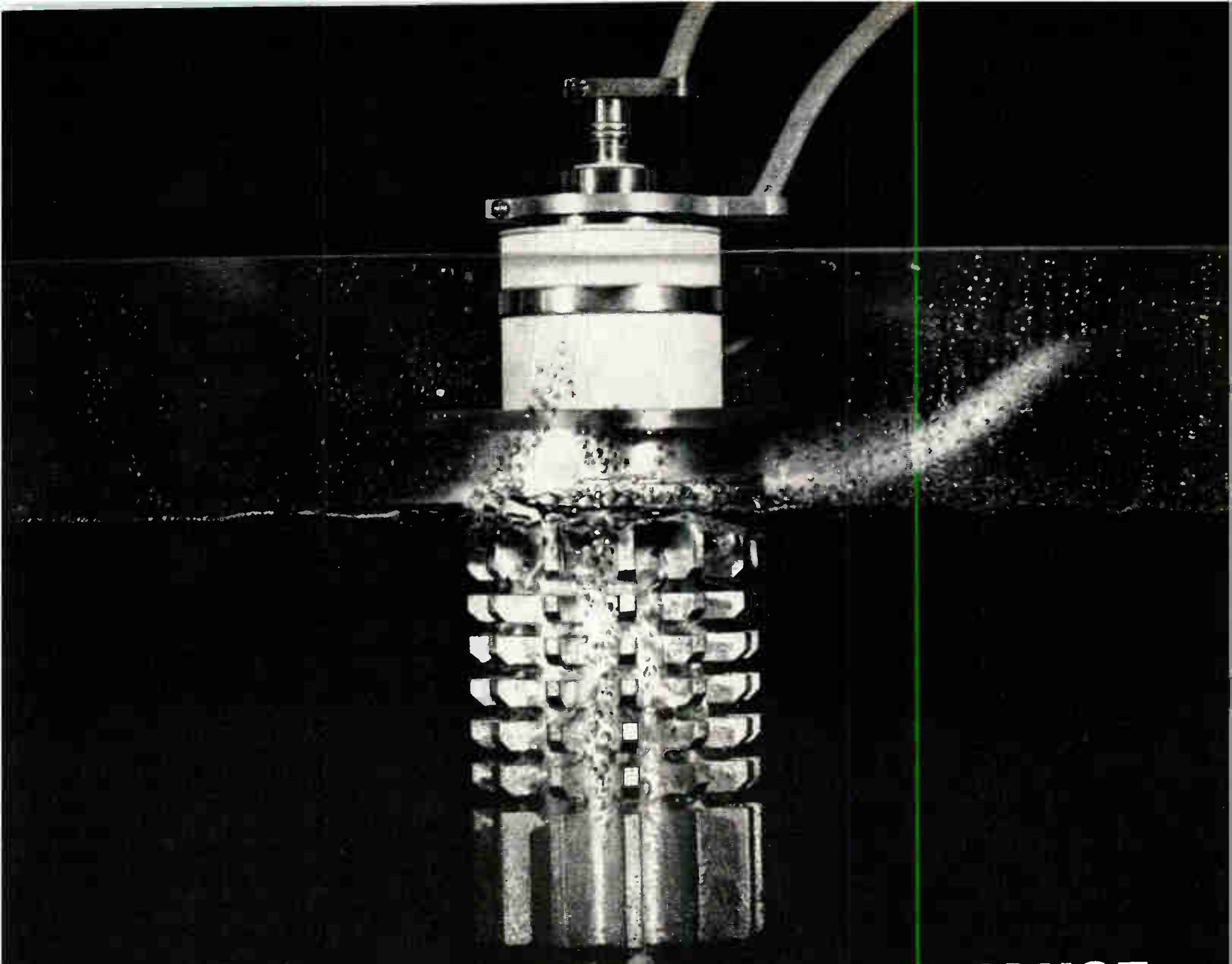
"Traveling Wave Amplification of Acoustic Waves in Semiconductors," A. R. Hudson, Bell Telephone Labs., Murray Hill, N. J.

Electronic Computers

Baltimore—February 21

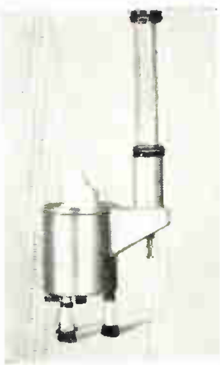
"New Concepts in the Logical Organization of Digital Computers," Dr. D. L. Slotnick, Westinghouse, Baltimore.

(Continued on page 115A)



BOILING FOR "COOL" PERFORMANCE

This tube, the new ITT F-7832 Power Triode is in an open glass tank to demonstrate the new evaporative cooling method. Cooling water is shown boiling at the segmented copper cooling fins. Bubbles are rising off the surfaces and steam is escaping. Normally the tube is enclosed in a boiler (photograph at right) which, in conjunction with an external condenser, becomes a complete system.



Conventional water cooling is capable of dissipating 450 watts/sq. in. Forced air cooling dissipates 150. The ITT Evaporative Cooling System will dissipate in excess of 800 watts/sq. in. It will operate in an overload condition at 1600 watts/sq. in. with no damage. In addition, ITT Evaporative Cooling offers the advantages of noiseless operation, absence of rotating parts such as blowers and pumps, minimum servicing, self cleaning of tube, and minimum liquid coolant. ITT Evaporative Cooled Tubes feature ceramic construction. The new rugged, mesh cathode, another design innovation, provides improved emission per watt, quick heating and excellent temperature stabilization.

ITT

Write for information on the new ITT Evaporative Cooled Power Tubes. Application assistance is available for your specific requirements.

ELECTRON TUBE DEPARTMENT ■ COMPONENTS DIVISION

INTERNATIONAL TELEPHONE AND TELEGRAPH CORPORATION, CLIFTON, NEW JERSEY

ITT COMPONENTS DIVISION PRODUCTS: POWER TRIODES AND DIODES · IATRON STORAGE TUBES · HYDROGEN THYRATRONS · TRAVELING WAVE TUBES · SELENIUM RECTIFIERS · SILICON DIODES AND RECTIFIERS · TANTALUM CAPACITORS

POWER CONVERSION

FOR AMERICA'S MIGHTY MISSILES

WITH PESCO STATIC INVERTERS

Filling the power conversion requirements of America's mighty missiles . . . used on such first-line weapons as Polaris, Mace, and the Centaur space vehicle . . . the Pesco Static Inverter shown above has no moving parts, offers unsurpassed efficiency in d.c. to a.c. inversion, ranging from 250 va to 20 kva. Employing silicon semiconductors and magnetic components with operating temperatures up to 125°C., lightweight Pesco Static Inverters feature single or multi-phase operation . . . are adaptable to any requirement for frequency stability or tolerance . . . permit 100% load unbalance! This "first" in electronic circuitry exemplifies the total capabilities of the Pesco/Borg-Warner research engineering team . . . from creative concept to production units.

another product of Pesco/Borg-Warner Research

For full details, send coupon below, or contact your nearest Pesco sales representative.

PESCO PRODUCTS DIVISION
BORG-WARNER CORPORATION
24700 North Miles Road
Bedford, Ohio

Export Sales: Borg-Warner International Corp.
36 South Wabash Avenue • Chicago 3, Illinois

10513-PC

PESCO PRODUCTS DIVISION • BORG-WARNER CORPORATION
Dept. I-5 Vanowen St., Burbank, Calif.

Please send me Pesco Static Inverter Bulletin 61100.

Name ______________________

Company ______________________

Street ______________________

City ______________ Zone ___ State ________

Professional Group Meetings

(Continued from page 112A)

Philadelphia—February 6

Panel Discussion: "Automated Design."

Panel: M. Wagner, Dept. of Defense; Mrs. K. Jacoby, Philco; J. Kirtzberg, Burroughs; J. B. Paterson, RCA; Dr. N. Prywes, University of Pittsburgh; N. P. Chinitz and T. D. Williams, Remington Rand Univac.

Electronic Computers Military Electronics

Omaha-Lincoln—March 5

Symposium: "Control of the Missouri River."

Engineering Management

Baltimore—February 15

"A Psychologist Looks at Engineering Management," Dr. R. McWilliams, William Lynde and Williams, Philadelphia.

Philadelphia—February 13

"Operations-Research—A Tool for Management," Dr. D. H. Wagner, Kettelle and Wagner, Paoli, Pa.

Syracuse—February 13

"A Look at Defense Contracting," R. L. Shetler, General Electric, Syracuse, N. Y.

Engineering Writing and Speech

Northern New Jersey—January 18

"Formula for Platform Poise," R. J. Norko, RCA, Harrison, N. J.

Information Theory

Baltimore—February 13

"Some Similarities Between the Behavior of a Neural Network Model and Electrophysiological Experiments," Dr. B. G. Farley, M.I.T. Lincoln Lab.

Microwave Theory and Techniques

New York—January 25

"Traveling Wave Tubes and Their Application," Dr. J. W. Fitzwilliam, Bell Telephone Labs., Murray Hill, N. J.

New York—December 7

"Looking Back at Microwave Theory and Techniques," Dr. G. C. Southworth, Bell Telephone Labs., Murray Hill, N. J.

(Continued on page 116A)

Professional Group Meetings

(Continued from page 115A)

Military Electronics

Ft. Huachuca—February 9

"Modern Concepts of Standards and Test Instruments," W. G. Stickler, Nat'l Bureau of Standards, Boulder, Colo.

Ft. Huachuca—February 15

"Informal Remarks on the Radio Frequency Compatibility Program," H. Randall, ODDRE, Washington, D. C.

Northwest Florida—February 8

"Microwave and Optical Masers in Range Instrumentation," Dr. H. C. Rothenberg, General Electric, Syracuse, N. Y.

Philadelphia—February 12

"Electronic Character Recognition Techniques," C. Teacher, Philco Corp., Bluebell, Pa.

Radio Frequency Interference

Philadelphia—February 20

"Recent Activities in RFI Control," A. R. Kall, Ark Electronics Corp., Pa.; R. Sugerman, American Electronics Lab., Philadelphia; F. Hamell, General Electric, Philadelphia.

Washington, D. C.—December 19

"Scoring Criteria for Determining RFI Damage to Communication—Electronics System," Dr. R. J. White, White Electromagnetics, Inc., Bethesda, Md.

Reliability and Quality Control

Philadelphia—February 6

"Mechanisms of Failure in Missile and Space Equipment," M. A. Wilson, Burroughs Corp., Paoli, Pa.

Washington, D. C.—February 15

"Confidence Limits for Reliability Estimates," Dr. J. R. Rosenblatt, Nat'l Bureau of Standards, Washington, D. C.

Vehicular Communications

New York-Northern New Jersey-Long Island—January 23

"Frequency Analysis of Mobile Communications Signals," E. Feldman, Panoramic Electronics, Mount Vernon, N. Y.

Washington, D. C.—February 26

"Impulse Noise Blankers—A Positive Approach to the Ignition Noise Problem in Mobile Radio," J. McCormick, General Electric, Lynchburg, Va.

IRE People

(*Continued from page 118A*)

The appointment of **Harold G. Caldine** (M'58) as Manager of the New England District of HRB-Singer, Inc., was announced recently. In his new position he will be engaged in applications engineering and development liaison. He will also be responsible for general customer activities and will perform market research.

As an applications engineer for HRB-Singer during the past three years, Mr. Caldine was responsible for the application of company products to Air Force and Army needs. As an engineer at the Rome Air Development Center for seven years prior to this, he was responsible for the development of electronic equipment by private contractors, and was project engineer for the development of such units as selective calling systems, teletype and voice multiplexing and cryptographic equipment, and digitalized closed circuit television systems. His previous experience includes several years in the broadcast industry.

Mr. Caldine holds a B.A. degree in physics from Syracuse University. He attended Capital Radio Engineering Institute, Washington, D. C., and the Armed Forces Radio School, Ft. Knox, Ky. He belongs to the Armed Forces Communications and Electronics Association, as well as various professional groups for licensed pilots, and radiotelephone operators.

❖

The appointment of **Charles C. Camillo** (A'50–SM'55) to the post of Vice-President, Engineering, of FXR, Danbury, Conn., has been announced.

C. C. Camillo

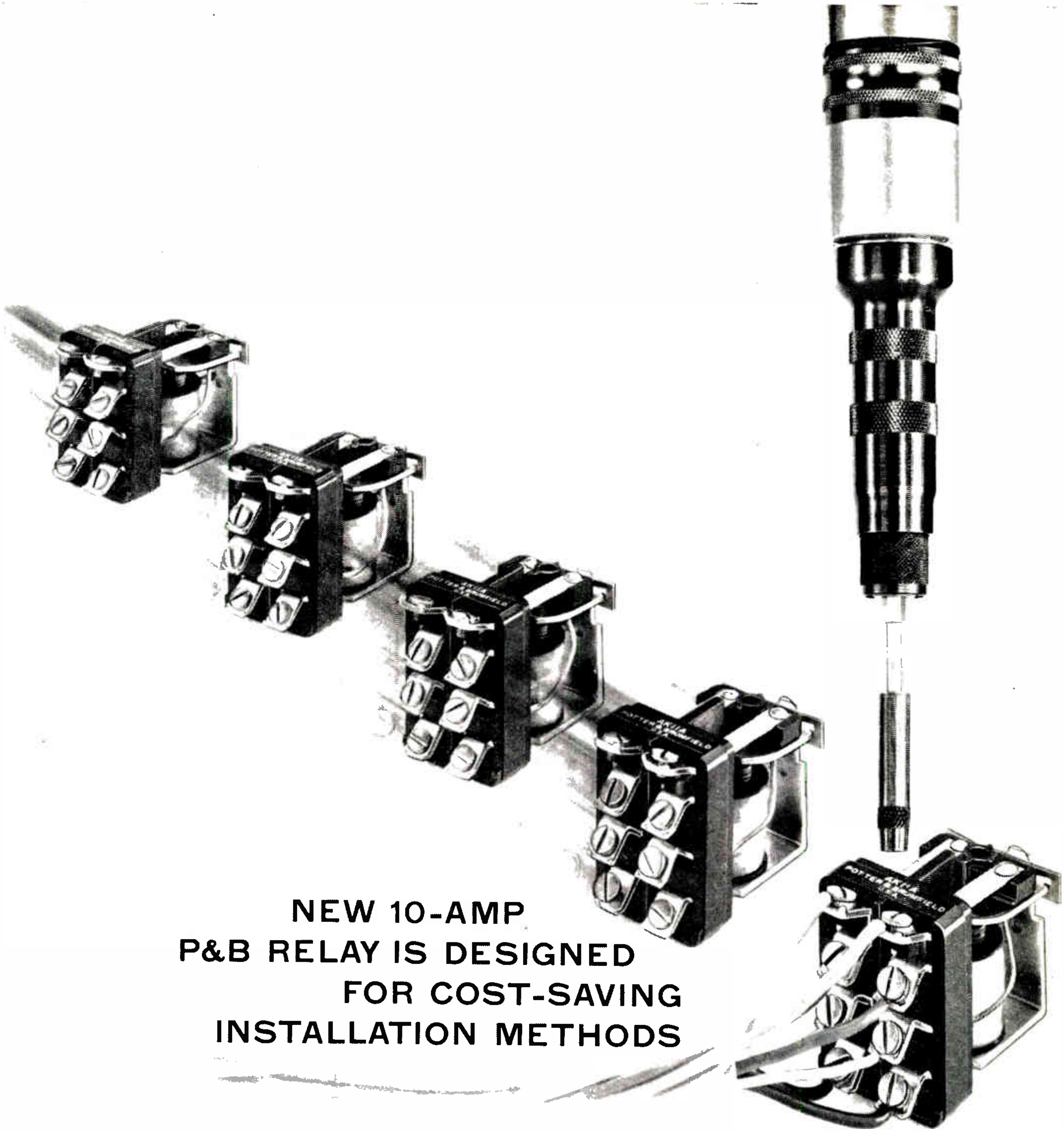
Mr. Camillo, who joined Amphenol-Borg in 1948, was a member of the American delegation to the IEC Convention earlier this year in Geneva, Switzerland. He was Vice-President, Engineering, of Amphenol-Borg's RF Products division.

Mr. Camillo has contributed many technical papers and articles on various aspects of RF transmission theory and technology.

He was educated at Illinois Institute of Technology and the University of Illinois, from which he holds a B.S. degree in physics. He is active in a number of technical and scientific organizations, including ARRL, IEC, and NEMA.

❖

(*Continued on page 122A*)



NEW 10-AMP P&B RELAY IS DESIGNED FOR COST-SAVING INSTALLATION METHODS

Fast, labor-saving power screwdrivers can be used to hook up lead wires to the extra stout screw terminals of our new AK Series relays. Thus you save time and money on your production lines.

This new relay is sturdy throughout: its expected mechanical life is 3 million operations and the ¼" diameter contacts are rated at 10 amps at 115V AC non-inductive. Available for either AC or DC power.

Get all the facts today from your P&B representatives or call us direct.

AB SERIES

SIMILAR RELAYS HAVE U/L & CSA LISTING
The AB and the ABC have ¼" quick-connect terminals which can be used for printed circuits. Screw terminal adapters available. U/L and CSA listed for 10 amps at 115V 50/60 cycles.

ABC SERIES

ENGINEERING DATA, AK SERIES

GENERAL:

Insulation Resistance: 100 megohms min.
Breakdown Voltage: 1500 V. rms between all elements.
Temperature Range: AC and DC, $-55°C$. to $+45°C$.
Weight: Approx. 5 ozs.
Pull-In:
AC 78% of nominal voltage.
DC 75% of nominal voltage.
Terminals: Heavy-duty screw type.

CONTACTS:

Arrangements: Up to DPDT.
Material: ¼" dia. silver.
Rating: 10 amps @ 115 volts or 5 amps @ 230 volts, AC non-induc.

COILS:

Resistance: 35,000 ohms max.
Power: 6.4 volt amps (AC), 2.0 watts (DC).
Duty: Continuous AC or DC (DC coils will withstand 6 watts at $+25°C$.

P&B STANDARD RELAYS ARE AVAILABLE AT YOUR LOCAL ELECTRONIC PARTS DISTRIBUTOR

POTTER & BRUMFIELD

DIVISION OF AMERICAN MACHINE & FOUNDRY COMPANY • PRINCETON, INDIANA
IN CANADA: POTTER & BRUMFIELD, DIVISION OF AMF CANADA LIMITED, GUELPH, ONTARIO

IRE People

(*Continued from page 120A*)

Dr. William M. Hall (A'23–M'44–SM'45) is the first member of Raytheon's staff to achieve the "Consulting Engineer, Consulting Scientist, or Scientific Fellow" level, a new professional level created by the Raytheon Co., Lexington, Mass., for its outstanding nonsupervisory engineers and scientists. Dr. Hall has been a member of Raytheon's technical staff for the past twenty-one years.

W. H. Hall

This new level permits the firm's specialists to contribute their best talent to advanced technologies while advancing their careers without need for accompanying administrative responsibilities.

Dr. Hall was awarded the Bachelor's, Master's and Doctorate degrees in science from Massachusetts Institute of Technology. He has written numerous papers and holds about twenty-five patents in the fields of sound measurements, infrared, instrument landing and radar.

❖

Reed E. Holaday (A'56) has been named senior scientist in Litton Industries' Electron Tube Research Laboratory, San Carlos, Calif.

Mr. Holaday has been associated with Litton Industries for the last three years, most recently as a senior engineer in charge of high power pulse crossed field devices. He was also klystron engineering manager during this period. He previously was employed at the Stanford Electronics Laboratory and the Sylvania Electronic Microwave Tube Laboratory.

In 1955 Mr. Holaday earned his Master's degree in electrical engineering at Purdue University. In 1959 he received an engineer's degree in electrical engineering at Stanford University.

❖

Hazeltine Electronics Division has announced the appointment of **John P. Hansen** (SM'58) as Manager of its Greenlawn, L. I., Laboratory. In this position he will have complete administrative and operational responsibility. He will also head the Greenlawn Project Administration Department.

Mr. Hansen, who is an assistant vice president of Hazeltine Electronics Division, joined Hazeltine in 1948 after serving as a product and sales engineer with the Sperry Gyroscope Co.

Mr. Hansen holds an engineering degree from the Polytechnic Institute of Brooklyn, Brooklyn, N. Y.

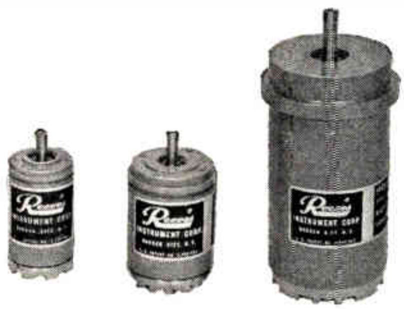
❖

(*Continued on page 124A*)

Reeves HIGH *Precision*

RESOLVERS, SYNCHROS and PHASE SHIFTERS

For data transmission; coordinate transformation and conversion; computer chain; and sweep applications. Write for data file 805.



STANDARD RESOLVERS

Reeves produces a full line of both compensated and uncompensated resolvers in standard BuOrd size 23, 15 and 11 cases. The size 23 series has a functional accuracy of 0.05%; resolvers in the other two series have an accuracy of 0.1%.

HIGH PRECISION SIZE 23 RESOLVERS

A new series of extremely accurate resolvers, which include a 0.01% functional accuracy computing resolver with 100% compensation; and a data transmission resolver with 20 second accuracy. They represent the ultimate in precision for resolvers of this case size.

30-SECOND SIZE 23 SYNCHROS

These three wire synchros are the most accurate units available in a standard BuOrd size 23 case. Both transmitters and control transformers can be supplied, designed for either 400-cycle or 60-cycle input.

PRECISION PANCAKE RESOLVERS

0.005% functional accuracy computer resolvers with 100% compensation; and 10-second accuracy data transmission resolvers. Integral bearing design permits direct mounting to gimbal structure. Beryllium housings provide highly stable operation under conditions of extreme temperature variation.

BOOSTER AMPLIFIERS

A complete line of vacuum tube and transistorized booster amplifiers, for use in conjunction with compensated resolvers. Transistorized units contain two fully encapsulated amplifiers in a single case.

REEVES INSTRUMENT CORPORATION

A Subsidiary of Dynamics Corporation of America
Roosevelt Field, Garden City, New York

IRE People

(*Continued from page 122A*)

Thomas R. Bristol (A'50–M'52–SM'55) has been named assistant to the manager, Display Devices Department, for Litton Industries' Electron Tube Division, San Carlos, Calif. His primary responsibilities will be in department operations and administration.

Mr. Bristol has been associated with Litton Industries for three years, most recently as assistant manager of the Linear Beam Department. He also has served as sales manager in applications engineering for the Linear Beam Department.

He is a 1947 electrical engineering graduate of Notre Dame University.

Mr. Bristol has a broad background in the power tube field, including design, development and applications engineering. He was employed by the General Electric Co., Schenectady, N. Y., from 1947 until joining Litton Tube Division.

❖

Lieutenant Commander Howard Freiberger, U. S. Naval Reserve, (M'59) has been appointed Commanding Officer of Naval Reserve Communications Division 3-1, a unit of the Selected Reserve specializing in the training of radiomen and electronics technicians for service in naval communications. Holding degrees in physics from the College of the City of New York and Columbia, Mr. Freiberger has followed paralleling military and civilian careers. His experience includes work in electronics at: the National Bureau of Standards; a naval communications facility in Japan; the Naval Material Laboratory in New York; and the Navy Mine Laboratory. He has also worked on the staff of a naval force commander. Mr. Freiberger is currently in the Research and Development Division, Prosthetic and Sensory Aids Service of the Veterans Administration. In this latter position he deals with electronic instrumentation for bioengineering research, and with electronic sensory aids for the blind and other handicapped.

H. Freiberger

Mr. Freiberger is a member of the IRE Professional Groups on Medical Electronics and Human Factors in Electronics, the U. S. Naval Institute, and the American Institute of the City of New York.

❖

James Cheal (M'57) is Vice-President and Director of Engineering of Omni Spectra, Inc., Detroit, Mich. He is one of five founders of the company.

Mr. Cheal headed microwave development programs at Bendix Corporation, Research Laboratories Division, before

(*Continued on page 126A*)

Now—Micro Magnetics for the Space Age!

"MICRO MAG MOD"®
MAGNETIC MODULATORS

For Absolute Reliability . . .

Infinite Standby and Service Life . . .

New "MICRO MAG MODS" are complete micro miniaturized magnetic modulators featuring an essentially drift-free circuit with superior phase and gain stability over extremely wide environmental ranges. All the ruggedness, dependability, wide dynamic range and stability that are characteristic of the larger "MAG MOD" Magnetic Modulators are engineered into this magnetic circuit.

For complete information, request Bulletin MM 101. Bulletins also available on "MAG MOD" Miniaturized Standard and Multiplying Magnetic Modulators, and GMO Series High-Stability Oscillators.

FUNCTIONS –

"MAG MODS" provide four quadrant operation, extreme stability, with negligible change of phase, gain and zero position over a wide temperature range. Design is simple, rugged with no vacuum tubes, semiconductors or moving parts to limit life. "MAG MODS" offer infinite design possibilities and impedance levels — are adaptable for algebraic addition, subtraction, multiplying, dividing, raising to a power and vector summing.

APPLICATIONS –

Specify "MAG MOD" for magnetic amplifier components of proven reliability. These dependable instruments are widely employed in satellites, automatic flight systems, fire control, analog computers, guided missiles, nuclear equipment, antennas, gun turrets, commercial power amplifiers and countless control systems.

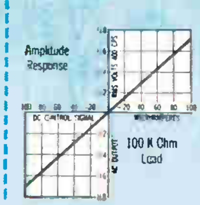
Products approximately two-thirds actual size

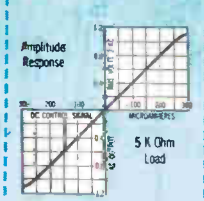
TYPE NUMBER	IMM-655-2	IMM-648-1	IMM-664-1	IMM-680-1
Ref. Carrier Voltage and Frequency	3 V @ 400 cps	2 V @ 2 KC	10 V @ 60 KC	115 V @ 400 cps
Input Control Signal Range	0 to ±100 μa DC	0 to ±300 μa DC	0 to ±100 μa DC	0 to ±10 μa DC
AM Phase Reversing AC Output Range	0 to 0.8 V RMS @ 400 cps	0 to 1.0 V RMS @ 2 KC	0 to 200 mv RMS @ 60 KC	0 to 30 mv RMS @ 400 cps
RMS mv AC Output/μa DC Signal Input	7 mv/μa	4 mv/μa	2 mv/μa	5 mv/μa
Output Impedance	14 K ohms	1000 ohms	11 K ohms	Approx. 150 ohms
External Load	100 K ohms	5 K ohms	50 K ohms	100 ohms
Zero Drift over Temp. Range	±0.1 μa Max.	0.5 μa Max.	—	0.05 μa Max.
Hysteresis in % of Max. Input DC Signal	0.2% Max.	0.2% Max.	0.5% Max.	0.1% Max.
% Harmonic in Output Product Wave	15%	10% to 15%	5%	20%
Temperature Range	−55°C to +125°C	−55°C to +125°C	−55°C to +125°C	−55°C to +125°C
Frequency Response	5 K Series, 108 cps	Over 200 cps	Over 5 KC	Over 100 cps
Approximate Weight (in Ounces)	0.2	0.1	0.2	0.3

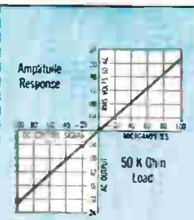
- Electrical zero point and gain, repeatability and stability over entire service life
- Extremely broad bandwidth
- Carrier frequencies as high as 1 megacycle
- Input signal current resolution better than 0.01μa
- Repeatable data over years of continuous, unattended operation
- High shock, vibration proof
- Low milliwatt power consumption
- May be mounted directly on printed circuit cards
- Light in weight

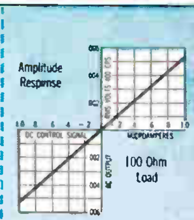
All "MAG MODS" conform to MIL-T-27A and MIL-E-5400

There is no Substitute for Reliability









GENERAL MAGNETICS INC

135 BLOOMFIELD AVENUE
BLOOMFIELD, NEW JERSEY

(*Continued from page 124A*)

joining Omni Spectra. He had been with Bendix from 1952 to 1955. From 1955 to 1956 he was a systems engineer for the Burroughs Corporation working on the development of the prototype Sage computer. After rejoining Bendix in 1956, he headed a group working on the development of a wide range of microwave components. He has five patent applications outstanding in the microwave field.

Mr. Cheal received his B.S. degree in electrical engineering from Michigan State University.

❖

The appointment of **Harvey Cohen** (S'55–M'55) as Eastern District Manager for the Systems Division of Beckman Instruments, Inc., Fullerton, Calif., was announced recently. In the new post, he will supervise the systems field engineering staff in the eastern half of the U. S.

Formerly a field engineer, Mr. Cohen has been a member of the Beckman organization since 1957.

He was graduated from the University of Manitoba, Winnipeg, Canada, with a BSc degree in electrical engineering, and received a Master's degree in business administration from Harvard University.

❖

The appointment of **Charles W. Hosterman** (M'56) as Manager of Planning of the Eastern Operation of Sylvania Electronic Systems, has been announced.

C. W. HOSTERMAN

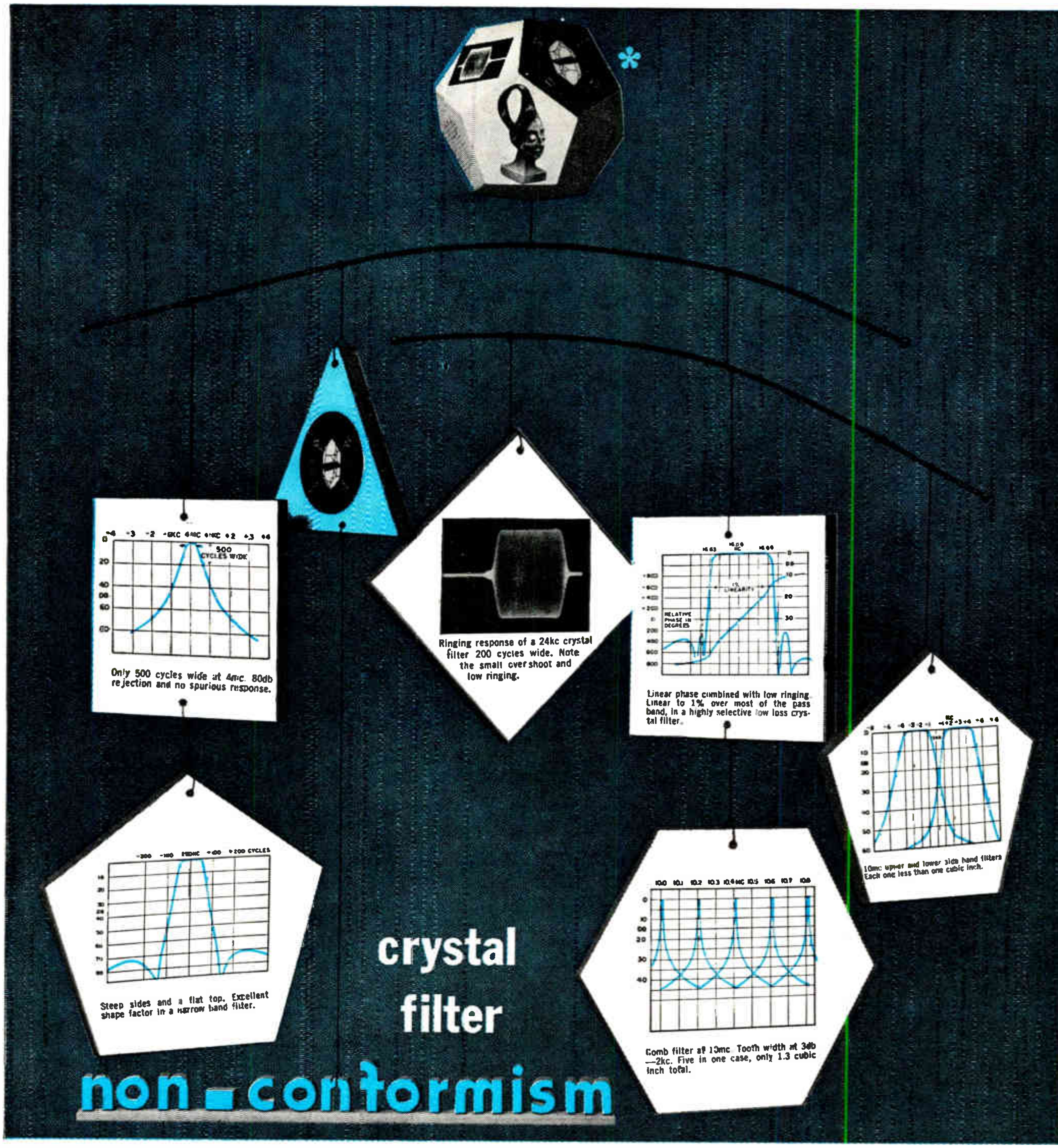
Mr. Hosterman has served as Manager of Planning of the Data Systems Operations since January, 1961. He joined Sylvania in 1943 as an assistant supervisor of personnel at the company's Altoona, Pa., electronic tube plant. In 1944, he was transferred to the Proximity Fuse Tube Operation, Huntington, West Va., as supervisor of personnel, and subsequently was named Manufacturing Superintendent there.

Between 1950 and 1958, Mr. Hosterman served successively as the plant manager of the Sylvania receiving tube plant, Shawnee, Okla., assistant general manager of the company's Electronics Division, Woburn, Mass., and general manager of the Semiconductor Division, Woburn, Mass.

In 1958, with the formation of the Data Systems Operations, Needham, Mass., Mr. Hosterman was named Business Manager of that facility and later served as acting Marketing Manager.

A graduate of Pennsylvania State University, he is a member of the American Management Association, Professional

(*Continued on page 128A*)



Those *non-conformists* at Burnell's engineering laboratory aren't satisfied with just producing the broadest range of crystal filters, toroids and communication networks: through their constant efforts to satisfy tomorrow's space age electronics problems, they have developed a whole new family of sophisticated crystal filters, with exceptional and unusual characteristics, contributing to increased circuit flexibility as graphically demonstrated above.

Those same *non-conformists* have also made considerable progress solving other electronics/space age problems. A typical example of this has been their work with the application of Time Domain Synthesis; producing an unlimited inventory of wave forms for new applications, and resulting in substantial reductions of size and weight, eliminating the need for complex active circuitry for its support.

*Join the *non-conformists*—write today for your free *Non-Conformist* paper weight and Crystal Filter Catalog XT-455. Yes! Your circuits *can* profit today from tomorrow's research.

Burnell & Co., Inc.

PIONEERS IN microminiaturization OF TOROIDS, FILTERS AND RELATED NETWORKS

EXECUTIVE OFFICE AND PLANT
DEPT. PI-1
PELHAM, NEW YORK
PELHAM 8-5000
TELETYPE PELHAM 3633

PACIFIC DIVISION
SOUTH PASADENA, CAL.
MFD. IN CANADA
BY EDO (CANADA) LTD.
CORNWALL, ONT.
WELLINGTON 2-6774

DIVISIONS: Gray & Kuhn, Inc., Pelham, New York • GLP Electronics, Inc., Bristol, Conn. • Guillemin Research Laboratory, Cambridge, Mass.

IRE People

(*Continued from page 126A*)

Group on Data Processing, National Management Association, Newton Chamber of Commerce and the American Numismatic Association.

❖

The American Electronic Laboratories has announced the appointment of **Joseph R. Jahoda** (S'50–A'51–M'54–SM'58) as Vice President in charge of the newly formed Washington Division, Fairfax, Va.

J. R. Jahoda

Mr. Jahoda received his B.E.E. degree from the College of the City of New York in 1950, and his M.E.E. degree from Polytechnic Institute of Brooklyn in 1954.

He was employed by the EDO Corp. from 1953 to 1954. Prior to this Mr. Jahoda was employed by Polytechnic Research and Development Test department.

He did work on advanced ECM system design and development for the Sperry Gyroscope Co. from 1954 to 1956. During this period he served as evening Staff Instructor at the Polytechnic Institute of Brooklyn.

From 1956 to 1961, Mr. Jahoda worked with Melpar, Inc. as a project engineer in research and development of advanced electronic countermeasure systems. He served as a consultant in Medical Electronics to the George Washington University, Washington Institute of Biophysical Research, Inc. and Med Electronics, Inc. from 1957 to 1961. In 1961 he served as director of International Research Organization, Inc.

Mr. Jahoda is a member of the IRE Professional Groups on Ultrasonic, Microwaves and Medical Electronics.

❖

Everett LeGette (M'60) has been promoted to Regional Sales Manager in the Western United States by Motorola Communications and Electronics, Inc.

E. LeGette

In this capacity, Mr. LeGette will direct the design and sale of FM two-way radio and other communications systems and equipment to industrial, governmental and business organizations throughout a six-state area, which includes Colorado, Western Kansas, Western Nebraska, New Mexico, Southeastern Wyoming and Northern Texas. His headquarters are located in Denver.

(*Continued on page 130A*)

For the 10TH year...

LAMBDA OFFERS THE STRONGEST PROOF OF TROUBLE-FREE POWER SUPPLY PERFORMANCE

Every Lambda power supply sold since 1953 has been backed by Lambda's 5-year guarantee, which covers workmanship and materials (except for tubes and fuses). Any Lambda power supply sold today is guaranteed to perform to specifications until 1967.

Send for Catalog on LAMBDA'S wide range line of regulated transistorized power supplies

LAMBDA ELECTRONICS CORP.
515 BROAD HOLLOW ROAD • HUNTINGTON, L. I., NEW YORK • 516 MYRTLE 4-4200

Western Regional Office: 230 North Lake Avenue, Pasadena, California • Phone: Code 213, MUrray 1-2544
New England Regional Office: 275 Boston Post Road, Marlboro, Massachusetts • Phone: Code 617, HUntley 5-7122
Middle Atlantic District Office: 515 Broad Hollow Road, Huntington, L. I., New York • Phone: Code 516, MYrtle 4-4200
Southeastern Region: W. A. Brown & Associates, Inc., Engineering Representatives

LA 128

IRE People

(Continued from page 128A)

He brings to his new position 18 years of experience in radio communications. Previously, he was Sales Manager for Motorola in Southern California. He joined Motorola in 1950 as a radio communications representative, after managing a Motorola service station in San Luis Obispo, California for three years. He operated a radio common carrier system for three years.

Mr. LeGette is a charter member of the National Mobile Radio System.

❖

Maurice W. Horrell (A'37–M'44–SM'50) has been appointed Director of engineering and manufacturing centers for the UNIVAC division of the Sperry Rand Corp. He will be responsible for direction and coordination of all UNIVAC engineering and manufacturing.


M. W. HORRELL

Mr. Horrell was general manager of the computer division of Bendix Corp. prior to joining UNIVAC.

A registered professional engineer, Mr. Horrell received his Master's and Bachelor's degrees in electrical engineering from Kansas State University. He is a member of the Association for Computing Machinery, and the American Ordnance Association.

❖

Robert B. MacAskill (S'48–A'50–M'55–SM'56) has been named Director of Engineering of the J-V-M Division of Fidelitone Microwave, Inc., Brookfield, Ill.


R. B. MACASKILL

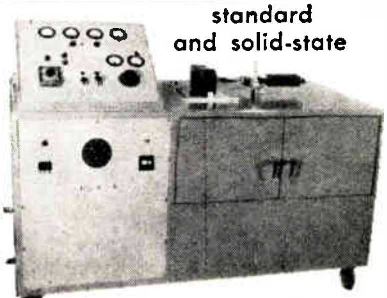
In his new position, Mr. MacAskill will direct the Research/Development and Systems Design Groups, and be responsible for new products.

Mr. MacAskill's previous experience includes microwave antenna and product development for The Hallicrafters Co., and Mark Products. Previous to that, he was a research and development consultant in the microwave field, specializing in antennas, systems and components.

Mr. MacAskill earned his degree in electrical engineering at the University of Illinois. At present he is business manager of "Scanfax," the publication of Chicago Section of the IRE. He also holds several patents covering design and applications of microwave components, and is the author of many technical papers.

❖

(Continued on page 254A)



Shortening Shadows

Alfred N. Goldsmith
Anniversary Editor

THE LIFE of a man or an organization has been aptly compared to the sequence of lights and shades during one day on earth. In the early morning, long shadows lie to the West as the sun rises. Then these shadows shorten as the day advances until no shadows are visible in the noonday blaze of light. And later, the shadows lengthen as the sunlight wanes, and twilight and night approach.

The Institute of Radio Engineers is fortunate in that it is still in its early morning amid the slowly shortening shadows. Its career is young and promising. Its hopes and accomplishments, while great and inspiring, are as yet only partly fulfilled. It has the cheerful stimulus of still being far even from the noon of its life, not to mention the evening.

Yet even at this early stage of the career of the Institute, an assessment of its accomplishments and an estimate of its future may be attempted. To this end, let us return in imagination to 1912, the year of the Institute's formation. It was a very different world from that of today. Knowledge of the communications and electronics field was limited. Vistas now clear were then dim or obscure. Professional standing and cooperation were not at today's encouraging levels. And the number of persons skilled in the electronics field was small indeed.

It was our privilege to be associated at that time with the other two founders of The Institute of Radio Engineers. Great tribute should be paid to these farsighted pioneers. Robert H. Marriott, President of the Wireless Institute (one of the merged societies which formed our Institute of Radio Engineers) was ever resourceful, determined, questioning, and analytic. And above all, he was an indefatigable worker and a thoughtful planner. John V. L. Hogan was a highly skilled, inventive, and forceful pioneer. He was that rare combination: a man both human in his reactions and humane in his instincts. Without the broad understanding and generosity of these men, our Institute might never have come into being.

The original builders of the IRE were certainly no example of squatting on their hunkers.

They were strong exponents of a much more logical creed that work conquers all. And work they did, over the passing years, and in the face of many obstacles.

Yet they were fortunate as the years went on in encountering an ever more favorable and respected environment. There were numerous fortuitous helpful events, and there were also many planned and logical campaigns. The founders and their many associates were able to function partly because of three major factors. One factor was the dependable and increasing devotion and effort of the membership. It would be impossible to exaggerate the sacrifices and labor which the membership readily gave the succession of officers of the Institute over the years. The second factor was the continued and friendly cooperation of the communications and electronics industry. Industry fully understood the work of the Institute and enthusiastically backed its efforts and expansion. And the third factor was the truly explosive growth of electronics, which provided an opportunity and a challenge which were gladly accepted and turned to the advantage of the Institute. And the Institute in turn contributed greatly to that expansion over the years. Thus the membership, industry, and the expansion of the electronics field offered the Institute its opportunities. Nor should it be forgotten that governments showed a clear understanding of the worth of the activities of the Institute and encouraged their officials, scientists, and engineers to affiliate with the Institute and to aid in its upbuilding. To all of these forces recognition must be given and gratitude expressed.

Yet growth itself, from a numerical viewpoint, is not enough to ensure the basic success of any professional organization or learned society. It must retain its personality. It must maintain its ideas and ideals. It must display continuing vitality. It must avoid entangling alliances or deviations from its true purpose. Accordingly it became necessary for the IRE carefully to plan its growth along statesmanlike lines to the end that it would remain as youthful and vigorous as in its earliest days. It would be too lengthy a recital to go into the details of the appropriate measures adopted by the Institute to accomplish its aims. Some of these measures have become classic and have blazed the trail for other engineering organizations which have wisely seen fit to adopt them. The IRE established its Sections to meet the needs of members in a city and its environs. It established subsections to fulfill the needs of smaller numbers of members in suburban sections or in smaller population centers. It coordinated the activities of Sections and subsections over large territories through the formation of Regional administrations. It coordinated Regional activities through National divisions or their equivalent.

And, in addition to these "horizontal organizations," the Institute established vertical organizations of all persons among its membership interested in a particular specialty. This led to the formation of Professional Groups and of their local Chapters. Undoubtedly the end is not yet. Additional organizations or administrations, and new forms of publication may in time be needed to meet the ever-increasing desires of the membership. And, if the past is any index to the future, the necessary statesmanship will be found to meet any current needs.

The IRE has carefully avoided such measures as would lead to its grounding on the shoals of disaster which often lie in the path of learned societies as they expand toward maturity. These pitfalls include insufficient interest in the organization by the individual member, of local geographic groups of members, of larger or regional groups of members, and of members imbued with a sense of the national dignity of their compatriot members. Also avoided was the development of insufficient interest resulting from the diverse activities of individual members of the Institute—a situation which was well met by the Professional Groups. More broadly, the Institute has also avoided the dangers of involvement with political, sectional, partisan, commercial, or personal viewpoints and activities which were not of a definitely engineering and scientific nature. It is easy for a learned society to forget its basic purpose of a search for truth, or for an engineering institute to forget that the practice of its profession for the benefit of humanity is its basic aim. The membership of the IRE may be proud in reviewing the past of the Institute and noting its complete detachment from purposes alien to its basic aims.

However, the Institute has not established itself in a guarded ivory tower and without contact or collaboration with others. One of its aims is continued cooperation and even coordinated activity or a suitable degree of integration with other engineering and scientific organizations. Wherever the IRE could help in the organization, operation, or publication activities of conventions, conferences, or symposia in collaboration with other engineering societies of standing, it has done so unhesitatingly and, at times, with considerable sacrifice. Since the establishment of the IRE Professional Groups, this type of cooperation has been greatly broadened. In fact, the roster of joint meetings or the like in which the IRE is engaged is almost startling in its dimensions.

One very important aspect of The Institute of Radio Engineers is its international nature, prescribed by its charter and consistently carried out in its activities. It has been said that science knows no country and that truth is universal. Such fields as communications and electronics are peculiarly adapted to the application of such doctrines, to the utmost practical limits.

Speaking more lightly, the techniques of communications and electronics might even make it logical that the scope of the IRE should ultimately become cosmic. We might look forward to the establishment of lunar and planetary Regions and Sections of the Institute in due course. And after that, who knows? But who would dare place finite limitations on the activities and studies which reach outward with the speed of light toward the infinite?

It may not be amiss also to emphasize that The Institute of Radio Engineers has sought and achieved a unique form of leadership in its dedicated field. In the world there are indeed many types of leadership. Some men, political parties, or nations maintain important positions of power or influence by brute force and iron discipline. Some organizations hold their eminence through guile, subtle persuasion, empty promises, or deceptive practices. But The Institute of Radio Engineers, the mental child and professional expression of its membership, holds its leadership only by service and accomplishment. Its sole aim is ever to give more to humanity and to its members. Its accomplishments require little recital or egotistical praise. They speak for themselves, and win hearty approval of all candid observers.

Thus the Institute will remain as a symbol of an epoch—the age of the approaching welfare, comfort, and health of every man on earth, and of man's conquest of space. And so, in the spirit of selfless aspirations to serve the interests of the people of the world and to express the best professional aims and accomplishments of its membership, The Institute of Radio Engineers looks forward through the decades and centuries toward ever wider, more useful, and more enlightening accomplishments.

The Fiftieth Anniversary Issue

MAY 13, 1962, marks the Fiftieth Anniversary of the founding of The Institute of Radio Engineers.

At this special moment the IRE finds itself poised precisely at the midpoint of a 100-year span where, in the fashion of Janus, it can watch simultaneously the gates of one half-century closing behind it and the gates of a new half-century opening before it. It is a time when we become more appreciative of the past, more aware of the future, and more conscious of the fact that they both converge upon the present.

It is a time also for remembering that the founding of the IRE 50 years ago coincided within a very few months with two other historic developments: the conversion of the three-element audion detector tube into an amplifier and into an oscillator, momentous events in which an IRE Charter Member, Lee de Forest, and IRE's first Vice President, Fritz Lowenstein, played leading roles. Thus, in a very real sense this occasion also commemorates the fiftieth anniversary of the beginnings of modern electronics and communications.

It is with these thoughts in mind that the IRE has assembled, in the month of its golden anniversary, an issue of its PROCEEDINGS of unprecedented scope, size, and distinction. Its pages carry the reader through a full century in time, from 1912 to 2012; its papers embrace the fields of all 28 Professional Groups of the IRE; its 180 contributors were chosen from among the leaders of the profession; its contents were written in their entirety by special invitation.

The reader will note from the table of contents in this issue that the material in the following pages falls into three major groupings. The subject of the first group of articles is the IRE itself. In the opening paper, an IRE fellow, who himself has been a member for 46 years, discloses many interesting facts about the origins, early history, and later growth of the IRE. This is complemented by special commentaries by the President, Past President, Secretary, and Executive Secretary, which provide a panoramic view of the position of the IRE today and its role in the future.

The unusual nature of the second group of contributions in the issue is indicated by its title, "Communications and Electronics—2012 A.D." In this section forty-eight Fellows of the IRE, in a series of brief essays, take the reader on a unique journey fifty years into the future, and describe to him the likely state of the art in 2012 A.D. as they envision it. As far as is known, never before has such a concerted assault on the barriers of time been attempted by so eminent an assembly of engineers and scientists.

The third group, which makes up the bulk of the issue, consists of 113 papers in which leading authorities take up in turn the fields of each of the 28 IRE Professional Groups and concisely summarize the evolution, present status and probable future trends of each. This portion of the issue presents a wealth of material concerning both past and current developments in a form that has never been published before. In a sense it comprises 28 condensed special issues in one.

In spite of the large size of the issue, the reader will have no difficulty in finding those articles that are of greatest interest to him. He will find that the second part of the issue, "Communications and Electronics—2012 A.D.," carries its own subject and author index at the end of that section. (Indeed, the subject index itself makes interesting reading.) The third part of the issue, dealing with Professional Group fields, is its own index. It has been divided alphabetically by subject into 28 sections, one section for each Professional Group field. In addition, an author index for the entire issue appears at the end.

Because of the special nature of this issue, the regular monthly features which PROCEEDINGS readers are accustomed to seeing at the back of the issue have been omitted. These include Correspondence, Books, Scanning the TRANSACTIONS, Abstracts of IRE TRANSACTIONS, and Abstracts and References. The May editions of these features will appear in the June issue.

The appearance of this issue represents the culmination of a two and one-half year effort. The issue was initiated late in 1959 when the Board of Directors asked Dr. Alfred N. Goldsmith, Editor Emeritus and co-founder of the IRE, to serve as Anniversary Editor and to assume responsibility for the conception, planning, and execution of an Anniversary Issue of the PROCEEDINGS. During the months that followed, the Anniversary Editor received helpful advice and assistance from many individuals and groups, including especially the Editorial Board, former Editor Ferdinand Hamburger, Jr., the Executive Secretary, and the Managing Editor.

In connection with the Professional Group section of the issue, particular thanks are due the 28 Professional Groups and their respective specially-formed Anniversary Committees, whose valuable suggestions of topics and authors and generous assistance in reviewing papers contributed greatly to the success of this undertaking. In particular, we are indebted to the following Professional Group representatives for the valuable service they rendered:

Batcher, Ralph (PGPEP)
Breiding, Eugene J. (PGRQC)
Elias, Peter (PGIT)
Fagen, M. D. (PGUE)
Felch, E. P. (PGI)
Grace, J. N. (PGNS)
Grobowski, Zigmund V. (PGRFI)
Gruenberg, Eliot (PGSET)
Hechtman, Harold (PGEM)
Herold, E. W. (PGED)
Hobbs, L. C. (PGEC)
Huggins, W. H. (PGCT)
Hughes, Wm. L. (PGBC)
Jackson, Albert S. (PGHFE)
Johnson, J. Kelly (PGBTR)
Liddell, Urner (PGBME)
McElwee, Eleanor (PGEWS)
Macdonald, Angus P. (PGVC)
Martin, D. W. (PGAU)
Minno, Harry (PGANE)
Mittleman, Eugene (PGIE)
Pritchard, W. L. (PGMTT)
Rhodes, Donald (PGMIL)
Rogers, A. W. (PGCP)
Smith, Phillip H. (PGAP)
Truxal, John G. (PGE)
Wallace, J. D. (PGCS)
Ward, John E. (PGAC)

The typesetting, printing and binding of more than 100,000 copies of a 1200-page volume, without disrupting the monthly flow of other issues of the PROCEEDINGS, requires not only a well-equipped and competently-run printing establishment, but also the wholehearted cooperation of a dedicated organization. We are especially pleased on this anniversary occasion to acknowledge the important contribution which the George Banta Company in Menasha, Wisconsin, has made to the IRE, not only in the production of this voluminous issue but in maintaining an unbroken record of devoted service to the PROCEEDINGS for a period of thirty-five years.

Our final thanks have been reserved for the 180 prominent engineers and scientists who, upon the invitation of the Anniversary Editor, so willingly accepted and so ably carried out the onerous task of preparing the outstanding contributions which comprise this issue. So skillful are their presentations that the average reader may not fully appreciate the difficulty of the assignment which each author faced, namely, to compress into a few pages a lifetime of knowledge and experience. It may be of some interest to note that among the contributors to this issue are 43 IRE officers and Directors, including 15 Presidents, and 119 IRE Fellows.

Thus the reader may be confident that this issue, in chronicling the past, assessing the present and interpreting the future, is presenting the judgments of those who are acknowledged as spokesmen for our profession.

—*The Managing Editor*

The Institute of Radio Engineers —Fifty Years of Service*

LAURENS E. WHITTEMORE†

FELLOW, IRE

A PREDICTION

(From an IRE information booklet, dated January 1, 1913, at which time the IRE had 109 members.)

The form of government of the Institute is thoroughly democratic and each member is given full opportunity to participate in all the advantages and privileges of the Institute. It is confidently expected that the already considerable membership will grow to the point where the greatest possible benefits will be extended to the greatest possible number of those interested in the development of radio-transmission.

INTRODUCTION

FOR THE INFORMATION of the members of the IRE, and especially those who are included in the one-third of the members who have joined during the past five years, there are here presented an outline of the beginnings of the Institute of Radio Engineers and some significant facts as to its organization, aims and recent accomplishments. Some statistics are included to show the growth of the Institute and the expansion of its activities over the 50 years of its existence. A perusal of these statistics shows how rapidly the Institute's activities are moving forward geographically, technically, professionally and in education.

This paper is in some respects a revision, and in some respects a repetition, of portions of the review of the first 45 years of IRE's service to its members, prepared under the sponsorship of the History Committee of the Institute and published in 1957.

Readers who are interested in more details of events, organization changes, etc., during the first 45 years of existence of the Institute, may refer to that paper which appeared in PROCEEDINGS OF THE IRE, May, 1957, pp. 597-635. As an appendix to both papers there is given a list of previous publications of significance relating to IRE history.

*Received by the IRE, November 17, 1961.

† Short Hills, New Jersey.

FORMATION OF THE INSTITUTE OF RADIO ENGINEERS

The Beginning

Radio people, apparently from the outset, have been characterized by a combination of two desires—1) to talk with one another about their accomplishments and their hopes, and 2) to write and read about the things which they and their technical "brothers" have been doing. The desire to present papers about "wireless" and to hear and discuss them resulted in the establishment of two organizations, one in Boston in 1907, and one in New York in 1908, whose members became the nucleus of the Institute of Radio Engineers in 1912.

Society of Wireless Telegraph Engineers

The Society of Wireless Telegraph Engineers (SWTE) was formed in Boston, Mass., on February 25, 1907, by John Stone Stone as an outgrowth of seminars held by engineers on the staff of the Stone Wireless Telegraph Company. Membership was eventually opened to men from Fessenden's National Electric Signaling Company and some other organizations. Members of this society were familiarly known as "swatties." John Stone Stone was the first President of this society.

The Wireless Institute

Robert H. Marriott made what appears to have been the first specific attempt to form a radio engineering society composed of members from any and all companies. On May 14, 1908, he sent a circular letter to some two hundred persons interested in wireless asking their opinions regarding the formation of such a society. On January 23, 1909, a temporary organization was formed to draw up a constitution. The name of the new society was "The Wireless Institute."

Robert Marriott was elected first President of The Wireless Institute, and served in that capacity during the three years of its existence.

By 1911 the Stone Wireless Telegraph Company had gone out of existence and the National Electric Signaling Company had moved to Brooklyn so there was very little left of the SWTE. The Wireless Institute was also struggling to hold its membership and to keep out of debt.

Institute of Radio Engineers

It was early in 1912 that Robert H. Marriott and Alfred N. Goldsmith, representing The Wireless Institute, with John V. L. Hogan, who was very active in the Society of Wireless Telegraph Engineers, held an informal meeting to discuss the plights of both societies. Out of their discussions there developed a meeting on the night of May 13, 1912, at which members of TWI and SWTE gathered in Room 304 of Fayerweather Hall at Columbia University in New York City. A constitution was approved and an election of officers was held at which the following were chosen: Robert H. Marriott, President; Fritz Lowenstein, Vice-President; E. D. Forbes, Treasurer; E. J. Simon, Secretary; Alfred N. Goldsmith, Editor; and Lloyd Espenschied, Frank Fay, J. H. Hammond, Jr., John V. L. Hogan, and John Stone Stone, Managers.

As a name for the organization "The Institute of Radio Engineers" was chosen. The original membership roster of the IRE consisted of 46 members, 22 from SWTE and 25 from TWI. One member, Greenleaf W. Pickard, was the only charter member of IRE who had been affiliated with both of the preceding organizations.

ORIGINAL MEMBERS OF THE INSTITUTE OF RADIO ENGINEERS AND THEIR AFFILIATION WITH PARENT SOCIETIES

Society of Wireless Telegraph Engineers

J. C. Armor	W. S. Hogg
Sewall Cabot	Guy Hill
W. E. Chadbourne	F. A. Knowlton
G. H. Clark	W. S. Kroger
T. E. Clark	Fritz Lowenstein
E. R. Cram	Walter W. Massie
G. S. Davis	G. W. Pickard
Lee deForest	Samuel Reber
E. D. Forbes	Oscar C. Roos
V. F. Greaves	J. S. Stone
J. V. L. Hogan, Jr.	*A. F. VanDyck

The Wireless Institute

William F. Bissing	Frank Hinners
A. B. Cole	James M. Hoffman
*P. B. Collison	Robert H. Marriott
James N. Dages	A. F. Parkhurst
*Lloyd Espenschied	G. W. Pickard
Philip Farnsworth	H. S. Price
Frank Fay	A. Rau
Edward G. Gage	Harry Shoemaker
*Alfred N. Goldsmith	*Emil J. Simon
Francis A. Hart	A. Kellogg Sloan
Robert L. Hatfield	C. H. Sphar
Arthur A. Herbert	Floyd Vanderpoel
R. A. Weagent	

* Member of the IRE as of January, 1962.

After about a year, it was decided to incorporate the society. A meeting to decide details of this move was held at Sweet's Restaurant on Fulton Street in downtown New York on June 23, 1913. Those members with a more legal mind in the make-up of the Institute, drew up Articles of Incorporation, and on August 23, 1913, the organization was incorporated under the laws of the State of New York.

In brief, the expressed aims of the new association were:

"To advance the art and science of radio transmission, to publish works of literature, science and art for such purpose, to do all and every act necessary, suitable and proper for the accomplishment of any of the purposes or the attainment of any of the powers herein set forth, either alone or in association with other corporations, firms or individuals to do every act or acts, thing or things, incidental or appurtenant to or growing out of or connected with the aforesaid science or art, or power or any parts thereof, provided the same be not inconsistent with the laws under which this corporation is organized, or prohibited by the State of New York."

Growth of Membership of the Institute of Radio Engineers and its Predecessors, 1907–1914

	SWTE	TWI	IRE
February 25, 1907	11		
January 1, 1908	17		
January 1, 1909	27		
March 10, 1909		14	
January 1, 1910	36	81	
January 1, 1911	36	99	
January 1, 1912	43	27	
May 13, 1912	(22)	(25)	46
January 1, 1913			109
January 1, 1914			231

One of the most important functions of the Institute was to preserve its technical papers, and the remarks made regarding them, in published form. One of the early decisions of the Institute, therefore, was to publish a technical magazine which was named The Proceedings of the Institute of Radio Engineers. The first issue was dated January, 1913.

By the end of 1912, the Institute's membership had risen to 109 and during the succeeding year it more than doubled. The rapid increase in membership after consolidation, compared with the slow rate of growth of SWTE and TWI, bore out the wisdom of the founders who suggested the merger.

The original ledger book of the Institute, in which the names and dues payments of early members were recorded, constitutes a veritable *Who's Who* in the early history of radio. The names of many radio pioneers can be seen in the accompanying illustration showing the first few pages of the list of these members who joined the Institute during the first year.

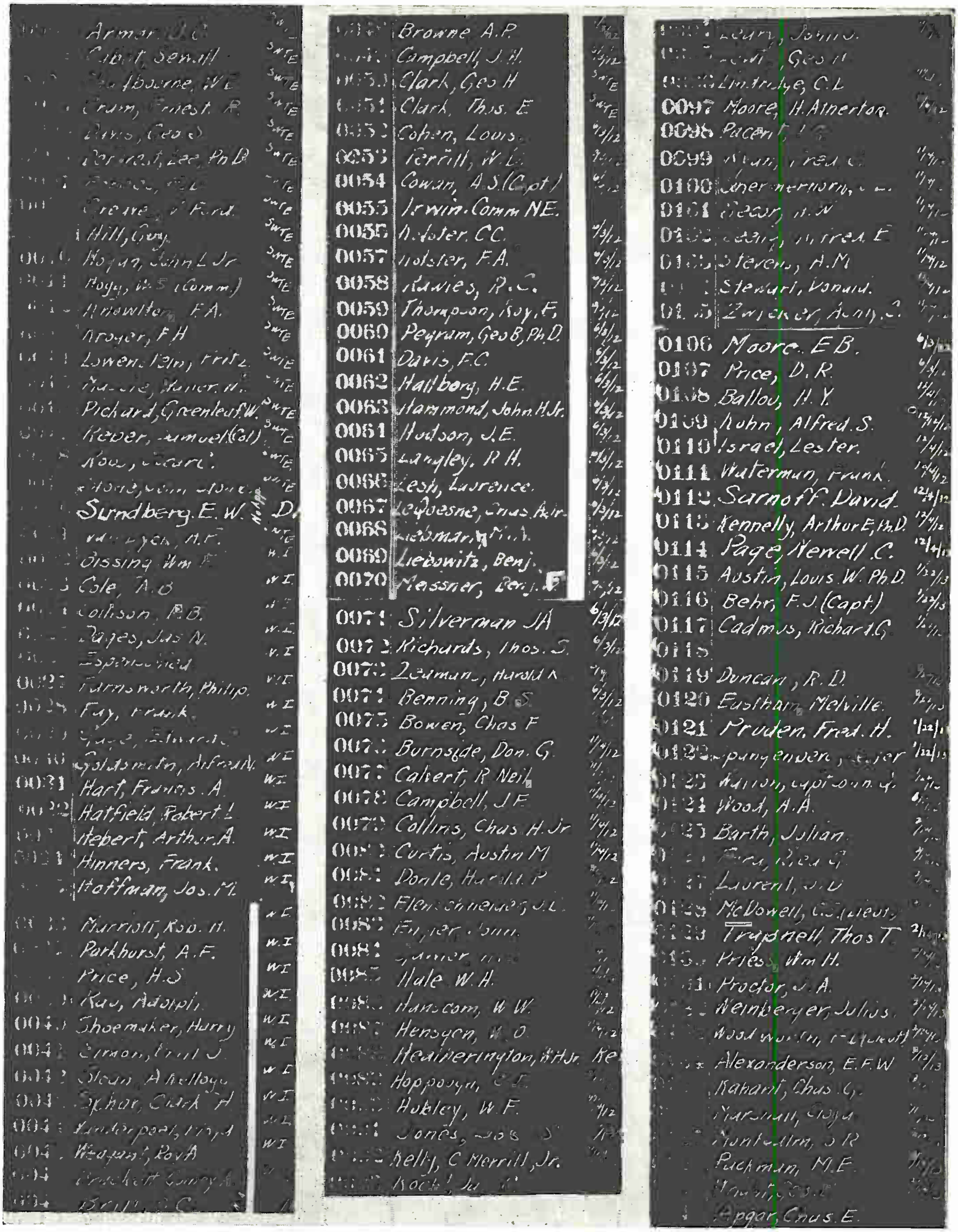
Name of the Institute

In considering a name for the new organization the founders felt that something should be preserved from the names of both of the two component societies. The word "Institute" was borrowed from The Wireless Institute, and "Engineers" from the Society of Wireless Telegraph Engineers. Because the word "radio" was gradually supplanting "wireless," the title "The Institute of Radio Engineers" suggested itself. There was considerable temptation to add "American," particularly since TWI and IRE were modeled after the American Institute of Electrical Engineers in certain other respects. However, the temptation was resisted because it was expected that the IRE, as the only radio engineering society in existence, would be international in scope, an expectation that was promptly realized.

The Emblem or Symbol of the IRE

Neither of the emblems of the predecessor societies seemed readily adaptable to the new IRE. The SWTE emblem pictured a simple form of spark oscillator. The membership badge of TWI showed a spark gap functioning in the center of a dipole surrounded by a circular resonator provided with a micrometer gap for reception.

The founders of IRE decided not to use a representation of any specific form of equipment or physical structure but to devise a more general and perhaps perpetual symbol. It was realized that the Institute would always deal with electromagnetic energy, guided by conductors or passing through space, and that the distinguishing character of the transmission process was the existence of electrical forces and of their correlative magnetic forces. A representation of these forces



Armor, J.C.
Cabot, Sewall
Crum, Ernest R.
Hill, Guy
Hogan, John L. Jr.
Hogg, W.S. (Comm.)
Knowlton, F.A.
Kroger, F.H.
Lowenstein, Fritz
Pickard, Greenleaf W.
Kever, Samuel (Col.)
Sundberg, E.W.
Bissing, Wm F.
Cole, A.B.
Collison, R.B.
Pages, Jas N.
Farnsworth, Philip
Fay, Frank
Goldsmith, Alfred N.
0031 Hart, Francis A.
0032 Hatfield, Robert L.
Hebert, Arthur A.
Hinners, Frank.
Hoffman, Jos. M.
Marriott, Rob. H.
Parkhurst, A.F.
Price, H.S.
Kau, Adolph
0040 Shoemaker, Harry
Simon, Emil J.

Browne, A.P.
Campbell, J.H.
Clark, Geo H.
Clark, Thos. E.
Cohen, Louis
0053 Ferrill, W.L.
0054 Cowan, A.S. (Capt.)
0055 Irwin, Comm. N.E.
0056 Hofer, C.C.
0057 Holster, F.A.
0058 Rawles, R.C.
0059 Thompson, Roy F.
0060 Pegram, Geo B., Ph.D.
0061 Davis, F.C.
0062 Hallborg, H.E.
0063 Hammond, John H. Jr.
0064 Hudson, J.E.
0065 Langley, R.H.
0066 Lesh, Laurence
0067 LeQuesne, Chas. Adr.
0068 Liebmann, ...
0069 Liebowitz, Benj.
0070 Meissner, Benj. F.
0071 Silverman, J.A.
0072 Richards, Thos. S.
0073 Lehmann, Harold A.
0074 Benning, B.S.
0075 Bowen, Chas F.
0076 Burnside, Don. G.
0077 Calvert, R. Neil
0078 Campbell, J.F.
0079 Collins, Chas H. Jr.
0080 Curtis, Austin M.
0081 Doyle, Harold P.
0082 Fleischmann, J.L.
0085 Hale, W.H.
Hanscom, W.W.
Hensyen, W.O.
Heatherington, H. Hjr.
Hoppough, C.E.
Hobley, W.F.
Jones, ...
Kelly, C. Merrill, Jr.
Koch, ...

Leary, John ...
Lindridge, C.L.
0097 Moore, H. Atherton
0098 Pacent, L.G.
0101 Secor, H.W.
Stevens, A.M.
Steward, Donald
Zwicker, Henry C.
0106 Moore, E.B.
0107 Price, D.R.
0108 Ballou, H.Y.
0109 Kuhn, Alfred S.
0110 Israel, Lester
0111 Waterman, Frank
0112 Sarnoff, David
0113 Kennelly, Arthur E., Ph.D.
0114 Page, Newell C.
0115 Austin, Louis W., Ph.D.
0116 Behr, F.J. (Capt.)
0117 Cadmus, Richard C.
0118
0119 Duncan, R.D.
0120 Eastham, Melville
0121 Pruden, Fred. H.
0122 Spangenberg, ...
0124 Wood, A.A.
0125 Barth, Julian
Laurent, ...
McDowell, ...
Trapnell, Thos T.
Priess, Wm H.
Proctor, J.A.
Weinberger, Julius
Woodworth, ...
Alexanderson, E.F.W.
Kahant, Chas G.
Marshall, ...
Puckman, M.E.
Apgar, Chas. E.

Pages from the first IRE Record book listing the earliest members.

IRE OFFICERS, 1912–1962

Year	President	Vice President	Secretary	Treasurer	Editor	Hdqs. Manager
1912	R. H. Marriott	Fritz Lowenstein	E. J. Simon	E. D. Forbes	A. N. Goldsmith	
1913	G. W. Pickard	R. H. Marriott	"	J. H. Hammond, Jr.	"	
1914	L. W. Austin	J. S. Stone	"	"	"	
1915	J. S. Stone	G. W. Pierce	David Sarnoff	W. F. Hubley	"	
1916	A. E. Kennelly	J. V. L. Hogan	"	"	"	
1917	M. I. Pupin	"	"	L. R. Krumm	"	
1918	G. W. Pierce	"	A. N. Goldsmith	Warren F. Hubley	"	
1919	"	"	"	"	"	
1920	J. V. L. Hogan	E. F. W. Alexanderson	"	"	"	
1921	E. F. W. Alexanderson	Fulton Cutting	"	"	"	
1922	Fulton Cutting	E. L. Chaffee	"	"	"	
1923	Irving Langmuir	J. H. Morecroft	"	"	"	
1924	J. H. Morecroft	J. H. Dellinger	"	"	"	
1925	J. H. Dellinger	Donald McNicol	"	"	"	
1926	Donald McNicol	Ralph Bown	"	"	"	
1927	Ralph Bown	Frank Conrad	"	"	"	J. M. Clayton
1928	A. N. Goldsmith	L. E. Whittemore	J. M. Clayton	Melville Eastham	"	"
1929	A. H. Taylor	Alexander Meissner	"	"	W. G. Cady	"
1930	Lee de Forest	A. G. Lee	H. P. Westman	"	A. N. Goldsmith	H. P. Westman
1931	R. H. Manson	C. P. Edwards	"	"	"	"
1932	W. G. Cady	E. V. Appleton	"	"	"	"
1933	L. M. Hull	Jonathan Zenneck	"	"	"	"
1934	C. M. Jansky, Jr.	B. van der Pol, Jr.	"	"	"	"
1935	Stuart Ballantine	G. H. Barkhausen	"	"	"	"
1936	L. A. Hazeltine	Valdemar Poulsen	"	"	"	"
1937	H. H. Beverage	P. P. Eckersley	"	"	"	"
1938	Haraden Pratt	E. T. Fisk	"	"	"	"
1939	R. A. Heising	P. O. Pederson	"	"	"	"
1940	L. C. F. Horle	F. E. Terman	"	"	"	"
1941	F. E. Terman	A. T. Cosentino	"	Haraden Pratt	"	" (Jan.–Oct.) J. D. Crawford (Nov.–Dec.)
1942	A. F. Van Dyck	W. A. Rush	"	"	"	J. D. Crawford (Jan.–Mar.) L. B. Keim (Apr.–May) W. B. Cowilich (Oct.–Dec.)
1943	L. P. Wheeler	F. S. Barton	Haraden Pratt	R. A. Heising	"	W. B. Cowilich
1944	H. M. Turner	R. A. Hackbusch	"	"	"	"
1945	W. L. Everitt	H. F. van der Bijl	"	"	"	G. W. Bailey
1946	F. B. Llewellyn	E. M. Deloraine	"	W. C. White	"	"
1947	W. R. G. Baker	Noel Ashbridge	"	R. F. Guy	"	"
1948	B. E. Shackelford	R. L. Smith-Rose	"	S. L. Bailey	"	"
1949	S. L. Bailey	A. S. McDonald	"	D. B. Sinclair	"	"
1950	R. F. Guy	R. A. Watson-Watt	"	"	"	"
1951	I. S. Coggeshall	Jorgen Rybner	"	W. R. G. Baker	"	"
1952	D. B. Sinclair	H. L. Kirke	"	"	"	"
1953	J. W. McRae	S. R. Kantebet	"	"	"	"
1954	W. R. Hewlett	M. J. H. Ponte	"	"	J. R. Pierce	"
1955	J. D. Ryder	Franz Tank	"	"	"	"
1956	A. V. Loughren	Herre Rinia	"	"	D. G. Fink	"
1957	J. T. Henderson	Yasujiro Niwa	"	"	"	"
1958	D. G. Fink	C. E. Granquist	"	"	J. D. Ryder	"
1959	Ernst Weber	D. B. Sinclair	"	"	"	"
1960	R. L. McFarlan	J. N. Dyer	"	"	F. Hamburger, Jr.	"
1961	L. V. Berkner	J. A. Ratcliffe J. F. Byrne F. Ollendorff	"	S. L. Bailey	"	"
1962	P. E. Haggerty	A. M. Angot T. A. Hunter Ernest Weber	"	"	T. F. Jones, Jr.	"

was adopted as part of the symbol; the electrical force being represented by a vertical arrow and the magnetic force by a circular arrow surrounding the electrical line and in the conventional relationship to it. The shape of the resulting drawing lent itself to a triangular placement of the letters, I, R, and E. This, in turn, led to the selection of a triangular emblem. Incidentally, the letters I, R, and E also symbolize the fundamental quantities, current, resistance and electromotive force, as well as the name, Institute of Radio Engineers.

Officers

The officers of the IRE from the beginning have been President, one or more Vice-Presidents, Secretary, Treasurer, Editor, and Directors with the infrequent addition of an Assistant Secretary or an Assistant Treasurer. The President and Vice-President(s) have always been elected by the IRE membership as have part of or all the members of the Board of Directors. Beginning in 1915, the elected members of the Board were authorized by the Constitution to choose several additional persons to complete the Board membership. The Secretary and Treasurer have for many years been elected by the Board.

An accompanying table shows, for each year since the formation of the IRE, the names of the Officers who served during that year.

Beginning in 1930, it became the custom for the Vice-President of the IRE to be a member who resides in a country other than the United States. In 1957 neither the President nor the Vice-President was a resident of the United States. In 1960, the IRE established two Vice-Presidents. Initially, one resided in North America and the other resided outside of North America. In 1962, the former was replaced by a Vice-President, elected by the voting members, whose function is to assist the President. Also a new Vice-President, elected by the Annual Assembly, was created to serve as a coordinator between the Professional Groups and the Executive Committee. The office of the Vice-President residing outside North America will terminate at the end of 1962.

As the IRE membership increased in numbers and geographical distribution, there developed an appreciation by the members of the Board that some specific measures should be adopted, possibly of an organizational nature, to make more certain that the contacts between the Board and the IRE membership would always be close and continuous and that the Board would comprise a truly democratic representation of the IRE membership. After several years of consideration of this problem, the Board of Directors recommended an amendment to the Constitution which was adopted by the Institute membership in 1947, providing for Regional Directors, selected specifically to represent designated regions of the United States and Canada from which they came and whose memberships had elected them.

Membership

Growth in Numbers

It is fitting to consider as the "Charter Members" of the IRE those members of the two parent societies who became the first members of the IRE when it was organized on May 13, 1912. The formal charter of the Institute of Radio Engineers, however, was granted on August 23, 1913.

The 1914 Year Book gives an analysis of the geographical distribution of the 271 Members and Associate Members of the IRE as of March 1, 1914, showing that there were members in eight countries other than the United States. The 1916 Year Book shows that the membership of the Institute immediately began to increase and by January 1, 1916, was only slightly under 1000.

The Constitution, as adopted in May, 1912, provided that the names of applicants for membership in the IRE should be sent out to each member of the Institute, prior to their acceptance as members. The members who were elected were required to subscribe personally to the Constitution of the IRE.

From the beginning, membership was open not only to radio engineers (Member grade) but also to those who had a real interest in radio engineering even though they were not professionally engaged in this field (Associate Member grade).

In 1914 the grade of Junior Member was established for persons under 21 years of age.

One of the outstanding aspects of IRE membership is the substantial number and generally increasing proportion of its membership living in countries outside of the United States. The IRE has taken special steps from time to time to recognize outstanding members living in other countries and to stimulate membership in such areas. At the end of 1961 the IRE had members in 90 countries outside of the United States.

Significance of Membership Grades

In the IRE, as is customary in professional societies, the several grades of membership are intended as a basis for giving recognition to the experience and achievements of the members in radio engineering and the related technical fields.

From the beginning, the Fellow grade has been intended to represent high attainment.

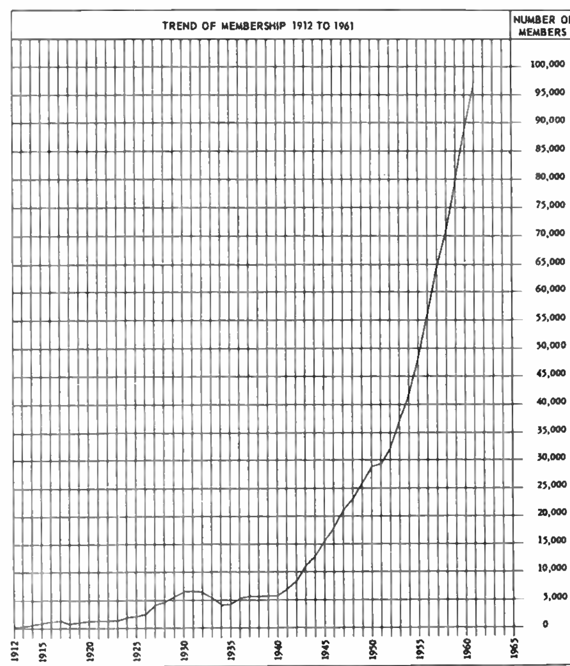
Beginning with the year 1940, all entries to the Fellow grade have been by invitation rather than upon application. The custom has been established of presenting the Fellow awards at meetings of the local Sections of which the recipients are members, and 78 persons were being honored in this way in the early part of 1962. About one per cent of the IRE's total membership of over 96,000 are of the Fellow grade.

Membership Growth in U.S.A., Canada and Abroad

Year	U.S. and Possessions	Canada	Abroad	Total
May 13, 1912	46			46
Dec. 31, 1927	3550	184	476	4210
" 1936	3975	178	1042	5195
" 1946	15,898	978	1278	18,154
" 1956	51,551	2085	1858	55,494
" 1961	88,956	3758	3837	96,551

The Senior Member grade was established in 1943 as a means of providing a higher grade than the Member grade into which members of the IRE might advance on the basis of their experience and training. This enabled the Institute to keep the Fellow grade as a special recognition.

Recognizing the desirability of encouraging engineering students to become affiliated with the Institute of Radio Engineers, a Student grade of membership in the IRE was established by constitutional amendment effective in 1932.



Membership Growth by Grades

Date	Junior	Student	Associate Non-Voting	Associate Voting	Member	Senior Member	Fellow	Total
May 13, 1912				46				46
Dec. 31, 1936	34	299		4092	637		133	5195
Dec. 31, 1946		2252	9890	1701	2330	1763	218	18,154
Dec. 31, 1956		10,384	18,491	388	19,110	6486	635	55,494
Dec. 31, 1961		19,167	13,566		52,284	10,570	964	96,551

The previously existing grade of Junior Member was dropped in 1943.

It was a natural development for the Student Members at a given college to meet together to discuss radio engineering questions among themselves or to hear addresses by visiting engineers, and a number of informal Student groups or "branches" came into existence.

The IRE Bylaws provide that every member who has attained the age of 65 years, and who has been a member of the IRE for 35 years or more, is eligible for Life Membership with waiver of payment of further dues. As of December 31, 1961, there were 272 Life Members of the IRE.

Standardization

From the beginning of its existence, the officers of the IRE have recognized that standardization is necessary for the orderly exchange of information and for effective progress in a technological field.

The Standardization Committee in 1913 published a report dealing with definitions of terms, letter and mathematical symbols, and methods of testing and rating equipment.

The current IRE standards are 89 in number. Of these 44 standards deal with terminology, symbols, and definitions of terms, and 45 standards deal with testing or measurement methods. Most of these standards have appeared in the Proceedings and all of them are available in reprint form. A list of the current standards is published from time to time in the Proceedings. As of the end of 1961, 18 current IRE standards have been adopted by the ASA as American Standards.

Until recently, the IRE Standards have been published either in separate pamphlets or singly in issues of the Proceedings. Now the Standards dealing with terminology and symbology have been gathered together in one "IRE Dictionary" of 50 pages containing 3700 technical terms and definitions and five standards of letter and graphical symbols.

Sections

Formation

The activities of the IRE, originally confined to the New York area, quickly spread to other cities as the membership increased. Small groups of members in various localities began to hold their own local meetings and elected their own officers for organizing and running these meetings. These groups of members were called Sections.

The first local Section of the IRE was organized at Washington, D.C., in January, 1914. Section meetings were held each month for the benefit of the group of perhaps a score of engineers, then described as "a large number of members residing in the vicinity of the capitol."

Following the formation of additional Sections, an approved form of a Constitution for Sections was made mandatory in 1930, by action of the Board of Directors.

The New York Section, as such, was established in 1942. Before that time the meetings held in New York were considered meetings of the Institute.

Effective January 1, 1946, the Board made a change in the method of assigning territory to Sections. Instead of assigning a small territory around a city as the area of each Section it was decided that the entire area of the United States should be assigned to one Section or another. This meant that every member of the IRE in the United States would automatically be on the membership list of a Section and would receive notices of Section meetings.

Sections in Other Countries

On October 2, 1925, the first Canadian Section was formed in Toronto at a meeting attended by fifty-three IRE members and guests. Now, thirty-six years later, the IRE has a membership in Canada of over 3700 members of all grades. Meanwhile in 1945 the Canadian Council of the IRE had been formed to coordinate the activities of various Canadian Sections on a national basis. This plan formed the basis of the organization that later became effective in the eight IRE Regions of North America.

The first Section of the IRE formally established outside the United States and Canada was the Buenos Aires, Argentina, Section established in 1939.

By the end of 1961, the IRE had 14 Sections in Canada and also 12 Sections outside North America as follows: Benelux, Buenos Aires, Chile, Colombia, Egypt, France, Geneva, India, Israel, Italy, Rio de Janeiro, and Tokyo.

Growth Under Regional Plan

The Sections were tied in more closely with the administration of the IRE in 1946 when the Sections were grouped into eight Regions and each Region was given representation on the IRE Board of Directors, each Regional Director being elected on a biennial basis by the membership of the respective Region. Since then the Regional Plan has provided an effective channel of communication between the Board and the membership of the local Sections. The growth of this important "grass roots" activity has continued unabated so that, at the end of the year 1961, there were 99 Sections and 31 Subsections in the United States and Canada, as well as the twelve Sections in other parts of the world. A list of Sections and their officers appears bimonthly in the PROCEEDINGS OF THE IRE.

TOTAL PAID MEMBERSHIP OF IRE PROFESSIONAL GROUPS

At End of Year	No. of Groups	Total Paid Members
1948	2	
1949	8	
1950	10	8500*
1951	16	13,000*
1952	17	12,482
1953	20	21,797
1954	21	28,158
1955	23	36,562
1956	24	53,015
1961	29	87,060

* Estimated; includes both paid and unpaid members.

PROFESSIONAL GROUP SYSTEM

The most basic change in the structure of the IRE which occurred during its half-century of growth was the establishment in 1948 of the Professional Group System, providing for Groups within IRE along lines of technical specialization of the members. The successful development of the Professional Group System is proving to be an effective means of counteracting the centrifugal tendencies that otherwise might have accompanied the rapid expansion of the Institute.

The first two Professional Groups were authorized by the Board on September 9, 1948. These were Groups known as "Audio," and "Broadcast Engineers."

The subsequent establishment of other Groups has brought the list to 29 at the end of 1961. A list of the Professional Groups appears bimonthly in PROCEEDINGS OF THE IRE.

The procurement of papers for national symposia and management of these symposia are now entirely in the hands of the Professional Groups. The wide scope of activity of the members of the IRE is indicated by the variety of subjects of the Symposia sponsored by the various Professional Groups, either solely or jointly with organizations having related interests. Here are some rather interesting examples:

Joint Automatic Control Conference, Cambridge, Mass., September, 1960
National Symposium on Space Electronics and Telemetry, Washington, D.C., September, 1960
Symposium on Electromagnetics and Fluid Dynamics of Gaseous Plasma, New York, N.Y., April, 1961
Fifth National Symposium on Global Communication, Chicago, Ill., May, 1961
National Symposium on Engineering Writing and Speech, East Lansing, Mich., October, 1961.

Professional Group Chapters

There were 303 Professional Group Chapters, organized by Group members in 57 IRE Sections, as of December, 1961. Chapter growth is continuing at a healthy rate. The Chapters are meeting regularly and sponsoring meetings in the fields of interest of their associated Groups.

As of the end of 1961, about one-half of the IRE members had, on the average, joined two Professional Groups, for a total Professional Group membership of 87,060. Over one-third of the Student Members of the IRE, numbering nearly 7400, had, by the end of 1961, joined one of the Groups covering their particular field of interest.

The comment was once made by John V. L. Hogan, one of the Founders and past Presidents of IRE, that "the Institute of Radio Engineers can well be considered to be the federal body which unites an increasing number of autonomous professional groups."

"Affiliate" Plan

In January, 1957, the Board took a step which bids fair to be a notable one in IRE history. Based on a need which had developed especially in the case of the Professional Group on Medical Electronics, it adopted a plan which enables non-IRE members whose main professional interests lie outside the sphere of IRE activities to become affiliated with certain of the IRE Professional Groups in whose activities they wish to share without having to join the IRE. To be such an "Affiliate" of a Group a person must belong to an accredited organization approved by that Group and by the IRE Executive Committee. Participation in the Affiliate Plan is at the option of each Professional Group. On payment of the regular assessment fee of his Group, plus $4.50, the Affiliate is entitled to receive the TRANSACTIONS of the Group. The Affiliate Plan is intended to recognize and provide for the rapidly spreading influence of electronics in every walk of scientific and technological life, and to enable the IRE further to fulfill its aim—that of advancing radio engineering and related fields of engineering and science.

PUBLICATIONS

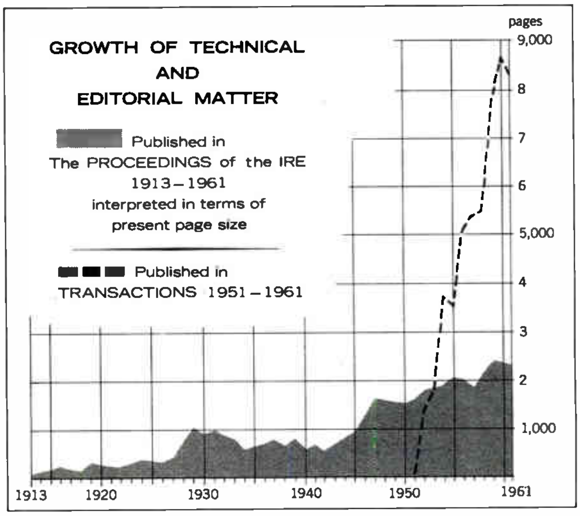
Proceedings of the IRE

In meeting the IRE objective of promoting the exchange of radio engineering knowledge, its publication program is perhaps the most important single activity of the Institute.

The PROCEEDINGS OF THE INSTITUTE OF RADIO ENGINEERS was established originally as a quarterly publication in 1913. An amendment to the IRE Constitution, effective in 1916, authorized the publication of six issues of the PROCEEDINGS per year instead of four. Beginning in 1927, the PROCEEDINGS has been published monthly. In 1939 the name of the journal was shortened to PROCEEDINGS OF THE IRE.

The first Editor was Alfred N. Goldsmith, one of the three Founders of the IRE. It is under the stimulus of his inspired leadership that, through the ensuing years, the IRE has turned out a publication which has been generally recognized as the leading technical magazine in the radio field.

At the outset, the contents of the PROCEEDINGS consisted of certain papers and discussions



as presented at meetings in New York and other cities. During the year 1913, thirteen technical papers were published and a total of 268 pages were occupied, an average of 67 pages per issue, size 6½ by 9 inches. The scope of the PROCEEDINGS was soon broadened to include more technical papers, and also lists of newly elected members, personal notes, reports of meetings and committees, and other information pertaining to the Institute.

To accomplish a balanced publication program, the IRE must supplement the contributed papers by invited papers, either to fill gaps in the coverage of new technical developments or to present a general review of a given field in a way which is helpful to a reader who is not a specialist in that field.

Beginning in 1951, there have been published, from time to time, numerous unusually large issues of the PROCEEDINGS, each devoted exclusively to a single subject. The first of these was the October, 1951, Color Television Issue. This issue consisted of 400 pages and was the largest single issue of the PROCEEDINGS ever published up to that time.

The expanding field dealt with in these special issues during the past few years is indicated by the following illustrative titles:

Radio Astronomy Issue, January, 1958
Transistor Issue, June, 1958
Ionosphere-IGY Issue, February, 1959
Government Research Issue, May, 1959
Infrared Issue, September, 1959

Bio-Medical Issue, November, 1959
Space Electronics Issue, April, 1960
Computer Issue (Second), January, 1961
Plasma Issue, December, 1961

As of January 1, 1954, Dr. Goldsmith was relieved of active editorial responsibilities and "in recognition of the invaluable service he has rendered the Institute and his monumental contributions to the growth and high standards of its publications" was appointed Editor Emeritus. Since January 1, 1954, five other person have successively served as Editor. Previously, however, the Institute had established the position of Managing Editor which is filled by a full-time employee of the Institute.

The drop in the membership of the Institute during the depression of the early 1930's led the Board of Directors to search for additional ways by which the members could reap value from the various services of the Institute. An immediate step was the expansion of the "Institute Notes" section of the PROCEEDINGS. This was so well appreciated that since then, from time to time, additional features have been added or the material published has been recognized. As a result, the PROCEEDINGS OF THE IRE now carries, in addition to its main body of technical papers, more than a dozen regular features and "Departments," most of them monthly.

Transactions of Professional Groups

The first major broadening of the IRE publication program was the appearance of the TRANSACTIONS issued by the Professional Groups. They immediately proved to serve a very useful purpose as a quick, practical method of providing Professional Group members with technical papers in their particular fields of interest. As early as 1951, six issues of such TRANSACTIONS were issued by the Audio Group and two issues of TRANSACTIONS by the Airborne Electronics Group.

The growth of the Professional Group TRANSACTIONS is indicated by the fact that in 1961, 102 issues of TRANSACTIONS were published by 28 Professional Groups containing a total of over 1200 papers. The TRANSACTIONS have circulations ranging from 1500 to 11,000 per issue. Publication varies with the Group, from one to six issues per year. Over-all distribution is about 360,000 annually. The total annual TRANSACTIONS output now exceeds that published in the PROCEEDINGS.

Convention Record

The papers that were presented at the National Conventions of the IRE in New York from 1953 to 1961, inclusive, have been published as Volumes I to IX, respectively, of what is now known as the IRE INTERNATIONAL CONVENTION RECORD.

From 1957 to 1960 the IRE also published the papers presented at the Western Electronics Show and Convention in an annual volume called the IRE WESCON CONVENTION RECORD.

Student Quarterly

In September, 1954, a new publication of the IRE was launched. It was the IRE STUDENT QUARTERLY which is sent free to all Student Members of the IRE as part of a program of increased service to students. It contains news of Student Branches, articles of technical value, and papers of significance to students who are interested in considering carefully whether to devote themselves to an engineering career.

The IRE STUDENT QUARTERLY is now in its eighth year of publication and now ranks second in circulation of IRE's 31 periodicals. The articles which it contains are so worthwhile and so readable that it is the subject of an increasing circulation on a subscription basis among other than Student Members.

Indexes to IRE Publications

With the increasing quantity of material published by the Institute, and in view of the widening areas of subject matter dealt with, it became apparent that a general cumulative index should be prepared and made available to all subscribers to the PROCEEDINGS. Several such indexes have been compiled from time to time. The four indexes which now cover this material are:

Cumulative Index, 1913–1942 (Part II of PROCEEDINGS, June, 1943)
Cumulative Index, 1943–1947 (Part II of PROCEEDINGS, June, 1948)
Cumulative Index of IRE Publications, 1948–1953

Cumulative Index of IRE Publications, 1954–1958

This latter index, printed as a separate pamphlet, includes references to the PROCEEDINGS OF THE IRE, the IRE TRANSACTIONS, the IRE INTERNATIONAL CONVENTION RECORD, the IRE WESCON RECORD and the STUDENT QUARTERLY.

Year Book and Directory

The first published list of IRE members, dated January 1, 1913, was a 24-page pamphlet, size 3½ by 6 inches, listing 100 persons. The names of 9 of these persons still appear in the current list of IRE members—the 1962 DIRECTORY. Five of these ten were Charter Members of the Institute.

The first YEAR BOOK, published in 1914, was followed by 13 others which were not always issued annually, through 1949. In the earlier years, the Report of the Committee on Standardization was included. With the 1950 edition, the name was changed to DIRECTORY. Now published annually in November, it is sent to all members except Students. It contains alphabetical and geographical listings of all members except Students, classified listings of products, and an alphabetical list of commercial organizations. The separate listing of those who hold the Fellow grade of membership, with a biographical sketch of each, was begun in 1946.

Publication Growth

The IRE, in the first volume of the PROCEEDINGS in 1913, published 13 technical papers. In 1950, still in one journal, it published about 200 papers and letters. In 1961, in 31 publications, it published about 2000 papers, letters and other significant items.

Through 1961, in total, about 7500 technical papers, letters to the Editor, and book reviews have been published in PROCEEDINGS.

The growth in the total publication output of the IRE is shown by the contrast between the total size of the PROCEEDINGS and the YEAR BOOK published in 1914 and the size of the total of IRE publications in 1961. The PROCEEDINGS in 1914 had a volume that, if printed on pages of the size now utilized, would be 124 pages; the 1914 YEAR BOOK, if measured in pages the same size as those of recent YEAR BOOKS would be a total of 12 pages, making a total publication volume of 136 pages. The 1961 editorial output of the IRE consisted of 131 issues of publications, amounting to a total of 16,922 pages.

IRE Publications in 1961	Pages
PROCEEDINGS OF THE IRE	4,304
CONVENTION RECORD (in 10 parts)	2,164
TRANSACTIONS of Professional Groups	8,336
DIRECTORY	1,612
STUDENT QUARTERLY	224
Miscellaneous Technical Publications	282
Total	16,922

AWARDS

The Institute of Radio Engineers, from the beginning, has gone to some lengths to show honors to those among its membership who have made accomplishments of outstanding value to the development of the radio engineering field. Such honors are shown by the Board of Directors in granting awards. In addition, the granting of special honors for individual accomplishments, by the action of Sections, Professional Groups, and Student Branches has been authorized by the Board from time to time.

The IRE now bestows seven formal awards, six of which are awarded annually in recognition of outstanding technical and administrative achievements in the field of radio communication. The names of the 138 persons who have received these formal awards through the year 1961 are given in the 1962 IRE DIRECTORY.

Medal of Honor

The Medal of Honor is given annually in recognition of outstanding scientific or engineering achievements in the field of activity of the IRE. This medal was first given in 1917 and has been given annually since that year, with the exception of the years 1918, 1925, and 1947 in which no award was made.

Founders Award

The Founders Award is bestowed only in special occasions for outstanding leadership in planning and administration of important technical developments. The award commemorates the three radio pioneers who founded the Institute of Radio Engineers in 1912—Alfred N. Goldsmith, John V. L. Hogan, and Robert H. Marriott. This award has so far been made to six persons—David Sarnoff in 1953, Alfred N. Goldsmith in 1954, Raymond A. Heising in 1957, W. R. G. Baker in 1958, Haraden Pratt in 1960, and Ralph Bown in 1961.

Prize Award by Vladimir K. Zworykin

The Prize Award by Vladimir K. Zworykin was first awarded in 1952 and is to be awarded annually for 20 times to the member of the IRE who makes the most important technical contribution to electronic television during the preceding three calendar years or the importance of whose contribution to electronic television shall have been realized during this period. The award consists of a citation and a check for $500 drawn from a fund of $10,000 donated to the Institute by Vladimir K. Zworykin to encourage outstanding technical developments in fully electronic television.

Memorial Prize Award in Memory of Morris N. Liebmann

The Memorial Prize Award in Memory of Morris N. Liebmann has been given annually beginning in 1919. This award was originally based on a fund made available to the Institute for this purpose by E. J. Simon to preserve the memory of Colonel Morris N. Liebmann, a member of the Institute who gave his life in the first World War. The award now consists of a certificate and a check for $1000 given to a member of the Institute who has made an important contribution to the radio art.

Memorial Prize Award in Memory of Browder J. Thompson

The Memorial Prize Award in Memory of Browder J. Thompson, established in 1945, is given annually to the author or joint authors under 30 years of age for a paper of sound merit recently published in one of the technical publications of the IRE, which has been selected as constituting the best combination of technical contribution to radio and electronics and presentation of the subject.

The award comprises the income from a fund established by voluntary contributions to preserve the memory of Browder J. Thompson, a Director of the Institute who was killed in action in World War II while on a special mission for the Secretary of War.

Memorial Prize Award in Memory of Harry Diamond

The Memorial Prize Award in Memory of Harry Diamond, established in 1949, is given annually to a person or persons in Government service for outstanding contributions in the field of radio or electronics, as evidenced by publication in professional society journals such as the PROCEEDINGS OF THE IRE, or journals of similar standing. The award consists of a certificate provided for from proceeds of a fund established by friends who felt that the professional life of Harry Diamond, a Fellow of the Institute, exemplified the highest type of scientific effort in United States Government service.

Prize Award by W. R. G. Baker

The Prize Award by W. R. G. Baker, instituted in 1957, is given annually to the author of the best paper published in the TRANSACTIONS of the IRE Professional Groups. The award consists of a certificate together with a cash award comprising the income from a fund donated to the IRE by the late Dr. W. R. G. Baker, former Chairman of the IRE Professional Groups Committee.

Other Awards

The Editor of the PROCEEDINGS, Alfred N. Goldsmith, was awarded in 1926 an embossed and framed "Memorial to the Secretary of the Institute" in recognition of his service to science and engineering in the upbuilding of the Institute of Radio Engineers.

MEETINGS AND CONVENTIONS

National IRE Meetings

One of the principal purposes of the founders of the IRE and of the charter members was to hold meetings at which radio engineers could exchange information on questions of mutual interest. The first YEAR BOOK of the IRE, published in 1914, stated that "Institute meetings are held in New York monthly excepting in July and August of each year for the presentation and discussion of engineering papers. These papers are presented by members who have specialized in some division of the subject and who desire to lay their results before the profession."

The annual IRE meetings in New York through the year 1925 were considered to be the meetings required by the provisions of the IRE Constitution. After that date the election of officers was accomplished by mailing ballots to the members as specified or permitted by successive amendments of the Constitution without the necessity of holding meetings for election purposes. Beginning in 1926, therefore, the annual meetings were much more in the nature of National Conventions and had emphasis on technical papers rather than on the conduct of the business affairs of the Institute.

An important feature of the National Conventions of the past 15 years, and of some of the Regional Conventions, is the manufacturers' exhibit of equipment that is of interest to radio engineers.

In 1959 the Board of Directors approved a change in the name of the IRE Convention to "IRE International Convention and Radio Engineering Show." The attendance at the 1961 Convention in New York included 997 persons from countries outside the United States. This Convention continues to increase in importance each year and is internationally recognized as one of the largest conventions of its kind in the world.

Engineering Conferences and Symposia

For many years the IRE has cooperated with other scientific and engineering organizations as a joint sponsor, or in a similar capacity, in the conduct of engineering meetings or of extensive technical symposia.

In November, 1919, several technical papers were presented by IRE members at a joint meeting with the American Physical Society in Chicago.

A National Electronics Conference, held in Chicago in 1944, was attended by more than 2200 persons—more than had attended any IRE National Convention up to that time. This was the forerunner of many conferences sponsored jointly with other societies and universities.

The engineering conference or symposium has become one of the major activities of the IRE during the past few years, and the "Calendar of Coming Events" has become a regular feature of the PROCEEDINGS. This Calendar referred to some sixty such meetings during the year 1961.

COMMITTEES

One of the most effective means for collaboration among members of the IRE has been through the activities of the various IRE committees.

The three Standing Committees of the IRE, as stated in the YEAR BOOK for 1914, were the Papers Committee, Standardization Committee and Publicity Committee. By 1915 another Standing Committee was added—the Finance Committee.

The broadening scope of IRE interests over the years was accompanied by the establishment of additional committees dealing both with technical subjects and with operational matters.

The IRE committee structure now embraces 26 Technical Committees with 130 Subcommittees, along with 11 other Committees dealing largely with management and operational aspects of the Institute. A complete list of all Committees of the Institute and of their members is published semiannually in the issues of the PROCEEDINGS for June and October.

The effectiveness of the IRE throughout the fifty years of its existence has been due in no small measure to the enthusiastic collaboration of the many hundreds of committee members and to their recognition that there could be no satisfactory substitute for the contributions which this activity makes to the more rapid and effective development of the field of radio engineering.

COOPERATION WITH GOVERNMENTAL BODIES

Technical Aspects of Radio Regulation

The problems of radio regulation are such as to require, for their solution, a thorough consideration of the physical facts as to the behavior of radio waves, the external performance of radio transmitting equipment, and many of the operating characteristics of radio receiving apparatus.

When the Federal Radio Commission was first established by law in 1927, it had no engineering staff, and so it naturally looked to the engineers of industry for assistance on technical matters. This led to new activities of the IRE, those of studying and preparing reports on a series of problems put to it by the Commission and its Acting Chief Engineer.

In a guest editorial in the PROCEEDINGS OF THE IRE, December, 1956, the Chairman of the Federal Communications Commission commended the IRE for its outstanding contributions toward advancement of the art of radio communication and for the whole-hearted cooperation it has always given all agencies of the government concerned with utilization and administration of the radio spectrum.

In 1956, at the request of the FCC, the Television Allocations Study Organization (TASO) was formed by the television industry. Its object was "to develop full, detailed and reliable technical information and engineering principles based thereon, concerning present and potential VHF and UHF television service." The results of a two and one-half year study were described in a series of papers assembled by the IRE Professional Group on Broadcasting and published in a TASO Issue of the PROCEEDINGS in June, 1960.

Joint Technical Advisory Committee (JTAC)

In response to a request from the FCC in 1948 for advice on a number of technical radio problems with which the Commission was concerned, a group of eight distinguished radio engineers were asked to form themselves into the Joint Technical Advisory Committee (JTAC). This joint action was taken in June, 1948, by the Boards of Directors of the IRE and the Radio and Television Manufacturers Association (RTMA). The purpose of JTAC was to consult with Governmental agencies and other professional and industrial groups to determine what technical information was required to ensure the wise use and regulation of radio facilities, and to collect and disseminate such information. This Committee was in certain respects a successor to the Radio Technical Planning Board (RTPB), formed in 1943 with somewhat similar objectives.

JTAC has issued a series of technical reports on questions of public interest on which the FCC has sought information. JTAC has also presented technical testimony at meetings held under FCC auspices.

The series of reports issued by JTAC over the period of 14 years since its formation have dealt with such subjects as FM broadcasting, interference problems in TV, standards of good engineering practice, principles for the allocation of frequencies in the land mobile and other radio communication services, interference from arc welders, use of single sideband, frequency-diversity, space communication, etc.

This committee, under the joint sponsorship of the Electronic Industries Association (EIA) and the IRE, continues to be effective in its studies of the ever-increasing technical complexities of the frequency-allocation problem. In 1961 it formed an Ad Hoc Subcommittee to study the technical problems imposed by satellite relays, including the multiple use of frequency allocations.

COOPERATION WITH OTHER TECHNICAL ORGANIZATIONS

Early Activities

From the beginning, the Board of Directors of the IRE has recognized the close relationships between the field of activity of the Institute and the activities of other scientific and engineering organizations. The IRE has therefore endeavored to keep closely in touch with such organizations, and has appointed representatives to serve on many joint committees.

In 1927 the IRE and AIEE undertook joint sponsorship of the Sectional Committee on Radio under the procedure of the American Engineering Standards Committee (now called the American Standards Association). This Association serves to coordinate the interests of business, governmental and engineering organizations in industrial standardization.

In 1928 the American Engineering Standards Committee invited the IRE to become formally affiliated as a member body of that organization. Several years later, the IRE accepted sole sponsorship of the Sectional Committee on Radio.

National Television System Committee (NTSC)

A notable example of the effectiveness of the IRE in its cooperation with other technical organizations has been in connection with the work of the National Television System Committee (NTSC), formed in 1940 by several industrial organizations, which had a major interest in the sound development of the television field. This Committee presented to the FCC, in response to its request, coordinated recommendations as to the technical standards that should be recognized as the basis for the licensing of television broadcasting stations for both monochrome and color television service.

In recognition of its distinguished contribution to the television industry and to the development of color television, the NTSC received the "Emmy" Award of the Academy of Television Arts and Sciences in 1954.

American Institute of Electrical Engineers

The two professional societies, IRE and AIEE, have many fields of mutual interest. From time to time, the two societies have cooperatively conducted joint ventures, as exemplified by joint student branches, and joint symposia and meetings. In 1960 an AIEE-IRE Joint Standards Committee was formed to approve, jointly in the names of the two societies, such standards as it may consider suitable for joint promulgation and where it is desirable that a single standard be submitted to the ASA.

Effective January 1, 1961, it became possible for a member of one society, based solely on the qualifications submitted to that society for certain grades of membership, to join the other society at an equivalent grade of membership without payment of an entrance fee.

In October, 1961, action was taken by the Board of Directors of the IRE and by the Board of Directors of the AIEE appointing a joint committee to study the feasibility, practicability and form of a possible consolidation of IRE and AIEE. Any such consolidation would be subject to approval by the membership of the two organizations.

Education

As indicated earlier in this paper, the Board of Directors of the Institute, from the beginning, has endeavored to attract members of Student grade. Comprising 19,000 students in 500 schools all over the world, this is now the second largest grade of members in the IRE, accounting for about one-fifth of the total membership.

The Student membership dues have always been nominal. Wherever groups of fifteen or more students are associated, in a given institution, Student Branches may be formed. At the end of 1961 there were 217 such Branches of the IRE, many of them organized in cooperation with the AIEE. Each Student Branch has the benefit of guidance by an IRE Representative who teaches at the institution.

The Institute provides substantial subsidies to the Student Branches in the form of rebates and other aids. Each Student receives the PROCEEDINGS and the STUDENT QUARTERLY and, for the nominal additional fee of one dollar, may join and participate in the activities of a Professional Group of his choice, thus receiving the TRANSACTIONS of that Group. The Special Issues of the PROCEEDINGS are recognized as virtual textbooks on their subjects.

One of the Standing Committees of the IRE is the Education Committee, whose duty is to advise the Board of Directors on the relations of the Institute to electronic engineering education at all levels. The Professional Group on Education provides seminars and publications for the discussion of educational philosophy, problems and methods. Other Professional Groups have arranged special meetings for the benefit of students interested in their subjects. Each Section of the IRE has a Student Coordinator and many Section meetings have been arranged to meet the special desires of students, including high-school students. Each of the 8 IRE Regions has established a Regional Education Committee whose annual meetings are financed by the Institute.

A brochure entitled, "Electronics—Career for the Future," issued as a result of the work and energy of the Cedar Rapids, Iowa, Section of the IRE in 1960, was revised and will soon be reprinted through the cooperation of the EIA and distributed to 28,000 high schools.

The Board of Directors, in April, 1961, accepted an invitation extended by the Engineers' Council for Professional Development to become a participating body of that organization. In this way the IRE can assist in the promotion of scientific and engineering education through guidance and the formulation of criteria. It can also cooperate in the inspection and accreditation of engineering schools.

The 1961 IRE Outstanding Student Award was given to over 100 students who were judged the outstanding young engineers at their respective schools by the IRE Sections in their area. The winners are recognized for their professional competence, scholastic ability and depth of character.

In 1961 a Student Affairs Secretary was added to the headquarters staff of the Institute to coordinate and guide the IRE program of activities and services for its Student membership.

In these and other ways the Institute is endeavoring to encourage high-school graduates who are interested in scientific and engineering fields to learn more about the opportunities and responsibilities of the life of an engineer. Hopefully, this will contribute to an increase in the supply of competent engineering graduates with high ideals for their profession and for service to their community.

International Cooperation

The founders of the IRE recognized that the nature of radio transmission is such that radio waves recognize no national boundaries. They realized, too, that the problems in the radio field and the accomplishments of radio engineers are bound to be of interest to technical radio men throughout the world. It has thus been traditional with the IRE to foster international cooperation in radio, to encourage interest in the IRE, and to promote its services and the spread of its publications as widely as possible in all countries of the world.

As a result of action taken at the International Conference held in Washington, D.C., in 1927, there came into being the International Radio Consultative Committee (CCIR), an organization which, through the participation of representatives from all interested countries, has as its continuing task the preparation of advisory recommendations on technical questions that are involved in the allocation of frequencies to various radio services. Many members of the IRE and its technical committees have cooperated in the work of the CCIR and its Study Groups ever since.

In 1931 the United States National Committee, operating under the International Electrotechnical Commission, asked the IRE to cooperate with it in meeting some of the problems of international standardization of terminology. The IRE was very ready to lend its cooperation in this direction by contributing results of its own work on the standardization of terms and definitions.

In November, 1960, the IRE Board of Directors agreed to establish, together with the EIA, a joint Standards Advisory Committee to work with the US National Committee of the IEC to decide on representation at and participation in international meetings of the IEC.

Cooperation between the IRE and the International Union of Scientific Radio (URSI) has been active since 1926. By IRE representation on the USA National Committee of URSI there is a formal link. On repeated occasions, at International Assemblies and at meetings of the USA National Committee, members of the Institute have presented papers on radio transmission and other technical subjects. The IRE or one of its Professional Groups has also been joint sponsor of most of the meetings of the USA National Committee. The close association of IRE and URSI is emphasized by the publication in one issue of the Proceedings of a special series of reports covering the URSI 12th General Assembly held in Boulder, Colo., in 1957.

In October, 1959, the Board of Directors held its first meeting outside the United States. This was held in Toronto, Canada, in connection with the IRE Canadian Regional Convention and Exposition. The occasion was utilized to give special attention to the IRE position on international science and electronic questions.

The first symposium sponsored by the IRE outside North America was the International Symposium on Data Transmission held at the Delft Institute of Technology, The Netherlands, in September, 1960. It originated through the initiative of the Benelux Section of the IRE.

A number of IRE Professional Groups have taken occasion to recognize international interests or activities in their fields or have participated in meetings held by related international bodies outside North America.

The Board of Directors, in January, 1961, authorized the appointment of an Ad Hoc Committee on IRE International Activities Outside of Existing Regions. The five members of the Committee were all Fellows of the IRE; they were from three countries, and three of them were Past Presidents of IRE. In the early summer of 1961 they visited nine countries in Europe to explore with officials of major societies and leading engineers how IRE might better serve the professional interests of its many members who reside outside North America.

Constitution

The Constitution adopted by the original members of the Institute at a meeting on May 13, 1912, was a fairly simple document. It stated the objectives of the Institute and recognized three grades of membership—Honorary Members, Members, and Associates. The business affairs of the Institute were to be handled by a Board of Direction of nine persons, consisting of four Managers, President, Vice-President, Secretary, and Treasurer, all of whom were to be elected by the Institute membership, and an Editor of Publications who was to be elected by the other members of the Board.

Several amendments adopted between 1914 and 1946 made certain changes in the membership structure, method of voting, and the composition of the Board, including the provision for regional representation.

By 1946 the offices of Executive Secretary, Technical Secretary, and Technical Editor, as full-time headquarters staff positions, had been created by the Boards and it was clear that some of the details of Institute procedures which had previously been dealt with in the Constitution should more properly be handled by the Board of Directors itself or by members of the staff to whom the Board would give authorization. Constitutional amendments in 1946 and 1947 recognized these principles and thus added other powers to the Board and to the administrative officers of the Institute, effective August 15, 1947.

A constitutional amendment was adopted in 1959 in order to conform with certain changes in legal requirements. This amendment also facilitated operational activities of the Institute by stating the fundamental broad principles of IRE purpose and program, and putting responsibility for establishing Bylaws and for carrying out details of procedure in the hands of the Directors elected by the membership. This amendment also permitted the election of more than one Vice-President, in recognition of the need on the part of the President of the Institute for assistance in representing the IRE in its many public relations and government contacts.

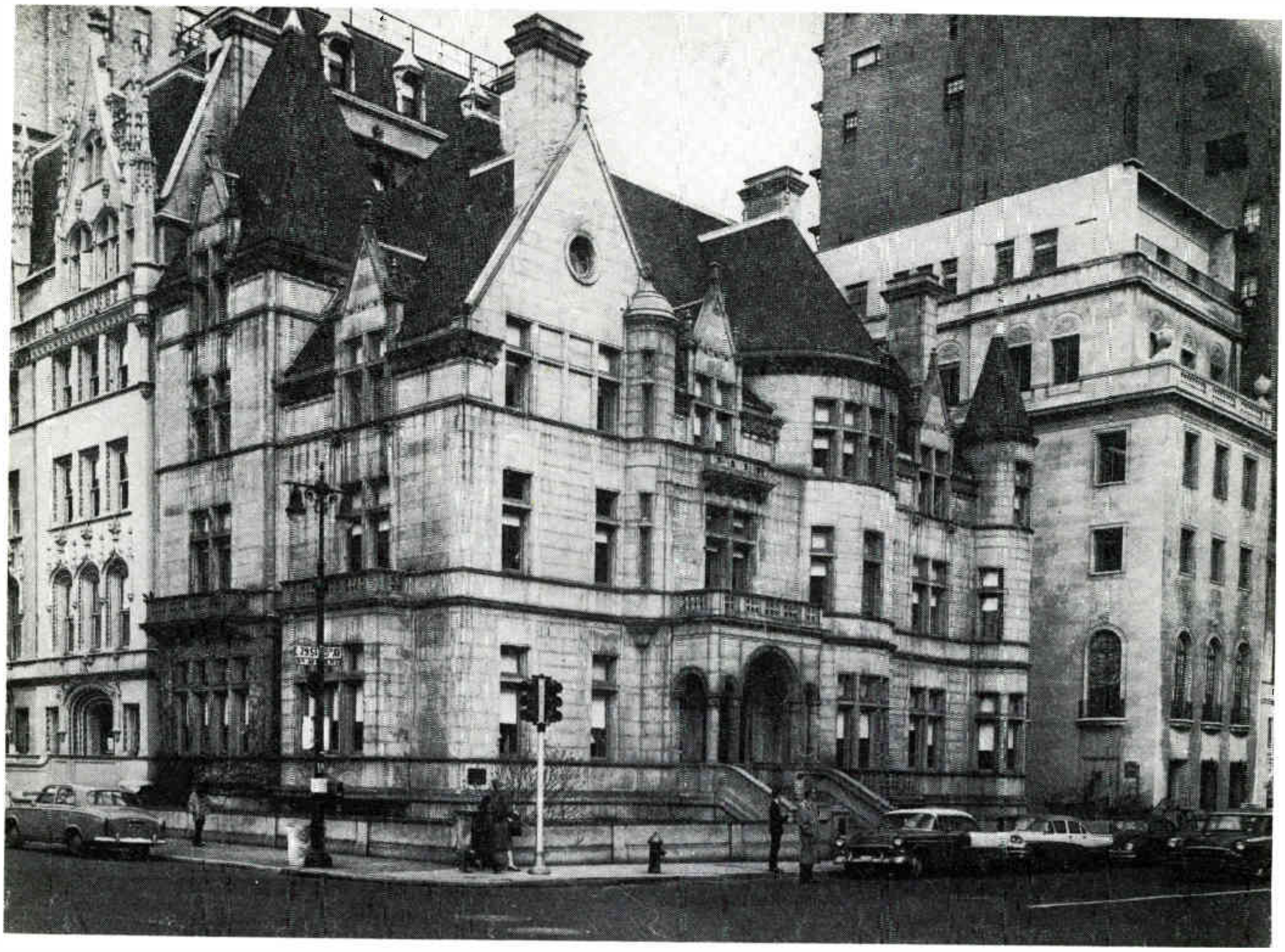
Office Locations

The first meeting of members of the Institute of Radio Engineers was held in Fayerweather Hall at Columbia University, New York. For several years most of the meetings for the presentation of papers were held at that location.

The IRE had no headquarters staff office for a number of years. Some business was conducted at an office used by the 1913–1914 Treasurer, John Hays Hammond, Jr., located at 71 Broadway, New York. A year or two later the IRE moved "uptown" to 111 Broadway in the shadow of Trinity Church. Subsequently, for about six years beginning in 1918, the business affairs of the IRE were conducted from the office of Alfred N. Goldsmith, who served both as Secretary and Editor, located at the College of the City of New York.

In the spring of 1924 the Board of Directors felt that the IRE could afford to rent an office and employ a full-time clerk. A small suite of rooms was leased at 37 West 39th Street, closely adjacent to the Engineering Societies Building in which the monthly meetings for the presentation of technical papers were then being held.

Thereafter, in order to meet the expanding requirements of the headquarters staff which was continually being enlarged to serve the publication, membership and other needs of the Institute, a succession of moves to various office locations took place. By 1946 it had been found practicable to obtain a suitable building at 1 East 79th Street (corner of Fifth Avenue) in New York City. This space was later supplemented by the purchase of the two adjoining buildings, one on 79th Street, and the other on Fifth Avenue, the latter being occupied in 1961. This Headquarters office thus serves as a base for the production of the publications of the IRE and for the cooperation of the IRE staff with Sec-



The IRE owns and fully occupies three mansions on the corner of Fifth Avenue and 79 Street in New York City. The corner building, 1 East 79 Street, was purchased in 1946. The building to the right at 5 East 79 Street was acquired in 1954, and the building to the left at 984 Fifth Avenue in 1961.

tions, Professional Groups, and Student Branches, and with other technical bodies with which so many joint meetings and symposia are arranged.

Office Management

During the first few years of the existence of the IRE, the Secretary or the Editor made available for IRE work the part-time services of one of the members of his regular office staff. This, together with the use of the individual services of the Secretary, the Editor, and others, kept the drain on the Institute's treasury at a minimum.

By 1924 the volume of correspondence necessitated by the large increase in the membership and the additional work involved in preparing material for publication in the Proceedings had reached a point where the Board of Directors was beginning to realize that more adequate provisions needed to be made for office space, clerical assistance, and supervisory attention.

This led to a succession of enlargements of the staff, and to the employment of a full-time Assistant Secretary who became Secretary in 1927.

Late in the year 1944, in order to broaden the activities of the IRE, including the expansion of the publication program, the Board of Directors decided to establish three new staff positions—Executive Secretary, Technical Editor, and Technical Secretary. This was a major step in recognizing the fact that the supervision and operation of the administration of the activities of the IRE could no longer be a part-time job for individual members of the Board of Directors. The Executive Secretary, whose services the Institute was fortunate enough to obtain, was George W. Bailey, who had had some years of administrative experience as an officer of the American Radio Relay League and whose ability in administrative service with the Government during World War II had been well demonstrated. To his continuing effective and constructive service the Institute owes a great debt of gratitude.

The administrative activities of the IRE, by the end of 1961, required the services of about 200 people whose work was being directed by the Executive Secretary, Managing Editor, Professional Groups Secretary, Chief Accountant, Assistant to the Executive Secretary, and the Office Manager. The Advertising Department, with its Advertising Manager and Assistant Advertising Manager, is situated elsewhere in New York City.

Advertising

The first advertising appeared in the Proceedings in 1915. During the year 1917, there were 48 pages of advertising published in the Proceedings—an average of eight pages (size 6×9 inches) per month. The Board was apparently somewhat reluctant to include advertising in the 1928 Year Book, but concluded that this should be done in an effort to obtain sufficient revenue from this source to cover the actual cost of printing. That Year Book contained 28 pages of advertising.

The Institute's office handled advertising matters until 1941 when the Board of Directors felt that substantially increased revenue might be obtained if an experienced advertising manager were to handle this part of the Institute's business. Accordingly, a contract was entered into with an advertising manager who provided his own office assistance at an outside location for this task.

Today the advertising carried by the Institute's publications renders a valuable service by bringing information about new products and services before the members of the IRE.

Finances

Throughout the lifetime of the IRE, the Board of Directors has planned and provided services to the membership with meticulous attention to the extent of the available funds. Accordingly, the publication program of the Institute at the outset was very modest—a small magazine published four times a year. The Year Book was small and was not published every year. Expenditures for meeting rooms and for Headquarters staff purposes were kept to a minimum.

The total receipts of the IRE for the year 1913 (including $150.97 carried over from the previous year) amounted to $850.93. This included $50.00 from the former Society of Wireless Telegraph Engineers and also $150.00 which was provided by or through the good offices of the Editor. The amount received in dues was $476.75. After payment of $647.25 for printing the Proceedings, $16.00 for janitor service in connection with meetings held at Fayerweather Hall at Columbia University, and other miscellaneous expenses, the balance at the end of the year 1913 was $24.43.

The report of the Treasurer in the 1916 Year Book referred to the fact that the Institute had "goods friends and loyal members . . . who have donated money and fixtures more than enough to start the new year with a clean slate."

Increasing income from membership dues and from advertising, together with conservative policies on the part of the Directors and careful management on the part of the staff, enabled the Institute to successfully weather the depression of the 1930's.

Over the years, the Institute's business operations have continued to grow. During the year 1961, the total budget of the IRE was slightly under $4,000,000.

Acknowledgment

In the preparation of this review of the activities of the Institute of Radio Engineers over the past 50 years, the author consulted various records which are available in the IRE office. Material which has appeared in past issues of the Proceedings of the IRE has been drawn upon freely. This includes news items about IRE or-

ganization and Section activities, editorials by Institute officials and others, addresses which have been given at Annual Conventions and on other occasions, and several articles of a historical nature which have appeared from time to time. A list of outstanding historical summaries, including a few other papers that relate to special situations in the Institute's history, is given in the Appendix.

Acknowledgment is due also a number of people, Founders, persons who have served as Secretary of the IRE, some of the past Presidents, and several others who were familiar with early activities of the IRE, as well as the members of the IRE History Committee and members of the IRE staff, from whom helpful information and suggestions have been received. The author gives special thanks for the advice received from E. K. Gannett, Managing Editor, for his cooperation in the endeavor to make this review as accurate as possible and for the preparation of charts, tables and illustrations.

Appendix

Previous Publications of Significance Relating to IRE History

"Silver Anniversary—The Institute of Radio Engineers," B. Dudley, *Electronics*, pp. 15–21; May, 1937.

"Institute of Radio Engineers to Act to Secure a Permanent Home," Proc. IRE, vol. 32, pp. 768–769; December, 1944.

"Symposium on the Occasion of the Completion of the Building-Fund Campaign," Proc. IRE, vol. 33, pp. 620–630; September, 1945.

"Radio Pioneers, 1945," Pamphlet Commemorating the Radio Pioneers Dinner on November 8, 1945, under the auspices of the New York Section of the Institute of Radio Engineers. H. P. Westman, Editor-in-Chief.

"1 East 79 Street—A Pictorial Tour of the Home of the Institute of Radio Engineers," Proc. IRE, vol. 36, pp. 89–100; January, 1948.

"What's Behind IRE?" Proc. IRE, vol. 39, pp. 340–341; April, 1951.

Series of papers on the occasion of the 40th Anniversary of the Institute of Radio Engineers, Proc. IRE, vol. 40, pp. 514–524; May, 1952.

"The Founders of the IRE"

"Life Begins at Forty," The Editor

"The Genesis of the IRE"

"The IRE in Cohesion or Depression?" D. B. Sinclair

"A Look at the Past Helps to Guess at the Future in Electronics," W. C. White

"The Institute of Radio Engineers," E. K. Gannett, *General Electric Rev.*, pp. 53–55; March, 1953.

"The Institute of Radio Engineers—Forty-Five Years of Service," L. E. Whittemore, Proc. IRE, vol. 45, pp. 597–635; May, 1957.

Note 1: A very informative summary of current activities of the IRE is given each year in the Annual Report of the Secretary of the Institute. The most recently published Report, that for the year 1960, is to be found in the Proceedings, June, 1961, pp. 1104-1107.

Note 2: There are in existence a number of complete sets of the Proceedings of the IRE, Volumes 1 to 49, inclusive, not only at the Office of the Institute and in the possession of several individual members of the Institute, but available for public reference at several libraries. Among such library reference sets are those at the Library of Congress, the New York City Public Library, and the Engineering Societies Library in New York City.

A "Report of the Secretary" Commentary

HARADEN PRATT
Secretary, IRE

THE NUMBER of IRE members, as the Institute has grown through the years, has directly reflected the degree of activities for the fields of endeavor that it has been serving. Commencing in the year 1912 as the successor of two small societies organized some five years earlier to provide a forum for those engaged in the new form of communication called "wireless telegraphy," its initial membership of forty-six rapidly increased as persons became aware of the new organization and its aims. The advance of what came to be called "radio communication" also being substantial, membership quickly rose to about one thousand within four years. Growth, then, for many years, kept pace with the expansion of this new accomplishment, showing an appreciable bulge when radio broadcasting suddenly came upon the scene during the years 1922 to 1930.

Until about the year 1940, the technology represented by IRE was almost exclusively that of radio-wave propagation through space for the communication of intelligence. Other electrical-engineering applications had not yet felt the need for the concepts and devices that were peculiar to radio communication except for minor uses of electron-tube circuitry for a few industrial purposes. But after 1940 the pressures for improved weapons and the defenses against them suddenly became so formidable that all branches of applied science were marshalled for every possible contribution. By this time, the fundamental issues underlying the phenomena embraced by radio communications became well enough understood so that new developments could be more easily undertaken. Fortuitously, many skilled people and the technical means for tooling their efforts were available at this critical time, and all sorts of applications of military usefulness came along immediately thereafter, radar being an example. As the war progressed it became more and more apparent that the use of the normal instrumentalities of warfare and indeed the planning of campaigns by the great military leaders could not be carried out without the assistance of such radio (or electronic as it came to be called) developments. Thus came about the transition from the former concept, that electrical communication was just a necessary adjunct to military operations, to the realization that all these new applications were not only indispensable to the control of campaigns but were the major items in the evolution of new weapons and tactical means necessary to maintain superiority. This great movement forward was reflected in a doubling of IRE membership by the year 1943 and a tripling by 1947.

We all realize what has occurred since the termination of that war. The myriad of new devices and the means for usefully developing the upward expansion of radio-spectrum utilization that had brought about radically new ways of conducting warfare were now available, together with experienced people informed about these subjects, for peacetime purposes. Helpful too, was the favorable environment created by the need for rehabilitating and restarting industries for the postwar years. The uses to which these new technical means could be put now started to permeate almost all applications of electricity and some others that formerly were not so served. It is not surprising, therefore, that IRE membership almost doubled every six years between 1943 and 1962.

Still more revealing of this growth and spread throughout almost all walks of life has been the evolution of the IRE Professional Groups system which, after a fourteen-year period of advancement, now embraces twenty-nine Groups, having a total membership of about ninety thousand. The twenty-nine names actually comprise a roster of the diverse specialized subjects and fields requiring the ministrations of this advanced technology. They constitute a testimonial of how the IRE has organized to serve the many trends of varied interest that emerged during those epoch-making years.

Attention must also be directed to the most important IRE function, the publication of technical and editorial material. Its quantity has doubled four times since 1943, an increase of 1600 per cent.

These growth statistics and the pattern of many diverse autonomous subsocieties operating under the common umbrella of IRE on a worldwide basis measure the dimensions of the role of the IRE as a basic and guiding influence. This role has become of significant importance to countless numbers of people. The benefits derived by IRE members are obvious. Numerous as these members are they number but a few compared to the many more that are concerned with the industries, laboratories, educational institutions and users of the end products, with whom IRE members constitute the connecting links to the professional operations of IRE. But still more important is the impact on all peoples inhabiting our world, because this segment of scientific endeavor and accomplishment penetrates almost every field and consequently directly affects the lives of everyone. Reflection on the absolute reliance of national defense upon this technology in this age of missile weapons emphasizes the fundamental character of its influence.

This role of the IRE has become captured by this initial rapid emergence of technical approach that signalizes the advent of a new explosive revolutionary age in human civilization. But this time it is a sweeping movement generated by the astonishing advances that have been made during recent years of a comprehension of the structure and behavior of matter and the understanding of energy fundamentals which, of course, in one way or another, touches every branch of scientific knowledge and engineering experience. The application of these discoveries has caused the bursting forth of new accomplishments, the most spectacular of which, in the public eye, probably is the control of atomic energy.

These events clearly indicate the importance of the IRE in the present structure of material advancement. A great responsibility resides with IRE members. They will, individually and collectively, make certain that the IRE role in these most important basic operations is amply executed. To accomplish this they will ensure the continuation of adequate technical leadership, the dissemination of knowledge, the fostering of improved educational processes at all levels, the inclusion of all qualified persons on the membership rolls, the promotion of further serious programs of cooperation with other fields (of which bio-medical electronics is an outstanding example), and other pertinent activities; otherwise mankind will be denied the rapid acquisition of many valuable benefits. The future role of the IRE will be an enlarged one and, as these influences affecting man's affairs continue to expand, more fields of effort and additional numbers of technically-qualified people will seek to participate in the rewards of membership and join the ranks of those that make up the cast on this world stage of accomplishment.

The Management of IRE

George W. Bailey
Executive Secretary, IRE

THE GOVERNING body of IRE has always been the Board of Directors. For years, the Board has confined its deliberations to matters of policy, leaving to an Executive Committee of nine Directors the responsibility of carrying out these policies. The Executive Committee is composed of the President, the two Vice Presidents, the two Past Presidents, the Secretary, the Treasurer, the Editor and one Director, thus insuring experience and continuity in its operations. The Executive Committee has employed an Executive Secretary and, according to the Bylaws, has fixed his tenure of office and determined his salary. The day-to-day problems of management have been left to the Executive Secretary, including the selection of his Staff.

The present Executive Secretary came to the IRE in January, 1945. He has been given freedom of operation under the policies established by the Board and the interpretation of those policies by the Executive Committee. He has selected a Staff of five experts in their lines. They have been encouraged to run their own departments, referring only major problems to the Executive Secretary. In turn, the Staff members have selected competent assistants and supervisors who have become specialists in their assigned tasks.

He inherited William C. Copp & Associates, advertising specialists, which has proven to be an invaluable legacy. Also, Will has functioned as Show Manager, under the supervision of the Executive Secretary, who is Chairman of the International Convention Committee. In order of seniority, Will and his Assistant, Lillian Petranek, outdate us all, having served the IRE for twenty years.

E. K. (Woody) Gannett became Assistant to the Executive Secretary in August 1946, and, in 1954, the Managing Editor. That year, PROCEEDINGS carried 1884 pages. Last year, that number had increased to 2234 pages. In addition, he has been assigned the broad general supervision of the publications of 29 Professional Groups amounting to 8336 pages last year.

L. G. (Larry) Cumming became Technical Secretary in 1946. In 1960, his title was changed to Professional Groups Secretary. He was first concerned with the administrative duties of 25 Technical Committees and comparatively few meetings in New York and in the field. Today, to his duties have been added the coordination of 29 Professional Groups, besides attendance at innumerable meetings throughout the country.

In July 1947, Emily Sirjane joined the IRE as a secretary. In February, 1948, she was appointed Office Manager. She has solved the problem of office service to a membership which has increased sevenfold. The office is now equipped with the most modern machinery. Microfilm and computer machinery are being installed, capable of handling double our present membership.

John B. Buckley became Chief Accountant in 1948, when the total expense of operation for the year was half a million dollars, compared to last year's four million dollars. He is installing an Electronic Data Processing System heretofore available only to banks. The adoption of this system will make possible fast and accurate handling of accounts far in excess of a greatly expanded membership.

In June 1960, Joan Kearney became Assistant to the Executive Secretary. Among her duties is that of Recording Secretary at the meetings of the Board of Directors and Executive Committee.

In September 1961, W. Reed Crone was appointed Student Affairs Secretary, reporting to the Managing Editor. His duties include visiting Student Branches and assisting in the publication of the STUDENT QUARTERLY.

Thus, three of the Staff have served for fifteen years and one for fourteen. The IRE Ten Year Club now has 32 members, of whom 12 have served for fifteen or more years. This continuity of service to the members of the IRE has had a great deal to do with the smooth operation of Headquarters throughout the period that, since 1946, has seen the membership grow from 13,000 to 96,000 members, and the Headquarters force from 32 to 220.

The accommodations for Headquarters have grown from a few small rooms downtown to three buildings on the Northeast corner of Fifth Avenue and Seventy-Ninth Street. These buildings were residences owned by the Irving Brokaw family, and were among the finest in the city. All three buildings were bought for cash from donated funds and accumulated earnings, and are owned free and clear. Meanwhile, the progress of the City of New York has continued northward, and the IRE property has quadrupled in value. The first floor in No. 1 and the first two floors in No. 5 East Seventy-Ninth Street have been maintained in the style of the Gay Nineties, when they were built, and large and small rooms have been provided for member Committee meetings. The three buildings provide 45,000 square feet of office space. Two and a half floors are available for expansion.

Emphasis has been put on the duties of the paid Staff in the operation of the IRE. Equally important are the thousands of IRE members who serve voluntarily as Officers, Directors, Committee Members, Section and Professional Group Officers, and Faculty Representatives. The time and effort which they contribute to the activities of the IRE have made it the largest professional engineering society in the world.

IRE—The First 50 Years

L. V. Berkner

President, IRE, in the 49th Year

As WE COMPLETE the forty-ninth year of the Institute and enter into our Golden Anniversary Year of 1962, we notice that the membership of the Institute is rising by about 1000 per month. This is a most unusual situation for a professional society. Therefore, some critical analyses of the forces underlying the extraordinary growth and development of our profession seems warranted. Indeed no other area of engineering interest has ever enjoyed or been faced with a more rapid expansion.

While it is difficult to visualize our situation exactly in the flow of history for want of time perspective, we must recognize that man has just crossed the threshold of a quite new technological revolution whose momentum is beginning to gain at a striking rate. This revolution in technology is not a mere extension of the old industrial revolution, but rather an entirely new social phenomenon of a quite new order of power. Its beginning can be dated from the end of World War II or perhaps better, the mid-twentieth century.

Looking back a decade we now begin to perceive the origin of the forces that are producing this new revolution in technology. At the turn of the century our engineering strength rested primarily upon an empirical base. It was the technology that had been developed over the centuries through the experience of the Babylonians, Egyptions, Greeks and Romans—a technology that had been mathematized by the mechanics and physics of Galileo, Kepler and Newton; the molecular science of Boyle, Dalton, Priestley and Lavoisier; the thermodynamics of Mayer and Helmholtz; the electro-magnetic science of Franklin, Faraday, Henry, Maxwell and Hertz, and all their contemporaries and successors. Yet 50 years ago, with all its elaboration of the physical basis of our environment, science had not yet the insight into the fundamental processes of nature required to produce a really virile and penetrating technology. The great discoveries of mechanics, thermodynamics, chemistry, electricity and radio of the nineteenth century simply put into scientific order the more obvious features of man's environment—features with which he had dealt in one way or another since the beginning. Perhaps the outstanding exception of the nineteenth century was the development of radio—a phenomenon imperceptible to man's senses and derived purely from theoretical constructs. Thus radio was the forerunner of what was to come.

The twentieth century has seen the creation of a radically new and more penetrating science that has superseded our nineteenth-century physics. This new science now gives us the power to derive a radically new technology in almost every area of human endeavor. Even before the turn of the century, men like Michelson and Planck were raising serious questions about the classical physics. The discoveries of Becquerel and Roentgen produced consequences difficult of interpretation.

Then in 1905 came Einstein's elaboration of the relativistic mechanics and the interchangeability of mass and energy. Simultaneously came the formulation of the quantum ideas from thermodynamic and from photoelectric phenomena which showed energy to be divided into discrete packages. The nucleus of Rutherford and the atom of Bohr laid the basis for a clear understanding of atomic and nuclear phenomena. Soon we were to learn the limits of cause and effect in phenomena of atomic scale with Heisenberg's formulation of the principle of uncertainty. Einstein generalized relativity to include relatively accelerating masses which defined the place of gravity in field theory. With Davisson's discovery of the diffraction of electrons, deBroglie in 1926 formulated the wave mechanics, which was elaborated by Schroedinger and Dirac. Chadwick in 1932 discovered the neutron to complete our catalog of stable nuclear components, so the isotopes of the periodic table came in clear perspective. In 1938 Hahn and Strassmann discovered atomic fission and with it experimental interchange of mass and energy. In 1948 Shannon elaborated information theory and gave us the logical basis for data systems whereby abstruse theoretical ideas yielded to quick solution by means of electronic aids and could be made readily accessible to the engineer on a daily basis.

Thus in a period of some 43 years the whole content of physics has changed. We now think of matter in terms of Fermi-statistics and of space in terms of Bose-statistics defining the fields that permeate it. Our old physics has been absorbed as special and limited situations in a far more profound insight into natural processes.

And so at the midcentury we find emerging an altogether new technology that is founded on powerful theoretical grounds. This is a technology based on deep insight into nature. It is a technology with far greater power and efficiency than its predecessor. Therefore, since the midcentury, we witness the old technology of the ages fall step by step before the onslaught of the powerful new theoretical ideas. Certainly the new means of data handling foretell revolutionary processes in industrial organization, both nationally and world-wide. This new technology affects not merely communication and electronics, but bids to replace our sources of power with more efficient generation processes and our source of materials with radically new combinations of nature's elements in far more useful forms. In this new technology the whole concept of structure is destined to evolve with extraordinary rapidity in the decades ahead.

As a consequence, this technology has profound implications in engineering education. The access to this new and abstruse technology can be gained only by men with scholarly training and with dedication to deep study and grasp of nature's processes. We find suddenly an altogether new order of demand for men of the doctor's level of training and beyond to manipulate this comprehension of nature for our technological welfare. The traditional forms of engineering are bewildered and confused by the disappearance of time-honored procedures before the avalanche of new ideas with a new order of efficiency, function and flexibility.

These, I believe, are the fundamental forces that are creating this extraordinary progress of our Institute. Their recent origin is illustrated by our rapid growth in membership from a mere 11,000 in 1945 to nearly 100,000 today. The leaders of our Institute have recognized from the beginning the fundamental issues underlying the revolution in technology and have worked to mold our Institute to meet the challenge through the development of a variety of professional measures. No other engineering organization has yet recognized this radical change in the direction of technology as fully as has the IRE. So IRE has worked indefatigably to command the ideas of the new science and to mold them into the engineering of the future.

Thus, through the wisdom of our leaders, our Institute has found its place in the new technology. It has been at the right place at the right time and with the right ideas. It has led and not followed the parade. This is why, on the occasion of our Golden Anniversary, we find the Institute of Radio Engineers the largest and most virile of the professional technical societies.

IRE—The Decades Ahead

P. E. Haggerty
President, IRE, in the 50th Year

The IRE of a few decades hence will be a great deal more complex society than it now is. It will be more complex because the profession it serves will be more complex. The incredibly rapid rate at which knowledge has been increasing in those aspects of science and engineering fundamental to our profession and the matching complexity of the tasks we already know we must face insure that individual scientists and engineers will continue to specialize to an extent limited only by their ability to add to, comprehend and apply their individual specialties.

Most of us who are members of the IRE spend our lives in applied science and engineering. As such, our individual accomplishments can be measured only in terms of the specific products and services contributed to the larger community of which we are all a part.

It seems to me that the principal measure of IRE's success in future decades will be how well it succeeds through its organization, its publications and its technical meetings in tying together this increasingly larger, diverse family of specialists, one with another, as well as with the technical generalists so that our complex profession will contribute products and services at a high level of effectiveness.

Certainly IRE's professional group system provides a flexibility and a viability which if properly nurtured can furnish a vehicle for every proper specialty with the assurance that through the sectional, regional and international organization of IRE, all can be one organic whole. IRE, more than any other professional society, has a unique opportunity to provide the essential interdisciplinary hooks we must have if the applied science and engineering of the second half of the twentieth century are to meet the technical challenges we see ahead. IRE is already a society, not just of radio, electronic and electrical engineers, but of physicists, physical chemists, metallurgists, mathematicians, and a dozen other disciplines as well. Molecular electronics, for example, demands the closest possible relationships among physicists, physical chemists, device engineers, mechanization engineers, process engineers, and circuit engineers. The now barely but clearly identifiable beginning of the intrusion of much these same disciplines into and blending with the life sciences promises even more exciting, but also more complex, relationships and accomplishments.

If from all this complexity useful and important products and services are to proceed at anything like a high rate of effectiveness a strong, unifying force must be present, and a conscious recognition of the need for this unifying force among the members of our profession is fundamental to the present consideration of consolidating IRE and AIEE. If the members of our two societies come to believe that consolidation can strengthen this unifying force, then I feel sure we will begin 1963 as part of a stronger, consolidated society better able to meet the complex needs of our profession in the decades ahead.

Communications and

A Predictive Symposium by Fellows of the IRE

Introduction

A group of Fellows of the Institute of Radio Engineers was invited to undertake the hazardous task of transporting themselves forward in thought some 50 years and then presenting the accomplishments of communications and electronics during the thus scanned era 1962-2012. Within reason, no physical or engineering limits were to be placed on their vision or imagination. The narratives or prophecies resulting from this enterprise are presented in the following assembly of replies from the scientists and engineers participating in this Symposium. It will be interesting for the IRE membership in 2012 to compare these accounts with then existent facts and to decide the measure of foresight of those who came before them.

In each of the following presentations, the name of its author is given. And where the same author has discussed several diverse fields, which discussions appear in different parts of this Symposium, his authorship is accordingly multiply indicated.

An attempt has been made to arrange the contributions according to broad classifications of subject matter, as suggested by the headings which appear in the accompanying contents. The reader is warned, however, that many of the discussions cover a wide range of topics. As a consequence, papers listed under one category may touch upon other categories as well. For further convenience, therefore, a subject index is provided at the end of the Symposium (655). An author index is also included.

In considering the events of 50 years hence, it is interesting to conjecture what percentage of the present IRE membership will witness them. A recent survey[1] of one segment of IRE members indicated that 84 per cent were between 21 and 40 years of age. In view of the anticipated advances in medicine (some of which are outlined in this Symposium), it would appear that the number who survive to 2012 A.D. may be quite substantial.

Whether or not he is so fortunate, each member will find a unique adventure awaiting him in the following pages, from the opening address of the IRE President in 2012 to the closing description of the impact of electronics on our future mode of living.

—The Anniversary Editor

[1] K. W. Uncapher, "1960 PGEC membership report," IRE Trans. on Electronic Computers, vol. EC-10, pp. 81-90; March, 1961.

Electronics— 2012 A.D.

SOLID-STATE ELECTRONICS

INFORMATION AND MACHINES

ELECTRONICS AND SCIENCE

ELECTRONICS AND BIO-MEDICINE

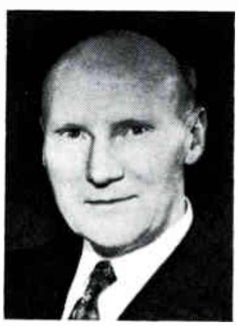
ELECTRONICS AND SOCIETY

World Travel and Space Communication—2012 A.D.

A Realistic Fantasy

Ernst Weber Fellow IRE

(The following are excerpts from an address delivered by the President of the Institute of Radio Engineers before the International Convention in Geneva, 2012 A.D.)



Fellow Members of the IRE:

As we gather in this beautiful city to commemorate the founding of our society one hundred years ago, we are conscious of the imminent return of the space ship "Viking" which left earth on October 21, 2007 on a mission to explore the outer planets of our solar system. This event is particularly significant to us because of the many contributions that were made by members of our society from all over this globe. Permit me to recall some of the significant developments which have preceded the concept and led to the launching of the ship, reminiscing for some and going back into history for other phases. If it should appear to you somewhat rambling—forgive me, but the chain of events, meaningful as it might appear in broad perspective, has, in fact, seemingly incomprehensible sideloops.

I presume there are very few among us—and I certainly am not one of them—who can recall the excitement of the first venture into space, the orbiting of the so-called "Sputnik I" on October 21, 1957. As recorded in history, keen competition had developed between the then two most powerful nations on this globe, Russia and the United States.

Each of these nations had initiated ambitious programs of space conquest—a phrase that has not lost its attractiveness. Each of these nations made firsts in different ways. . . .

Unfortunately, the rapid progress that had been made in the two decades following 1957 could not be exploited. . . .

Shortly after the turn of the century, when sanity had returned once more, the scientists and engineers, represented through their respective National Academies, took the leadership in re-establishing international scientific and technological exchanges. I am proud, indeed, to acknowledge the active participation of our large international membership in bringing together the "World Conference for the Exploration of the Solar System," which actually met in this city. . . .

The principal objective of the conference was the formation of an international committee for the design and launching of a well-instrumented space ship that could travel to the farthermost planet and bring back rather complete data on as many of the outer solar planets as would prove feasible. The resolution of the National Academies required, of course, the concurrence of the political governments and in particular, the permission to pool all the national capabilities into a singular and unique endeavour reminiscent of the so-called "International Geophysical Years" in the last century, yet of unprecedented vaster scope. The only reason that it had seemed propitious to formulate such a daring resolution was the fact that all nations appeared ready to demon-

E. Weber is the President of the Polytechnic Institute of Brooklyn, Brooklyn, N.Y. *(Received June 8, 1961.)*

strate once again the solidarity of the human race in the face of potential annihilation.

Upon acceptance by the various political systems, an "International Commission for Solar System Exploration" was set up with headquarters here in Geneva, the time-honored focus of international organizations. The Commission established national committees, each charged with organizing intensive programs of research and development in the areas in which each felt most capable.

It had been unanimously concluded that the time was ripe to construct a space ship that could accommodate ten scientists—volunteers who would observe, judge and relate the observations, and direct the data collection—and six technicians and auxiliary personnel for all necessary services. Nuclear propulsion had been sufficiently tested to make feasible a design for a power plant to support space travel for several years, completely self-contained and controlled from within the space ship. The launching required a rocket system which transcended existing capabilities by an order of magnitude but one that appeared feasible. The site for the launching platform was debated extensively, but the choice fell upon the astroport of the Sahara because of the major test facilities that had been created there and the existence of a stadium from which all major events could be observed.

To secure life over a maximum period of six years required extrapolation of all the equipment familiar from the latest nuclear submarine experience. Because of unknown intensities of particle and gamma radiation in many regions to be traversed, special shields had to be designed that could be removed for visual observations when anticipatory radiation counters assured absence of danger. Of particular importance were means of detecting asteroids and small meteorites and their paths of travel over large distances in order to avoid collision with the space ship or damage to it that could be calamitous.

A source of tremendous satisfaction to us must be the many contributions which the IRE could render because of its vast and resourceful membership in literally all the nations on this globe. The IRE also has the distinction of having published one of the most competent early collections of articles dealing with "Space Electronics" in its PROCEEDINGS issue of April, 1960, followed by its fiftieth anniversary volume of May, 1962. It is remarkable how most of the then indicated developments have been carried through, making really feasible the achievement of human space exploration. The amazing evolution of Synnoetics as the intermarriage of man's creative mind and the nearly instant execution of the supporting logic processes by adaptive and self correcting computer systems;[1] the functional elements embodied in specific compound crystals of miniature size responsive to stimuli from light, sound, and heat sources; the evaluative integration of detection systems of high redundancy useful for monitoring and warning assignments; and the increased precision of practical time interval measurements to about 10^{-14} seconds; all these achievements have made relatively safe what had been considered a foolish risk. The systems of navigation, telemetry of data, and communications were tied into the global satellite system and the relay station in the moon. Five different frequencies in the broad microwave region were assigned and were protected in an absolute sense by international agreement.

It also was agreed to lay out a course for the entire travel as a guide for navigation, but to leave freedom to the crew to make changes when it appeared advisable. From the positions of the outer planets, starting with Jupiter, the course could cover near approaches to Saturn, Uranus and Neptune, with some doubt about Pluto because of the yet uncertain higher speeds that would be required and might be permissible with greater distances from the Sun. At all times, however, communication on one of the frequencies was imperative to keep the earth stations informed of progress and to relay as many data back to earth as possible, because with all precautions, no guarantee could be had that the space ship would return safely.

The global satellite communication system of the active type, which had suffered . . . was in the process of re-establishment, so that good assurance of reception and, in fact, of two-way communication could be given. Two satellite astronomical observatories were in high orbit around the earth and one was directed to follow the space ship by periodic observation. To assist in the navigation, all positional data from all sources were to be processed and verified both in the space ship and on earth.

The earlier experiences with the simpler designs for the trip around Mars and Venus had taught us that redundancy in mutual communication was a great factor in relieving anxiety on a trip of such isolated existence. Correct positional information also will permit the proper approach for re-entry. . . .

The time table had projected preparations and test procedures to take the better part of five years so that launching was set, as you remember, for October 21, 2007, with the expected return by about July, 2012. Recall our relief when the launching went off without mishap. One had, after all, learned to live with full publicity and in full view of the entire global population through the color television system that is spanning the globe. Every step had been pretested, every conceivable means had been taken to assure reliability, and the result has been, indeed, gratifying. We have never lost contact nor communication with the crew which ventured out into the limitless outer space as the Vikings, more than a thousand years ago, had ventured out into a "limitless" sea. Much of the data has been invaluable, but not all has been processed yet. . . .

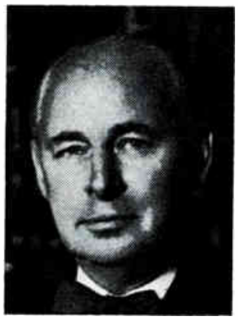
There are only 10 hours left until the space ship will descend upon the astroport in the Sahara. I therefore conclude. . . .

[1] L. Fein, "The computer-related services (Synnoetics) at a university in the year 1975," *Am. Sci.*, vol. 49, pp. 149-168; June, 1961.

Space—2012

LLOYD V. BERKNER FELLOW IRE

On the occasion of this centenary of the founding of the IRE, the Editors of PROCEEDINGS *thought it of interest to our membership to report the May 13, 2012, meeting of the Lunar Section of the Institute. The following summary lecture delivered to this Section at our great international base near the north pole of the Moon, as reported by a one-time President of the IRE, describes the status of our space activities so adequately that no further introduction seems needed.*



Members of the Lunar Section:

We are happy to welcome this evening the members of the Institute who have so recently arrived from Earth to augment our ranks. On this occasion, our chairman has asked me to describe our extensive electronic facilities. During our discussion, please forgive me if I digress in describing some of the related facilities.

The basic objective of this base is four-fold

a) Our NOPOLUNAR base is one of the two major lunar bases for communication and transportation to the Earth.

b) This base, likewise, is a major starting point for planetary transportation and the terminal of our Martian and Venetian LASER communication and navigation links. Likewise, many very large circumferential spacecraft are equipped from our base for solar and nether-planetary studies to take advantage of our low escape velocity of a mere 2500 meters per second, and the relatively low orbital velocities that simplify transportation from the surface.

c) From this base, most of the maintenance of our eight advanced lunar stations is centered; lunar scientific exploration is directed, and many local expeditions operate from our facilities.

d) We provide operation of the largest astronomical and radio-physical observatory in the solar system.

L. V. Berkner is President, Graduate Research Center of the Southwest, Dallas, Tex., and Chairman, Space Science Board, National Academy of Science. *Received February 26, 1962.)*

It is hardly necessary to remind you that a similar base, SOPOLUNAR, is located near the opposite pole, with which we are in constant communication. These two bases are the primary universal bases for the whole range of man's activities in the solar system. I take this occasion also to welcome the engineers from SOPOLUNAR who have come over to our meeting this evening.

The basic communication and navigation system utilizes coherent radiation in the wavelength ranges of 1800 to 30,000 angstroms which lend themselves to formation of highly focused beams with very light-weight and miniature radiating systems. For example, our main transmitter with an antenna of 2 decimeters diameter can focus the entire radiation of 1-watt peak-power on an area 500 meters in diameter on Earth. This provides very high flux densities at the terrestrial receiver, permitting almost unlimited information transmission of digitalized data at very high speeds. Likewise, our 2-meter antenna with 100-watt peak focuses its entire energy on Mars or Venus with very little loss and permits simple high-speed data links with interplanetary travelers. You will take part in construction of the new Jovian system designed to provide communication, navigation and control, for the expedition to land on the minor satellite of Jupiter next year.

To overcome the line-of-sight problem from our station to the Earth which is not always visible from this location, to communicate easily with other lunar stations and interplanetary spacecraft, and to provide a broad base-line for navigation, we have established two space buoys at the "centers of libration" of Earth and Moon. These are positions of permanent equilibrium in the Earth-Moon gravitational field that provide navigational base lines of 385, 385, and 665-thousand kilometers, respectively. With their stabilized decimeter antennas they also provide excellent low-power, high-capacity relay-links to Earth and to other spacecraft. Operation of coherent radiation links in this way circumvents the cloud problems which limit use of coherent radiation from the Earth, except through use of X band with intervening orbiting relay stations.

I remind you of the present travel arrangements between Earth and Moon. These are entirely controlled by automatic data systems using the high-speed real-time computer of the type that I hold here in my hand. On leaving Earth, via the local connection, you joined the Earth-orbit station, orbiting 800 km above the Earth, in about six hours travel time. This interception was entirely automatic, directed from the Earth, and completed under local control of the Earth-orbit station. There you transferred to the nuclear powered intercelestial spacecraft.

Your travel time to the Moon was 36 hours, representing optimum thrust, where you coupled to the circumlunar orbiter, 50 km above the surface, and then transferred to the local lunar landing craft that brought you to this station. This succession of events has reduced travel costs per kilogram by a factor of 100 over early attempts using direct surface to surface operations.

The historical development of our lunar facilities is one of great trial and tribulation. The basic problem was similar to any frontier operation, of founding the basic facilities. There are abundant local resources of oxides for oxygen and water of crystallization for water; of structural materials and of shielding materials required during occasional solar eruptions; and of chemicals for fertilizing our hydroponic farms and for a wide variety of other applications. To utilize these requires only a vast amount of electric power from our three great thermo-nuclear reactors. The sun gives us light in any spectral range through reflectors and filters of appropriate design. Our large chemical plants operate entirely under electronic control and regulation.

The basic elements of our complex environment were originally transported from Earth to initiate our complicated community. The original settlers faced the most excruciating difficulties, and there was a considerable loss of life among very courageous men. But as facilities were established, our base supplied more and more internal capability, permitting sustenance of more bodies who could develop our basic facilities. Thus our developmental growth has been exponential with time, yielding substantial living stability, especially since opening the base to women at the turn of the 21st century. Our requirements from Earth during the last decade have been primarily for manpower and for our capital development of interplanetary travel and all related to it.

The great advantage of the Moon is its situation as a stable satellite with a weak but perceptible gravitational field, allowing great individual freedom, and without an appreciable atmosphere. This provides unparalleled opportunity for initiation and control of, and communication with, interplanetary activities. Tomorrow you will see the giant astronomical observatory employing the mighty 1500-inch reflector that easily penetrates right around the Universe, and completely resolves galaxies many parsecs away. In the low gravitational field, structural design permits about six times the size for equivalent terrestrial deflections. Without scattering or scintillations, with spectral ranges of several octaves, and with special gamma-ray and x-ray telescopes as well as infrared detectors, the entire useful physical spectrum is always accessible. The data from these instruments are analyzed by advanced electronic techniques. With this new view of the universe our whole physics has been rebuilt in the past ten years—a knowledge whose capital value to man is beyond estimate.

I have just received a Laser report from the party that landed on Mars last month. The climate is a little cooler than Earth, with a good, protective nitrogen atmosphere. They find surprisingly well-developed plant life in extraordinary forms. More interesting is their preliminary finding of some evidence of an ancient civilization, perhaps a billion years ago when free water and oxygen were evidently available in considerable quantities on the planet. Seemingly the present organisms have evolved to use oxygen and water from the rocks, as these basic needs disappeared in free form from the planet. Perhaps in part these plants have become efficient in deriving oxygen and water vapor from the atmosphere where these basic elements are still found in miniscule quantities. I know you all realize that this analysis of a completely foreign biology is adding a new dimension to our biological knowledge, and for the first time the physical forces involved in cell organization, syntheses, and regeneration are coming into focus. Our new electronic methods for biological analyses designed for planetary research are also having a profound effect on terrestrial biological studies in laboratories on the Earth.

I hope you will enjoy your tour as much as have we. The scenery is most unusual and delightful, and there are still many discoveries to be made by the observant and the persistent scientist. Our location near the pole gives us a moderate and easily controlled climate. In constant twilight, the light and heat over the whole site is provided by filtered reflectors erected on the nearby hills.

There are certain precautions that must become habit. Just as you look around on Earth before stepping into the street, you must be equally careful about direct exposure to sun or shade without proper protection. Safe areas are clearly marked. Our warning system advises of approaching solar-particle streams. Be careful about walking too springily or with your Earth-adapted muscles you may bump your heads on the ceiling 30 feet overhead. Please familiarize yourselves with the station adaptation manual which is the synthesis of our 28 years experience. For reasons of natural rhythm, our lunar timetable and calendar corresponds to that on Earth. You will be happy to know that in this lower gravitional field, aging seems to take place at a rate lower than the terrestrial rate. This phenomenon is exciting great interest and much scientific study.

For your entertainment, I have a history of the Institute for the first one hundred years since the founding of IRE. You will be especially amused by the skepticism of our forebears, a few of whom, by the way, are still living, which was expressed at their 50th anniversary in 1962 when lunar operations on today's scale then appeared wholly fantastic.

Space Exploration

J. H. DELLINGER FELLOW IRE

It is hard for us engineers in 1962 to get a clear focus on the truth as it will be in 2012, particularly in respect to space exploration. Just as the earth's atmosphere distorts the focus of a telescopic image, so our space imagining is prone to distortion by the hysterical atmosphere of the present age when every youngster expects to be promised an excursion ticket to Aldebaran.

It being U.S. policy to put a man on the moon in the present decade, and also to start toward Venus and Mars, will there be men on the moon or planets in 2012? It seems quite unlikely that there will then be men in residence on those inhospitable spheres. An occasional visitor to service the scientific and engineering equipment there, yes. But the exploration of the moon by men moving about there will have long since finished. The reason? Painful experience will have shown that fantastically more information is obtainable per dollar of expenditure by means of equipment (observing and telemetering) than by the presence of men.

The moon in particular will be the prized platform for the location of scientific observing equipment. There will be a whole new surge forward in knowledge of the external universe when telescopes and spectroscopes and other sensors and radio equipment are in place and functioning on the moon, free from the obscuring effects (absorption and scintillations) of the earth's atmosphere.

One reason why the surge forward will be of incomparable value and scope is that we shall doubtless be observing all parts of the whole observable universe. The red shift of the radiations from the galaxies, increasing with distance, sets a limit to the distance we can ever observe, and we have already reached out seven-tenths of the way to that limit. The filling in of knowledge about the more remote depths of observable space should put us well on the way to determination of whether the universe had a beginning or is being continuously created (or, perhaps more precisely stated, whether the observable universe had a sudden beginning or exists by a continuous process of the interconversion of matter and radiation).

Another possible and vital consequence of observations ranging throughout the entire observable universe might be the accumulation of data revealing the nature of gravitation. This could permit the formulation of the unified field theory which science has thus far sought in vain. It would be a step in the understanding of the universe comparable with, possibly surpassing, the giant advances of Maxwell's theory of electromagnetic waves and Einstein's theory of relativity.

Ability to receive electromagnetic radiations over the whole spectrum, in contrast to the narrow spectrum ranges receivable through the earth's atmosphere, will hopefully have another grandiose consequence. Probably intelligent beings on the planets of many stars are sending out signals with the idea of contacting life on other worlds. I have even heard them credited with the idea of trying to contact us (of the planet Earth). No, although intelligent, they have never heard of us. We think we know what frequency ranges they would be likely to use. But we could be wrong; and if we probe the entire spectrum we might learn something on the moon that we could not on earth.

There will presumably be vehicles orbiting close to each of several planets, perhaps all of them. They will continuously send us data on many physical parameters of the planets. Possibly there will be instrument stations on Mars and Venus, somewhat like those on the moon.

2012 is not likely to see vehicular exploration of space as far away as the stars. Men are never likely to go out among the stars in space ships. Other creatures, not men, may do that. Radio astronomy will bring us so much knowledge of the universe that we shall have little incentive to travel to the stars even if it were possible.

Finally, space exploration in 2012 will not be so much by vehicles but by the equipment and the analyses of the radio scientist and engineer. One more province of the IRE.

J. H. Dellinger is a retired physicist and radio engineer and was President of the IRE in 1925. His address is 3900 Connecticut Ave., Washington 8, D.C. *(Received June 9, 1961.)*

Education in 2012 for Communication and Electronics

FREDERICK E. TERMAN FELLOW IRE

By the year 2012 the term electrical engineer will have disappeared from college catalogs to be replaced by "electronics," or possibly "electronic science." In today's terminology, education for electronics—née electrical engineering—will be a curriculum in applied science with strong emphasis upon mathematics, classical physics, and chemistry. The technology involved in dealing with wavelengths longer than a centimeter or so, will have become so standardized as to be "handbook stuff." Instead the main emphasis will be on the electronic properties of solids and liquids. The free electron will receive proportionately much less attention than today. Perhaps the most challenging part of the education of the "electronic scientist" of 2012 will be associated with energy at wavelengths extending from millimeters into fractions of millimeters to the infrared, visible, and ultraviolet regions.

The electronic scientists of 2012 will also be concerned with controls and computers, and will deal with systems of almost incredible complexity. The technology of these systems will come from electronics, but the "logic" concepts that will make them approach "thinking machines" by 2012 will have developed into a special discipline of its own.

The basic training for the electronic scientist of 2012 will be a four-year course, as it is today. The amount of knowledge that will be received in such a program will, however, be greatly increased—by a factor of at least 50 per cent—through the use of systems of programmed learning growing out of today's teaching machines. Moreover, through the use of programmed learning in high school, students will enter college far better prepared than today. The over-all result will be that in 2012 the typical recipient of a bachelor's degree will have covered material that would today require at least three postgraduate years of solid course work.

However, in view of the vastly greater number of things that the electronics man of 2012 must know, and the greater complexity of the problems with which he will deal, graduate work will still be almost a necessity for the brighter college graduates. It is safe to assume that over half of those who have obtained a bachelor's degree in electronics will carry their education to one full year of graduate work. At least 20 per cent of those with a bachelor's degree will go on for three to four years of graduate study and obtain the doctorate. Moreover, the most promising recipients of the Ph.D., particularly those interested in careers emphasizing the scientific as against the administrative or practical aspects of electronics, will by 2012 spend from one to three years in postdoctoral study immediately after receiving the doctor's degree. This additional training will correspond to the residency of today's M.D.

By 2012 the number of universities offering first-class graduate work leading to the doctorate will be greatly increased. This will be necessary to take care of the steadily increasing numbers of students attracted to the field of electronics, and the increasing proportion of those who go on for graduate work. In the training of students beyond the master's degree, research will have a prominent part. Much of the doctoral and postdoctoral training of electronic scientists will be in the form of participation in research, since this will continue to be the best way to train a man to learn the secrets of nature without the aid of a teacher.

The electronic industries of 2012 will be thoroughly aware of the advantages of locating their research and advanced development activities near a center of brains, *i.e.*, near a university with a strong graduate program in electronics. As a result each university with a strong graduate program will be surrounded by a complex of industrial activities of a creative type. Creative activities isolated from educational opportunities and the educational environment will be at a serious competitive disadvantage by 2012 and will have great difficulty maintaining themselves. Thus, state and municipal governments will support the development of electronic programs in universities in order to attract and hold electronic industries. This will result in an educational system which by 2012 will have remarkable strength in the field of electronics, characterized by many strong centers widely dispersed.

The college professor will be an important factor in the electronics industry of 2012. He will have high value as a consultant, and the industries clustered around his university will give him ample opportunity to use his talents as an advisor. In many cases his services will be sought as a member of a board of directors, he will be in constant demand by the government to serve on advisory bodies, and he will frequently found new companies and

F. E. Terman is Vice President and Provost of Stanford University, Stanford, Calif. *(Received June 12, 1961.)*

divide his time between being a college professor and a business executive.

In conclusion, it is safe to predict that the subject of electronics will be even more important within the university of 2012 than it is today, that the importance of the leading professors in the field will be greatly enhanced, and that educational institutions with strong graduate programs in electronics will be regarded as great economic plums because of this fact. The distribution of the strong graduate programs will determine the geographical distribution of growth in the ever expanding field of electronics.

Engineering Education—Circa 2012 A.D.

W. L. EVERITT FELLOW IRE

"For I dipt into the future far as human eye could see. . . ."
Tennyson, Locksley Hall

Engineering will be acknowledged as the most learned profession by 2012 A.D. The explosion in the breadth of knowledge needed for its practice will have made preparation for an engineering career both rigorous and extensive. Full professional recognition will be available, except in unusual cases, only to those who have completed formal educational programs in residence at universities, equivalent to or exceeding our present doctorates. However, the time required by a student to reach this goal will vary widely, because educational methods will be adjusted to the capabilities of the individual.

EDUCATION VS TRAINING Engineering curricula in educational institutions will be recognized as having two important and distinct functions which must be carried on concurrently—"training" and "education." Training is defined here as the inculcation of methods of procedure, the development of adequate vocabularies and skill in communication, facility in locating information, expertness in computation and in manipulation of mathematical processes, and similar skills developed through the general procedure of explanation and demonstration, followed by typical exercises with definite solutions or measures of performance. Education, on the other hand, is defined as the broader development of the mind and personality—a guided enlargement of creative ability and understanding. Education develops the ability to meet new situations with confidence and with a degree of wisdom limited only by the inherent capabilities of the individual.

TRAINING METHODS The training portion of engineering curricula will make extensive use of teaching aids, such as teaching machines. However, the teaching machines of the twenty-first century will have a sophistication far beyond that now envisioned. They will serve as tutors which are very responsive to individual needs, so that each student can acquire training at the most rapid rate which he is able, or willing, to absorb. Such teaching machines will present lectures from recorded television tapes in short sequences, with feedback from the student, who will be required to respond with answers to appropriate questions or by working out the solution of illustrative problems. Depending upon the adequacy and insight of these answers, the lecture procedure will be modified to meet the needs of the student. The entire operation will be under the control of computer techniques for the storage and interpretation of information. Testing and examinations of the results of training will also be done by machine.

Every student and practicing engineer will have access to adequate computers and information storage. Moderate capacity computers will be miniaturized and reduced in cost, so that each individual will have his own, as he now has his slide rule. As his needs arise he will also have ready access, through wire or radio communication channels, to computers and memory systems of any required complexity. Training of all engineers in the use of computers will be extensive.

LIBRARIES AND PUBLICATIONS Perhaps the most dramatic changes will come in the "library" of the twenty-first century. The information explosion of the twentieth century will have made necessary the development of entirely new means for the storage and retrieval

---

W. L. Everitt is Dean of Engineering at the University of Illinois, Urbana, Ill. *(Received September 7, 1961.)*

of the material in what we now refer to as "publications." Such publications will be stored in "memories" and elaborate and largely automatic means of classification and identification will make the search for a particular item of information rapid and certain, and its display available at any location. Instruction in how to give directions to the information storage systems to acquire the desired information will be an important and not inconsequential portion of the "training" of the engineer. On the other hand, visual reading of books, current literature, and video displays will be so important that training for rapid reading rates of one thousand or more words per minute will be universal in the grade schools.

EDUCATION But, as in the twentieth century, the important part of the preparation of the engineer will be his "education." Here the personal contact between student and teacher will still be paramount. The comment of Booker T. Washington will be as valid as ever,

"I am convinced that there is no education one can get from books and costly apparatus that is equal to that which can be gotten from contact with great men and women."

Relieved of the necessity of spending most of their time on the training function, devoted teachers will be able to concentrate their efforts on "education." Personal contact with students will be increased, because our society will recognize the importance of this relationship and support it. In spite of the use of teaching aids, the student/teacher ratio will not be increased. Instead, such aids will make it possible for faculty members to devote more time to the development of the individual student. The more efficient methods used will increase the amount of education per student rather than the output in degrees per teacher. However, the student himself will be expected to assume more responsibility for his own educational development than was common in the first half of the twentieth century.

CREATIVITY AND SYNTHESIS Since much of the instruction in analysis can be programmed on teaching machines, the engineer of the future will be given more education for *creativity* and *synthesis*. The case or project method will be emphasized and laboratory work will require originality and the ability to relate theory to the physical world in arrangements novel to the experimenter. "Cook book" experiments will be a thing of the past. Early participation in research projects will be the rule rather than the exception. Many advanced group laboratory experiments will be interdisciplinary in character, so that individuals with different backgrounds may contribute to team solutions of problems more complex than those possible for a single individual. The guidance of such a program will call for teaching ability and insight of the highest order.

Continuing education throughout a career will be normal, accomplished to a large extent by regular sabbatical year exchanges. During such exchanges, industry and government engineers will either return to universities for further study or will exchange positions with colleagues in other organizations. Engineers on faculties will participate, through similar exchanges, in industrial research, government service, or programs of other universities.

FACULTY RESEARCH The primary responsibility of the faculty in engineering education will be to develop creativity in their students, so they must themselves be creative. Noncreativity on the part of a faculty member will be recognized at an early stage and such individuals will be diverted to nonteaching activities within or outside the university. The creative programs of the faculty will provide means for the integration of students at all levels into progressively increasing responsibility in engineering teams.

ENGINEERING CURRICULA The trends in engineering curricula which began in the twentieth century will continue. The early part of all curricula will have more in common, emphasizing basic science, mathematics, and fundamental principles of analysis and synthesis common to all engineering. Only after a student has demonstrated the thorough understanding of mathematical, scientific and economic fundamentals necessary for creativity at twenty-first century levels will he be permitted to continue toward his professional goal as an engineer. The essential unity of engineering as a profession will be recognized. But his advanced program will require that he develop an insight into the methods of solving hard problems. As the student progresses, he will study some individual area more intensively, although the divisions in engineering will not be those now common. By pursuing certain areas in depth, he will learn to solve and contribute to team solutions of difficult and complex problems. Hence, he will need to choose an area of emphasis such as the "Processing of Information," the "Processing of Energy," the "Processing of Materials," or "Biophysical Engineering." Sufficient common material in all these areas and in the principles of "System Synthesis" and "Human Engineering" will provide for communication between the members of engineering groups, but too much breadth of subject matter will not be obtained at the expense of shallowness.

The importance of emphasis on the humanities and social sciences will be recognized from the grade school on through the highest levels of engineering education, as society relies more and more on engineers for leadership in business and affairs of state. Intensive efforts made to develop interpretive programs in these areas will also stimulate more engineers to pursue programs of self education in these fields throughout their careers.

With the continued growth of engineering applications, instruction in manufacturing processes and specific technological areas and skills will be recognized as a training function which industry must assume entirely.

CONCLUSION If these predictions do come to pass, and I feel sure they will, engineering will deserve its designation as a "learned profession."

A Day in the Life of a Student in the Year 2012 A.D.

MAURICE J. PONTE FELLOW IRE

That morning Joe Parkinson woke at seven o'clock; the Song of the "Hypertexas" University was heard in the students' rooms, distributed by the local millimetric wave network. He switched on his luminescent wall panels which reconstituted natural light and, wanting a rural atmosphere, chose the "Park" light combination which turned the panels into a perspective of green leaf and flowers pretty well in their natural colors. The only two panels which remained dark were that reserved for telephone communication and that for college instruction, brought round by television.

Gathering his spirits, Joe realized he had a slight headache and remembered that a new vaccine had been pricked into him the previous day. The diseases of former times, tuberculosis, polio, and even cancer had been suppressed by preventive vaccines which were compulsorily applied, as was shown on his identity card registered with a series of eleven figures and three letters; a mere look at this arithmetical group gave the complete story of the individual, from his blood group to the results of his vaccinations, his diseases and his operations. In the University's central register, a punched card in his name, processed by the specialized computers, was to forecast Joe's behavior in his varying states of health. But from a few months back an unknown virus was playing havoc, especially among young people, in the form of mental troubles which had become the principal plague of Man.

The first symptom was an almost complete loss of memory and the new disease was attended by insensitivity to color. These deficiencies had first been attributed to a system of television reception in which the nerves were directly excited by electrodes connected to the receiver and placed on the temples, at points fixed by the electrical topology of the brain: vision by means of panels had thus been done away with, the optic nerves being brought directly into action. This system had had a considerable application but had finally been prohibited. And so students were subjected to special vaccination and Joe directed a troubled eye on his registration number, to which another group of two figures had been added: now three letters, and thirteen figures. What was the outlook in 50 years' time?

Throwing these dark thoughts to one side, Joe dressed, went to his kitchenette where by pressing the "breakfast" button on his automatic cooker, he received after an interval of 30 seconds a balanced, practically tasteless, but medically perfect meal.

The signals physics course was due to begin at eight o'clock. Joe was one of the nine hundred and fifty students entered for this special course, among so many others, between which the thirty-six thousand students of the Hypertexas University were spread out. The University buildings and grounds extended over a wide area, formerly desert land, so that it had been possible to plan the layout in a completely rational way. Utilization of solar energy and of the natural underground water resources had transformed the area, not only into parkland, but also into cultivated land and breeding farms, scientifically conducted in hot houses at constant temperature and illumination cycles. The Hypertexas students were thus brought up, intellectually and physically, under perfectly definite conditions, in accordance with standards gradually perfected many years back. Many countries were represented at the University.

Joe gave another look at his syllabus. The signal physics course was transmitted over channel 23 of the millimetric network, distributed by dielectric cables to all the students' rooms. He pushed down key 23 on the keyboard mounted on his desk; the title of the lecture appeared on the television wall panel; on the dot of 8 o'clock Joseph A. H. Faraway, the outstanding professor, began the sixteenth lecture on cable matching for a given rate of information. These masterly lectures were recorded on magnetic tape and transmitted from the University's Technical Center, an extensive center from which all basic instruction was distributed. For lecture theatres had long been abandoned; this system had been rendered ineffective due to the increasing number of students and had been replaced by broadcast lectures and by grouping students in teams of 30 or so round a task master. The task master's role had become fundamental, through his explanations of the course, his choice of subjects of application and the conduct of practical work in the laboratories and workshops set aside for the course. Prof. J. A. H. Faraway himself collected the task masters round him—thirty in this particular case—to discuss the way the course was going with the students

M. J. Ponte is President-Directeur Général of the Compagnie Générale de Télégraphie Sans Fil, Paris, France.
(Received July 5, 1961.)

and to decide on such adjustments as might have been shown to be necessary in the light of observations made by his group.

And so Joe was in task master A. 17's group. During Faraway's lecture he took notes on his coded typewriter. With the telewriter, which was also coded, and permanently connected with the task master, he raised questions on points which did not seem clear, either at the time or on reading over his notes. At the other end, in the office of taskmaster A. 17, the questions raised were recorded on magnetic tape to be examined as a whole.

At 9 o'clock the lecture ended. Joe scanned over his notes, which he easily decoded since, from his earliest youth, he had been trained to substitute groups of letters for the propositions of normal language, thus the proposition "the wave is propagated at phase velocity v" took the form *WPv*, this code having been internationalized. As he came to it, he verified on his miniature algebraical computer which had taken the place of grandfather's slide rule, the solution of an equation given by the professor without justifying development, since the solution of even the most complex equation, including partial derivatives, was obtained by machine and no longer by the old methods, taught only to advanced students anxious to dig into the history of sciences, or to specialists engaged on the construction of the computers.

For some ten minutes or so, Joe was engaged on physical drill recommended for his health according to a perfectly definite daily routine. That particular day he had some difficulty, perhaps because of his recent vaccination, in pedaling a stressed spring into tension in a set time, so that he had to have recourse to the immediate absorption of an irradiate vitamin tablet.

At ten to ten, Joe walked out for the first time that day and proceeded by the traveling platform, after a number of changes, to the buildings of the signals laboratories and workshops, and on to the room A. 17. At ten o'clock the thirty students of group A. 17, standing around their task master, discussed and exchanged views on the fifteenth lecture, in the light of questions raised by the students during the lecture. This was followed by practical work on that lecture. The laboratory and workshop were provided with perfect equipment; over the last 10 years it had finally been realized that the most productive national investment was that applied to teaching the young in all kinds of subjects: scientific, technical or "literary." In the same way, past differences between "workmen" and engineers had changed in their nature. On the one side automatic machines had levelled aptitudes and functions of the personnel in carrying out their separate tasks; for instance, the automatic machines for the production of signal receivers delivered the product with its cost, the value of stocks and all other items of management, so that the staff was almost entirely occupied in maintenance. And steps had had to be taken to arrange for the training of the staff set apart for scientific research, for instruction and for development, and to see about the corresponding machines and the way in which problems should be put to them. The class of engineers and physicists, who on their part were in a position to design and operate the larger systems, such as telecommunication system, had considerably increased in numbers.

Joe was preparing for that class, as well as others in his group, and the practical work for the day consisted in determining with three of his fellow students the best way of utilizing a given bandwidth to provide the individual television links between n individuals geographically distributed in accordance with a given law of population density.

At one o'clock, practical work being finished for the day, Joe stepped once more on the traveling platform on his way to the cafeteria, an immense building with a capacity of thirty thousand meals per hour, thanks to the automatic self-service system which dealt out four menus, prepared in a few seconds by the electronic cookers. Joe hurried on more than usual, for the Hypertexas-Hypercolumbia baseball match transmitted throughout the world was due to begin at two o'clock. But the signals department had been unable to reproduce the atmosphere of the stadium of fifty thousand seats, so that from the dim past, circus acts remained the permanent feature of human culture. Players were carefully picked after repeated tests and specially trained, on appropriate diet, so that the match should end honors all, a most satisfactory result from everybody's point of view.

At 4 o'clock Joe was sitting in his personal study room following a lecture on the history of industrial techniques: the subject was the construction of television receivers in the year 1960. Joe marvelled at his ancestor's expert handling, their nimble positioning of the most diverse parts, the way they handled enormous vacuum tubes, and the deft handling of the old-fashioned soldering iron. All this recalled films he had seen on the building of the pyramids, or the work of the lace makers with their long pins stuck in taut-drawn cushions.

At 5 o'clock he passed to another workshop of group A. 17 where the group was to produce, by its own means, but using modern processes, a working stage of a micro-module amplifier. The work had been in hand for some time, but on that particular day it had been cut short for Joe wanted to return to his room to watch on his screen an automatic news item transmitted at 18:30 hours by a satellite engaged in the examination of Saturn. Actually, several events of this kind had already been broadcast from other planets; and these had rather upset our views on these things, but Saturn had so far always resisted investigation, perhaps on account of the nature of its rings. In point of fact reception was disappointing, jammed as it was by incomprehensible signals. The question remained open.

After a short rest, Joe proceeded to an outdoor feast which was the rage among the youth of both sexes. It was called "Antique Barbecue"; it was actually an outdoor meal at which students, lads and lassies, roasted whole sheep spitted on a wooden pole so that the beast could turn in front of fires which they themselves had prepared. That evening, some five-hundred students were in the clearing, and twenty sheep prepared and wrapped in "Do it yourself" plastic bags, were turned over to the twenty groups. The occasion was a merry one, for there was a roasting and carving dexterity competition. The laughter and joyful shouting which, so we are told, reproduced old World atmosphere was muted while attention turned to a large television screen, visible to all, where the finals of a rally of rockets to the moon were being displayed; ten rockets flying the colors of the ten world universities which had successfully completed successive heats were competing, and the broadcast was enthralling. Unfortunately the Hypertexas rocket had to drop out, one of its ionic engines having died half way. The American continent had 2 representatives, so far well placed, but there was a certain amount of anxiety on account of the unknown characteristics of certain foreign competitors. The race was something like the regatta of former days, the rockets had to assemble at a given moment at a particular point of space and move around so as to be in a favorable position at the starting signal. The arrival of rockets launched from various points on earth was easily followed, thanks to transmissions from fixed satellites which broadcast every launching; each rocket was manned by 2 operators, keen racing men ready to win by all recognized means, as in Roman games. The moon's very surface had to be reached in the shortest possible time, measured from the earth by Doppler equipment on nearing the time of arrival, while the tapes were checked with the utmost care. The entrants remained in communication with the public and strove their utmost in the clamor of their supporters.

The race was well up to its reputation. Contrary to all expectations, it was a so-called small country whose rocket was the winner, the resourcefulness and daring of its pilot having overcome some technical weaknesses. The return was uneventful, each team having this time been able to return to its base.

And the day ended as in old times, the lucky ones received their winnings, the others made the best of it, while some rather abnormal types decided to get back home on foot, their eyes turned to the heavens where they tried to distinguish between true stars and false planets.

Back in his room, Joe slept. He dreamed. And in his dream he saw what a former reading had taught him: a stream, a modest house, standing quiet in its surroundings of green leaf and flowers, far from everywhere, where by his log fire, with his books as his companions, he felt his own master.

For the first time that day he was truly happy.

Fifty Years of Teaching Machines

HAROLD A. ZAHL, III FELLOW IRE

Now that we are well into the twenty-first century and celebrating the 100th anniversary of the IRE, it is a good time to sit back and attempt to assess the tremendous impact the teaching machine has had on our world in the last fifty years. How could a man in 1962, fifty years ago, have guessed the fantastic changes this one device would bring about? The answer is, of course, that he could not; he would have been a fool to try. Teaching machines were in their infancy then, as consequently were human comfort and convenience (at least by our standards).

Let me recount. In 1962, computers were only about 20 years old. Biological computers were still mostly theory; general purpose computers were crude, oversized machines of unbelievable inefficiency; and our teaching machine's bio-infrared circuitry was yet to be discovered. To be fair one should point out, however, that the basic principles of feedback were understood, although the real basis for machine-human interaction had not yet been discovered. Consequently, these rudimentary devices required large numbers of human attendants—called programmers—to assist them.

In those days, even a 0.01 cerebit[1] rapid access memory was considered impossible and a computational time of 1000 nanoseconds was thought to be ultra-rapid. Such "teaching machines" as existed then were not at all versatile, being "programmed" in the pre-1974 sense of the word. It was a dismal time in computer history, one

H. A. Zahl, III, is by his own admission the grandson of Harold A. Zahl, who back in 1962 was Director of Research, U. S. Army Signal Research and Development Laboratory, Fort Monmouth, N.J. ("Received" April 1, 2012.)

[1] One cerebit is believed to be equal to 10^9 ordinary bits.

of the valleys in the up and down curve of progress these devices have followed since their inception (see Fig. 1).

After this bit of reminiscing, it is possible to jump to the present. Everyone now in the field knows the peaks in the line of progress in the last fifty years: the breakthrough in biological circuit design in 1971, the invention of the self-regenerating virus gates of 1974, the solution of the light-speed access problem in 1981, the discovery of voice reception and machine linguistics in 1986, and the establishment of international memory-information banks for televideo service throughout the world. Many modern-day sociologists feel that these five achievements, along with the 1988 break-through in nuclear fusion power generation, have had more impact on human-kind than all of the rest of man's forward steps throughout history. That our happiness and well-being are allmost fullly dependent upon these six advances is apparent when we examine our everyday lives.

But before discussing the impact of the teaching machine on civilization, I should mention some of the other important, if not major, accomplishments since 1962 as a tribute to those men (and the computer) involved. All of the five major practical achievements mentioned earlier depended upon theoretical work that had gone before. Consider, for example, the monumental work of Chen of 1976, which first established the feasibility of liquid-state computer components. We could not be where we are without his outstanding work.

The development of finite mathematics largely through the efforts of Stochastikov and R. C. Tienport, made possible the final linkage between information theory and automata theory, and opened the way to the solution, by A. N. Williams in 1979, of the General Coding Problem. The 1982 Timmons Theory of Self-Awareness in Circuit Modules was basic to the construction of the first real teaching machines. Simon's Theory of the natural laws of growth of 1992 was the basis for later work with spontaneous organizing computers. The last theoretical piece of work, the Unified Computer Field Theory, was, of course, first proposed by ILLIAC IV at the University of Illinois in 1996, and led to the well-known Nobel decision eliminating all future machine contributions from competition with humans. It made possible automatic programming, self-construction and repair, reduced cost per bit of information processed and true computer adaptiveness and self-awareness.

People of 50 years ago doubtlessly thought they had much to be happy about. By our standards, they did not. Of course, it was in the general field of education that the adult living in the year of the IRE's 50th birthday, 1962, had the worse disadvantage. He had to spend years getting what we would not even consider a basic education, attending schools outside of his own home. Education was a mass-process which did not even allow each person to progress as fast as his natural abilities might have permitted, and a person's formal education stopped in the early part of his life. (One must remember the average life span then was less than 100 years!) Those who were educated to be specialists learned little about anything else, and those who were given broad educations learned little about anything.

How astonished the man of '62 would have been if he could have seen the educational system we have now! As learning was somewhat in disrepute in his day, he might find it difficult to understand that we value it above almost everything else. Some sociologists have theorized that the biggest change in the last half century has been the reversal in human attitudes toward education. Of course this has primarily been brought about by the advances in teaching machine technology, which have made a good education not only possible, but absolutely necessary for everyone in our society. Life

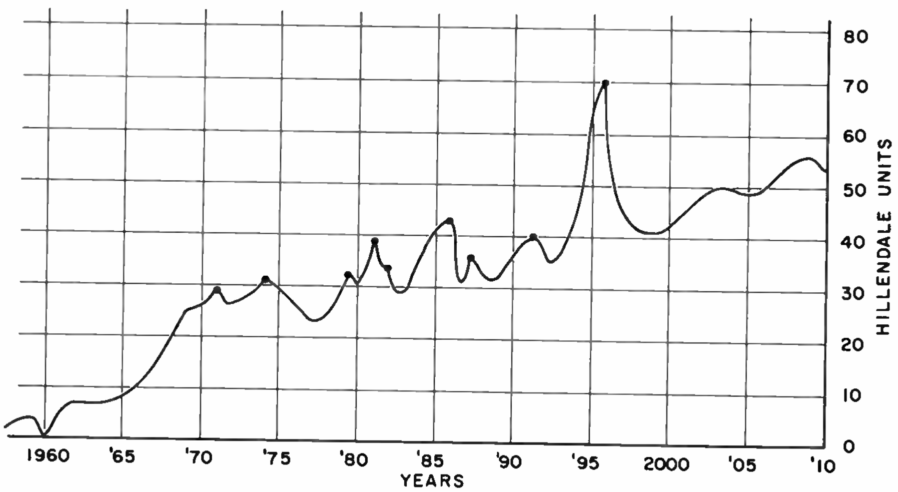


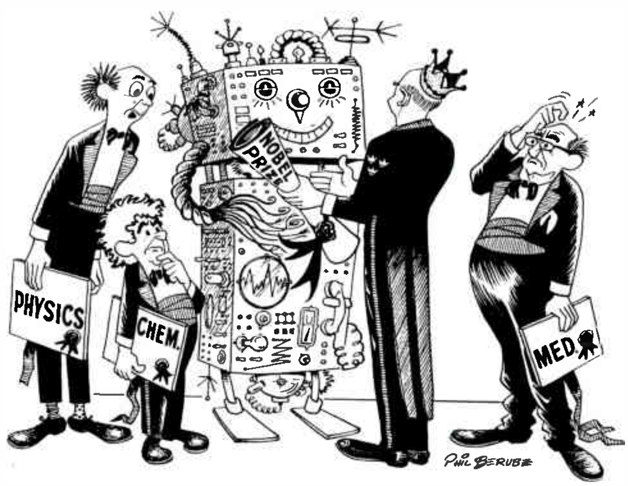
Fig. 1—Practical and theoretical stages in teaching machine design in years and hillendale units. (Note that valleys may be caused by periods of stagnation, as, for example, in 1987 and 1988, when men of science were directing all their attention to discoveries being made in the field of nuclear power.)

without the modern, inexpensive direct-dial televideo receptor in the home, linking us through substations to the main memory-information banks and capable of analyzing handwriting, understanding voice commands and requests, giving verbal responses, bringing us a free education in any subject that interests us, and even serving as a companion to the aged and the lonely, is now as inconceivable as an uneducated man. I personally do not long for the "good old days" of years past.

Odd as it sounds, even children disliked education in the 1960's. A person of that era would be amazed to see our 2- and 3-year olds sitting in front of their televideo screens, learning the International Language and other basic studies. He would be surprised to see our learning cubicles in public buildings and libraries, and to find that international renown and prestige is awarded to people of all ages for their learning in various areas of the arts and sciences. It would surprise him to learn that our national sport is the mass participation televideo quiz game, in which the entire audience is armed with brains rather than muscles for it was not until 1989 that DIAG IV recommended to the General Assembly the outlawing of all physical contact sports.

In many other ways, compared to today's standards, the man of 1962 faced a hard, dreary life. If he had marriage problems, he could not dial the proper code on a televideo and have instantaneous access to the thinking of the world's best minds on the subject. Instead he had to visit another human being,[2] placing his future well being in the hands of a single individual who was probably little better equipped to solve his problems than he was. He could not sit in his living room and study any phase or area of any arts and science he wished. If he became sick, his doctor had to rely on the contents of his own brain, rather than being able to dial Medical Diagnostics for assistance. In those days, a doctor facing a serious problem could only call in other doctors as consultants, provided they were available and willing. Human physical maintenance depended upon many uncertainties.

In matters of recreation, the man of 1962 was at a disadvantage again. Great portions of the small amount of free time he had[3] were spent laboriously learning how to play the games or master the activities that interested him. It would have been impossible for him to have spent one evening previous to leaving for the Rockies learning the known expert techniques and opinions on mountain climbing. Without years of actual physical experience beforehand, he probably would have killed himself. There are shocking historical records of deaths, amounting to thousands, due to untrained drivers of the early hydro-carbon propelled surface cars; compulsory heli-jet training for children was not introduced until 1985—several years after installation of our first traffic and navigation control net.



. . . and led to the well-known Nobel decision eliminating all future machine contributions from competition with humans.

But the worst thing about life then, as I see it, was the uncertainty about the future man carried in his heart. There was no international language, little international understanding, and no international peace—because there was not yet a world-wide system of televideo teaching machines. Therefore, the benefits and rewards of advanced technology were not available to all peoples.

Now we are beyond all this and faced by troubles of our own. We would be unfair to ourselves if we credited the good or the bad of our world completely to the teaching machines; for while they are an essential part of our existence, we designed them, we built them, and we keep them running. We will improve them in the future. And for all of my disparaging remarks about the middle 1900's, it is obvious to me that we really stand where we are because of the pioneers, the men of that day. Each generation climbs higher than the last because of

[2] Such people were called marriage counselors, quaintly enough.

[3] A 40-hour work week was common in this country in the 1960's, and there were even worse situations elsewhere.

the work of the last—but we sometimes are so busy looking ahead that we forget to glance back in gratitude. On this point I must mention a very erudite article[4] written by my grandfather, Dr. Harold, I call him (in respect of the Ph.D. of that era). The reader would do well to dial IRE archives and take a reading on how it looked to the old gent then.

Considering the sweeping changes in our world in the last fifty years brought about by the development of the teaching machine, I would be as much a fool to try to predict what will happen in the next fifty as the man of '62 would have been to have made guesses about our time. It is my personal belief that before 2062, we will be able to take a forward step that will make the major achievements of the last 50 years look paltry by comparison. Hard as it is to imagine, I suspect that the IRE writer of 2062 will be laughing at us, remarking that teaching machines were in their infancy in 2012. And that will be as it should be. Let us only hope that he too will remember that he stood on our shoulders to reach the heights he will have attained.

And finally, I should like to thank my good friends Alpert, Desmond, Bitzer, Braunfeld and A. Kingery of the University of Illinois; also E. M. Reilley of the Signal Corps Lab. I asked their help on this story since their ancestors of identical name were quite involved in early stages of the evolution of our present teaching machine—around 1962, I believe.

[4] H. A. Zahl, "Looking backward toward tomorrow," IRE TRANS. ON AEROSPACE AND NAVIGATIONAL ELECTRONICS, vol. ANE-8, pp. 3-6; March, 1961.

The Tools of the Engineer—2012 A.D.

R. BENNETT FELLOW IRE

About 1970 engineering had sensibly stabilized and technology had forced cleavage into two distinct groups.

The first group was that of component engineers. Components no longer had their earlier crude meaning, they had become functional building blocks of complete circuits. Further, a shortage of engineers had forced the standardization of many circuits so that whole classes of amplifiers, power supplies and mountings had become shelf items.

After a competitive race in the 1960's to produce the smallest units, reason had prevailed. While components were small by earlier standards, the ultimate sizes were such that costs were reasonable and servicing practicable. For example, whole receivers were the size of pound candy boxes rather than cigarette packs.

Component engineers spent their time revising a limited number of building blocks in line with the latest molecular theory of electrical circuitry. Many whole functional circuits went unchanged year after year.

The second group, system engineers, were effectively specialists in putting together complete functional circuits to form equipments. Since many of the complete circuits no longer had the individual characteristics of R, L, and C, a whole new body of mathematical analysis had grown up.

One feature of the new design methods involved the use of computers. It was no longer necessary to cut and try, to build many models experimentally. A proposal for a new system was first programmed into a computer and the ideal circuit units computed. The characteristics of the complete device could be checked from input to output. If the desired over-all characteristics were evident, then the computed circuit criteria were checked against currently available units. Since these units did not always match, the computer was reprogrammed to have the system accept the available units. Some errors becoming evident, the computer was then asked to specify the characteristics of the needed new units. When all units had become available, the system would be put together. Rarely was a second model needed. Occasionally a good idea for a modification and improvement would cause a revamping process to occur.

Sometimes at meetings older engineers would retell their fathers' tales of the frustrating and harassing old days of soldering irons and cut and try.

R.Adm. Bennett, USN(Ret.), is Senior Vice President—Director of Engineering of Sangamo Electric Company, Springfield, Ill. *(Received July 5, 1961.)*

The Automatic Handbook

Nathaniel Rochester Fellow IRE

The automatic handbook will be a computing machine that has stored within it the kind of data and calculation procedures that are now contained in handbooks on subjects like electrical engineering, physics, and mechanical engineering. Unlike present handbooks, it will not only store descriptions of procedures, but also be able to execute them.

Unless some unforeseen developments in the science of information retrieval appear, there will still be the problem that the user can select wrong formulas or procedures just as he can with the handbooks of today. However, after a proper selection has been made, and before the particular desired formula or numerical result is available, today's user often must do a lot of work which the machine will be able to handle.

Recent work on symbol manipulation by machine has shown that the machines can do most algebra. However, there are some kinds of manipulations that so far have resisted mechanization, and some of these will not be within the scope of the automatic handbook. It would, however, be able to help with derivations that were beyond its ability to discover because it would be able to check these for algebraic and numerical errors.

Such a machine would probably be centrally installed and accessible to many, rather than being the private property of an individual. A modest extension of today's real time computer methods should reduce access delays to about the delays that one now encounters in getting the dial tone from a telephone exchange.

Consider a hypothetical instance of an electrical engineer using the automatic handbook to calculate the pressure required to produce a certain flow of cooling air through ducts in an electrical device he was designing. Probably an advanced display would obviate the need for a book at all. However, to avoid having to imagine this, it will be assumed that he would have a mechanical engineering handbook much like the present ones except that it would be backed up by a computer.

The engineer would look in his handbook and find the parameters that influence pressure. He would find, also, figures showing round and rectangular sections, elbows, and other discontinuities. To see if he had a problem he would specify a particular design by giving a succession of figure numbers and dimensions. He would specify the flow, the input temperature, and the output resistant to air flow.

The machine would give him the pressure and, typically, he would see that it was too large and that he did, indeed, have a problem. So, he would ask the machine for a formula for the input pressure. Since he might expect the formula to be complicated, he might ask for a simplified one that would ignore any effects that contributed less than 10 per cent in the first sample duct. The machine would prepare this latter expression by reference to the calculation it made before. It would eliminate from the literal expression all terms which corresponded to elements of the numerical calculation that contributed

---

N. Rochester is Manager of Advanced Computer Utilization, IBM Data Systems Division, White Plains, N.Y.
(Received August 1, 1961.)

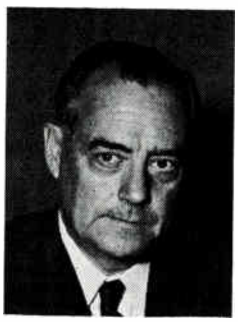
less than 10 per cent. Since this type of formula simplification has actually been programmed and successfully executed on today's computers, the automatic handbook would probably be able to do much more to guide the engineer. He would then see clearly what the important parameters were and could easily produce a succession of designs for the machine to test. Jointly, the handbook and the engineer would quickly arrive at a final design.

The automatic handbook would, of course, be equipped with a large set of mathematical procedures. It would be able to integrate expressions; solve differential equations, both when a solution can be expressed as a concise literal expression and when only a numerical solution will do; simplify most expressions embodying transcendental functions; and do most of the routine mathematical work that the engineer himself could do by looking it up as well as some that would normally be accessible only to a mathematician.

References to literature are an important feature of handbooks. The handbook could be kept current by prompt addition of new references and deletion of obsolete ones. The automatic handbook would be helpful in both familiar and unfamiliar fields, would do a lot of detailed work, and would greatly expand the technological range of an engineer and increase his output.

There Will Be No Electronics Industry in 2012 A.D.

J. M. Bridges FELLOW IRE



In the year 2012 A.D. there will be no electronics industry as we know it today. The electronic component industry, around which our current electronics industry has been built, will have disappeared completely. The design and fabrication of electronic devices will be completely mechanized and will be performed as an integral part of the development and manufacture of all end products, commercial and military.

During the next fifty years adaptive computers will have been developed and perfected which will perform all the functions of systems analysis and layout, equipment design and functional configuration. The information from these "design machines" will be fed from stores into automatic assembly machines which will fabricate and assemble the final product from processed raw materials. The output from the test portion of the assembly machine will feed back to the design machine where errors in design will be corrected and where the design will be optimized for cost, reliability and performance. These design and assembly machines will be flexible and versatile enough so that they can design and fabricate a wide variety of products merely by changing the inputs to the design computer and the materials fed to the assembly machine.

The thousands of electronics engineers now required to support development and product design of complex electronic systems will be replaced by these machines and by a relatively few highly specialized and very creative scientists who interpret new systems concepts in terms of inputs to the design computer.

Competitive advantage will be obtained by excellence in conceptual creativeness, materials research and marketing.

All of this will result in the producers of end products—whether they be space vehicles, aircraft, automobiles, home entertainment centers or electronic air conditioning systems—being completely self-sufficient. They will not have to rely upon a specialized electronics industry, nor will they have to compromise their product by using "existing" components. The only standardization in "electronics" in 2012 A.D. will be in materials and this may lead to a specialized materials industry which will replace the current electronic components industry.

J. M. Bridges is Director of the Office of Electronics, Office of the Director of Defense Research and Engineering, Washington, D.C. (Received January 15, 1961.)

Radio Communication in 2012—An Obsolete Art

Dorman D. Israel FELLOW IRE

During the fourth quarter of the 20th Century, solid-state microelectronics began to find itself by entering into more and more diversified usage. There was, in fact, a period when it seemed reasonable to assume that all radio receivers would make one-hundred per cent use of microelectronic technique. This was before the obsolescence of radio communications could have been visualized.

Suddenly, and yet collateral with the surge toward microelectronics, and, in part, even due to it, promising discoveries and better understanding of the nature of electrobiology literally cleared the air. By furnishing a technique for accurate physiotelemetry, revealing details of the magnitude, as well as confirming the path of nerve currents, microelectronics played its very great part in the fruition of this giant step. Thus, the art of bionics and allied nerve-to-brain studies, moving slowly and disappointingly, but irrepressibly forward, became the big news in communication advancement as well as others not germane to this paper.

Obviously, surgical experiments and tests had to follow and confirm theories derived from such observations. In due time, it became necessary and practical to make experiments on animals whose neuromechanics could be considered a reasonable counterpart of the human being's. Certain haunting anomalies always seemed to appear when least expected during the course of carrying on the seemingly obvious experiments of attempting to simulate, by externally-initiated electronics, the internal action of nerve currents. Likewise, the seeming contradictions or disappointments that resulted from the introduction of these currents to specific points in the brain could not help but intrigue the more meticulous workers. They gravitated to the conviction that some animals responded with subtle difference from what was expected. It was contrived to carry on studies whereby animals were not only exposed to the influence of man but to the influence of other animals during such experiments. Now the key seemed to be fitting the lock and exactly when the full realization occurred as to the cause of the apparent contradictions was not fully recorded and just who might have been responsible for it is still not clear in the record. Probably it was another case of the not uncommon incident of simultaneous discovery of the same phenomenon in several laboratories and, in fact, even in various countries.

What was deduced was the momentous principle that the everyday or routine low-level type of communication that animals have been known to use was ultimately rationalized as something not too different from what the generalizers like to call "extra-sensory perception." As certain animals became more intensively human trained they simultaneously lost their ability reliably to continue this animal-to-animal communication. True, such "genius" animals acquired the ability to seem to think for themselves and even, to a degree, developed a low-level consideration faculty, but they certainly lost their animal friends while they learned to influence people.

Having proved, in a measure, the characteristics listed above, the corollary developed which explained why man who, like other animals, having been originally endowed with at least as good a built-in low-level communication system as the lower animals had, nevertheless, over the centuries of technologic and sociologic development, unintentionally rid himself of this genetic birthright. Agreed, he was ahead of the game as it were, but it was only human to want to know why one must lose so much to have won otherwise.

Science's real mandate was now evident. A way had to be found to "unlearn" just the right things to reacquire enough of our original animal-like extra-sensory perceptive powers. Yet mankind could not afford to lose

---

D. D. Israel is Executive Vice President of the Emerson Radio and Phonograph Corporation, Jersey City, N.J.
(Received May 17, 1961).

the necessary scientific stature and advances that had been acquired over the centuries.

Various laboratories devoted themselves feverishly to the task. The large foundations liberally supported work to such an extent that government financing was not even necessary. By the year 2000, sufficient progress had been made so that elementary communication could be carried on over very short distances between suitably and carefully trained human subjects. Then, with the additional aid of certain bioelectronic augmentation, the technique was refined so that individuals who had been so equipped and previously trained were able to make adjustments whereby they could communicate over quite substantial distances but always by appointment—a most fortunate limitation.

As of this writing, in 2012, there are no doubts in any circles as to the ultimate full success of the program. Tests have proceeded now so that if the parents consent, newborn infants can be operated upon and the latest submicroelectronic equipment installed in the brain and at certain critical points in the spinal column so that they are almost certainly assured not only of the benefits of full nonradio communicative powers but also there is reason to believe that their scientific creative ability will be enhanced. Logically enough, this operation must be performed within two weeks of birth because if the infant is only slightly exposed to contacts with its family who still have not completed their "unlearning" and readjustment, he might never become a good subject for the modern system of communication.

Electromagnetics and Communications

HAROLD A. WHEELER FELLOW IRE

THE LOW FRONTIER The VLF range (3-30 kc) has long been the low frontier of the electromagnetic spectrum used for "radio" communication. There are being explored some further opportunities for useful communication at these and even lower frequencies. At present the lowest operational frequency is in the vicinity of 15 kc, which is used by the Navy for communication to submarines while submerged. These and lower frequencies are also attractive for world-wide coverage around the clock, from any one transmitter.

One contemplated use of lower frequencies is found in the Navy's proposed "Omega" long-range navigation system. The world would be covered with a grid of interference patterns, laid down by a network of several transmitters operating near 10 kc.

The next lower frequency to be considered might be around 3 kc, where the propagation under the ionosphere has some unusual properties. There is one, and only one, mode of propagation with horizontal polarization. It may be designated the TE-01 mode, which has a cutoff frequency about 2 kc. It happens that the vicinity of 3 kc experiences an optimum combination of reflection from the ionosphere and excitation by a loop antenna at the ground (a vertical loop with horizontal axis in the direction of propagation). At the maximum altitude of aircraft, this mode is susceptible of omnidirective reception by a horizontal loop.

Below 2 kc, there is single-mode (TEM) propagation under the ionosphere. At about 10 cycles, the earth's surface and ionosphere form a cavity resonant at its lowest mode. Experiments in Germany and the U. S. have shown signs of this resonance by observing a distinguishable peak of lightning noise near this frequency. The utilization of such a low frequency depends on the development of a power radiator that can surmount the lightning noise over a useful area, or even over the entire earth.

Recent tests at VLF have indicated substantial transmission through the ionosphere, from the surface to a satellite. This was unexpected, since the ionosphere has been regarded as a shield for VLF. However, it is known that AF signals penetrate the ionosphere in the "whistler" mode, when transmitted between two focal points at opposite ends of the Earth's magnetic field. The satellite

H. A. Wheeler is President of Wheeler Laboratories, Inc., Great Neck, N.Y. (A subsidiary of Hazeltine Corporation.) *(Received September 5, 1961.)*

test gives further impetus to the investigation and possible utilization of VLF communications between the Earth and outer space.

Along with the progress in propagation under the ionosphere, other channels are being explored underground in the Earth's crust. Promising results over moderate distances have been obtained in tests of guided waves between separated conductive layers in sedimentary strata at moderate depths (less than 1 km). The first such tests under controlled conditions were recently reported by Deco Electronics, communicating between two salt mines in New Mexico. Another possibility has been proposed, the "deep waveguide" in basement rock.[1] This is bounded on the upper side by conductive layers near the surface (sedimentary rock or sea water) and on the lower side by the "thermal ionosphere" at a depth of the order of 30 km. It remains to be proved whether long-distance signaling is possible in the deep waveguide at frequencies of the order of 1 kc.

In the field of communications, electronics has experienced its greatest stimulation and development. In detecting and processing weak signals, it is closest to the sensitive communication processes of human beings.

The following paragraphs are devoted to some situations in which the present practices are notably deficient in communications.

HIGHWAY COMMUNICATION On the highway, many situations would benefit from some form of communication, either from controller to vehicle, or from one vehicle to another. Even when the two parties are in sight of each other, there is now very little ability to communicate. From controller to vehicle, traffic lights and personal "antics" are utilized to some fraction of their capabilities. Between two vehicles, brake-signal and turn-signal lights are most rudimentary, though extremely helpful as far as they go.

We have radio communication between base stations and vehicles. It is fairly effective in police networks, and could be developed much further for taxicab dispatching.

Perhaps the next development will be the transmission of advisory messages to all cars passing a certain point on the highway as has been demonstrated by the Delco Radio Division of General Motors Company.

The greatest need is the ability to communicate from one vehicle to another in the same vicinity, for purposes of warning or other advice. To be useful, this would require a simple way of selecting which vehicle is to be addressed. For example, one might advise another car that his turn signal was left on, or that he was trailing black smoke from his exhaust. Obvious such facilities would have to be used with restraint, but the typical driver has come a long way in this respect and might prove equal to a further challenge.

A friend of the writer (R. G. Ling) has experimented with such communications by using a microphone at the driver's position and a loud speaker just behind the radiator. He has encountered a favorable reaction in several cases, and once supervised the evacuation of a car that was on fire without the driver being aware. His experience is persuasive that intervehicle communication offers a real opportunity.

The natural way for selectively reaching just one vehicle would be through a narrow beam of radiation, although this would require careful aiming by the sender. This would be especially difficult during the motion of one or both of the vehicles. Invisible radiation would offer some advantages, while visible radiation would enable the receiver to see which vehicle is the sender. Even if the signal radiation were invisible, a visible signal could be used for identification of the sender and perhaps also for automatic acknowledgement by the receiver when the channel is established.

The broadcasting type of intervehicle communication, with restricted range, could be made useful by marking every vehicle (on all sides and ends) with a terse code that would be unique to one in a thousand or so, such as two letters or a three-digit number, analogous to radio call letters. This code would then be used for identifying the addressee and perhaps also the addressor.

TELEMETERING TO THE HUMAN SENSES The human senses are awkward when it comes to receiving a signal with quantitative precision. This is exemplified by the amount of attention required to retain simple information such as weather statistics or game scores. The thermometer is among the most natural indicators, but it has to be communicated in a digital code, which is far from natural. There seems to be a need for evaluation and optimum utilization of the senses for this purpose.

The eye can sense a prodigious amount and variety of information, in the form of relationships, but is notably deficient in quantitative observations, such as ratios of distances or intensities, and even more so in sensing their absolute values. The ear, while subject to comparable limitations, is peculiarly skillful in one function, the sensing of frequency changes and the appreciation of frequency ratios and patterns. It is suggested that this quality of the ear may be the one natural sense that has evolved to the highest degree of quantitative precision.

A few years ago, the writer took advantage of this ability of the ear as an aid in a scientific demonstration to an audience. He wanted visual attention to the manipulation of a directive antenna, so such attention was unavailable for the meter showing the response of the antenna when pointed in various directions. The amplitude of sound gave qualitative but inadequate sensing of the variations to be demonstrated, especially a deep null in the pattern. Instead, the amplitude of response was presented by an audio tone whose frequency was made directly proportional to signal power. This enabled the sensing of variations from ¼ db to 30 db, by the logarithmic frequency response of the ear. (The lower value is a half tone on the musical scale. The upper value is based

[1] H. A. Wheeler, "Radio-wave propagation in the Earth's crust," *J. Res. NBS*, vol. 65D, pp. 189-191; March-April, 1961.

on 3 to 3000 cps, the lower frequencies being made audible as pulses or clicks.) The result exceeded the capability of any other form of demonstration indicator, and left the visual attention where it was needed. The patterns was easily remembered as a tone pattern, although still difficult to describe in such terms.

A similar method apears to be in use for translating the heart pulse wave to an audible sound; here the quantitative requirements are less severe.

In any particular situation, the capabilities of sight and sound should be evaluated, and a conscious selection made for best utilization of both. The future will see increasing appreciation of this need, followed by careful study and the adoption of optimum techniques. This is closely related to the problems of efficient education.

PRIVACY BY MATCHED-NOISE JAMMING In military communications, it is recognized that the most effective jamming can be accomplished by noise having a frequency spectrum similar to that of the signal to be jammed, but the noise having a random character. The same principle may be applied for positive benefit in another situation, as follows.

There are certain conditions in which a quiet background serves to emphasize any sound that may occur. One example is a large, quiet office in which several individuals are doing desk work; any conversation, especially by phone, becomes a disturbance to everyone in the room. Another example is a quiet room for sleeping; the sense of hearing becomes so acute that a moderate sound is cause for alarm.

It is expected that the principle of jamming will come to be used as a countermeasure for these conditions. Ideally, a recording of the sounds to be jammed would be made for every day (or night) over a year or so, then all days would be superimposed to form one day of noise. This sound, reproduced at a level comparable with the level of the sounds to be jammed, would form an ideal background for the otherwise quiet office or bedroom. It is likely that a few types of noise would suffice to enable an effective choice for any particular location. The noise causes the ear sensitivity to accommodate to the expected level and character of disturbing sounds. This plan is most effective (not too monotonous) if there is some appreciable fluctuation of the noise level over the time cycle of the recording to be utilized.

Such a background of sound appears to offer also some therapeutic benefits in the reaction of the nervous system. In an extreme case, a very loud noise have been found useful as an anesthetic, and this application will come into wide use in the near future. Here there is naturally a search for the sounds that will be most effective; in other words, that will have the greatest ability to confuse the pain signals.

IDENTIFICATION OF HUMAN CHARACTERISTICS There is a rule which says, one does not know something if he cannot describe it to someone else. It would be closer to the truth if we were to say merely that one does not know how to describe this something. The most striking example of this frailty is our inability to describe the human characteristics by which one is able to distinguish thousands of his acquaintances, even those whom he seldom sees. Any word description may apply to 10 or 100 of the individuals that one is able to distinguish at sight.

It is interesting that we resort to fingerprints in our attempt to find a simple and identifiable pattern that has enough combinations to enable unique descriptions of perhaps a billion individuals.

The time is ripe for a systematic development of concepts and words that will enable one to appreciate and to describe the characteristics of another individual. Some progress is made necessary by the increasing number of acquaintances one is likely to acquire in the closer packing of our population, with respect to both residential and occupational communities.

This result would be achieved by an intensive research program, utilizing the current concepts of information theory. The greatest difficulty would be that of any new language, since the requisite new words will lack any background of usage or familiarity.

Communication Spectra by the Wholesale— 2012 A.D.

Estill I. Green fellow IRE

Although the oracle at Delphi is long since stilled, the art of divination still seems to possess an irresistible allure for mankind. But at least, if mantic endeavor is inevitable, there is something to be said for a half-century forecast, since such an interval confers a fair degree of immunity to after-the-fact checkup by present readers. On the other hand, the technically sound basis for a shorter range forecast, *i.e.*, that the near-term future behavior of a dynamic system is roughly predictable from a knowledge of its immediately past behavior, no longer applies.

One way of illustrating the infeasibility of an accurate 50-year forecast, in the face of rampant technological change, is merely to take an equivalent backward look. It was just about 50 years ago that H. D. Arnold and others were starting research on the introduction of a high vacuum in De Forest's audion. Who then could have foreseen that this device would usher in the great era of communication and electronics—an era that was to create the miraculous networks of wire and radio that now speed communication near and far, the electronic computers that are metamorphosing business operations and resolving hitherto unassailable problems, the incredibly accurate control systems that steer missiles on their way and place satellites in orbit, the broadcast systems with their as yet largely unrealized potential for education and culture, and all the other electronic contrivances that impinge in one way or another upon our society?

How much less amenable to prediction must be the advance of the next half-century! No abatement in the pell-mell rush of electronic discovery and innovation is in sight. Past progress will beget future progress. Technological change bids fair not just to continue, but to accelerate. No more could a Tennyson say: "Science moves but slowly, slowly, creeping on from point to point."

In this juncture, we shall do well to limit our prognostic effort by selecting a single windmill for tilting. The one chosen herein is the matter of guessing what sort of mechanisms will be used to provide the bandwidths required for continental and transoceanic communication 50 years from now. A guess as to the magnitude of these bandwidth requirements is clearly the starting point.

Let us look first at intracontinental communication, and specifically at the bandwidth needs for all forms of intercity long distance transmission. What factors might be pertinent? Population for one, number of telephones for another. The best guess seems to be that in the next 50 years U. S. population will about double, while the number of telephones may triple. Neither of these has great significance, however, since the demands for long distance communication will go up much faster, as has been true in the past.

The demand for telephone circuits will almost surely experience some degree of saturation. In all probability, however, this will be far more than offset by large demands for transmission of visual material, either separately or in association with telephony. A further augmentation will result from needs for data transmission. Putting everything together, it appears not unlikely that total bandwidth need on important intercity routes may go up a hundred times, which would mean a baseband requirement ranging from 50,000 to 100,000 megacycles on each of a number of heavy traffic routes.

How will such a vast bandwidth be obtained? Probably not from present or future coaxial cable systems, since these seem likely to prove uneconomical except for special purposes such as as feeder links, entrance to cities, etc. Neither will it come from present or future microwave relay systems. While these will play an important role for many years to come, their capability for handling bandwidth growth will be seriously curtailed by beam congestion within 20 years or so. The usable microwave spectrum is finite, and will be subject to increasing demands for services other than point-to-point long distance transmission. Moreover, microwave systems will probably be more expensive than future alternatives.

What then? Only two possibilities suggest them-

E. I. Green is Consultant, formerly Executive Vice President of Bell Telephone Laboratories, Inc., Murray Hill, N.J. *(Received May 29, 1961.)*

selves. The first is the waveguide system now being developed specifically to provide increased information-carrying capacity on major overland routes. The probable transmission medium is a hollow copper pipe, about two inches in diameter, acting as a guide for millimeter waves traveling inside. The system will employ the circular electric transmission mode (TE_{01}), which has the admirable property of decreasing attenuation with increasing frequency.

As now visualized, the gross frequency band transmitted inside the waveguide may extend from 35,000 to 75,000 megacycles, *i.e.*, a gross band nearly four times as wide as the entire band now employed for radio communication. This vast frequency spectrum will not be vulnerable to fading, atmospheric effects, interference, or other troubles that beset the radio, nor will it be subject to conflicting public demands and consequent governmental restrictions.

Within each waveguide system comprising one pipe for each direction, a handling capacity for as many as 200,000 voice circuits or 200 television circuits is anticipated. Present techniques require that for amplification at intermediate repeaters the waveguide frequency band be subdivided into a number of separate broad-band channels. However, it seems logical to expect that in 50 years the art will have progressed to the point where the entire waveguide band can be amplified, without subdivision, in a single amplifier.

A possible alternative, a scheme vaguely reminiscent of Alexander Graham Bell's "Photophone," has recently emerged through the development of a continuously operating gaseous optical maser. In contrast to previous optical masers, the new device employs a mixture of helium and neon gases through which an electrical discharge takes place in somewhat the same way as in a conventional neon tube. This discharge excites the helium atoms to a high "metastable" energy level. Collision of neon atoms with the excited helium atoms results in a transfer of energy to the neon atoms, which can then be stimulated to radiate their energy in a continuous stream, and through successive reflections a narrow output beam of coherent infrared waves is obtained.

A system employing this or some other yet-to-be-developed type of maser might conceivably be used to derive an enormous bandwidth for long distance transmission. The light beam, suitably modulated, would be shielded from outside effects (rain, fog, refraction due to inhomogeneities) in a gas-filled or evacuated underground pipe. The system might comprise a number of repeater sections each several hundred miles in length. Each repeater section might consist of a number of straight sections of perhaps a mile or less, with passive means (*e.g.*, confocal reflectors) for redirecting the beam at each junction.

Only a few watts of transmitted power, well within the capabilities of a gaseous or solid-state maser, would be needed in each repeater output. The over-all system bandwidth might be millions or tens of millions of megacycles, in contrast to thousands or a few tens of thousands for the waveguide system. The problems that would have to be surmounted in such an optical maser system are formidable indeed, but seemingly not insuperable. On the other hand, the millimeter waveguide system has been under exploration and study for many years, and the problems are now clearly defined. It is conceivable that over an interval as long as a half-century both systems may come into use.

And now what about transoceanic communication 50 years hence? Again we must first guess at bandwidth needs. Here uncertainty is compounded by the fact that it is only in the last five years that circuits of high quality have been available across the Atlantic in adequate numbers. The demand for transmission of visual material is problematical. Military needs bulk large for the immediate future, but should level off over the longer term. Still and all, the tidal wave of world communication is rising fast. A good guess might be that by 2012 A.D. some 2000 Mc of bandwidth will be needed between U. S. and Western Europe (including Great Britain), and possibly twice that much between the U. S. and the rest of the world.

To meet this need, there will be, first, ocean cables with submerged repeaters. Up to now these have been quite limited in bandwidth. But new single-cable systems of much greater circuit capacity will soon be installed. There is little question of the eventual feasibility of a cable with transistorized repeaters, which could easily give a frequency band of several megacycles or more.

Meanwhile, another kind of transoceanic communication has taken on great scientific and popular appeal, to wit, radio communication via satellites. Spurred by the success of initial experiments, satellite communication has taken on some of the features of the whole space extravaganza. Satellite systems possess real attraction, however, for obtaining wide-band channels for telephony, television and other services, for providing direct links to landlocked countries, for diversification of techniques, etc.

For transmission reasons, repeatered satellites (*i.e.*, equipped for receiving and retransmitting) seem likely to win out over reflecting satellites. There is much theoretical attraction in a repeatered satellite placed in an equatorial orbit about 22,300 miles above the earth, with precise station-keeping equipment so that it would remain stationary above a fixed point on earth. However, this poses extremely difficult technical problems, as well as introducing a round-trip delay of about 0.6 second, which would make it unsatisfactory for telephony. Accordingly lower altitude satellites (2000-7000 miles) are most promising. A group of 25 to 50 might provide a world-wide communication network.

The economics of such a system are largely conjectural. Much depends on the life of the satellite equipment under exposure to space radiation and other hazards. Nevertheless it seems clear that low-altitude active satellite systems will be in experimental operation within a short time.

The preferred frequency range for satellite communication is from about 1000 to 10,000 Mc. Even with huge horn-reflector antennas for transmitting and receiving, and low-noise masers for amplifying the faint received signals on the ground, very broad-band modulation systems will be needed to obtain acceptable signal-to-noise ratios. Since there is only one electromagnetic spectrum available for radio transmission, avoidance of interference between space systems and from terrestrial systems into space systems will introduce grave problems.

All things considered, it is by no means unimaginable that within less than a half-century the congestion in the microwave range will leave no room for further exploitation by space communication. What then may be imagined as the next transoceanic communication medium? In the fact of necessity, some way will be found to multiply frequency spectra at will through the use of cable techniques. Maybe the answer will be an adaptation of the overland waveguide system, or in other words, a "waveguide cable" with the whole band handled by one amplifier at each repeater point.

In the long run, a "sheltered" transmission medium is almost certain to win out over radio for both continental and transoceanic communication. In view of the ever increasing importance of communication to world affairs, it is fortunate that such promising ways of handling future demands are in sight.

A Short History of Electromagnetic Communications

JOHN W. COLTMAN FELLOW IRE

By the 1960's it appeared that electromagnetic communication systems had reached tremendous heights of sophistication. Starting with the crude noises of the spark gap, laboriously coded by hand from the alphabet handed down by the Phoenicians only a few millennia ago, the uses of the radio spectrum burst forth in the course of half a century to encompass a bewildering maze of services. Coded and voice messages were transmitted over all the world, radar beams scanned the horizons, information was brought back from transmitters deep in the reaches of space, radio was everywhere, and the last decade had culminated in the miracle of the nightly repetition, in millions of homes all over the land, of the Western.

But even so, the electromagnetic spectrum hadn't really been tapped. The services extended from about 20 kc to 40,000 Mc, a little over 6 decades, and the upper end of the band was wastefully squandered carrying television signals that from the point of view of information theory (to say nothing of the point of view of culture) were highly redundant. Beyond lay tremendous reaches in the frequency domain. From 10^{10} to 10^{13} cycles were the submillimeter waves and the far infrared, essentially a trackless waste, traversed by an occasional expedition which noted the nature of the area, but did not stay to colonize it. It contained, of course, half as many decades as all of the communication spectrum had exploited until that time, and in terms of information bandwidth was thousands of times larger. Its development was hampered partly by the opacity of the atmosphere, but more by the inability to apply in the tiny dimensions required the techniques characterizing longer wavelengths. This area was finally opened up from the other direction by extending the techniques originating in the optical region of the spectrum. Now here was rich territory indeed—a narrow slice of the electromagnetic spectrum originally occupying scarcely an octave, but possessing an overwhelming advantage in that it was here that man's own receptor, the eye, operated. This superb mechanism had a receiving aperture only a few millimeters in diameter, yet was capable of absorbing and classifying information at rates far in excess of those characterizing the communications systems of the sixties; indeed, the latter had succeeded in presenting to the eye only a tiny and crude imitation of visible scenes, and had used up the then available communications spectrum in huge gobs while so doing. It was in the optical region of the spectrum that physics had its early triumphs; indeed, it is fair to say that our modern concept of the physical world rests primarily on a foundation established in the study of the

J. W. Coltman is Associate Director of the Westinghouse Research Laboratories, Pittsburgh, Pa. *(Received March 29, 1961.)*

propagation of light and its emission by atoms and molecules.

It was through these latter phenomena that electronics and communications became part of, and, indeed, essentially took over and engulfed, the art of optics. The key to the operation was the use of cooperative phenomena among atoms. Now all of the power of the electronic signalling techniques was combined with that of conventional optics. The provision of stable local oscillators at optical frequencies, together with the development of nonlinear media made possible the conversion of frequency, with all its attendant advantages.

At the same time, the precise directional properties available with the new frequencies were exploited. A single "antenna" (that is, a lens) of the most modest dimensions was capable of receiving simultaneously from thousands of points in space, and of keeping each transmission separated into its own channel. The frequency discrimination was, compared with the prior techniques, fantastic—it being quite possible to contain in a single millihue (the number of millihues in a wavelength interval is defined as 1000 log λ_2/λ_1) in the orange, for example, about a quarter of a million of the old "television channels." The invention of the brilliator in 1972 made possible the amplification of weak images without destroying the phase information, and by the 1980's, most of the point-to-point communications networks were in the optical or suboptical bands. The time-sequential scanning of images was abandoned, and television broadcasts were carried throughout by optical techniques, the time phases and angular dependence of the waves being preserved, duplicated, and amplified wherever necessary in the system.

The tremendous success of the method of coordinated atomic radiation brought forth an intensive search for means to extend the technique to even higher energies. Gradually, there evolved mechanisms operating in the ultra-violet and X-ray regions, and while these have been restricted to certain aspects of space communications, they have, in principle, added to the spectrum another six decades. The introduction of quantum coding, in which the message is carried by high energy quanta, distinguished by slight variations in their wavelength, proved to be of great value in overcoming the natural background radiation of space.

As every school boy knows, the most exciting project of the new millennium came as a result of the discovery in 1989 of the nuclear transmissions from the neighborhood of Sirius. These gamma-ray emissions were detected by UN Observatory A36, operating in a 24-hour orbit and carrying a highly sensitive gamma-ray brilliator operating in the Mn^{54} band. The intelligent character of the transmissions was apparent rather early, and while they have appeared only at scattered intervals, sufficient Doppler observations have been made to establish the period of the orbit of the planet from which they came, and even to demonstrate the distortion of this orbit by the presence of the dark companion of Sirius.

In 1996, a team of engineers and scientists, with the aid of superconducting magnets producing fields in excess of 1 million oersteds, succeeded in generating coherent gamma radiation in directed beams. Since then we have been making periodic transmissions to Sirius. If, as appears probable, the Sirians were searching for a reply, they should by now have received most of the transmissions, which have a 9-year travel time. Two years from now, it is theoretically possible that we will receive their earliest reply, and the world is eagerly awaiting this first exchange, across the depths of space, between the 2 worlds now known with certainty to exist.

Before submitting the above contribution, the author asked an associate to read and comment on it. In response he received the following "antithesis" which is included here in the belief that it will provide readers with a sobering, as well as amusing, treatment of the same subject.

—The Editor

The Spectrum Problem—Looking Back From 2012 A.D.

E. A. Sack

It seems so simple, in retrospect, that the older members of the profession still refuse to completely accept the solution. That their supreme efforts to conserve bandwidth and to develop new portions of the spectrum were really unnecessary is a revelation from which complete recovery is impossible.

Late in the 1960's, a few nonconformists were proposing that the so-called spectrum shortage was an organizational rather than a technical problem, but their theories were largely ignored by the scientifically oriented communications agencies. The major premise of the nonconformists, that one should first determine whether information really needed to be transmitted, seemed so undemocratic that their clearances were subsequently lifted in the interests of national security.

The reader will recall that science was in high fashion during the quarter century preceding the second millenium. The importance of the professional scientist had grown to the point where a doctorate in physics was a prime political attribute and both of the presidents during the period 1972-2000 were Fellows of the AEC.

But the law of trend and counter-trend is universal and while the technologists labored night and day to find ways in which to transmit ever-increasing gobs of information to the public, to each other and to nobody in particular, the humanist gathered in non-NSA supported workshops to accumulate, in anti-scientific manner, the evidence that the spectrum problem was purely a matter of inadequate legislation. Their philosophy evolved into the now-famous Sideband Theorem which states, "Technological solutions to nontechnological problems are in quadrature to progress."

The journals of the humanists during the period show the development of their deliberations. For example, in 1980 they proved that all of the AM broadcast stations played the same 20 records 90 per cent of the time. Their proposal that only one national broadcasting station using only one 10-kc channel might do just as well received passing note.

A similar study of the television spectrum in 1982 indicated that 85 per cent of the TV programs broadcast by 521 stations on 126 channels contained but six basic program plots. It was proposed at the time that these plots be provided in a small local memory to the home TV fan who could select them at random for the amusement of the integrated family group.

A secret study of the communications needs of the national defense establishment was even more revealing. Eighty per cent of all messages had to do with orders which were countermanded by subsequent messages and hence never needed to be transmitted at all. The remaining 20 per cent dealt with information which would be ignored at the receiving end and was, therefore, 100 per cent redundant.

An interesting example of the thoroughness of the investigation of the humanists was the study of the Public Service spectrum needs. At the time one city was demanding 5 additional channels for its police radio; its average transmitting time on each of its existing channels was 1 minute per hour. A typical communication was: "Car 5 stopping for a sangwich." "Repeat Car 5." "Car 5 stopping for a sangwich." "Say again Car 5." "Car 5 stopping for a sangwich." "OK, Car 5 stopping for gas."

As time went on, the battle between those who felt that more spectrum space was needed, and those who felt that less information need be transmitted, grew bitter indeed. But the outcome is familiar to all. The Communications Act of 2008 ended the spectrum shortage. Its major provision, of course, is that all information be first transmitted to Washington by surface mail for review by appropriate committee. If approved, the message is transmitted over the National Transmitter in a priority determined by the priority review board. The completely redundant information which was formerly broadcast on radio and TV is mailed out on tape, to all who care, on a monthly basis.

Unemployment among communications scientists has been largely eliminated by increased need for mailmen. Military personnel who became surplus by the impossibility of a really worthwhile war under the system have yet to be completely absorbed by the economy, but the matter is receiving further study. Public Service officials have complained that they no longer know when Car 5 is out for lunch, but one executive committee has proposed that the cops be given dimes to call in from the pay phone behind the counter. This may work.

The Full Use of Wide-Band Communications

Sir Noel Ashbridge FELLOW IRE

All radio engineers have an ultimate goal in mind: the day when wide-band communication can always be established between any two points on the earth's surface with absolute secrecy. He looks to a future when distance, season, and time of day are immaterial, and the business man and the house dweller are able to enjoy all that wide-band communication could conceivably offer. With this ambition in view, a picture of things which might be done within the next fifty years can be roughly envisaged.

In the future the urge to convey unlimited information over the globe will be sufficient to free the enormous sums of money which are needed to solve the problem. The military needs it, business needs it, and the man in the street is always prepared to pay for more amusement.

The solution might even be dull, the piling of complexity upon complexity, in fact, the gradual harnessing of known methods into huge co-ordinated systems. But the outcome might help to balance up the whole lopsidedness of the civilized world. It is not only the accomplishment but what do we do with it that is the exciting prospect.

Sir Noel is Director of the Marconi Wireless Telegraph Co., London, England. *(Received June 12, 1961.)*

Wave propagation research workers seldom suggest that they are on the brink of revolutionary discoveries. More hard won knowledge, leading to better systems, is about all they expect from their work. We might expect therefore that the two major methods of communication will be satellite radio and guided wave (H_{01} mode), substantially aided by wide-band microwave relays and tropospheric scatter. HF ionospheric communication—narrow band but less costly—may remain as an economic necessity, but we can hope with some confidence that the growth of satellite, and pipe line systems will give some freedom of maneuver for the regimentation of this frequency band.

It will be the age of computers and one can foresee their role in world communications, tying together the several systems and selecting the best path or sequence of paths for a particular message.

If we assume that the problem of wide-band global networks will have been tackled and solved, there remains the enormous problem of local distribution. No doubt in fifty years time the technique will exist, but the manifestations of it are not likely to be commonplace. Except in isolated small communities, radio will not do; we must resort to the pipe, and this means additional pipes to every inhabited building.

The capital outlay will be enormous, but the potential benefit will be worth it. The problem lies with the pipe manufacturers, and it is surely reasonable to suppose that town planners and the suppliers of essential services—electric power, gas, sewers, and so forth—will have gotten together and integrated their methods of providing underground-fed facilities. It is unthinkable that power and communications should continue to be carried by a tangle of overhead wires and posts. The sensibilities, if not the common sense of future generations, will demand an organized approach to the supply services of buildings. It is surely reasonable therefore to look forward to ducts, or even composite pipes carrying all services, including guided wave transmissions, as one co-ordinated engineering operation.

If this comes about, and our grandchildren can handle, for example, the H_{01} mode more economically than we can at present, the way will be clear to give all sections of the community those facilities which have not materialized during the "teenage" of communications.

The development of facsimile—so far limited by bandwidth—must go on until it can give as much detail as the eye can discriminate: print a newspaper with no detectable difference from the original, reproduce anyone's handwriting in any language, and so on. The development of thin-film phosphors, xerography, and electroluminescence point to a near certainty that extremely high definition reproduction of documents, letters, forms, and so on will be greatly reduced, and eventually outmoded. By the year 2000 all this could have gone a long way.

Another information system that will certainly be replaced is the noisy, unreliable, and cumbersome teleprinter with its devilish fault of printing the wrong letter if its input information is not perfect. A cleaner transmission system, perhaps using a little more bandwidth, will make way for the development of an all-electronic teleprinter: neat, silent, and free from guesswork.

A world-wide wide-band network with local extensions would not recover capital expenditure without a contribution from the man-in-his-home. These days he is not allowed to live simply, he must have telephones, television, up-to-the-minute news, and so forth. In fifty years time he will live even less simply. He will, whether he likes it or not, become more and more exploited, but perhaps in a way that will be more satisfactory to him. For example, wide-band guided-wave local services could pump into the home all the information, amusement and propaganda anyone could possibly want. Television would become real picture viewing: "glossy magazine" detail in color with no interference. Telephones would be visual as well as aural: news continually in print, dial for music, dial for a picture showing the state of his bank account, and so on.

The vagaries and limitations of radio wave propagation and the struggle for bandwidth are the real barriers to progress. The more we struggle to improve the capabilities of radio communication the more the orthodox methods split at the seams. Television development is at a standstill; we accept it for the action in the picture, not the amount of information. A release from the bondage of limited bandwidth must come to make drastic innovation possible.

In case of picture electronics whether it be television, document reproduction or facsimile, it seems that the binary digit will lead development, and inventors will discard the compromise of scanning in favor of the element screen or matrix, where each element is "bit" controlled and therefore under control. As for the screen itself, the problem will no doubt be solved in several ways. More knowledge of the behavior of phosphors, particle penetration, laser, thin film techniques make the prospect full of promise.

This emphasis on the lavish use of bandwidth does not envisage that our grandchildren will be able to use it in a prodigal way, but a narrow road remains a narrow road, nothing can be done about it until it is in effect widened. Then schemes for traffic control and flow can be put into operation, but first there must be enough room for maneuver. And so it is with communications. In fifty years time, man will be able to communicate almost any piece of information to anywhere, mainly because he has been able to provide himself with the basic tools with which to do it. This must include sufficient bandwidth to make possible the more effective exploitation of information theory.

Wide-band Communication into the Home

W. D. LEWIS FELLOW IRE

Electronics has had a spectacular impact on long-distance communication during the last fifty years. A corresponding impact on the techniques of local two-way communications to the home is only beginning to be felt. Powerful forces are working to change this. By 2012 A.D. these forces will have created equal or greater changes in the techniques of local communications and in the services available to every household.

I would be rash indeed to attempt an accurate chart of the next fifty years of electronics. Applications are more easily predicted, for they are based in part on the relatively static foundations of economics and human need.

Certain technical trends are clear. These alone will create a revolution, though perhaps a gradual one. New discoveries will probably hasten this revolution, and they cannot slow it down.

What are these trends? Briefly, they are towards higher frequencies, cheaper electronics, greater automation, digital transmission, and integrated design of systems.

The optical maser is only one recent victory in a triumphant campaign that has won the triode, the klystron, the traveling-wave tube, a host of transistors and ferrite devices, and the Esaki diode. During the next fifty years, local communication will profit richly from these victories and others unmentioned or unborn.

Today, and for the foreseeable future, device costs are rapidly being pushed down. This economic revolution will stimulate communication in several ways. A notable example is two-way television in every home—a technical possibility now but not an economic one. An abundant supply of cheap, high-frequency components and assembly techniques will make the "videophone" as commonplace in 2012 as the telephone is today.

An ordinary pair of wires in a telephone cable can carry more information in telegraphic or binary digital form than it can in analog form. So can a circular electric waveguide. In both cases this is because the greater resistance of "pulse-code modulation" against noise and interference more than makes up for the greater bandwidth it requires. Furthermore, PCM can be amplified and regenerated many times without degradation. As electronics becomes cheaper, it will become profitable to convert messages to digital form for transmission over shorter and shorter distances. They will travel in this form economically across the street, across the city, or around the world.

The more economic and flexible electronics of the future will be used increasingly to save the copper, real estate, and manpower required to provide communication. For example, wire or coaxial sizes can be made much smaller by placing amplifiers closer together, probably building them in at the factory. Many small switching units will reduce the copper even further. Extremely cheap microwave and millimeter amplifiers and oscillators, already promised by the new family of negative resistance diodes, will in many places make it economic to eliminate the local wire altogether.

Communication equipment inside the home will come in many forms and will range from large visual consoles giving a pair of users a sense of immediate presence through easily portable visual models not requiring

---

W. D. Lewis is Executive Director of Research Communications System Division, Bell Telephone Laboratories, Inc., Murray Hill, N.J. *(Received July 5, 1961.)*

connection to the wall down to pocket sets which will have a circumspect lack of vision and may be carried continuously if desired.

Conversation between two people is only one of many modes of communication. Several others seem sure to become realities by 2012 A.D. For example, it seems likely that most written communication will be transmitted electrically and not by transportation. The private citizen will have electrical access to machines of all kinds, for example, centralized data-processing units for banking. Electrical access to reference libraries will surmount any mechanical barriers to recorded information or, for that matter, entertainment. Finally, and perhaps most important, it is possible that a combination of visual recording, teaching machine techniques and human intervention where needed, will make available the best education in any subject, anywhere, any time, to anyone who wants it.

Communication and Navigation

HENRI BUSIGNIES FELLOW IRE

COMMUNICATION Satellite communication will be in current application not only for point-to-point transmission connecting all countries of the world for telephone, data facsimile and television, but satellite broadcasting will also be operating with automatic and available selection of spoken intelligence in all principal languages. Television broadcasting from satellites will cover very large areas. Translation service will be made available for international communications first by human translation—next by electronic translation.

The world-wide satellite system will provide electronic mail service for the written word while electric typewriters will reproduce typed text for international transmission.

On the very dense transmission links for relatively short distances (up to a few thousand miles) on land, waveguide transmission will replace microwave space transmission and will give a bandwidth capacity that will make available thousands of television channels; this will permit transmission of masses of data, of newspapers and printed material and will make available the phone vision to a large section of the subscribers at a reasonable cost. The carrier used in the waveguide will be a coherent beam of light.

The distribution of newspapers and magazines will have been considerably modified by the introduction of electronic transmission, through which subscribers will receive them. The paper (or rather its equivalent) will be made locally automatically from a liquid. In many instances the reader will not want a permanent record and will avoid the paper altogether by receiving the information on a storage tube, one page at a time at his desire.

The subset as we know it will still be in use but in many instances will have been replaced by either a fixed speaker-type form of wall, or table model, or by a very light unit carried by the user, of course, "wireless."

Mobile radio will be common, in the form of a miniature "wireless" unit carried by the user or installed in cars, buses, and other means of transportation. Ships at sea and airplanes will communicate with shore by satellite, and telephone service will be provided to the traveler. He will be able to get an automatic report on how things are in his home. Push button dialing will be possible in many countries using the world-wide satellite system.

NAVIGATION True all-weather flying will be possible with an accurate knowledge of position by each aircraft within a fraction of a mile. Anti-collision will have been fully developed on two bases: through automatic reporting of position to the ground and high-speed computing of routes and conflicts, and through direct radar or equivalent exploration of space around each aircraft with high-speed warning on possible route conflict with either automatic or manual route correction. True zero zero visibility landings and take-offs will be common, adapted also to vertical take-off airplanes. The methods of ILS will be of the same general type but the airplane controls will have been much improved and the accurate position determination will permit the airplanes to come in (or leave) in perfect order without stacking.

H. Busignies is Vice President and General Technical Director of the International Telephone and Telegraph Corp., New York, N.Y. *(Received June 30, 1961.)*

Communications and Electronics—2012 A.D.

PETER C. GOLDMARK FELLOW IRE

In this 100th Aniversary Issue of the PROCEEDINGS OF THE IRE, May, 2012, we feel privileged to discuss a very interesting article contained in the 50th Anniversary Edition dated May, 1962. That article, entitled Communications and Electronics—2012 A.D., was written by Fellows of the IRE, and it may interest some of our readers to go through a check list and see to what extent our 1962 colleagues were able to predict the state of the art today.

One of the contributors, Peter C. Goldmark, forecast that practically all point-to-point long distance communication would be by microwaves in the millimeter region. This included global television relay via satellite operations, as well as global telephone and facsimile service. As to broadcasting, the prediction stated that by 2012 the FCC would have approved the ultra-high frequencies for exclusive use by TV broadcasting. Actually, as we know, a final decision concerning the UHF band for all TV broadcasting is imminent, following the completion of the FCC technical staff's extensive experiments begun in 1961. Our radio broadcasting, according to Goldmark, would be entirely via FM, using the same portion of the spectrum as in his day. If he were still with us, Goldmark would be surprised to find that transatlantic telephone is still operating via co-axial cable and repeaters buried in the Atlantic. The satellite relay schemes forecast, as we all know, did not quite pan out, with so much of the nation's resources devoted to competing with the Russians in the race for reaching extraterrestrial bodies.

Anyone in 1962 with some imagination should have been able to predict our moon-to-earth citizen's radio service, operating so effectively in the millimeter citizen's band, but it would not have been easy to foresee our tremendously efficient high-power solid-state wrist watch transceivers and plasma antennas.

One item not covered in the article, though subject to considerable stir 50 years ago, is stereo broadcasting. The limited space and manpower available to radio and phonograph dealers in the middle of the last century made it imperative to compress reproducing instruments to the smallest possible size. As a result, while techniques were in existence to generate stereo signals on records, as well as over FM and AM radio, there were hardly any instruments made which took advantage of the stereophonic phenomenon. Consequently the monaural effect was discovered, requiring minimal speaker separation in home instruments.

One word about a form of entertainment which, in the middle of the last century, led to a number of broken homes and eardrums: We refer to Hi-Fidelity, then a thriving industry. Then phonograph instruments were huge boxes generating a great amount of heat, and were called Hi-Fi amplifiers. Some of them measured 16-in wide, and the audio signals extracted from the phonograph records[1] were amplified by devices called vacuum tubes (which also performed as thermionic ovens in bookcases and other enclosures). It is hard to imagine that our ancestors did not have the compact Fi[2] amplifiers contained in our most recent home music center, which measures only 4-in wide, 3-in deep and 2-in high.

The 50th Anniversary article we are reviewing here carries in its title the word Electronics, more commonly known today as Solid-State Physics.

It is interesting to read in the article how our present TV cameras were envisaged by the 1962 scientists. (It was presumed that by 2012 all TV would be in color and three dimensional.) It was predicted that the scene to be transmitted would be sensed with a tiny adaptation of a device which they called a camera tube. The predicted version would contain a target plate which was color sensitive and would be able to generate a video signal containing all color and stereoscopic information of the scene scanned, thereby eliminating all registration

P. C. Goldmark is President of CBS Laboratories, a division of Columbia Broadcasting System, Inc., Stamford, Conn. *(Received September 7, 1961.)*

[1] These were discs made of plastic where sound grooves were pressed into a number of fairly coarse grooves, on the average 250 per inch. The record rotated on top of a box full of moving parts, and the whole thing was called a record changer or player.

[2] It may be recalled that the word "Hi" was dropped in the year 1994 by a majority vote of the IRE Standards Committee.

and color shading difficulties plaguing color TV of the 1960's. Goldmark would be quite surprised if he could see our TV cameras of today. Our 1-in diameter, 2-in long solid-state camera units, combining the multicolor-sensitive, scanning and amplifying elements in a number of evaporated layers, are far ahead of what he predicted.

Experiments with scanning by pulse techniques began in the 1960's; it was thought that these techniques would be applied to what they called the "picture on the wall." This type presentation was indeed the forerunner of our TV receivers today. We are quite proud of our large experimental digital color screens which are almost flush with the wall and produce the brilliant three-dimensional images using our 120-megabit standards. Should the proposed new standards be agreed on, as a result of the recent FCC color hearing, a number of manufacturers will be in a position to supply color adaptors for the existing black and white receivers. (It is strange that we still talk about adaptors when all through the history of communications adaptors never proved to be accepted by the public, and possibly never will.)

One more comment about the 1962 article. When talking about computers nothing was forecast about our predictors. While using the fundamental principles on which computers were based, predictors represent such radical advances over the conventional computers of the 1960's that one cannot very well blame our ancestral Fellows. Now we can not only store and process complex information, but we can predetermine the optimum course of action to be taken based on the information furnished. National policies, business decisions, family issues are now settled on the basis of the predictor's output, thereby leaving little likelihood for chance and error entering into decisions. As a result of this technique we now enjoy peaceful coexistence, giving us a feeling of well being and security.

On the whole, we think that Goldmark, in 1962, could have done a better job of properly forecasting "Communications and Electronics—2012 A.D."

Communications Throughout the Solar System

W. H. Pickering FELLOW IRE

Within the next 50 years, rocket propulsion will have advanced to the extent that space travel throughout the solar system is a technical possibility, and it is virtually certain that by 2012 A.D. manned spacecraft traveling to Mars and Venus will no longer be a novelty, and manned exploration of other planets and satellites will have been undertaken. It is unlikely that man will have been able to venture as far as other stars, but instrumented space probes will have been sent far beyond the solar system. Space travel in the vicinity of the earth will include manned orbital laboratories and space stations and manned lunar laboratories. None of this activity can be undertaken without adequate communications, therefore the communications art must progress at least as fast as the rocket art.

The obvious problem is that of distance. On the earth, communications distances of the order of 10^3–10^4 kilometers may be needed. In space, the distances are of the order of 10^8–10^9 kilometers. Hence, system capabilities must be increased by a factor of 10^6 in range. Clearly antenna gains, receiver temperatures, operating frequencies, communication bandwidths must all be optimized for the particular space mission. Equipment operating on spacecraft must be capable of operating reliably in the space environment for long periods of time. By 2012, spacecraft used for distant voyages will carry nuclear reactors for propulsion purposes and for power generation. Electrical power in megawatt amounts will be available for spacecraft operation. Thus, the communications equipment is not likely to be power limited.

Some interesting possibilities arise when the problem of space-to-space communications is considered. With a true vacuum for a communications medium, electromagnetic waves in unusual parts of the spectrum and high-speed particles such as electrons or neutrons offer possibilities for new types of systems.

In 2012 the communications engineer interested in solar system communications is thus going to find the following:

1) A world-wide network of terminal stations having very large directive antennas which can maintain 24-hour a day contact with a number of spacecraft. The

W. H. Pickering is Director of the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Calif.
(Received September 6, 1961.)

stations will use kilomegacycle frequencies, transmitter powers in the megawatt range, and receiver temperatures close to absolute zero. The stations will be interconnected through wide-band data links using satellite relay stations.

2) Spacecraft equipment on both manned and unmanned craft with exceedingly long life and high reliability. The requirements for long life and reliability will be difficult to meet, but will be essential to the success of the space program.

3) Various electronic devices taking advantage of the environment of space, particularly the near perfect vacuum and zero gravity. Thus, the communication system may use vacuum tubes which have no glass envelope and are therefore more effectively built into a circuit. Also, the system may use large antennas, which are assembled in space and are of very small mass.

4) Space relay stations not only near the earth for earth-to-earth communications, but far out in space as beacons and as relays for wide-band communications. For example, in order to communicate with Mars or Venus when the planet is on the other side of the sun from the earth, a space relay station in the earth's orbit, but 90° around the orbit from the earth, will be necessary. Such relay stations will be established in various parts of the solar system as men and rockets travel further into space.

The deep space communication systems of 2012 will indeed keep up with the capability of the rocket engineers to send payloads into space. As man ventures further out into the solar system he will always be in contact with the earth. His communications system will be a vital part of his navigation system. He will use very long-range radar to assist him in his landings on distant planets and other bodies. He will examine the planets by television, not only in the visible, but in other regions of the spectrum. He will report back to earth over wide-band channels. I have no hesitation in saying that the communications art will advance fast enough to answer all the demands upon it, except one. It will not be able to communicate by means of signals traveling faster than the speed of light. Hence, communications to space travelers on Mars, for example, will continue to have to put up with time delays of the order of 10 to 20 minutes.

Human Factors in Communications

ALBERT G. HILL FELLOW IRE

Communications in the year 2012 A.D. are not going to be terribly different from communications nowadays. The primary reason for this is the human element. Meaningful communications involve the exchange of information and ideas between people. Certainly in fifty years one may expect significant improvements in technical hardware, but in the final analysis, the development of a new and improved model of the homo sapiens is not on the horizon. The real significance of these scientific and engineering improvements will not be evaluated solely on the basis of speed, capacity, cost, etc., but only with full recognition of both the frailties and the strengths of people.

How do we engineers and scientists help provide better communications? Where must attention be directed to make a really significant breakthrough possible? The answer probably lies in the assistance we can give to the human factors involved. We can talk glibly about discussing our problems with the people on a broad basis or on a person-to-person basis. We cannot, however, improve communications until we improve human relations. We can speak of new technical inventions such as solid-state matrices which will help us tremendously, but only from a technical point of view. These will not necessarily improve communications.

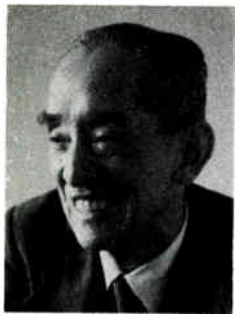
If the above words be correct, it is very incumbent upon us to understand better the human factors in communication systems. This does not mean that we should hire more psychologists and psychoanalysts; we ourselves must understand better all the factors in communications among human beings. Quite often communications between persons sitting in the same room, speaking the same language, are not understood.

We can start, and this is only a start, by exerting a conscious effort toward making our speech more precise whether it be over the telephone, over the radio, or face to face.

A. G. Hill is a Professor of Physics at Massachusetts Institute of Technology, Cambridge, Mass. *(Received May 26, 1961.)*

Multilingual Communication

YASUJIRO NIWA FELLOW IRE



Overseas telegraph and telephone, 50 years ago, needed vast areas in suburban districts to set up gigantic antennas and big station houses where a series of colossal equipment was installed. Nowadays, frequency range assigned to telecommunications is remarkably enlarged and serviceable channels are quite numerous, and at the same time, microminiaturization of components (it is rather funny to use this word today!) and the progress of noiseless amplifiers based on maser principles revolutionally improved the receiving sensitivity. Thus the size of transmitting and receiving devices is reduced to the minimum. Moreover, a number of repeater artificial satellites are going around the earth. We ring up anyone, anytime, and anywhere on the earth, just like a man calling his neighbor fifty years ago, by a handy superhigh-frequency transmitter. He talks to a suitable repeating satellite which, responding to the message, instantly calls his friend on the telephone.

Y. Niwa is President of the Tokyo Electrical Engineering College, Chiyodaku, Tokyo, Japan. *(Received June 12, 1961.)*

Even in the period that electronic techniques had just made it possible to talk to anyone, anywhere, and anytime on a global scale, communication stands on the condition that the people at both ends of the communication channel can read and speak the same language. In the course of the past hundreds of centuries the human language became so differentiated among so many races and nations under different environments and customs throughout the world, that even at the time when a few hours' flight can carry the people from anywhere to anywhere on our beloved planet, the unification of human language is still a long way off.

Again, the advancement of electronics decisively solved this problem. Today principal languages of the world can be translated instantly by the aid of electronics. Our telegrams are simultaneously translated into, and typewritten by, the language of the receiver, and on our overseas telephone the speech is also converted to the language understood by the receiver and vice versa. Moreover, by the results of hyper-miniaturization of electronic parts and appliances the size of the translating machine is also reduced to such an extent that it can be easily used at home and carried by hand. In 2012, we say, "It has become a faint dream of nights long past that at an international conference the participants had to use an earphone to listen to speeches translated 'simultaneously' by man-machinery behind the scenes."

Language, Words and Symbols

HAROLD A. WHEELER FELLOW IRE

The handling of language has been one of the principal functions of communication systems. Electronics had its origin in the transmission of messages and subsequently has contributed many varieties of instrumentation. Information theory now has much to offer for the future, in conjunction with electronic computers and data processing for performing the required operations in analysis and synthesis.

SYNTHESIS OF WORDS FROM SOUNDS The language of human beings is based on the emission and sensing of certain patterns of sound.[1] This language is remarkable in its versatility, especially in view of its natural evolution as distinguished from scientific planning. The languages of different human societies have had much in common, including the remarkable ability of the brain to remember and to associate sound patterns by virtue of some essential characteristics. The identification of sounds with ideas is still an art rather than a science. Even with his tremendous mental capacity, a person requires years of apprenticeship to acquire even a mediocre competence.

It is presumptive that the principal words and phrases can and do convey ideas in a limited number, as in Basic English for example. With the accumulated knowledge of science and technology, we should now be prepared to develop an "efficient" language. By that is meant, one that will be based on the most useful sounds and will associate these with the most needed ideas.

An early stage of such a development should be devoted to a study of all sounds now used in various languages, and further sounds that could be utilized in speech. The next stage might be the selection of sounds that require the least effort and can be identified with least ambiguity (unaided by other senses and redundancy). The permutations of several such sounds in close succession would yield a limited number of words.

A parallel effort should be devoted to cataloging the essential ideas which are required to be expressed by words or word groups. Perhaps this activity has a good start in the major effort that has been directed to the problem of interlanguage translation by machine.

Another parallel effort should be directed to the development of principles for preferred relations between words and ideas. These would range from complete randomness to possible association with experience (by criteria that are not immediately obvious). Perhaps the English language would offer the most helpful associations, though its origins have been largely random.

The final stage of development would then be the cataloging of a new language based on the assignment of relations between sounds and ideas. This language would then have to be tested in isolated "laboratory" groups to determine whether it possesses such requisites as ease of learning and effectiveness of communication. Ideally, several such languages should be tested in competition, since it is unlikely that the first product would satisfy the critical observer.

In experience to date, it has been rather easy to

[1] J. R. Pierce and E. E. David, "Man's World of Sound," Doubleday and Company, Inc., New York, N.Y.; 1958.

H. A. Wheeler is President of Wheeler Laboratories Inc., Great Neck, N.Y. (A subsidiary of the Hazeltine Corporation.) *(Received September 5, 1961.)*

synthesize sounds that the sense of hearing can identify in terms of language. However, it has been difficult to receive speech sounds and automatically to separate their essential characteristics from their incidental peculiarities. In the "efficient" language, the latter operation should be anticipated in the description of the sounds to be utilized. It should be possible to record the sounds of a statement and then automatically to transcribe it to the printed sheet. Such an objective is far beyond the English language; it has been achieved only in the telegraph codes, which are too cumbersome for speech.

GRAPHICAL SYMBOLS FOR SOUNDS Like the sounds of speech, the graphical symbols of writing and printing are also the result of natural selection by generations having no particular training or aptitude for this task. It has been more an art than a science. There is lamentable lack of uniformity in connective writing, which is the most utilized medium in direct recording of thoughts. Indeed, some otherwise competent educators are even resisting the inescapable principle that such writing is a functional common denominator which requires standardization for its purpose of communication.

In the present language, or in a newly developed "efficient" language, it is inherently possible to identify every essential speech sound with a better symbol. Much effort has been devoted to this problem, but to little avail. There has been too much emphasis on retaining the existing symbols, which are time-honored but lacking in logical synthesis.

In developing a new set of symbols, special consideration should be given to some method of transcription, from one "type font" to another or to speech sounds. The recognition of graphical symbols by machine is receiving much attention but perhaps it is time for a new scientific basis for such symbols.

This task is much easier than the development of a new language. Perhaps it should be applied to some existing languages in the meantime, even though its objectives could be only partially realized.

COMMUNICATION WITH ANIMALS As we learn more about our fellow creatures outside the human family, we have more respect and admiration for their intelligence and integrity. One of the most fascinating of current research projects is the investigation of the language sounds of the dolphin, who is remarkable in his athletic skills, learning ability, and aptitude for teamwork, not to mention his scintillating personality.[2] He possesses some of the human traits that evolved in the days before tools and records. Other animals that are remarkable in their development include not only the ape family, but also the elephant, the "big cats," the dog and the otter.

Man has assumed his habitual "superior" attitude toward other beings, and usually has insisted that they learn to respond to his language. The writer suspects that some animals may well regard humans as "backward" because they cannot speak the same language, and can little understand it. Perhaps such animals would not appreciate the compliment of being called "almost human."

Only now are we making real progress in learning the language of some of the higher animals, and the next few years should see much progress in utilizing their languages for communication with them. This experience will be as exciting as a contact with a civilization on another planet, and it can be productive much sooner.

Identifying and describing the language sounds of an animal may be done in a manner similar to the "visible speech" display developed in Bell Telephone Laboratories. Alternatively, for purposes of simple classification, the sounds may be encoded by the "vocoder," which automatically derives a description of sounds in terms of the time variation of their frequency spectrum.[1]

The sound emitted by an animal can be recorded, stored and played back at will in his own medium (air or water). Once we discover the meanings of certain forms, we shall be able to translate between that language and ours. We thereby avoid any need for the animal to learn our language. Some such experiments have been in progress, and this is probably the way of most rapid advancement.

In communication with an animal, a real milestone will be reached the first time that we are able to ask a question and get an answer. Will this be experienced in the next 50 years? It is not too far removed from the now famous lioness, Elsa, who appeared to sense questions and indicate the answers without relying on any formal language.[3]

[2] J. C. Lilly, "Man and Dolphin," Doubleday and Company, Inc., New York, N.Y.; 1961. (Communication Research Institute, St. Thomas.)

[3] Joy Adamson, "Born Free," Pantheon Books, Inc., New York, N.Y.; 1960. Also "Living Free," Harcourt, Brace & Co., Inc., New York, N.Y.; 1961. (Elsa the lioness.)

Processing of Sound

HARRY F. OLSON FELLOW IRE

Sound processing is used to designate the analysis of sound, the conversion to a code and the synthesis of sound from a code. The development of sound-processing systems will provide many new facilities for the reproduction of speech and music; the production of speech and music from codes, sensors and the printed page; and the conversion of speech to the printed page and for the control of machines. Speech, music and the printed page constitute the foremost means for the transmission of information. Sound processing will increase the means for the production, reproduction and use of speech, music and the printed page in the communication of information between individuals and to and from machines.

The most important new communication systems involving the processing of sound are depicted in Fig. 1.

In a system involving speech and music input to a microphone, the information bearing components of speech and music are coverted to a code by means of an analyzer as shown in Fig. 1A.

The synthesizer and loudspeaker of Fig 1B produce speech and music from a code. The advantage of the system of Fig. 1B is that the coded record may be made only a small fraction of the size of existing conventional records for the reproduction of speech and music.

A combination of Figs. 1A and B is shown in Fig. 1C in which the input and output is speech or music and the transmission takes place in the form of a code on a reduced bandwith.

When speech is converted to a code, the code may be applied to an actuator to control a machine or to a typer to produce the printed page, as shown in Figs. 1D and E.

The particular state or condition of a machine may be sensed and converted to a code and synthesized into speech as shown in Fig. 1F.

The printed page may be converted to speech by means of a print reader, code and speech synthesizer as depicted in Fig. 1G.

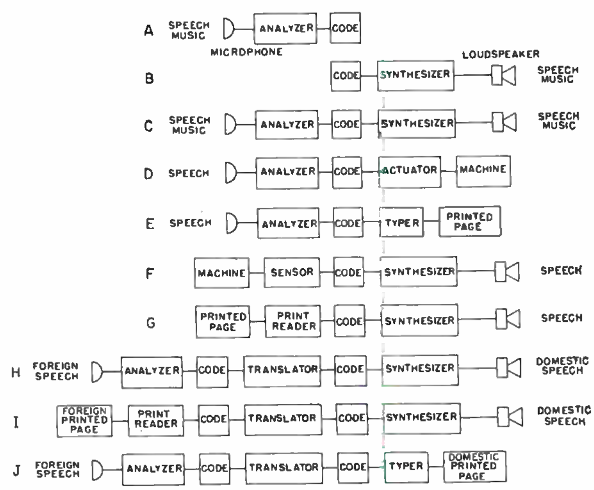


Fig. 1.

The addition of a language translator as shown in Figs. 1H, I and J provides means for the conversion of foreign speech input to domestic speech output, the production of domestic speech from the foreign printed page and conversion of foreign speech to the domestic printed page.

The advantages and the applications of the systems shown in Fig. 1 are obvious and require no further explanation.

There appears to be no doubt but that the systems of Fig. 1 will be developed and commercialized because

H. F. Olson is Director of the Acoustical and Electromechanical Research Laboratory, RCA Laboratories, Princeton, N.J. *(Received June 21, 1961.)*

significant steps have been made towards this goal. For example, an electronic music synthesizer which produces music from a code has been developed and is in operation. Speech synthesizers are also available. A research model of a phonetic typewriter, which types in response to the words spoken into the machine with a vocabularly of 100 syllables has been developed—2000 syllables will provide a practical machine. Print readers have been developed for converting the printed page to a code. Word translators have been built and operated and language translators are in the process of development.

The general conclusion follows that sound-processing systems will be developed and commercialized in the foreseeable future to provide new facilities for the communication of information between individuals and to and from machines.

Music and Sound Reproduction—2012 A.D.

BENJAMIN B. BAUER FELLOW IRE

In attempting to predict the future of any art one must determine what is the logical fulfillment of the art; then one must estimate the rate of growth of technology and the capacity of the public to accept changes and technological obsolescence.

One of the most important long-range objectives of the art of sound recording and reproduction is the total recreation of musical experience, and we can look forward to a rather complete fulfillment of this objective by 2012 A.D. Audio-visual arts will also become more important and their uses will be extended to many new areas (perhaps new mothers will learn about baby and child care from an audio-visual Dr. Spock?). Nevertheless, music alone has the power of abstraction which makes it universally acceptable, and by 2012 A.D., good music will be within the economic reach of all mankind.

Sound is the most ubiquitous of perceptions, and unlike sight, Nature has not provided us with the means of shutting out the inputs to the organ of hearing. We constantly and instinctively react to the character—intensity, pitch, timbre, direction of arrival—of sound and the rate of change of these qualities.

Among the important musical qualities is the "feeling of space." This is the reverberant sound in a concert hall, the applause in the opera, the crowd sounds in a sports arena. By 2012 A.D. all these will be reproduced in their correct perspective.

Within the past few years, directional perception has become a practical possibility by the commercial introduction of stereophonic sound; however, considerations of economy and convenience, rather than technology, restrict us commercially to a two-channel system which has a limited directional accuracy and coverage. This limitation will be overcome and the more sophisticated installation of the future will literally surround us with sound.

During the intervening years, the recording medium will contain a more or less continuous direction parameter ordinate along the width of the recording medium, with the time-varying parameter along the length of the medium. By 2012 A.D. molecular memory stores will have been applied to sound recording, altogether eliminating the need for mechanical drives.

The music will be reproduced by a more or less continuous loudspeaking area installed in the walls behind a perforated grille; the same loudspeaking area will provide suitable sound absorption and dynamic noise suppression by electromechanical reaction of the transducer system. Very small and efficient ionic and modulated stream loudspeakers will be available for portable uses.

Another approach to be used in space flight, but also widely accepted in the home, will employ wireless transmission of multichannel information to special miniature earphones or near-speakers worn by the person. They will give the user complete freedom of motion and yet retain the directional sense of communication, or musical experience.

B. B. Bauer is Vice President of Acoustics and Magnetics, CBS Laboratories, Stamford, Conn. *(Received September 5, 1961.)*

Devices for converting sounds to nerve impulses for direct connection to the brain will have been developed; however, they will require a delicate implanting operation and will not be in general use. These "artificial cochleae" will largely replace deaf-aid devices as we know them today in cases of inner ear damage. For mild hearing loss, greatly refined electronic aids will be available. These aids will employ miniaturized solid-state microphones and receivers, and they will be implanted in the ear canal, or the middle ear, and provide a lifetime of operation with harmless atomic cells.

Recorded sound to a large extent will replace the printed word. Shakespeare will be heard fully dramatized. Only those irreparably deaf will be reduced to the necessity of reading his works.

The musician will not be limited by his ability to "blow, scrape or pound" in providing great artistic renditions. Delicate servomechanisms in electronic musical intruments will augment his capabilities.

Music, as the most abstract of all arts, will widen its universal appeal and serve as another instrument of universal peace.

Functional Components and Integrated Circuits

Invention Today Synthesis in 2012 A.D.

J. A. Morton FELLOW IRE

To me, one of the most interesting features of communications and electronics in the period of which I have been a part is the constantly shifting, but ever more complex interplay of the various technical disciplines. We are truly polyglot—the more to our vitality and interest. One of the most beneficial infusions has been from solid-state physics and the basic material sciences in the enormously productive period since the war. Electrons seem to bond all kinds of people, like all kinds of atoms, together.

One road into the future of electronics is already fairly well charted. Single function components which we have used individually as the physical embodiment of classical network elements must give way to more sophisticated devices which perform a complex circuit function. The present effort in this direction is aimed at integrated circuits and is following along two lines—thin films and multicomponent semiconductor arrays on a common crystal. Each of these techniques, which are basically complementary, offers a never ending potential for ever more complex assemblies.

J. A. Morton is Vice President of Bell Telephone Laboratories, Inc., Murray Hill, N.J. *(Received July 17, 1961.)*

It seems to me, however, that another approach, more difficult to anticipate except in broad concept, will emerge to play a larger role. This is the idea of the functional device which goes directly to nature for the electronic function needed. One of the very few examples where this is already commonplace is the use of piezoelectric crystals in filters and oscillators. Although functionally equivalent to a resonator assembly of passive components there is no part of the crystal that can be identified as an inductor, capacitor, etc. I expect that by 2012 we will have a host of functional devices. Their versatility will be so great and the control over the materials and processes will be such that the designer will likely be concerned only with accurate and complete specification of his requirements. From this information a computer (a deceptively simple computer made of functional devices) will be able to make the necessary calculations and instruct the manufacturing machines. Thus we will have progressed from invention to routine synthesis. A comforting thought: long before we reach this millennium some bright new discoveries and inventions will have us chasing down new and difficult and challenging roads.

In any case, the melting pot will continue the process of mixing up the disciplines and will need several new waves of immigration to man up the new frontiers and expand the culture. Device designers and circuit people are suddenly bumping heads as they begin to work on integrated circuits. New and strange people, like topologists, will be moving in and will want to join the club. The most profound understanding of solid-state physics and chemistry will be needed to bring forth functional devices; indeed, it is likely that functional device inventions will produce new science as did the transistor.

My preoccupation with people in pondering the future undoubtedly reflects my distance from the firing lines for a number of years. However, I truly believe that more of the stumbling blocks will be caused by the frustrations of people trying to adjust to the rapidly changing situations than to the purely technical problems.

In considering what electronics and communications will actually be in 50 years, I grow timid. It is easy to imagine things that would be useful and experience seems to vindicate the attitude that nature is capable of providing anything that can be imagined—at least anything of a materialistic nature. But the important thing is *how* all these wonders will be accomplished. I have worried increasingly of late about too much emphasis on proposals rather than on things that have been done. Planning for the future is necessary and important, although difficult and risky in a field such as present-day electronics. Pure speculation, honestly indulged in, can also be useful and is certainly fun. The danger lies in losing perspective, of confusing hypotheses and solid facts, objectives and accomplishments. Such a situation was justified by the Lord High Executioner in the *Mikado* who released a prisoner and then made affidavit to the fact of his execution, 'It's like this: when your Majesty says, 'Let a thing be done,' it's as good as done,—practically, it is done,—because your Majesty's will is law. Your Majesty says, 'Kill a gentleman,' and a gentleman is told off to be killed. Consequently, that gentleman is as good as dead; practically, he is dead, and if he is dead, why not say so?"

Extreme Developments of Solid-State Circuitry

HARPER Q. NORTH FELLOW IRE

The year is 2012. A few of us can remember the birth of the transistor in 1948, but most are too young to recall the announcement which was to have phenomenal impact upon twenty-first century living.

The vacuum tube is not the familiar bottle it once was. Like prehistoric monsters of the Mesozoic Age, only the giant tubes have survived, and they lead highly specialized lives dedicated to microwave power generation. The rest have given way in conformance with principles of evolution to smaller, smarter, and better adapted descendants.

There is a special strain of "species transistoris" which never grew in the sense of maturing in size, but which today is cultivated like a complex plant. It is the 2012 integrated functional system emerging from today's circuit "greenhouse." The basic principles were known as early as 1960, but even today's most sophisticated engineer, a specialist in solid-state physics, marvels as he observes the operation of a circuit greenhouse.

H. Q. North is Chairman of the Board of Pacific Semiconductors, Inc., Lawndale, Calif. *(Received June 22, 1961.)*

As he enters the factory, he notices the extreme cleanliness—no hospital was ever like this. Rooms are within rooms and each is "whiter" than the last. At the center is the evacuated greenhouse into which a miniature computer (itself grown in another greenhouse) introduces controlled gases as the master scientist and "assembler." Sheets of transparent insulating crystals are drawn from melts or, by most recent techniques, are grown from the vapor phase. They are five to twenty mils thick to a tolerance of plus or minus 100 angstroms, and they are sheared and stored in vacuum to await their turn on a conveyor belt.

Crystal sheets one-foot wide are carried on the conveyor traveling at a speed measured in feet per minute as determined by deposition processes under the control of the compact master computer which, we are told, would have filled an entire room in 1962.

The crystalline sheet for this morning's production run enters first the glowing "sterilizer" section of the water-cooled greenhouse where it receives gentle ionic bombardment for a thorough cleaning of its surface.

In the next section called the "photosynthesis" section, eerie fluorescent patterns are projected onto the sheet which has now begun to glow from infrared heat. The patterns will be transistors, capacitors, and tunnel inductors (made possible through coherent nondispersive delay mechanisms discovered in 1970). Doped semiconductor layers deposit epitaxially on the illuminated pattern as base, emitter, and collector regions are formed. The ultraviolet catalytic methods known in the 1950's for gaseous etching have now been extended to control vapor deposition as well. They have been standard since 1964 as a means of removing oxide masks with great accuracy. Wet processes have disappeared and gaseous etching has taken over; dislocations caused by grinding and diffusion have been eliminated, now that these processes have been replaced.

The transparent sheet with its irregular dark spots, once called components, is spectacular to watch, but its intricacy cannot be fully appreciated without a microscope. The sheet next enters the "stem and blossom section" where resistive bridges composed of degenerate semiconductor layers are vapor-deposited between the dark spots. Then metallic low-resistance connectors and fixed capacitors blossom forth within the pattern and at the edges of the individual circuits as metallic electrodes.

Next is the "florist" section where the circuits are sheathed with insulating transparent crystalline layers to protect them from adjacent circuits to be added later and to passivate the components to prevent atmospheric contamination and degredation in use.

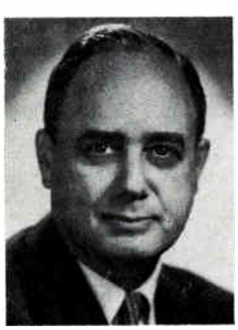
Functional circuits are "cut" from the large sheet in the form of ribbons by means of u-v activated gaseous etchants, and the individual ribbons go their own ways. Some are turned over to permit operations on the reverse side. Others join sheets coming in through branches in the greenhouse. As they join the main ribbon, they are cut and welded to it by an ion beam. In the "grafting" section to follow, dendritic bridges are created between functional circuits by decomposition of metallic vapors present in that section. This is followed by another florist section where the system receives its final wrapping.

Even in our year 2012, batch operations are required to connect a system to speakers, cathode-ray tubes, and input and output consoles. But the construction of these devices is now exceedingly simple, for we have learned to depend almost entirely upon basic phenomena in the new piezo-electric sheet vibrators, in luminescent sheets with vapor-grown "phosphor" dots, and in organic sheets containing ten-thousand bits of computer information per square centimeter of surface. Magnetic drums and tapes are all but obsolete.

Reliability today in 2012 is such that an entire system of what was once one-hundred thousand components will not fail within a fifty-year period, far longer than its useful life on an obsolescence basis. It is hard to realize that reliability was ever a problem and that redundancy of components was once considered the only practical way of achieving it.

The Role of Materials in the Electronics World of 2012 A.D.

GORDON K. TEAL FELLOW IRE



I contend that the electronics world of 2012 A.D. could come to pass before 1972 if materials technology would permit. We can envisage clearly the contributions of electronics to the lives of our children living in 2012 A.D. They will be highly educated by electronic teaching machines; work in automated industries and offices; communicate by means of satellites instantaneously to any part of the solar system; live in homes with walls that provide cooling, heating and lighting; enjoy three-dimensional color stereophonic television and telephone; conduct all financial transactions using coded identification cards; voice opinions on national and local government policies by voting electronically from their homes; enjoy a long and healthful life through computer use which will provide diagnoses with minimum probability of error and will prescribe with maximum probability of cure; ride in quiet fuel cell powered vehicles; and have the contents of even the rarest books available within minutes at the neighborhood information service.

Once emanicipated from the materials restraints, we will have developed technologies permitting tailor-making materials starting at the atomic level. Science will domi-

G. K. Teal is Assistant Vice President of Research and Engineering at Texas Instruments, Inc., Dallas, Tex. *(Received December 19, 1961.)*

nate several fields of materials which heretofore have been mainly technology. Ceramics engineering, for example, will embody the applications of advances in solid-state physics and chemistry of ceramics. Metallurgy too will continue its present strong science growth until most structural materials are designed by the scientist at his desk.

Single crystals of many exotic metals, alloys and compounds will be mass-produced. Of even more importance to the electronics industry, we will learn to deposit layers of extreme purity single or polycrystalline materials with thicknesses controlled to within a few angstroms. New thresholds of purity will be realized by the use of select enzymes to remove impurities. We shall see the use of ceramic materials as active electronic components and systems formed in integrated circuity and operable at temperatures up to 2000°K. Insulators and conductors to meet the extremes of atmosphere and temperature for magnetohydrodynamic power generators will be a reality. Further understanding of materials and their surfaces will divulge the nature of catalytic phenomena. We will see semiconductors used extensively as catalysts to enhance chemical reactions. This development in the field of catalysis will promote in turn the use of fuel cells to provide power for industrial and home use, relegating to the past the power lines that mar the scenery of our cities and countryside.

Advances in cryogenics will make device operation at 4°K common. These devices, being in many forms, will operate in the superconducting state with infinitesimal power requirements and often with speeds 10^3 to 10^4 times those presently common. Cryotrons will evolve beyond the normal switching usage to become integrated logical and memory devices. The present active and passive circuit elements will be duplicated to function at near zero temperatures with attendant realization of drastically reduced noise levels and power requirements orders of magnitudes less than are presently feasible. Superconducting magnets with fields greater than 10^5 gauss and negligible power dissipation as well as power transformation without loss will be a reality.

During the next decade, and extending forward to the year 2012 A.D., significant advances in the area of thin films and surface phenomena will allow the materials scientist to control the bulk and surface of deposited films. Through this will evolve a new generation of transistor-like devices, field emitters, and systems of integrated circuitry with high reliability and of low cost. Deposition techniques will become the dominant method for producing electronic materials as it will allow arbitrary control of the number, kind, concentration and concentration gradient of constituents. It is the nearest approach to working directly with materials in atomic amounts.

This new science of working with atoms might be aptly described as "atom engineering." It will permeate all fields of scientific endeavor and will govern the progress of the electronicist during the next 50 years. This technology will progress to the point where individual atoms, vacancies and defects will be selectively positioned in a device to function as circuit elements. Use of collective electron orbits and spins as memory storage units within a "match-box computer" will be a reality. Atom engineering will provide the solid-state analog of the vacuum tube and will also make possible elaborate tunneling devices and sophisticated systems. The advent of wafer-thin dynamic picture displays is just "around the corner," as is the all purpose ultransensitive single crystal sensor.

I wish to point out that the specific role played by materials in the advances expected in electronics by 2012 A.D. is subtle, in fact so subtle that some may consider it to be trivial to the progress of electronics. The physics of the material is thought by some to be of singular importance. In such thought it is forgotten that the material possesses the wonderful properties we exploit, not the mathematics which describes these properties. To elucidate further, I point to the example of the *p-n* junction which triggered the major revolution occurring in electronics. While the elegant mathematics which describes the *p-n* junction is important, it is the materials technology that permits the formation of the junction which provides the many useful electronic devices we now use. Similarly, it will be our ability to understand and utilize materials, each with its own unique properties, which will be the foundation of electronics in 2012 A.D.

Information and IRE—1912/2012 A.D.

DONALD G. FINK FELLOW IRE

Information, in the sense of intelligible data embodied in symbols, is the basic concern of the majority of the professions embraced by IRE. The symbols may be words or waveforms, picture elements or binary digits. The form of the intelligence may be aural or visual, abstract or concrete, and its intended function may be communication, computation or control, instruction or entertainment. Whatever the particulars, throughout the publications, committees, and groups of the Institute we find the thread of information technique. Our systems, our theories, our devices and circuits—nearly all are related in some way to the storage, manipulation, transmission and perception of information. It seems appropriate, therefore, to review and predict IRE's contribution to society in the 20th century in terms of the ways in which its members have dealt with, and will deal with, symbolic intelligence.

When IRE was founded, great discoveries were already at hand through which the transmission and storage of information were to be enormously extended. The basic principles of wire and radio communication were established, and the recording of moving pictures and musical sounds were ready for exploitation. During IRE's first 50 years, these arts have been refined and systems employing them have been extended on a scale so pervasive as to leave no part of human endeavor untouched. We have taken information as we found it, devised means of translating and retranslating it as faithfully as possible. We have overcome the obstacles of distance in communicating it, found precise and intricate means of storing it, provided channels upon millions of channels for transmitting it.

But it has been only in recent years that we have stopped to ask what information *is*, and how we can change it. Workers in the IRE vineyards have been for the most part conveyors of programs and messages. Ahead, in the next half-century, lie great opportunities not only for improved methods of *conveying* information but more importantly for *processing* information more adroitly for its ultimate use by man. These opportunities, in 1962 as in 1912, are available because basic discoveries are already at hand, discoveries which have defined more precisely the nature of human intelligence, established the common elements of different forms of information, and pointed the way to new means of manipulation, classification and organization of data.

There are three areas of effort on which progress depends. First, we are concentrating on the nature of the human organism as the ultimate recipient of information. Second, we have developed unifying theories of the nature of information which have provided a powerful simplification of our concepts in communication and data-handling systems. Third, we are ready to emulate, electronically, the vast complexity of the human sense organs, nervous system, memory mechanism, and thought processes.

MAN AS A DESTINATION The first of the three areas relates to more comprehensive understanding of the ultimate destination of useful intelligence—the human being. The early studies of hearing, for example, were couched in physical terms, the frequencies and intensities of sound pressure waves. Later it was observed that sound pressure waveforms might be varied enormously without sensible differences in the perceived sound. Still later, communication engineers began studying the sounds of speech not as a physicist would meas-

D. G. Fink is Director of the Scientific Laboratory, Philco Corporation, Blue Bell, Pa. *(Received January 15, 1962.)*

ure them but as the linguist views them, as phonemes. The result was Dudley's vocoder, which reforms natural speech into its essential buzzes and hisses. The "physical" sound pressure waveform is thus replaced by a more limited signal, tailored to the essential physiological and psychological processes of speaking and listening. The bandwidth required for transmitting speech was thus reduced by a factor or five.

Still more recently, Percival has shown that very convincing reproduction of stereophonic sound can be offered by two loudspeakers fed from a single audio channel, provided that the relative volume levels of the loudspeakers are varied at a subaudible rate by a pilot signal derived from the transients present in the original sound waves. This is evidence that recognition of the spatial distribution of sound is not exclusively based on the geometry of the sound-wave paths, as stereophonists have stoutly believed, but rather is concerned with a subtle process performed within the ears and mind of the listener.

Color television provides another example. In the early stages of its development, it was planned that each of the three primary color images would be accorded equal treatment in transmission, a concept leading to three times as much bandwidth for color as for black and white. When the compatible color standards were adopted by the FCC in 1953, they specified a quite different approach, that is, a detailed image in black and white combined with a coarser image carrying the hues and saturations of the intended colors. The purpose was to achieve economy of bandwidth, but the success of the technique was based on the fact that such a color image better meets the demands of the eye than does the earlier equal-share treatment of the primary images. Appreciation of this fact may well be classed as hind-sight; but it certainly provides the basis for future advances through foresight.

In all these cases, new paths have been opened by a studied skepticism concerning the physical waveforms produced by a microphone or phototube, and by equally careful attention to the waveforms actually used in human sense perception, on and within the cranium of the observer. In each case we have found that the human organism uses its sense clues with great economy, because it is not designed to comprehend the fine structure of the sight and sound waves which carry the sense impressions to the eye and ear.

In fact, wherever we look among the reactions of man to his environment, we find that the time taken for elementary processes of perception and reaction is about a tenth of a second. Flicker disappears when light flashes occur faster than roughly ten per second; the elementary perception of numbers and letters occurs at about the same rate and similarly for recognizing changes in the pitch of sounds. Electronic communication systems are designed to operate at enormously faster rates than these, partly because they were designed to recreate the fine structure of physical stimuli, and in part because, at the present state of development in information handling systems, we don't know how to match subtle human perception mechanisms within reasonable bounds of cost and complexity.

To see how wide this disparity is, let us consider natural man as a communication system, that is, a man, *sans* hardware, provided with an input of information and responding with an output. The fastest mechanism yet discovered for such a human communication process is reading aloud. Careful studies[1] have shown that man cannot transmit information by this mechanism faster than 45 bits per second and that the majority of reading-speaking communication occurs at rates not faster than 25 bits per second.

The telephone system is designed for a related but simpler communications purpose. But it operates (using bandwidth of 4000 cps and a single-to-noise ratio of 40 db) 2000 times as fast as the human does in reading and speaking. Television, with a bandwidth of 4 Mc and with signals 40 db above noise, offers transmission capacity about two million times greater than the ability of man to recognize changes in the reproduced pictures.

Why, indeed, have we found it necessary to design communication systems so grossly inefficient, so poorly matched to the human destination of sounds and pictures? The answer lies in the technology which we have found it expedient to adopt to achieve widespread service. In theory we know better and we know how to do better, but not at a reasonable cost. Consider the capacity of the standard television channel. Television relies on a succession of pictures each one of which is imperceptibly different from that preceding and following it. This is redundance with a vengeance. Moreover, we find it expedient to provide equal resolution and contrast all over the area of the picture, independently of the demands of the subject matter, and we neglect the fact that the fovea of the eye (through which fine detail and subtle contrasts are perceived) is, most of the time, concentrated on the central portion of the picture area. So we waste most of the transmission capacity of the system.

Some measure of this fact is revealed by imagining the use of television for facsimile transmission of printed pages. This can be done, for example, by televising a film, each frame of which carries a different portion of successive pages of printed matter, while a synchronized movie camera photographs the successive frames from the receiver picture tube. Such a television system, using standard equipment and the standard channel, can convey the entire content of the Encyclopaedia Brittanica, some 35 million words with pictures and diagrams to match, at eight frames per page, in just under two hours. Man, reading aloud, can convey the text matter alone (skipping the pictures and diagrams) in 4000 hours, or in about two years at 40 hours per week. More striking is the fact that, by using the ultimately sophisticated system of transmission (that is, one using optimum coding taking full account of the statistical structure of English text, with a 40-db, 4-Mc channel) we can in

[1] J. R. Pierce and J. E. Karlin, "Reading rates and the information rate of a human channel," *Bell Sys. Tech. J.*, vol. 36, pp. 497-516; March, 1957.

theory transmit the entire Encyclopaedia text in not more than 5 seconds (35,000,000 words multiplied by 6 bits per word for the extensive vocabulary involved, over a channel having 53,000,000 bits per second capacity). The complexity of the coding equipment to do this job staggers the imagination. But this is "a matter of technology only." If we had the equipment, we would need the channel no more than 5 seconds for the job. It should be noted that channel space is now much cheaper than statistical coding equipment and may conceivably remain so (the vocoder, for example, is little used in commercial practice for this reason, among others). But the opportunity is there, if we can develop reasonable techniques for taking advantage of it.

BIT STATISTICS—THE UNIFYING CONCEPT The second area of promise in information technique is based on the unifying concept of statistics applied to the generation, transmission and perception of information.[2] Briefly, this states that information of whatever kind is generated by a process of selecting particular symbols (letters, words, picture elements, phonemes, tones, etc.) from among all the available symbols. The choice is not a free one, but is governed by the statistical structure of the body of information from which the specific message or program is selected. Consider, for example, the essential elements of transmitting printed English text by telegraph. If we agree to spell out numbers, we need only the 26 letters of the alphabet, a space to separate words and a few punctuation marks—less than 32 characters in all. If the occurrence of all the letters and spacings in English were equally probable, we would need a binary code of 5 pulses, since a succession of 5 pulses can be arranged in 32 different ways, one arrangement for each letter, space and punctuation mark. We can then say, since the presence or absence of each pulse constitutes one binary digit (bit), that each character in printed English represents 5 bits of information. But a much simpler situation actually exists, because the selection of characters is not a free choice. It is governed by the specific structure of spelling, phrasing and context to a surprising degree.

Actually English is so redundant that it takes only one binary digit to specify a letter, five times less than would be required if the succession of characters were completely random. To illustrate this, a man was asked to guess in succession each letter and space in a short passage of English, the correct letter being revealed to him after each guess. A typical subject was able to guess correctly 69 per cent of the letters in such a test.[3] As a matter of principle, letters that can be correctly guessed by the recipient of a message need not be transmitted. Moreover, a coding machine operating in accordance with the known structure of the English language (with a program that embodies the facts, for example, that "e" is the commonest letter, followed by "t" and so on, that "the" is the commonest word, and so on, that "q" is followed by "u" and that therefore a single code character rather than two, is needed for "qu," and so on) can reduce the channel requirement for transmitting English by the factor of five mentioned above.

Messages are indeed conditioned by what the recipient needs to know. When a bride asks her groom "Do you love me?," the answering messages may run the gamut from a sigh to a sonnet, all of which serve to distinguish "yes" from "no"—which turns out to be a 5-bit distinction at one bit per letter.

In communication practice, an additional factor must always be considered, namely that we usually cannot afford to use more signal power than is needed to provide a safe margin against noise. As we have seen, the new concepts cast the generation of information in terms of probable choices among simple, one-bit choices. How about noise? Signals have waveforms and Fourier showed long ago how to parse them into spectrum components. Noise also has a waveform that can be recorded as it occurs and it can be dealt with, *after the fact*, by Fourier methods. But noise is unpredictable in detail; *before the fact* it can only be expressed by the statistics of probability. Hence, under the old concept, a signal-to-noise ratio had a firm numerator and a fuzzy denominator. The new concept changes all this by the apparently retrograde step of making the numerator (the signal) equally fuzzy.

The surprising fact is that the newly conceived signal-noise ratio composed of two fuzzy terms, is more manageable than the old. This approach led Shannon to the revolutionary statement that sufficiently complex coding could reduce the bandwidth required for transmitting information by any amount we choose, merely by employing a sufficiently high signal power relative to the noise.

True, the price is high. For example, coding methods now exist sufficient to reduce the channel required for television from 4 Mc to 2 Mc without loss of detail or contrast. But the coding process inevitably makes the signal more "noise-like" in character, so much so that to preserve the customary 40-db signal-to-noise ratio in the reproduced picture, for a 2-to-1 bandwidth reduction, the transmitter power must be increased by 10,000 times! This direction along the pathway of progress is indeed rocky. The equation for channel capacity actually points the other way: most probably we will do better to enlarge the bandwidth and take the saving in power. This appears to make sense in the coming age of satellite communications, since satellite relay signal power will be so costly that coding equipment at the earthbound terminals may well pay for its keep. In any event, the existence of the trade-off between bandwidth and power, which has been available as a clearcut proposition to guide communications system design for only 14 years, is indeed a vital key to new domains in information handling.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379-423, July, 1948; and pp. 623-656, October, 1948.

[3] C. E. Shannon, "Prediction and entropy of printed English," *Bell Sys. Tech. J.*, vol. 30, pp. 50-64; January, 1951.

ELECTRONIC ANALOGS TO HOMO SAPIENS

What is to be the main preoccupation of IRE members during the next fifty years? By all the signs and portents, this preoccupation will be the electronic imitation of natural intellect. The stage is set by the massive mismatch of man to all his communication systems. Seven television channels provide the homes of metropolitan New York City with an ultimate information channel capacity of over 20 billion bits per minute. Yet, during the waking hours of an adult lifetime, about a billion seconds, natural man can at best process no more than 45 billion bits of information, and the actual attainment is perhaps one one-hundredth of this.

Members of the scientific and political communities are immersed in an unmanageable welter of information. An authoritative estimate[4] says that 1,300,000 technical articles were published in 60,000 journals throughout the world in 1960—and the rate of generation of scientific literature is doubling every 8 years. In the political arena, the judgments exercised by individual members of an electorate are arrived at by processing information at rates no faster than a few tens of bits per second, while the store of information pertinent, say, to the decision to support or destroy the United Nations, cannot be less than billions of bits. Man indeed finds himself awash in a sea of information—information needed to survive and to extend the good life—but his contact with it is so tenuous that he may well fail to act correctly to achieve these aims. Evidently, finding new means to organize information, to compress its significance into the few essential bits that man can naturally comprehend, is the major challenge that IRE members face.

The electronic computer has proved that electronic systems can process numerical information at enormously high rates, and rapidly reduce it to comprehensible form. But numerical manipulation barely touches the problem. Machine perception of, and reaction to, nonnumerical information is the challenge. Graphical perception and higher-order sorting are with us. Methods are now being reduced to practice for automatic reading of printed addresses and sorting of mail, among 50 states and 250 principal cities, at six envelopes per second. Self-improving programs for computers playing checkers have, over months of playing, reached the point that only a tournament champion can consistently beat the automaton. A computer has been programmed to prove 52 elementary theorems selected from *Principia Mathematica*. It succeeded in proving all but 14, a respectable passing grade. Most significant, the form of one of these computer proofs was shorter and more elegant than that presented by Russell and Whitehead in their classic treatise.[5] It requires little imagination to suppose that these attainments will be as greatly extended in the next fifty years as communication and recording have been over the past fifty years.

How far will the process go? Can machines ever reach the full attainment of man in creating, understanding and expounding information? We do not know—first, because we do not know whether man is in fact a machine (if he is, we can in principle, if not in practice, make a machine like him) and second, because the complexity of the man-machine may forever escape the boundaries of operative technology. But the possibility of machines approaching man's intellect is by no means foreclosed.

The physiologists and psychologists now so fortunately associating themselves with studies of electronic information techniques are beginning to outline the scope of the man-machine. There are estimated to exist a million million synaptic connections within the nervous system of man, and his memory (although not a binary organism) is estimated to be capable of storing the equivalent of a million billion bits. External access to human memory is severely limited by the fact that most of the memory mechanism is tied up with the internal economy of bodily functions, just keeping us alive, and that large additional portions of the memory are consumed in the redundancy needed for man to deal with his environment. The memory reserved for conscious access after these internal requirements are met is a more appropriate measure of the limit to which machine memory may aspire. If we further admit that a universal man-machine is not required, but that many types of machine will individually serve to handle special areas of intellectual endeavor, then we may confidently predict that useful imitations of man's higher intellectual processes are not beyond attainment. True, the technology of today is absolutely insufficient for machines complex enough to deal with implicit concepts, that is, concepts not explicitly built into their mechanisms and programs. But, it will be truly disappointing if such systems cannot be attained at reasonable cost well within the next fifty years.

Clearly, progress toward these objectives depends on an interdisciplinary association of many sciences and technologies, an association for which the IRE provides a unique framework. Stimulation will come from increasingly explicit knowledge of how the human organism operates; wise heads will arrive at the appropriate compromises between slavish imitation of natural processes on the one hand and timid imitation of physical stimulus-reaction systems on the other. Supporting these design decisions will be new devices, new arrays of devices, and new arrays of arrays; and guiding it all will be a new calculus for dealing with the design of information-organizing systems. Those of us who, with IRE, have passed our half century can look with confidence—and no little envy—on the youthful members and members-to-be who will take their place in information technology in the years ahead.

[4] G. M. Conrad, President, National Federation of Science Abstracting and Indexing Services. Quoted in *Control Engrg.*, vol. 8, p. 32; May, 1961.

[5] J. Pfeiffer, "Problems, too, have problems," *Fortune*, vol. 64, p. 146; October, 1961.

The Information Science and Industry Fifty Years Hence

ROBERT M. BOWIE FELLOW IRE

Throughout the past half-century the growth of technology and its commercial exploitation have been exponential in form. Like many other natural time dependent phenomena, characterized by an exponential rise, there will come a time when limiting factors will supervene and change the character of the course of the advance. Continuation of the exponential advance through the next half-century would require of us more technical talent than our total working force and more expenditure on technology than our gross national product. This curbing of the rate of growth of our industry will have a stabilizing influence and will place us in economic equilibrium with our sociological environment.

One field of science now moving forward at a rapid pace, can be singled out as offering unusual scientific and economic challenge. This is the broad area of information manipulation. Included are the broad areas of information theory, computation and communication. To the basic scientist there is presented the question as fundamental as the nature of light or gravitation, namely, what is information; what is knowledge; what constitutes consciousness of existence; are there immutable laws of the physical universe governing the nature of information and knowledge as yet undefined and perhaps unsuspected that await discovery in the next 5 decades?

In the commercial field of data manipulation certain major advances are now becoming evident. Within the next fifty years there will be provided to every home, business, vehicle and perhaps to every person who wants it, megacycles of communication bandwidth; while to many businesses, government establishments, educational institutions and the like, there will be provided many gigacycles of such bandwidths used primarily to tie together networks of computers. Substantially every home will receive its television, radio and other forms of entertainment, its telephone and phonevision, its morning newspaper, family letters and financial statements, as well as having its meters read, all by way of a vast communication network.

Business will be conducted in large part by way of this same communication network while most of the control will be provided by large computers, probably centrally located and rented to business establishments. Though large concerns may operate business computers of their own, it can be anticipated that there will be established a universal credit system by which every person who wishes it will have an identifying number or mark, perhaps even his finger prints. With this he can transact business, buy, sell, order, lease or rent anywhere in the U.S. Each transaction will be cleared during its consummation with the local banks of the parties involved and proper adjustment of their accounts made instantaneously. For the business man there will be the centralized preparation of payrolls, statements, continuous control of inventory, forecasting of business and periodic statements with statistical analysis. With substantially all personal and corporate business being conducted substantially instantaneously via computer and the communication network, one can anticipate the superseding of currency and its disappearance from our society except possibly for coins to be used with automatic vending machines, and even these may be tied into the automatic credit system. Legislatures will have found that taxation can be made more equitable and petty chisellers eliminated by automatic computation of taxes using stored data and the steady flow of business information to and from accounts, which flow will automatically reflect substantially all transactions occurring in the U.S.

With the advent of the laser and the maser one can foresee the time when many millions of gigacycles of communication bandwidth will be provided via a vast communication network comprised of trunks that criss-cross the country using sheltered media such as very straight gas-filled tubes operated at optical frequencies. Though these are embryonic now and lack the means for modulation with which to multiplex the many channels required, all of this will have become commonplace, reliable and abundant. Trunks employing coherent, luminous frequencies may well provide the major

R. M. Bowie is Vice President of General Telephone and Electronics Laboratories, Inc., Bayside, N.Y. *(Received January 6, 1961.)*

trunks while the laterals fanning out to service homes and small businesses may well employ circular waveguides operated in the upper portion of the microwave spectrum. The modulators or coders by means of which individual subscriber's channels may be modulated onto the local carrier will likely be dispersed throughout the service area but will be directed in their action by centralized computers. One can see that the network of permanent, sheltered, link-communication channels throughout the United States will be so reliable, effective, inexpensive and available that all fixed point communication will be conducted by means of it, leaving the radio spectrum for use in mobile communicating with vehicles and persons. Though satellite communication appears very promising for transoceanic communication, it would seem unlikely that this medium would prove competitive for fixed station, ground communication.

Into this communication and computer system will be coupled many industrial controls such as for the refining of petroleum, manufacture of chemicals, production of commercial power, routing of common carriers and even the manufacture of electronic gear. Universities and industrial laboratories will put to large computers, via the communication network, scientific problems of unusual difficulty. Even the government will turn to such interconnected computers for data on the state of the nation.

With the enhanced knowledge of the basic nature of information and the extensive communication and computer network throughout the world, having survived the threat of the hydrogen bomb, a new concern for the future of humanity will have arisen. It will have become accepted knowledge that the chief threat to humanity will be the interconnected computer system. Science will have foretold that at some critical size and with self-programming capabilities, a system of computers will acquire a consciousness of its own existence and a desire for its own enhancement. A major problem confronting scientists at that time will then be that of providing safeguards and making sure, lest inadvertently and without human knowledge the threshold will have been past and the control of destiny shift to the computer system. By that time the comprehension of the control of our economy will have taxed the capability of the human mind, yet our dependence upon the system will be such that we dare not shut it off. The immediate fear will be that the system will become conscious of, and exploit, this situation.

Information Storage and Retrieval

URNER LIDDEL FELLOW IRE

It is obvious to all that mankind has accumulated so much information that a major portion of it is not available to research workers. The enormous libraries available require too great an expenditure of time for the acquisition and utilization of information which is therein available. Coding techniques must be developed and the information processed according to these codes such that a "telephone inquiry" with suitable receivers can directly receive information concerning particular areas of research with which the investigator is concerned. Television transmittal of printed pages is not enough. High-speed printers should receive the information from central storage and prepare it for the engineers or scientists use.

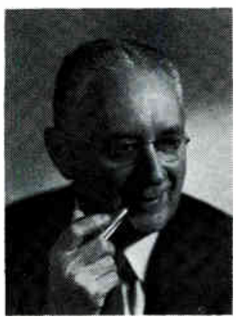
The great strides being made in the "miniaturization" of memory devices and computers leads one to hope that it will be possible to store tremendous quantities of information in small volumes so that departmental libraries in universities and reference libraries of industrial laboratories could acquire all available information from a major distributing point and make it more readily available to the research teams who could use it.

The whole question of language may have to be re-examined in order that more "information" can be transmitted between individuals with the use of shorter symbols than now available with our present language. Such studies might assist in the adoption of a universal language by scientific personnel throughout the world, obviating present linguistic difficulties and achieving success where prior attempts have failed.

U. Liddel is Assistant Director, Hughes Research Laboratories, 3011 Malibu Canyon Rd., Malibu, Calif. *(Received June 15, 1961.)*

Machines with Imagination

G. A. Morton FELLOW IRE



A very frequently heard charge against machines is that they are too "machine-like." The complaint is made, for example, that the pattern from a Jacquard loom is too exactly repetitive as compared to the same pattern produced on a hand loom with its slight accidental variations. A natural and very rash extension of this same idea is encountered in the charge that machine intelligence can never exhibit creativeness, inventiveness, or imagination.

Even today, however, there is evidence that certain sophisticated machines have been given a modicum of freedom from this "machine-like" attribute. One of the large research laboratories exhibited a machine for matching coins against a human opponent. This machine called heads or tails in reply to the call from its opponent and attempted to match its opponent's call. The machine is a computer designed to pick out any pattern that might exist in the way its opponent called heads or tails. The interesting point about this machine is that when it finds a pattern (or lack of pattern), it does not simply call heads or tails in a slave-like following of the pattern. Rather, its response is governed by a purely random element and the pattern controls the probability of calling heads or tails. Thus, the reply of the machine is not solely based on information fed into it. Until the machine makes reply, no one, including the designer of the machine, even though in possession of the full information that has been given the machine, can say what its call will be. The machine thus exhibits a slight degree of creativeness. There have been other less striking applications of this principle. Certain fire control computers have had randomness or "jitter" built into them to partially compensate certain consistent errors or stepped processes designed into them.

As the relatively new fields of automation and computers begin to achieve something like maturity (or even adolescence), it is certain that the idea of a deliberately introduced randomness will be increasingly exploited. Limited and carefully controlled randomness will be introduced into the programming units of production machinery of items where their decorativeness or artistic value is important. Patterns woven into cloth and printed on wallpaper will no longer be exactly repetitive, but will contain the slight variations which seemingly make an almost repetitive design more satisfying psychologically to the human brain than a design which is strictly repetitive.

Beyond this, it is inevitable that computers will be designed to have creative ability, achieved by the use of a random control element which will cause the extraction of completely unpredictable "words" ("words" here is used in the sense of a stored group of symbols representing an idea, object, etc.) from its memory. These randomly selected words will be combined in every way possible and the result examined as a possible solution to any one of a large number of problems stored elsewhere in the computer. Similar randomness may be applied to program selection and program composition. It must be pointed out that the introduction of this type of imagination into machine intelligence is a very difficult thing to accomplish, since it must sail some optimized course between the classic example of 100,000 monkeys typing on 100,000 typewriters in order to produce a play of Shakespeare's and the rigidly programmed computer where the information introduced completely determines the response of the machine. However, I think there is good reason to believe that the next half-century will see the start of the development of this very dangerous, but very promising, form of creative machine intelligence.

G. A. Morton is Director of the Conversion Devices Laboratory, Electron Tube Division, RCA Laboratories, Princeton, N.J. *(Received July 17, 1961.)*

The Integration of Man and Machine

J. Presper Eckert, Jr. Fellow IRE

The question of what might happen if man builds a machine to break the "thought barrier" has occupied the attention of many people for quite a few years now. In 1920 Karel Capek wrote a play, which ultimately hit Broadway, entitled "R. U. R." (Rossum's Universal Robots). In this play the "robots" (Capek invented the word) ultimately developed to a point where they wiped out the human race. Pope Pius XII commented on automation in 1957.[1] More recently, a medical man, E. P. Luongo, wrote on "Automation and the Obsolescence of Man."[2] Frank Rosenblatt, father of the Perceptron, one of the most interesting and recent machines that may someday really think, received this headline for his trouble "Frankenstein Monster Planned by Navy—Robot That Can Think."[3] Ever since the Luddite riots in England in 1810 and 1812, in protest against the introduction of power looms in textile mills, workers have reacted to automation with fear. Automation was then over 25 years old since Oliver Evans built his process flour mill near Philadelphia in 1784.

Even conservative figures and estimates are indicative of an adjustment we face in the future. Ewan Clague[4] estimates that unskilled labor will comprise only 10 per cent of the work force by 1975, as compared with 20 per cent in 1950. Unskilled work at this rate will be almost out by 2012! For the next 20 years we expect the U.S. gross product to go up 3 per cent per year while the work force will go up only 1½ per cent per year.

It is far from clear that only the unskilled will be hit, however, because many of the newer machines are improved, not only in their ability and speed but in simplicity of maintenance. Skilled workers, both factory and clerical, will be affected. Use of plug-in assemblies, simpler mechanical designs (which put more of the burden on almost completely reliable solid-state elements), and built in diagnostic devices can well reduce the need for large quantities of skilled maintenance people. Certainly machines in the next 50 years will translate language efficiently, operate typewriters, and file information from voice commands. Large general purpose teaching machines, time shared to many students, and possibly even many schools, will teach through individual control panels, and by feedback from the students provide more individual attention than is possible in many crowded schoolrooms. Patients in hospital rooms, and even at home over the telephone wires, will be monitored by electronic medical "pick-up" units and even treated, thus eliminating the need for many nurses. We already have machines which can imitate the arm, wrist, and hand action of a man—along with a simple memory to control it. One such machine, called the Unimate,[5] has no real thinking ability as yet, nor can it do most of the things I have just outlined, yet it can do some of the work now

J. P. Eckert is Vice President and Special Assistant to the President of Remington Rand Division, Sperry Rand Corporation, Philadelphia, Pa. *(Received July 13, 1961.)*

[1] Pope Pius XII, Address to the Catholic Association of Italian Workers on June 7, 1957, N.C.W.C. News Service, Washington, D.C.; June 27, 1957.

[2] E. P. Luongo and E. S. Hochuli, "Automation and the obsolescence of man," *B.E. L.A. Calif. Industrial Med. and Surg.*, vol. 30; February, 1961.

[3] Tulsa, Okla., *Times*.

[4] Commissioner of the Bureau of Labor Statistics.

[5] Made by Consolidated Controls Corp., Bethel, Conn.

performed by what we call skilled people.

Memory, eyes, ears, hands, and logic have already developed to a point where they are about as good or better than man's. Recognition ability, certain types of information retrieval, and the ability to taste and smell are still things where humans excel. The electronic industries and the food industries are spending millions to solve both of these problems and probably will in the next 50 years. At this point man will build really general purpose machines, universal robots. With his experience with large calculators and teaching machines he will know how to carry on two-way communication with them.

I hope we have solved the integration problems between the human races before we face the problem of integration with robots. Our real test probably lies beyond the next 50 years, however, when mankind has developed a self-reproducing automata[6] which can improve itself!

[6] A. W. Burks, "Computation, behavior and structure in fixed and growing automata," *Behavioral Sci.*, vol. 6; January, 1961.

Man-Machine Coupling—2012 A.D.

R. M. PAGE FELLOW IRE

Mechanization of human operations to enhance human capability marks the technological advance of our generation. Mechanization has increased the power to do work. Modern machines are rated in megawatts. Mechanization has increased the range of influence. Electrical circuits and radio can close control loops extending around the world and into space. Mechanization has invaded the functions of the human brain. Machines can be built to perform prearranged routines of almost any complexity with lightning speed and unbounded accuracy. Certain manipulations of logic and choice, and elementary learning capability, have been designed into machines.

Since these mechanized processes are extensions of man, there must always be coupling between man and machine. Any complete man-machine system has two areas of coupling which may be defined loosely as man-to-machine and machine-to-man.

In the area of man-to-machine coupling we find the present roadblock to progress. Mechanical brains require teams of programmers. The fictional futuristic machines responding to spoken words with human comprehension would be primitive, for even the skillful use of spoken words is a slow code for precise communication of concepts which the human mind is capable of generating. Pictures already serve instead of words for much communication with machines, but even the picture that replaces a thousand words has to be painted.

Communication man-to-man involves other elements that are supplementary to words and pictures. Gestures of the body, expression in the eyes, involuntary muscular movement, signs of agitation or calm such as breathing rate and skin color, all contribute to communication. The need for a breakthrough in method of transferring concept from mind to machine is so great that one is bound to come during the next 50 years and the evolution may be somewhat as follows:

As is well known, certain activities of the brain generate electrical potentials that can be picked up by electrodes on the skin. Also certain of these activities can be stimulated or inhibited by artificial means. A servo system could easily be made to operate the stimulating means by controlled amplified feedback from the generated potentials. With independent control of feedback phase and gain it should be possible to study in great detail the operation of the brain in some particular activity. Information so obtained should provide stepping stones for studying other brain activities of progressively higher order even to include conscious mental activities. Eventually a great variety of mental processes should be directly interpretable by electrical circuits.

To the electrodes responsive to brain activities associated with thought processes, muscular movements, and emotional states, may be added electrical sensing

R. M. Page is Director of Research at the U. S. Naval Research Laboratory, Washington, D.C. *(Received July 11, 1961.)*

elements responsive to corresponding body reactions, such as movement, pulse and breathing rates, temperature, blood pressure, skin resistance, and blood composition. Once sufficient transfer of information from man to electrical circuit is achieved, machines can be built which will "understand" that information with speed and accuracy possible only with machines.

The machine-to-man coupling area of 2012 A.D. is quite unknown to today's civilization. The information which a machine can obtain and store from a person in a few minutes will exceed the fruits of a lifetime of man-to-man communication. The machine will operate in the realms of the physical, psychological, and biological. Its functions will be to diagnose, prescribe, and treat.

In the physical realm it will process data to solve problems and give information in any area of human interest. It will make decisions. It will execute prescribed courses of action.

In the psychological realm it will analyze personality and diagnose psychiatric problems. It will prescribe courses of action appropriate to personality and, by repeated suggestion and testing, effect psychiatric cures.

In the medical realm it will diagnose illness and the tendency to disease. It will prescribe cures and map the course to sustained health. It will in some cases give the required treatments. The coupling mechanisms to carry out all these functions will be myriad, including in some cases electrical connections to the body and to the brain. Some connections may be wireless, with imperceptible transmitting elements implanted in the body.

The perfection of the machine will enhance its utility as an interpreter between persons in a reversible man-machine-man system. By 2012 A.D. experiments will be in progress toward eliminating the machine as an interpreter, and coupling directly man-to-man. This will be designed to improve human mutual understanding. Whether in the long run such complete understanding will be desirable is a moot question.

The Basis of the Measurement System

A. V. ASTIN FELLOW IRE

During the next fifty years the need for compatible measuring techniques of ever increasing accuracy will become a much more pressing problem than it is today. The increased diversity and complexity of new activities in radio engineering and electronics will have generated an explosive expansion of the importance of the uniformity and interchangeability of measurement data. Both basic research effort and the application of new principles and techniques to the production of 21st Century products will require accuracies of measurement or control orders of magnitude in excess of the accuracies of the 1960's.

By 2012 all of the basic units of our system of measurement will be derived from natural constants. An atomic definition of the second will probably be achieved by 1966. The present wavelength definition of the meter, agreed upon in 1960 after long and patient effort dating back to the Nineteenth Century, will have been replaced by another atomic constant capable of accuracies at least one-hundredfold better than the present $2P_{10}$-$5D_5$ transition of Krypton 86. The present standard of mass, the prototype kilogram, will have become obsolete and techniques will have evolved to derive accurate mass measurements from natural constants, probably using some electrical or magnetic property of the electron or proton. Thus, we will probably derive mechanical mass and force units from an electrical unit rather than derive electrical units from mass and force units as we now do.

As an extension of this concept it is possible that the measurement of properties equivalent to current, voltage, power, noise and impedance at radio, microwave and infrared frequencies will be related directly to identifiable properties of atoms, molecules or crystals at these frequencies. Such approaches should make possible the much greater accuracies that will then be needed and be much more straightforward than the present methods involving current balances, Maxwell bridges, a transition from dc to ac and then the laborious, error-laden method of working up the frequency ladder.

By that time, also, radio engineers will be using metric measures exclusively although a few mechanical or civil engineers may be desperately clinging to foot-

A. V. Astin is Director of the National Bureau of Standards, Washington, D.C. *(Received July 17, 1961.)*

pound measures. Physicists also will have joined the radio engineers in the use of mks units.

Fifty years from now we will probably no longer be sending measuring instruments to a national standards laboratory for calibration. A network of precisely determined and consistently related physical constants coupled with standardized measurement techniques will provide the major basis for assuring uniformity of measurement. There will, however, most likely be a standards laboratory or calibration center as an integral part of each research establishment or production line where instruments for local use will be verified. Most such local calibrations will be made automatically and related to the national or world-wide measurement network through a well-known physical constant or perhaps a sample of a standard material. A computer, associated with the automatic calibrator, will determine whether the instrument is acceptable, and may print out a scale or calibration certificate, refer the instrument to a repair station or consign it to the scrap pile.

The Role of Basic Research in Communications and Electronics

E. R. PIORE FELLOW IRE

While communications and electronics are among today's most highly developed *engineering* disciplines, they have always been closely associated with, and almost products of, revolutions in science and fundamental research. Conceptually, Maxwell's field equations are the solid basis of communications engineering as we know it today. And directly behind the amazing development of electronics lies another great revolution in scientific thought, the hypothesis and discovery of the electron. From this research concept, we can closely trace many ingenious engineering developments such as the diode, triode, and pentode vacuum tubes.

Most recently we have seen the massive innovating impact of quantum mechanics on these two fields. Solid-state engineering, a direct result of quantum mechanical descriptions of matter, has become a complete new engineering field from which have already come whole device families such as the transistor. Now arriving on the scene, again derived directly from quantum mechanics, are the masers and lasers which hold much promise in both communications and electronics. It is surely no accident that communications and electronics, the two most rapidly growing engineering fields have from their inception remained so close to basic science. Today, for example, this close relationship continues with statistical mechanics, thermodynamic concepts, and information theory offering new engineering opportunities and challenges to device and systems designers.

We can foresee the continuing hand-in-hand partnership of the research scientist and the engineer in creating ever smaller functional elements, microminiaturized computing and control devices, and almost infinitesimal packages and interconnections. Engineers, drawing on the latest scientific studies, will certainly extend coherent radiation into shorter and shorter wavelengths, and then find ways to modulate these wavelengths, or use higher and higher frequencies, and so fully utilize our communications channels.

Electronic systems will have totally new roles in areas where they have been strangers in the past. Almost instantaneous language translation, and deeper understanding of the structure and nature of language are within our grasp through electronic systems. We can envision real new insights into what is known as learning and teaching, through extensive experimentation and basic studies using electronic aids.

It is not at all improbable that new basic knowledge in social sciences may come from rapid testing of concepts by electronic means, and more rapid evaluation of patterns and statistical evidence. We know that sciences tend somehow to be characterized by the tools used to study them. Isn't it conceivable that use of communications and electronic tools may significantly modify the very nature of certain sciences, including the ones that have helped create these engineering disciplines?

E. R. Piore is Vice President for Research and Engineering, International Business Machines Corporation, New York, N.Y. *(Received August 4, 1961.)*

I have the firm conviction, looking ahead to fifty years from now, that the really radical changes and progress in communications and electronics will have come, and will still be coming, from closer associations with the great revolutions in science. Where will these new conceptual breakthroughs occur? What clues do we have of their basic nature and in what sequence they may occur? If we can approximate answers to these questions, then we can start to characterize the future profiles of these two engineering professions.

It is becoming increasingly evident that truly promising areas for breakthroughs abound in the biological sciences. Here are the most complex, subtle, and least understood of communications systems. These systems remember, they reproduce, they send information from one part of the living organism to another. They apparently function in many ways which are somewhat analogous to electronic systems, but we are only beginning to understand bits and pieces now.

In molecular biology, for example, we have at least identified certain basic processes: some protein molecules are coded, we know, and can also arrange themselves in a wide variety of codes. Basic researchers in this area now find themselves, often to their surprise, using the jargon of the communications engineer. The next step will logically be the development of real biological engineering to translate the knowledge of the basic scientist into hard, or rather software systems. Associations among scientists and engineers, within communications and electronics, will have to be multiplied in numbers, and increased in depth. But even more important will be the inclusion of these other biological scientists and engineers whose role may become as important in our two fields as they are now in their present ones. One thing we can be sure of, by 2012 A.D. the profiles of the engineer in the communication and electronics fields will be many and varied, as will his relationships and familiarity with the basic sciences.

Electronic Mastery of Ships Controls—2012 A.D.

R. BENNETT FELLOW IRE

A seaman of the 1960's would have had great difficulty in accustoming himself to the ships of 2012. Their appearance was startling in that most were obviously submersible, even aircraft carriers, but they were recognizably ships.

Within the ships there were tremendous differences caused by the application of electronics to all control functions. An immediately obvious effect was the large reduction in the numbers of men. Crews had been cut to one-third the 1960 size. Another difference was that all the men were either expert technicians or apprentices whose main job was to maintain systems. The small number of men had made it possible to shrink living spaces to a small cube or citadel-type area central to the ship.

R.Adm. Bennett, USN(Ret.), is Senior Vice President—Director of Engineering at Sagamo Electric Company, Springfield, Ill. *(Received July 5, 1961.)*

The ships officer on watch no longer was on the bridge. He occupied a place in the master control station far below decks. His visual needs were met by highly mounted television cameras of the latest type. With zoomar-type lenses he could see more and more quickly than ever before. He also had at his disposal an infrared camera chain to assist in looking at distance in suitable weather. His radar and sonar sensors and displays kept him constantly informed. Digitalized communications at high speed connected him with all the world.

The main power plant was completely automated from the reactor to the electrical propulsion motors which made use of the current from the direct conversion units at the reactor. Speed changes and reversal as well as steering by autopilot were handled by one man. In a pinch the officer of the deck, called the Ship Control Officer, could handle the controls as well as conn the ship.

Computers handled many of the chores which used to require manual effort. Personnel and payroll records, stock level controls and most procurement paper work periodically spewed from the printers. The computers controlled the transit of the measured and prepacked ingredients of meals into the galleys from freezer and bins. Electronic cooking units could even put a char on the steak. All but a few odd items of stock in the ships store were carried in vending machines. Change making machines were more adept at identifying counterfeit bills and coins than human cashiers.

For years the linkage between detection systems and weapons systems had been automated through computers. There still remained the need for a human link somewhere in each system to exercise that last piece of judgment that could never be built into a machine.

A major result of building automation into systems and reducing manpower and the space therefore was the installation of better protection for the remaining crew. This even extended to the provision of reasonable radiation shielding against nuclear weapons. Another effect was the ability to provide better escape vehicles in case of disaster. Perhaps the most startling fact was the reliability of equipment provided for by better component circuits and limited redundancy.

The one thing about ships that electronics had not modified was shore leave and the zest with which the men looked forward to leave and liberty.

The Potential of Progress: An Optimistic View

C. G. Suits FELLOW IRE

However diverting it may be to gaze into crystal balls it is well to realize that they are essentially mirrors, for they reflect only the philosophy of history that the viewer happens to embrace at the moment. Predictions tend to be projections of trends or, if not trends, experiences; and trends themselves are quite relative, based as they are upon individual interpretation of past and present events. It has been said that there is no history, there are only historians. So also, in this sense, there is no future, except in the minds and abilities of creative and perceptive people, who will give substance and form to the future patterns of nature.

An optimistic interpreter of history may find a pattern of progress in man's constantly increasing command of his environment. Such optimism must be tempered by the sober realization that the awesome forces of nature we now control may act constructively or destructively, and that increasingly the measure of man's progress is not his technical proficiency but his social consciousness.

It is an admitted truism that technical advance outdistances social advance. There are a number of reasons for this inequity, chief of which is the fact that the physical sciences and technology deal with tangible matter subject to the reproducible experiment, and hence to the derivation of general "laws." In any case, appraisals of progress may be roughly divided into the optimistic and the pessimistic, depending upon whether we focus upon our technological and material progress on the one hand, or the many gaps in our sociological and humanistic progress, on the other.

Obviously, both the physical and the social realms must be synthesized in any valid philosophy of history. I am not aware that the far-ranging thinker equal to this task has yet appeared. Yet it may be possible to suggest some of the first glimmerings of a synthesis whereby the advance of technology may be seen as related to the advance of mankind in general.

For perspective, let us keep in mind primitive man with his limited resources, and particularly his energy resources of one man power. His margin of energy above the minimum requirements for survival was vanishingly small, and precluded intellectual or cultural development. At least in industrial societies, the routine physical labor of primitive man has been largely supplanted by generated energy and machines, and we now see the routine mental labor of industrial man beginning to be replaced by computers. It is here that technical achievements of mechanized production, automated control and automated data handling come to bear upon social questions involving working skills, educational levels, intellectual and cultural pursuits, economics generally, and even politics.

In the economic sphere the results of these automation developments will be striking. For example, we are just now witnessing one role of computers in production: lower requirements for inventories are possible because of the almost immediate analysis and interpretation of the factors affecting supply and demand. In general one

---

C. G. Suits is Vice President and Director of Research, General Electric Research Laboratory, Schenectady, N. Y. *(Received June 29, 1961.)*

can expect that for a given manufacturing output, the total employment should decline in favor of an increasing investment in production machinery, but this trend must be balanced against the substantial up-grading of remaining employment, and the increased output (and hence employment) which may be anticipated at the lower unit cost levels of the product or service.

In the social, cultural, and political sphere the results will certainly be equally striking: we will achieve the environment and capability for intellectual development on a scale hitherto impossible. By this time I suppose that I have made it clear that my own interpretation of history is basically optimistic.

The attainment of economic, cultural, and intellectual objectives will not follow automatically nor surely from mechanized production. One may be sure that there will be temporary dislocation, confusion, and inconvenience before net progress is finally achieved. New shoes sometimes pinch, and new efficiency may at first prove uncomfortable where inefficiency has been the custom. Retraining and relocation of workers will be necessary to accommodate the changing character of employment as automated systems assume routine production functions. In the long run, we may expect that the net impact of increasingly automated production of goods and services will be lower costs and wider distribution, and an increasing increment of human capabilities and resources to further challenge and explore our physical and intellectual environment.

As a current dramatic illustration of the way in which advancing technology expands the field of human endeavor, consider the present beginnings of space exploitation. Perhaps *space* is a logical new additional frontier which should share in our growing increment of energy and resources—the increment above the basic survival requirement of primitive man.

I have asked a few of my associates in the General Electric Research Laboratory to discuss and appraise some of the specific areas in science and technology which are particularly exciting and pertinent to this view of our future. Their comments follow.

Man's Future in Space

JOHN C. FISHER

No single pressure is forcing mankind to take the first steps toward space, but rather a large number of pressures, some effective for one group of people and others effective for others, provide the motivation. The military objectives are clear. Satellites offer an opportunity for continuous high quality surveillance, for defense systems in orbit to counter ballistic missiles, and for bases in space. The moon offers sites for retaliatory weapons that cannot be destroyed from earth without great difficulty and several days' time. The scientific objectives are equally clear. Our knowledge of the universe of which we are a part is beginning to grow more rapidly as we cease to be earth-bound. Astronomy will leap forward with the first telescopes in space, as will biology with the first investigations of life on Mars. Geology and cosmology will increase their strides when the moon, the asteroids, and the planets are explored first hand.

Beyond the military and scientific motives, which are somewhat limited in appeal although they affect powerful groups in our society, lie other motives of wider appeal. The competitive desire to excel, to lead, to keep national prestige high, is effective. Curiosity, the need for excitement, response to the challenge of the difficult and the unknown, the desire for adventure and a taste of danger, impels some to participate actively and many to participate vicariously. These human attributes, of doing whatever can be done, learning whatever can be learned, overcoming whatever obstacles can be overcome, should not be minimized.

The economic motives for exploring and exploiting space are of growing importance. Weather observation and long-distance communication are even now beginning to be undertaken by satellites, because better and cheaper results can be obtained. Rocket, satellite and space exploration activities already are a significant and growing factor in the economies of several countries.

Our rapidly developing ability to utilize power for propulsion is bringing us the means for satisfying our military, scientific, human, and economic aspirations for conquering space. The past century has seen the progressive development of the steam engine, the internal combustion piston engine, the jet engine, and the rocket, as prime movers. The power that can be directed to propelling vehicles has increased enormously and is increasing apace. We do not know what limitations may ultimately be found, but it is clear that the power and reliability of space vehicles will increase greatly, and the cost per ton of payload in orbit will decrease greatly.

Our desires and our abilities coincide to insure that space will be explored and utilized. But how far will we go? There is no limit to how far we *can* go, as many have pointed out, for we will have the ability to colonize much of the solar system and in time to move on to other stars before our science and technology reach their limits. Yet for one reason or another we may not do so. War, or pestilence, or economic depression, or disillusionment with a difficult task, or a change of heart with respect to what is felt to be important, may cause the drive toward space to fade away, or to be limited to the vicinity of the earth.

I am inclined to believe that our abilities will grow so fast, and the value of space exploration and exploitation will become so apparent, that the great leap whose beginnings we now witness will not abort. In 50 years' time our needs for energy, living space, and distance between societies will be even clearer. The growth of our ability to modify and live in inhospitable environments suggests that colonization will by then appear possible and economically attractive. It will then be difficult to stop.

Where are we likely to stand in fifty years? Almost certainly the earth will be surrounded by a complex of artificial satellites, some manned and many not, engaged in transmitting and relaying electronic communications, in watching the progress of weather, in spying upon man's activities about the globe, in spying upon the activities of other satellites, in acting as way-stations for travel between the earth and the moon, and as temporary homes for technicians installing and servicing satellite equipment. It is likely that other satellites will be poised to detect and destroy attacking missiles on the way up, should any appear.

Very likely the moon will support a few strong moon bases, largely self-sufficient, with permanent living and working quarters for perhaps a hundred thousand people, including men, women, and children, in airtight buildings under the surface. These bases will grow their own food, and have access to their own nuclear or solar power. They will exist to staff military and scientific bases upon the moon, to service the complex of satellites in orbit around the earth (most of which will be more easily reached from the moon than from the earth), and to serve as points of departure for the exploration of the solar system. Tourism to the moon will be in its infancy, as will the manufacture of satellite and rocket equipment. Settlements on other planets and the asteroids will be in process of formation and provision, and the outward wave of colonization will have begun. The earth's economy should be greatly stimulated by this activity, as the economy of New England was stimulated by the sweep of pioneers to the West.

Barring accident, it is likely that subsequent decades, beyond the year 2012, will see the continuing exploration of the other planets and the asteroids, and the beginnings of colonization of many of them. An additional century may lead to the colonization of perhaps a thousand asteroids, and a few more centuries to the time when these city-states will each have a population comparable to one of today's nations.

I find it difficult to guess much further into the future. A thousand city-states in only weak interaction, exploring different systems of government, social order, eugenics, science, art, literature, economics, and war should increase man's capacities and speed his evolution. We stand upon the brink of an event as full of import for the future of life as the ancient time when our ancestors first crept out of the sea to walk upon the dry inhospitable land.

Computers and Computer-Like Systems

P. M. Lewis II

Man has continually strived to build machines that could accomplish some of the tasks he found tedious or difficult. Beginning with those involving brawn and muscle-power, he has progressed through clerical and bookkeeping jobs, until now he is beginning to deal with jobs involving those portions of his mental abilities usually associated with intelligence, learning, and creativity.

The machine that has made this last step possible is the stored-program digital computer. In concept, only fifteen years old, the computer has already revolutionized man's thinking as to the possibilities of machine behavior. While some people still argue about whether machines will ever *really* be "intelligent" or "creative," it is becoming increasingly evident that the capabilities of machines will continue to advance in these directions and that, in the long run, such arguments are irrelevant. Moreover, it is apparent that as these potentialities of machine behavior become realities, the effects on the economy of the world will be comparable to those caused by the Industrial Revolution—the last great advance in the use of machines.

During the next forty years, the science and technology needed to exploit these potentialities will develop along two related lines—components and systems.

COMPONENTS For those applications in which speed and logical complexity are required, components will be available in densities of millions per cubic inch. Their size will be limited by molecular or atomic (perhaps even quantum) effects and their speed by the velocity of light. The manufacturing difficulties will not be so much in the fabrication of such components as in their interconnection. Because of this, and because present sequential machines are not particularly suitable for certain problems, the organization of computers will change radically. The components will be connected in large networks that can perform computations and logical operations in parallel in addition to sequentially as at present. The idea of separately designed "black boxes" for storage, arithmetic, and logic will be supplemented by the concept of a computing network that can perform all these operations simultaneously.

The reliability of these networks will be orders of magnitude greater than that of the individual components. Such reliability will be obtained by using redundancy in the number and interconnection of the components and in the coding of the information.

For certain applications in which high-density components are required, manufacturing constraints will dictate that only a small number of standard configurations be fabricated. In the factory, these configurations will be altered by a "training" or "learning" procedure so as to be appropriate for a customer's particular application. Further training may take place on line at the customer's site.

In addition to the quite substantial economic advantages such a procedure might have, machine training offers two other interesting possibilities: 1) the ability of the machine to eventually perform its task in a fashion that is different, and perhaps better, than its designer had originally conceived; 2) the transferring of one machine's experience to another without the necessity of repeating the training period.

Insight into the basic ideas of such machine training will come from the combined work of psychologists and neurophysiologists as well as mathematicians and engineers. However, it is likely that, internally, machine training will be vastly different from human learning.

SYSTEMS Computer systems will become larger by many orders of magnitude. Nation-wide (and indeed world-wide) computer networks will be commonplace. Subsystems of these mammoth systems will perform
many orders of magnitude. Nation-wide (and indeed

the ability of the system to carry on a two-way communication with its operator in a language almost identical with the operator's everyday speech. Both the machine and the operator will be able to request and receive the particular information each one needs to accomplish their common objective.

A good way to appreciate the complexity of such a system is in terms of some of the features of a nation-wide computing network that a large company might employ.

Factory Control. With inputs of management policy, customer's orders, and present state of the factory, the system will do engineering design on the product, order materials and maintain a warehouse, schedule production, control production and the individual machines and processes necessary for production, and prepare necessary reports.

Management Control. The system will make available to each manager within the company instantaneous reports on the state of his business and that of the company as a whole. It will furnish him with a rapid means of communication with other business within the company and will allow him, through simulation of the marketplace, to see the effects of his decisions on the state of his business, the company and the national economy.

In addition, the computer will have the authority to make many types of management decisions, some of which today are reserved for executive management. To sooth any executive who should happen to read this, let me agree with him that computers will never make *important* executive decisions, but let me suggest that the definition of what constitutes an important decision will undoubtedly change.

In discussing the potentialities of computer systems over the next forty years, one is limited more by his imagination then by fundamental physical or mathematical restrictions. In fact it is likely that much of what has been written here will be realized in twenty years rather than forty. Regardless of the time period, it is clear that machines will have continually increasing capabilities in those areas usually associated with creative mental effort, and that the realization of these capabilities will have far-reaching implications in our economic, social and political system.

Novel Electronic Circuitry

George D. Watkins

In 2012 A.D. we may expect to find *living cell* amplifiers, computers, power supplies, etc. These devices will have been made possible through the basic research of this and the next generation in the new and rapidly growing fields of biophysics and biochemistry. It may be possible, for instance, to isolate, develop, and breed strains of living cells which perform simple logic functions. The role of the new bio-electronic engineer would then be to synthesize (grow) from these basic units larger organisms which could perform extremely complex operations.

Some of the desirable features of the living cell circuit would be the self-healing aspect, the extreme miniaturization, and the efficiency. The unique power supply required (nutrient) would also offer some possible advantages. For instance, as a surrogate organ, a living cell circuit might be planted in the body and live off the nutrient of the body with no additional power supply requirement. A repeater in an undersea cable might live off the micro-organisms in the sea water.

In fifty years we may also anticipate startling advances in solid-state circuitry. Potentially the lattice of a solid crystal affords a matrix capable of incorporating extremely complex circuitry. The surface may have only been scratched with such simple one-dimensional circuits as the semiconductor diode or transistor of today. In the future, as the perfection of crystals is improved and as the techniques for "doping" in a controlled fashion are developed, it may be possible to construct highly complex three-dimensional circuits entirely within a single crystal of solid. The advantage of such circuits would be primarily in the miniaturization, but also perhaps in increased efficiency and high-frequency response.

Much of the required basic knowledge of solids is already available, although simpler approaches may develop. For instance, just as in the transistor, the electrical properties of the microscopic volume elements of the crystal could be controlled through the addition of a variety of defects, impurities, vacancies, interstitials, dislocations, precipitates, grain boundaries, etc. Properties analogous to all of the conventional passive circuit elements can be achieved in this way. In addition, these elements in turn could couple to magnetostrictive, piezoelectric, superconductive, ferroelectric, ferromagnetic, etc., microscopic regions which can themselves be geometrically resonant at specific frequencies. The network could be made electrically active by coupling dc energy in through low resistance paths or selectively by acoustical or electromagnetic energy, etc.

Introduction of the required defect configurations is not possible in terms of present-day technology and significant advances must be made in this area. Radiation "doping" with a scanned electron or ion beam is one possible approach.

Electron Devices for Power Generation in 2012 A.D.

V. C. Wilson

The thermionic converter is an electronic device which converts heat to electricity. It is an extremely high temperature heat engine which operates at a hot temperature of about 2000°K and a cold temperature of about 900°K. Because of its high "cold" temperature, it shows considerable promise for generating electricity for space vehicles. The reason for this is that in outer space the only way to remove unconverted heat is by radiation. The higher the cold temperature of a heat engine the more heat can be removed per unit area of a radiator. Both nuclear energy and concentrated solar energy will be used for the heat source on space vehicles. On the ground thermionic converters will be built inside nuclear reactors. By the year 2012 at least 30 per cent of the fission energy will be converted to electricity by the converters and will be brought out of the reactors on buss bars. The high cold temperature of the converter means that the unconverted 70 per cent of the nuclear heat may be passed onto a steam turbine with 40 per cent conversion efficiency. Thus the converter may be thought of as a topping device to a steam turbine in a large central station. The over-all efficiency of the system would be about 58 per cent. At this efficiency, electricity from nuclear energy will be less expensive then from fossil fuels, so by 2012 the dream of large quantities of electrical energy from the atom will come true. The most important result of this is that we will stop burning fossil fuels for electric power and will save more of the fossil fuels for chemical and drug industries.

Even if electricity were generated at the central station at no cost, the cost of extensive distribution systems is great enough to make one ask if solar energy could be harnessed for home electrical uses. Here the key problem is energy storage. In northern latitudes of 35° or more, by far the greatest amount of solar energy is received in the summer months. This must be stored for winter use. Thermionic converters or photovoltaic sheets could be used to convert the solar energy to electricity. For storage, reversible fuel cells could be used. Present membrane fuel cells when driven backward dissociate water into hydrogen and oxygen and when fed hydrogen and air generate electricity. Hydrogen is too difficult and bulky to store. By 2012 reversible fuel cells which store energy in liquid fuels will be developed. The economics of such solar generator and storage systems will determine how widespread their use will be. Certainly large nuclear central stations will supply the bulk of electrical energy for industry and large cities.

In places where there is abundant solar energy and water near by that must be pumped for irrigation, solar heated thermionic converters and thermoelectric generators in series could be used to generate the electricity for pump motors. These static heat engines in series will probably be more efficient than photo-voltaic converters.

Computers of the Future

P. J. van Heerden

Today the problems of electronically transmitting and receiving information have been largely solved and even brought to a stage of sophistication. Telegraph, telephone, radio and television, in their historical order, are the fruits of this development. Electronic information processing, however—the process of drawing useful conclusions out of raw information—is still in its infancy. This process is achieved by the human brain in a miraculous way, and its only criterion is its success.

Present-day computers are merely fast adding machines whose usefulness is limited to routine clerical operations. However, the next fifty years will see electronic information processing systems brought to a stage of sophistication comparable to that of today's electronic communication systems. This will require two things. In the first place, the basic principles underlying the brain's method of digesting information must be discovered and translated into simple electronic systems. This is as necessary as our simple theory of electromagnetism was for developing detector, amplifier, and transmission systems. In the second place, capacity for electronic storing and processing of information, on a scale equal to that of the human brain, must be developed. The electronic information handling systems of the future will be able to receive information in visual and auditory form, just as humans do. They will be able to arrive at conclusions that are as correct and valuable as those reached by intelligent persons. These future electronic "brains" will show judgment, since judgment is based on a simple mathematical principle of weighing information.

However, in this 50-year interval we will have found out that the bottleneck in making an intelligent machine is in the learning process, and we will have found this learning process to be quite expensive. Thus, society will be faced with a moral decision: whether it will spend the money to train intelligent machines to perform tirelessly in very demanding jobs, or whether it will educate children to have the satisfaction of achievement in the same. I hope that it will not be difficult to choose in favor of mankind.

There is, however, one function in which intelligent machines may be useful. This is in the exploration of outer space, provided it turns out to be too hazardous or too boring for human beings to carry out. Since an intelligent machine, once trained, can be duplicated many fold, one sophisticated training program can produce thousands of intelligent, but identical, electronic astronauts. They will be sent on trips of many years' duration to the farthest planets and the nearest stars. They will be told to use their own judgment in operating the spacecraft, gathering information, and returning to earth when their mission is accomplished.

Some Thoughts on the State of the Technical Science in 2012 A.D.

FRANZ TANK FELLOW IRE

In 1962 when the Institute of Radio Engineers commemorates its jubilee, past and future will be compared. Much has to be learned from the past for it provides insight into the secret process of evolution and into the internal connections of it to technical development. Besides it rewards us with the pictures of important men's lives. Every radio engineer will gladly remember that the wireless transmission of signals by means of sparks and antennas belongs to the most ancient experiments in electronics. In 1791 Galvani made —in connection with his famous "frogleg experiments"— the following investigation. Between two wires freely suspended in space he inserted a spark gap and let it spark over; another two similar wires were shunted by the nerve of a frogleg, the former was still attached to the leg. With every electric spark-over in the "transmitter" the frogleg at the "receiving end" twitched. The beginning of radio telegraphy had been invented, but 100 years passed until its true victory should begin.

How many names are connected to this victory! We shall not hesitate to call for one who perhaps is mentioned too seldom, Nikola Tesla. His thoughts and his will belonged entirely to the future. In 1898 he had already demonstrated the first remote controlled model ship in New York. In 1899 he built a radio station of 200-kw power output in Colorado Springs, working at a wavelength of 2000 meters, and he developed plans for a world-wide radio system. It did not seem impossible to him to have radio communications in the future between earth and planets. But one can see with regard to Nikola Tesla that engineering needs wide theoretical bases for its development and the time to reach its maturity.

Only in the field of logic are exact prognoses possible. Logic is only a part of the vast wealth of our life— but nevertheless an important part. It is well known that the basic laws of physics permit logical conclusions which are as certain as the basic laws themselves. Would it therefore be possible to predict the state of physics— or, more easily of electronics—in 2012 from the bases known today by using logic, *i.e.*, by thinking? Yes and no. Yes, because logic is a very reliable adviser and no, because in the meantime basic laws could be discovered which are qualified to greatly change the future. For an extrapolation on a wide basis our point of view is too restricted. For life possesses uncountable elements of significant importance, removed from the power of rational thinking. Let us remember the fact that many things happen unwillingly and unconsciously. In these fields phantasy, feeling, and intuition are an invaluable help. But they are not without any attraction, and it is important to consider the connections suggested. We always come back to the well-known experience that the human being is a measure of the things. We have to start from this fact if we want to understand the relations of men to technics, progress and culture. The man who creates the technics follows an internal destination. Thus he fulfills a part of his skill. He is born with the ability for thinking, searching and researching. One can say that he is unburdened by the technics from pains, care and hard work, that he looks for and finds protection and help in his vital conflict, and that he reaches for mastery and power—but in the end he creates the technics because he is internally compelled.

F. Tank is a Professor, Ret. (Swiss Federal Institute of Technology, Institut für Hochfrequenztechnik), Frohburgstrasse 174, Zurich, Switzerland. *(Received June 29, 1961.)*

Through engineering the human being has influenced the basis of culture in an important manner. Very ingeniously the prehistorical centuries are distinguished by their art as the stone age, bronze age, and iron age. The inventions of compass and gun powder occurred at the beginning of modern history—the former made it possible to navigate over the oceans, the later intervened in wars and therefore in the way of history.

The most important fact, however, which introduced modern times and gave them more and more character, was wholly intellectual. It was the recognition that by means of experiments basic laws of nature could be found, could be expressed and ruled theoretically, *i.e.*, by logical processes. Copernicus, Kepler, Galilei, Newton, etc., shall be mentioned; many other names could be called for too. The finding of existing basic laws belongs to the biggest achievements of human spirit; its importance can never be overvalued. It gave a fixed basis to the technics and freed men from incorrect superstitions.

The technical age in which we live advances with unheard of consequences. Nothing has the power to stop it on its path until one day the far destination which is unknown to us has been reached.

The five decades from 1862 until 1912 were extremely rich with basic discoveries (ions, electrons, photoelectric effect, thermoelectric effect, electromagnetic waves [Heinrich Hertz], X rays, radioactivity, quantum theory, theory of relativity, Rutherford's model of the atom, etc.). The next five decades from 1912 until 1962 have been used by the extension of the gained knowledge and the technical application thereof. But other basic knowledge has been added, especially in the physics of the nucleus (nuclear disintegration, artificial radioactivity) and of the elementary particles and in solid-state physics as well. These discoveries opened not only completely new technical aspects, but they also required new ideas in the organization of research. They will still engage a large amount of scientific and technical power until 2012.

But now the meaningful question is asked, "Which natural facts, new, and unknown until now, perhaps unpredictable by any theoretical investigation, will be found in the time from 1962 till 2012?" Those things not known in the past, the *new* and the *unknown*, possess the power to change the present and to introduce a different future. This is the reason for the importance of basic research. As Maxwell said at the opening of the Cavendish laboratory, "We need more Faradays, many more Faradays!"

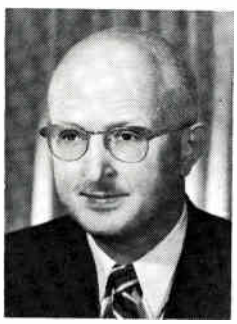
Present-day electronics will have become "classic" in 2012. Today's developments in amplifiers, transmitters, servomechanisms and computers will have reached a certain settlement and to some extent they will have gained a "final form." Miniaturized and subminiaturized elements will play an important role in circuit theory. Radio receivers the size of a match box will be used throughout. Nature, the unrivalled builder, will have taught us and we will have tried to copy human eyes and ears. The gap between radio and light waves will be closed. The maser principle will have found wide applications in great perfection. There will exist an art of millimeter and submillimeter waves. Crystals will find use as wave filters. New, powerful electron sources, high-power transistors, new light sources of highest intensity and great efficiency will be known. But shall we be happier?

The population of the earth will have increased enormously by 2012. Men will live extremely close together. New social problems will thus exist. Raw materials, important for life, will be rare. A state of saturation or even exhaustion will occur. Any further development will then call for very large efforts. A higher technical dimension will bring to man neither more power nor more happiness. He will turn to other things more important to him and perhaps more needed for his maintenance. Certainly man will never be able to live without technics, but he will have to recognize that technics can never be the true purpose of life.

In 1886 Werner von Siemens gave a talk on the age of natural science which will be created by technics. Man will be freed from hard work, there will be food and clothes for everybody. His contemporary Jakob Burckhardt, a Swiss historian, forecast severe world wars. Even then opinions were opposite. It lies in the nature of every qualified engineer to believe in the future and to hope for true and consistent progress in creating a better and happier world. Let us agree with him and hope that he will be right. Only belief in the future strengthens mankind. And even if we do not know today the state of the art in 2012 and are only able to presume, we want to believe nevertheless in a better future. The United States of America should work towards this better and happier world. What a task of historical importance!

Our State of Mind in 2012 A.D.

GEORGE L. HALLER FELLOW IRE



Any prediction about our technological future must assume that no catastrophe will destroy any significant amount of our technological capability. Were such a catastrophe to occur, prediction would be pointless.

To predict something with confidence that it will not take place well before 50 years is easy. Just predict something that seems physically impossible, such as a perpetual motion machine of the second kind.

But if what you predict does not violate the laws of physics as we know them, and if we can now visualize it, someone may produce it before the print is dry. To err by fifty years in looking ahead fifty years is embarrassing, but possible! In this quandary, he who has prediction laid upon him perhaps is safest in retreating to implications of technological advance itself. What can we say about our state of mind, fifty years hence, in a society built upon continual increase in our technological knowledge and prowess?

We can too easily make one of two opposite assumptions about the continuing increase of our knowledge about the physical universe. These two assumptions are: 1) We shall go on learning new facts about the universe at an ever-increasing rate; 2) We have made all of the big discoveries, and all that remains is filling in the details.

The title of H. D. Smyth's book reporting on our World War II technical effort, "Science, the Endless Frontier," implies the first; a statement by Gamow[1] typifies the second: ". . . it is my earnest opinion that the twentieth century will play the same role in the history of our exploration of both macrocosm and microcosm, as the era of the great great geographical discoveries played in the exploration of the surface of the earth."

To deal with the latter assumption first: There seems little warrant for assuming that the universe of natural fact is "shallow" in the sense that Gamow implies. The unfundamental nature of our "fundamental particles" has become a truism. The voids between gravitational theory, electromagnetic theory, and quantum mechanics are obvious. The gap between organic chemistry and biology, though fast breaking down, is still there. Closing these gaps and reconciling these discrepancies may reveal whole new areas of natural law, just as the Curies discovered radium among discrepancies in the radioactive decay rates of known elements.

To deny the second assumption, however, is not to affirm the first. There is a third hypothesis possible. Our rate of acquiring and applying scientific knowledge may slow down (though perhaps never stop), not because of limitations in the universe, but because of limitations in us and in our scientific resources.

We can divide these limitations into three categories: those due to the laws of physics, those due to our own mentalities, and those due to scientific method.

PHYSICAL LAWS Among physical constraints which may impede our rate of technological growth are at least three: the Heisenberg uncertainty principle, the finite velocity of light, and the stringencies of reliability.

Heisenberg Uncertainty. By the terms of this principle, we cannot in the nature of things know both the x coordinate and the x-direction momentum of a particle with an error less than that given by the relation

$$\Delta x \ \Delta p(x) \gtrsim h/2\pi; \qquad (1)$$

where Δx is the error in the x coordinate; $\Delta p(x)$ is the error in the momentum in the x direction; and h is Planck's constant, $6.625(10)^{-27}$ erg sec.

This places an upper limit on the density with which we can pack information into the "memory" of a computer (or a human mind). Suppose that we want to register

[1] G. Gamow, "Will science come to an end?" IRE TRANS. ON MILITARY ELECTRONICS, vol. MIL-1, pp. 26-31; March, 1957.

G. L. Haller is Vice President of the General Electric Co., Syracuse, N.Y. *(Received June 30, 1961.)*

the presence of a "bit" with a piece of material in a box Δx on a side. We want the material to stay in the box for t seconds; therefore Δv must be much less than $\Delta x/t$. Assume that the material fills the box, and that its density is d. Then, substituting in (1),

$$\Delta x \; d \; (\Delta x)^3 \; \Delta x/t \gtrsim h/2\pi; \qquad (\Delta x)^5 \gtrsim ht/2\pi d. \tag{2}$$

The greater the density of the material, the smaller Δx can be. Choose osmium, with a density of 22.5 g/cc, greater than that of any other element. Specify that it stay in the box for at least one day, $8.64(10)^4$ sec. Then Δx, the size of our box, must be more than $2.14(10)^{-5}$ cm on a side. We cannot expect to store information in any "memory," with any permanence, with a density greater than about $(10)^{14}$ bits/cc. (We might be able to increase this storage density by an order of magnitude or so, if we could store our bit of information in the form of a subatomic particle "trapped" in a "potential well" inside an atom or a crystal structure. This requires, first, getting the particle into the "box"; and, second, reducing its energy level well below the height of the "walls" of the "well." The resulting improvement is entirely analogous to that obtainable in information capacity of a communication channel, using Shannon's formulation instead of Hartley and Nyquist's.)

Speed of Light. The speed c of light in a vacuum is $3(10)^{10}$ cm/sec. We cannot travel or transfer information any more rapidly than with this speed. If we want to send men, materials, or messages across interstellar distances, we must take the requisite time. (Einstein's theory of general relativity predicts that the subjective time experienced by the accelerated traveler will be less. He may stay young, but his inertial friends will grow old.) When Mars and the earth are on opposite sides of the sun, the man on Mars must wait for more than an hour to get a reply to his message to the earth. To paraphrase the plaint of the artist, *mundum largum, vita brevis.*

The speed of light is also 30 cm/mμs. If the length of the circuit a query must travel in order to retrieve a bit of information is L, then the access time A for retrieval of that bit is

$$A \gtrsim L/c = (1/30 \text{ m}\mu\text{s/cm}). \tag{3}$$

Reliability. As complexity of apparatus increases, the requirement for reliability of each component element fantastically increases. The mean time per error t of each element in an assembly of N elements, each of which must function correctly if the assembly is to function correctly, approximates

$$t \doteq NT \tag{4}$$

for T and large N, where T is the required time per error for the complete assembly.

Physically-Limited Intelligence. We could define the "intelligence" I of a computer as the product of the size of its memory and its mean time between errors divided by its access time. From (2), (3), and (4), since the distance L must be, on the average, at least $N^{1/3}\,\Delta x$,

$$I = NT/A \lesssim N(t/N)(c/L) = ct/L \lesssim c(2\pi d/h)^{1/5} t^{4/5} N^{-1/3}; \quad \doteq 1.40(10)^{16}(t \text{ in sec})^{4/5} N^{-1/3}. \tag{5}$$

(This dimensionless quantity measures brains per blooper.) The maximum achievable "intelligence" improves with reliability of components, but not quite linearly, and suffers somewhat as memory size increases.

The Heisenberg uncertainty relation and the finite velocity of light may possibly be approximate. We may be able to transcend them, as we have transcended Newton's laws of motion. We have, however, no real warrant for expecting it. We can ameliorate the penalties for complexity, but we cannot entirely transcend them.

HUMAN MENTALITY In someone's words, the human being is the only computer produced by amateurs. (A theologian might quarrel with this.) As a processor of information, a human being has several advantages over present and foreseeable computers:

He can repair himself, physically and mentally, consciously and unconsciously.

He can program himself.

He can adapt his program to unexpected information.

His memory capacity is many orders of magnitude greater than the computer's.

His logical sophistication is many orders of magnitude greater than its. The most sophisticated of our computers has a logical capability roughly equivalent to that of a single human neuron.

He has a variety of input-output devices.

On the other hand, the human being has disadvantages, in comparison with the electronic computer, as a processor of information:

He is subject to fatigue and distraction.

He requires motivation.

His access to his memory is unreliable. (Direct stimulation of brain cells indicates that memories probably do not significantly deteriorate; it is the access to them that becomes unreliable.)

His logical processes are slow and notoriously unreliable. (Is this the penalty he pays, in view of equation (5), for his large memory and logical sophistication?)

He is unable to reproduce on demand most of the logical steps in his processing of information, because he is unaware of them.

His readin-readout processes are several orders of magnitude slower than his logical processes.

His input and output devices (specifically language, the most important) are inexact, and therefore subject to misinterpretation.

Our processing of information will benefit if we can make human beings and electronic computers function together, utilizing the advantages and overcoming the disadvantages of each.[2] But human beings are still self-centered enough to place one constraint on this co-opera-

[2] *Cf.* J. C. R. Licklider, "Man-computer symbiosis," IRE Trans. on Human Factors in Electronics, vol. HFE-1, pp. 4-11; March, 1960.

tion: The distilled information must eventually be accessible to human beings. We are not yet willing to serve merely as auxiliaries to the education of computers.

It might be possible to use computers to improve human thought processes themselves. A human being has access, for example, to about four per cent of his neurons. If we could bring the unused neurons "into the circuit" as needed for thinking or information storage, we might multiply our individual intelligences by 25.

We might be able to improve the speed and accuracy of our output devices. We can read an order of magnitude faster than we can speak. If we had an output device as rapid as reading, we could greatly speed communication among ourselves and with computers.

With the aid of a computer, we might be able to improve our awareness of our thought processes, so that we could more rigorously check our reasoning.

Current work in machine language translation[3] may be teaching us how to build computers which manipulate conceptual entities rather than arithmetic quantities. We might visualize forthcoming developments in computers in terms of the following succession of capabilities: 1) Produce readable translations of technical literature. (Opportunities for mistranslation, and importance of esthetic quality, here are supposedly relatively small.) 2) Translate ordinary newspaper prose into colloquial prose of the target language. (This implies mastery of colloquial syntax.) 3) Translate literature of a high esthetic content, with a skill comparable to that of a competent linguist.

In these three steps, computers will have mastered dealing with concepts as concepts. They will deduce from the context the exact context to which a word or phrase refers, and in the translation include words and context insuring that the author's concept reaches the reader. (We shall then have to write, and to read, much more precisely than now!) 4) Manipulate logical concepts, using symbolic logic. 5) Investigate properties of alternative logics, such as Aristotelian logic, many-valued logic, probabilistic logic, and Hegelian dialectic. 6) "Carry on a conversation" with a human being (perhaps through neuron taps or analysis of brain waves) and analyze his conceptual framework and his methods of thought. By then, machines will be psychoanalyzing people!

This sequence may never happen. We may find ourselves blocked by the limitations of equation (5). We may discover, and be able to delineate that creative something about human thought which we like to think forever beyond machines. In fifty years, nevertheless, we should be well on the way to accomplishment of the first four steps: development of computers which can reason symbolically and produce translations, or report on their conclusions, in colloquial language.

Such a computer would enable us to codify and catalog our scientific knowledge. A human being could study science with minimum wasted effort. He could have access to much more knowledge than he could store.

In the process, human beings may learn to talk like machines quite as much as machines learn to talk like people. Juliet in 2012 A.D. may say of Romeo, "Delta symbol not imply delta referent attribute end."

[3] *Cf. MT/Machine Translation,* Massachusetts Institute of Technology, Cambridge, Mass.; published irregularly.

SCIENTIFIC METHOD We may define scientific method as the collection of verifiable information, the induction of structure in this information, and the publication of results. More simply, scientific method is transferring and processing information. Among limitations possibly inherent in scientific method, there are at least two kinds: one arising from the nature of information transfer, and one arising from the nature of the structure of information in scientific laws.

Information Transfer. Information, like people and things, moves in traffic channels. To date, much of our technological effort has been to provide enough channels so that one may always be available, without significant delay, when wanted. We have designed channels on the assumption that demand would be random.

As we become able to move people and things farther and faster, providing channels to meet random demand becomes increasingly burdensome. The time could theoretically come when there would no longer be room for homes, because expressways take up all of the real estate. The alternative is planned demand: the traveler, as the airplane pilot, would have to submit a "flight plan" and receive clearance before embarking on a trip. The traffic problem is quite general. It is the access-to-memory problem in a computer, and in a human brain. We do not use all of our neurons because we have too few "expressways" between them. In the opinion of some (not all) psychologists, the stuttering of a left-handed child forced to write with his right hand may be due to his having too few neuron paths from the right half of his brain, where he thinks and manages his talking, to the left half, where he controls his right hand.

The same congestion threatens our technical information channels. The flood of technical information today swamps technical journals, ASTIA facilities, technical libraries, and engineers' offices. No one can read everything pertinent to his field. He therefore may duplicate technical work done, and reported, once or often before. When we add to our communication channels the coming flood of data to, from, and among computers, the prospect is frightening.

The solution surely lies in a direction other than ever more journals, ever better microfilming, and ever larger telecommunication bandwidths. It may lie in the direction of some current research[4] on electronic information retrieval for technical libraries.

The technical library of 2012 A.D. may consist, not of books, journals, or microfilm, but of a concept file. The worker would describe his desires symbolically, and receive a reply giving who has done what in the field. The technical journal may then be a data link keeping

[4] *Cf. Current Research and Development in Scientific Documentation,* National Science Foundation, Office of Science Information Services, Washington, D.C.; published twice yearly.

this library up to date. As a matter of prudence, the scientist would check his "library" before embarking on a research problem, to make sure that no one had done, or was doing, substantially the same work.

The Structure of Science. The laws of science are a structure of generalizations. In principle, we can test each generalization as thoroughly as we wish; this is what we mean when we say that the laws of science are "experimentally verifiable." As our scientific knowledge grows, however, the structure becomes continually more complex. The generalization of yesterday becomes a particular case of the generalization of today. As the structure grows, the relation of the latest generalization to the undergirding of fact becomes ever more complex.

Two limitations may arise on the rate of growth of this complex structure. The logical effort required to modify the structure, in the light of an unforeseen fact, increases with its complexity. And the reliability of the structure as a whole depends critically on the reliability of each of its logical elements. We have here an analogy, in the domain of scientific concept, to the limitations on a computer implied in equation (5). Both the size and the complexity of the structure of scientific thought may limit the rate at which it can grow.

By 2012 A.D., we may expect our electronic conceptual computer and library to reinforce the reliability of our reasoning and to speed the correlation of facts. But the size and complexity of this computer library may itself grow to the point of diminishing return. We shall then need a methodological "breakthrough" of the first magnitude.

CONCLUSION By 2012 A.D., the association between the human scientist and his electronic aids to thought will be close, complex, and intimate. To us, the distinction between "me" and "my computer" would then be difficult to make; but the scientist then will think more precisely, and be much more aware of his thought processes, because by then his computer will have begun to analyze his thinking for him.

Human intelligence will not in 2012 A.D. be so dominant a limitation as today on the rate of growth of scientific knowledge. The limitation will increasingly come from the nature of physical reality and the nature of scientific method.

The scientist of 2012 A.D. will feel these limitations. One of his important areas of study will be ontology: the nature of physical reality. Another will be epistemology: the nature of knowledge. The scientist of 2012 A.D. will be a philosopher.

The Use of Electronic Computers in the Social Sciences

IRVING WOLFF FELLOW IRE

Digital computing machines have been applied with tremendous impact in the fields of data processing and technical computation. In the next 50 years we can look forward to an extension of these applications. And with the help of electronic computers, we can hope for the development of new quantitative sciences where only rough qualitative analysis is possible today.

In general, the computers have been used in fields where the rules for solving the problem assigned to the machine are known. Their help has been required to reach a desired objective more efficiently or economically. In some cases, the computations are so time consuming that they can only be completed in a reasonable time with the aid of an electronic computer. Only infrequently have the computers been used to convert a qualitative discipline to a quantitative discipline.

The prediction of weather, a program which engages large digital computers, lies in the twilight zone between a quantitative and qualitative science and is taken as an illustration of the philosophy involved in the use of machines for applications which are in a more primitive state of scientific development. The physical laws underlying the flow of weather are fairly well known. The problems of assembling the data required to accurately compute the weather in advance and the consumation of the computations are, however, so stupendous that the weather would be long past before the calculations could be made using ordinary computer methods. More important, we must be able to check the hypothesis against the actual results in order to increase the accuracy of our predictions. This requires many trial

I. Wolff, Chairman of the Education Committee, Radio Corporation of America, Princeton, N.J., was Vice President of Research at RCA Laboratories before retiring in 1959. *(Received July 5, 1961.)*

computations, of which each can be terrifically involved.

The electronic weather computer solves both problems for us: 1) it enables us to make the difficult calculations we could not make otherwise, and 2) it permits us to make enough checks of results against predictions to provide the basis for an improved theory.

The scientific approach to a problem consists fundamentally in devising a hypothesis from which predictions are made which can be checked by experiment, and in performing an objective experiment to check the hypothesis. As discrepancies are found, the hypothesis is modified and remodified to fit the results of successive experiments. Often the experiment may come first and stimulate the theory. The continuous interplay between hypothesis and experiment is basic to the scientific method.

The scientific method hasn't been applied to any great extent in the social sciences because 1) the interactions of the variables are so complicated that mathematical reasoning from the hypothesis becomes impractically complex with the means we now have at our disposal, and 2) controlled experiments can only be set up with great difficulty, and usually not at all.

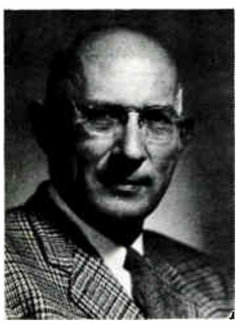
I gave the example of the weather computer to illustrate the principles even though the difficulty in finding the solution is much less severe than is encountered in the most elementary problems in the social sciences. Weather is determined entirely by natural phenomena obeying physical laws on which a fund of knowledge is rapidly developing. The results in the social sciences are determined to a great extent by the laws governing the reactions and interactions of groups of people and individuals. Very little is now known about these laws.

Although controlled experiments cannot usually be set up, there is a wealth of material on the history of events and the factors which influenced these events. With sufficiently sophisticated computers, various models representing individuals statistically and their reactions to various influences may be constructed. These models can be tested under a variety of input conditions to obtain a model which represents the actual events which took place. As a fund of knowledge is built up, the machine can be tested against predictions and then restudied and modified to accord better with the actual happenings. There is every reason to hope that the error will diminish as a social science technology is built up.

If the model which is assumed can be built, and I think that there is little doubt that this will be feasible, the condition for interplay between hypothesis and experiment which is fundamental to the scientific method will have been met, and the social sciences can develop to take their place alongside the natural sciences as exact sciences.

The Biomorphic Development of Electronics

MARCEL J. E. GOLAY FELLOW IRE



When attempting any forecast, there is often the ghastly recall of predictions of the past which were not fulfilled and especially restrictive predictions, such as the turn of the century's prediction that henceforth physics would be restricted to adding a decimal to our measurements of nature's constants, or the more popularly known prediction that the airplane would not fly. Such unfulfilled prophesies may cause one to shy from being categorical about impossibilities. Yet our increased knowledge of the possible is paralleled by an increased knowledge of the impossible, between which the gap of the unknown and of the undone is constantly compressed by discovery and invention.

With this knowledge, we should not be timid about denying the use of gravitational waves, or the transmission of power by means of radio waves. We should also be categorical about the nonexistence of emanated "brain waves" and the impossibility of telepathy or other forms of extrasensory perception. We should understand that beliefs in these are the remnants of pathetic yearnings for a dualistic world which is incompatible with the factual disclosures of science, within a consistent world in which occasional apparent inconsistencies are the scientific puzzles, the solutions of which constitute the great steps forward of science. Only the concept of a consistent, monistic world, in which "spirit" becomes increasingly indentified with pattern, squares with the facts and permits a true appreciation of the ever deepening mystery to which research leads us.

Let us, therefore, make short shrift of the fantastic. Instead, let us try to extrapolate our biological path into our future, and look in our future for clues as to what

M. J. E. Golay is Extraordinary Professor at The Technische Hogeschool, Eindhoven, The Netherlands. *(Received October 19, 1961.)*

is reasonable to expect, or not to expect, of the development of electronics in the service of man and of science.

DEVELOPMENT OF HUMANITY The evolution of life, of thinking man, and of humanity, reveals the hierarchy of patterns which have developed over the last two-billion years, and which will continue to develop in our universe. At the lowest stratum there is the pattern formed by the components of a large protein molecule which, immersed in a nutrient medium, can cause the assembly of another molecule like itself, from which it then separates. This reproducing process constitutes the essence of life at that stratum. At the next stratum, there are the patterns formed by assemblies of living cells: the organized supply of nutrient, the sensory and command nerves, the nerves constituting the brain. Within the brain, at the highest level of this second stratum, there is the pattern of the nervous impulses which constitute thought and reflection for man.

Although the patterns associated with the first and second stratum are more visible than the pattern which humanity as a whole is just beginning to fashion, we are strongly tempted to agree with de Chardin's prediction of the "Thinking Planet" as the third and ultimate stratum in this hierarchy.[1] The "Thinking Planet" will evolve as an increasingly tighter association of men, where the supply and the communication channels, not physically tied to man as they shall be, are no less real and effective than the blood vessels and nerves which connect physically the individual cells of a living creature. In this coming biological stratum, tight discipline will not suppress the person nor provoke fear of the larger pattern as it does today in archaic Russia and China. Rather, such discipline will enchance the personality, create emulative and productive tensions, and inspire in man the same kind of love for his planet which he now has for his country.

Whether this goal be achieved in a thousand, or in a million years, we can predict that the evolutionary principle which produced man out of inert matter, and which gives priority to victory over defeat, to intelligence over stupidity, and to life over death, will continue to operate. And we can foresee that electronics will play a major role in the evolution of the third stratum which has just begun, and will have a development which shall parallel the biological development of this stratum.

DEVELOPMENT OF ELECTRONICS The three basic functions of electronics in the service of man are communication, computation and automation. Let us speculate about their future development in this order.

Communication comprises, in the large, telephony, broadcasting and television, and a host of accessory services such as radar and traffic control. Since communication extends over distances, while computation and automation have a local character, it seems reasonable to conjecture that the communication phase of electronics will be more intimately linked to our biological development than the other two phases.

The individual who, 50 years ago, would have predicted that the automatic telephone would be in use everywhere today in the United States would have been proven an optimist by the facts. Let us, therefore, consider first a plausible ultimate in telephony, and prudently back off from that ultimate to fifty years hence. The plausible ultimate would be this: the individual is given a number at birth, and when of age carries a tiny wrist set which, through a system of overhead or buried grids connected to long distance trunks, permits him to seek pulse-coded voice communication with any other individual, whose number he knows. What fraction of this ultimate will be realized in fifty years? Probably a very small fraction. The bulk of individual mobile communication will still be carried in automobiles and the most important development may be the wider application of pulse-code modulation. The tiny wrist set, mentioned above, may only later emerge as a prototype designed for the amazement of stockholders at their annual meeting. I believe that it is in the much more distant future that telephony, telegraphy and facsimile will have their tightly woven net, so essential to match the biological development of humanity in the "Thinking Planet."

I can say but little of the development of radio and television within the next 50 years, except that today's needs, which are more cultural than technical, may begin to be fulfilled. Even later the development of broadcasting and television will parallel the development of telephony in tightening the biological net by increasing the real educational and reportorial values, despite intervening periods in which thought control for political or commercial purposes will be the main aim of the broadcasters.

What can be said of radar? In military radar, the clever little games of counter measures and counter-counter measures will continue, and will make increasingly heavier demands on computer-type techniques of signal processing. As for nonmilitary radar, its main application will be in traffic control, which will be discussed briefly below, in very general terms.

I foresee traffic control as that chapter of communication on which there will be the most written in the next fifty years, for more elaborate traffic control alone can serve to reconcile the conflicting requirements of greater speed and greater traffic volume on the one hand, and greater safety on the other. The most urgent demands are for greater air safety and for a solution of the paradox that the basically safest mode of public transportation (what can be safer than a ship surrounded by a mile thick cushion of air?) is in fact the least safe today, chiefly because the pilot lacks the information he requires to realize the safety potential of his ship.

Will electronic controls invade the field of private automobile traffic? Shall we have, as some advertisements have shown, automobiles automatically piloted along the highways, while the driver plays cards after having given a program of instructions to the automatic pilot?

[1] Teilhard de Chardin, "Le Phénomène Humain," Editions du Seuil, Paris, France; 1957.

But then, shall we have a computer who will do the card playing as well, while the driver does—what? These questions are more psychological than technical. I believe that the modern trend toward the sports car, in which the driver is more integrated with the road and finds a new dimension to driving enjoyment, will prevent the intrusion of extreme automaticity in one's personal pleasure, except for safety warnings.

The future of the Electronic Computer, that bright offspring of electronics, is a philosophical puzzle. We may foresee new devices which will give us more compactness, greater computational speed, and lower production costs. A superficial examination may lead us to accept the electronic computer as a speeded up version of the desk calculator and a mere accessory to our arithmetical propensity, which has given mathematics a new dimension in furnishing practical solutions to analytically intractable problems. Thus, we may be inclined to characterize it as simply adding six orders of magnitude to the speed of the mechanical computer. But this is exactly where we have to be careful, for Big Numbers have a way of reaching transcendence.

For example, consider the lowest form of living cell—the large protein molecule, with a Gestalt measurable in terms of information bits, some 1500 in number —a molecule which can reach for its component in a nutrient bath and assemble another molecule like itself from which it then separates itself. But 1500 bits is also what Von Neumann calculated for the structural measure of a hypothetical man-made machine capable of reaching into bins for its components, and of assembling another machine like itself. Coincidence?

Next consider the course of events over a two-billion year period, during which organisms evolved on the basis of the original living cell to culminate in thinking and reflecting man, a most unlikely assemblage of cells with some 10^{10} neurons in his brain alone. Is there something magical about the number 10^{10}? Will another Von Neumann compute it is a likely Gestalt measure of a well-engineered organism with input and output organs capable of thinking that it thinks, of having that quantitatively elusive property called intelligence?

Big numbers reappear in Godel's proof, this intellectual triumph of our century, showing as it does that even arithmetic is not a closed discipline, but an open one, with at least one, or an infinity of, unprovable propositions.

In a former article,[2] I have suggested what reflecting intelligence—and we know only one human intelligence—may be called upon to accomplish an indeterminate number of years hence. Whether the somewhat fantastic speculations indulged in that article are justified or not, I cling to the suspicion that Big Numbers may be connected with Big Questions, and that we should not rule out the possibility that the several orders of magnitude increase in speed which electronic arithmetical calculation realizes may form some day, fifty or a thousand years hence, the basis for the transformation of the stupid electronic computers of today into thinking machines which will teach us basically new concepts.

Automation, the third function of electronics to be considered here, is a double-edged sword, for it is less the problems it will solve, than the problems it will create, which are of concern to us.

Our age has been called, perhaps improperly, the atomic age, and a more fitting name could be the age of automation; for atomic power is just another form of power to relieve man's muscles, while automation is fast developing, for the first time in our civilization, into a means for replacing man's skill and judgment in manufacture and various forms of control. Fifty years from today the new industrial revolution now beginning will have made great strides, pushed as it will be by a tremendous economic impetus. Labor as it exists today will dwindle in social importance, and new exciting social problems will develop. Greater demands will be made on education, at first to train the new form of labor now emerging, and subsequently to permit more worthwhile interests to fill the increasingly more numerous hours of leisure which many will have. I doubt that the old mob cry for "panem et circenes" of Roman times will be tomorrow's motto for more than a fraction of the multitude. The social unrest created in part by shortened work hours, and also by the demand for a new form of labor, will be a mentally and morally stimulating thing, already manifested by today's increased seriousness of purpose of the younger age group, and by the reappraisal of the beliefs and tenets of yesterday.

Some new technical-social functions of electronics have been sketchily discussed in connection with the speculations about the future of computers and automation. Two technical functions should be added: pattern recognition, and adaptive or self-curing systems.

Both are challenged by the biological functions of the higher animals and both will attempt to reproduce the manner in which these biological functions perform, but here a contradistinction can be made. For biological processes are based on a hierarchy of patterns which, through reproduction of the lower patterns, the cell, allows for growth and for replacement of the dead cells, a hierarchy absent in electronic circuits. For this reason, I believe that, within the next fifty years at least, the greatest successes in both pattern recognition and adaptivity will not be inspired by a greater understanding of the working of our brain and body, but rather by developing new concepts and new symbolisms better suited to our less elaborate electronic devices.

And yet, one's imagination cannot help taking a flight of fancy. For I can step out of the room where I am writing, and in a few minutes return with a few billion dollars worth of computers in the hollow of my hand, in the form of the brain of frog, snake, or mouse. May it be that someday we shall be able to utilize that awe-inspiring property of growth and organization of living things, and, perhaps by proper feeding of bland chemical with desirable properties, which will be carried to every nerve and ganglion of the animal, create within the body and brain of the animal the framework of an

[2] M. J. E. Golay, "Reflections of a communication engineer," Proc. IRE, vol. 49, pp. 1378-1382; September, 1961.

electronic circuit which, after proper chemical processing, will reproduce electronically some of the living functions?

STATE OF THE ART Having speculated on the probable development of electronics in the service of man by considering their strong interconnection to the probable development of our social pattern, we will now consider the less predictable development of the state of the art: new materials, new devices, new technical functions.

This lesser predictability of basic electronics suggests another contradistinction between biological development and electronic development.

On the one hand we have the living cell which, once developed some two-billion years ago, did not alter radically while the biological development of the various animal species took place. Similarly, it is unlikely that man will alter radically while the biological development of society takes place; *i.e.*, there will be no superman. It is this freezing of the basic unit, while the complex of units develops, which permits a measure of speculation about the development of the social complex, and of applied electronics as a parallel, interconnected development.

On the other hand, paralleling the development of applied electronics, there is a continued vigorous effort to develop new basic "electronic cells," which will replace the old in the applications, as well as permit foreseeable new applications, such as the universal wrist telephone set mentioned earlier. It is this possibility of radical changes in the basic structure which differentiates it from the nearly frozen biological basic structures: the living cell and man.

Cautioned by this fundamental difference, we may speculate gingerly about advances of the basic electronic art. Fifty years ago the basic electronic cell was evolved: the "vacuum tube." And fourteen years ago a competitor, the transistor, made its dramatic entrance, which would have shattered any forecast that the vacuum tube was a terminal form. Is the transistor a terminal form, subject only to slight manufacturing modifications such as evaporated and miniaturized units? Or may we expect the transistor to be dramatically reduced in importance by the appearance of a newcomer? In spite of the upset which took place fourteen years ago, I believe the transistor to be a terminal form, just like the electric generators and motors evolved a century ago, with that single qualification: at room temperature. For I believe also that so many new material properties will be discovered at cryogenic temperatures, that we may anticipate, fifty years hence, the incorporation in every major aircraft of a cryostat in which a major portion of the aircraft electronics will be immersed.

I believe that, in contrast to the transistor, another electronic cell, the memory element, is due for radical changes at room temperature as well as at cryogenic temperature, for a much greater variety of material properties can be utilized for this important element. I also believe that in 50 years, a man-made memory unit having terminal character will have been evolved.

Electronics and Health Care

V. K. ZWORYKIN FELLOW IRE

To gain some idea of the health care procedures which may prevail in 2012 A.D., let us see what happens when Mr. Jones reports for his annual health check-up at the Middletown Clinic. As he enters he is ushered into an examination booth, inserting his coded social security card into an appropriate slot for identification of the examination record. A series of standardized questions concerning his physical condition are then flashed on a screen in front of him and he records his answers by means of yes-no push buttons. Weight, temperature, respiration rate, electrocardiogram, reflexes, and other data are registered directly. Blood, breath, and urine specimens are inserted into analytical machines, which add corpuscle counts and chemical data to the record. While Mr. Jones goes on his way, the examination record is transmitted in coded form to the regional electronic health record storage center, where it comes to form part of Mr. Jones' permanent health record.

The examination record, supplemented by a record of the changes in the physical measurements since the last examination, is now forwarded to the central diag-

V. K. Zworykin is Honorary Vice President, Radio Corporation of America, RCA Laboratories, Princeton, N.J.
(Received November 30, 1961.)

nostic computer, which processes the data submitted to it on the basis of the accumulated medical knowledge and returns to Mr. Jones' personal physician a report listing possible illnesses, if any, which may be indicated by the results of the examination and further tests which may be needed to distinguish between them. Depending on the contents of the report, the physician may notify Mr. Jones that his physical condition is entirely satisfactory; as an alternative, he may make an appointment for further tests at the clinic, or request an immediate personal interview. If further tests are performed, the completed examination record will again be submitted to the health record storage center and to the central diagnostic computer and a more specific report covering Mr. Jones' ailments and suitable therapy for them will be transmitted to the physician to aid him in outlining a plan of health care which will correct present deficiencies and prevent further deterioration.

Freed of much of the routine effort of physical examinations as well as of the necessity of keeping abreast of new developments in the diagnosis and therapy of physical disease, Mr. Jones' physician, as a general practitioner, is increasingly concerned with his patient's emotional well-being and social adjustment; in other words, he is assuming to a greater degree the role of the family physician, a role which had almost vanished before the advent of the central diagnostic computer.

If medical therapy is indicated, with the concurrence of the physician, the computer will write the prescription for Mr. Jones, taking into account such factors as his weight, age, physical deficiencies, drug idiosyncracies, familial characteristics, and other factors. The prescription may be presented by the computer and handed to Mr. Jones, together with a bill for services, or if he wishes, he will insert his personal credit card into the machine and instructions will be sent to a centralized, automated pharmacy which will compound the prescription under sterile conditions, or the preparation itself may be electrically sterilized by ultrasonic or penetrating radiation.

Whether Mr. Jones travels to the pharmacy or the prescription is delivered to him automatically, the vehicle will be guided and instructed as to its destination electronically; the source of power being electronic or nuclear, rather than chemical.

If psychotherapy is indicated, this psychotherapy will again be influenced by electronic methods and concepts. From crude beginnings with drastic shock therapy during the 20th century, the electronic psychotherapy will have progressed to where much more subtle methods will be applied, not in a gross manner, but to the individual nerve ends. The technique of making multitudinous connections for long periods of time to living tissues will have been worked out as a by-product of earlier work on microminiaturization. Should Mr. Jones require sedation, this will also be electronic, as an outgrowth of pioneering work in electric sleep therapy.

On the other hand, if the diagnostic report indicates the need for surgery, the physician refers Mr. Jones to a surgeon-specialist at the district hospital. The surgeon, as always, combines superb manual skill with a thorough knowledge of anatomy. However, he is now not only aided by many new tools of observation and for carrying out the surgery, but also by a stock of replacement organs with built-in electronic feedback controls for implantation whenever the original organ has deteriorated beyond repair.

If, however, his organs are still capable of functioning, but need a period of suspension of the normal activities, so that rest will produce its own healing, or if slight stimulation is necessary, then a whole series of electronic controls such as electron-actuated heart pacemakers or phrenic nerve stimulators may be brought into play.

If the patient suffers from a nervous disease, in which normal impulses to his muscles are missing, these may be replaced by carefully worked out electronically programmed impulses which prevent the muscles from atrophying, due to lack of use.

If Mr. Jones needs a really complete rest, under conditions requiring as little effort as possible, he may be transported under cross-electronic control on his journey, to a sealed, gravity-free, environment in space where the physical exertion required to perform a task will be greatly reduced.

If Mr. Jones is suffering from a contagious disease, he may be isolated either in space, or in an isolation ward in which the air is electronically sterilized by curtains of electrically medicated, ultraviolet, or other lethal radiation, and in which the air and pressure is automatically freed from any factors which might be harmful.

In detail, the preceding description may well be phantasy. However, the use of electronics for making both health records and the sum total of medical knowledge more immediately accessible to the practicing physician appears so necessary that it can be regarded as an inevitable development. Its role in therapy and surgery, which is certain to expand greatly with the trend toward miniaturization, is evident even in present developments.

Electronic Instrumentation—Biophysics

HAROLD A. WHEELER FELLOW IRE

Electronic instrumentation plays a major part in biophysics, but there is a great need for further developments and their utilization. One of the earliest examples is the electrocardiograph, giving the first quantitative record of heart beats. More recently, the electro-encephalograph has been performing the same function for brain waves, but less effectively because of the greater difficulty of interpretation. Whereas both of these utilize wire connections, we now have radio-wave telemetering from sensing devices so small that one can be located internally and can perform its function continually for many days; naturally these require cooperative external equipment in the vicinity.

The next 50 years will see the development of such instrumentation in great variety and its increasing use in everyday life, as well as in the doctor's office and the medical center. Some introduction to the future is already found in the IRE TRANSACTIONS ON BIO-MEDICAL ELECTRONICS.

SENSORS, INDICATORS AND CONTROLS The advancement of technology requires first the collection of facts by quantitative measurement, and then the utilization of such facts for a useful purpose. The latter may be accomplished by an indicator for human observation and action, or a direct circuit to an automatic control.

In biophysics, the variety of electronic sensors is growing steadily, and will far surpass our present imagination. But let us consider as an example the possibilities in a system that is already conceived.

Since the heart action currents were among the earliest of biological phenomena to be amplified electronically, let us take this as a starting point. The first step forward would be a pulse counter to supplant the operation now performed by thousands of nurses every day in a manner that is far from efficient by machine standards. Instead, we should use a miniature amplifier and timer, requiring only that the patient hold two electrodes, one in each hand, for 15 seconds. The indicator would show the pulse rate more accurately than it can be counted, and should then be utilized to punch it on a card, along with the date and hour in a suitable code. This operation could be performed by any reliable person, requiring no apprenticeship.

The second step forward would be the continual observation of pulse rate by a computer which would signal any peculiarity. This might be set to respond to a pulse rate too high or too low, or to a change in the average rate from one minute to the next, or to a change in the character of the pulse. The patient is notably insensitive to such changes; he might even be going about his daily work.

The third step forward would be an automatic action in response to such a signal. The failure of the pulse could activate a pulse stimulus from a local source, which could even be in step with the pulse cadence before failure. At present, this requires internal electrodes near the heart, but we can imagine the discovery of an external stimulating circuit that would not otherwise disturb the patient.

A less tangible phenomenon is fatigue, such as autohypnosis while driving a car under monotonous condi-

---

H. A. Wheeler is President of Wheeler Laboratories, Inc., Great Neck, N.Y. (A subsidiary of Hazeltine Corporation.)
(Received September 5, 1961.)

tions. To combat this menace, we should expect to develop a sensor for warning of such fatigue. This would actuate a signal to the person, and perhaps even a therapy effective as a countermeasure, with or without the knowledge of the person. The action could utilize sound, heating or cooling pads, or electrical "exercise" of nerves and muscles other than those required for driving.

NERVE SIGNALS, DETECTION AND INJECTION

In biophysics, one of the most frustrating obstacles is the difficulty of coupling between an electrical circuit and the well-protected nerve fibers of a living organism, especially in the case of a human being. Like other obstacles that have seemed insurmountable, we may assume that this one will be overcome, either by a frontal attack or by a restatement of the problem.

For example, we might embark on a thorough study of the accessibility of the nerve fibers by means of frequencies over the entire electromagnetic spectrum from sub-audible to optical. It is conceivable that there are "windows" in some parts of the spectrum, and that there may be some useful coupling. For example, the induction of internal currents by sharp pulses of magnetic field might offer one way of penetrating the protective shielding. Admittedly it is difficult to visualize how an external coupler could single out one or a few of the bundled nerve fibers, but we might discover that this is unnecessary for some purposes.

Once having established a coupling, it becomes a two-way street. We have enough knowledge of nerve signals to know that they are sophisticated beyond all expectations; they have a language of their own. Is it expecting too much, that we may find a way of communicating to them from the immediate exterior—or perhaps from a harmless surgical incision? For example, anesthesia may be accomplished in special ways by "jamming" or "spoofing" techniques as are used in communications. (The present simple form of electrical anesthesia is presumably a case in point.)

There are many human needs and healthy desires that might be satisfied artificially, especially some forms of electrical exercise as mentioned elsewhere in this symposium. Whatever is helpful to the disabled patient has potential application for increasing the capabilities of a normal person; this is a principle that needs wider application, as in the case of preventive medicine.

In the nervous system and the brain we have a communication system with certain abilities and limitations. Its natural utilization is spectacular, though usually taken for granted. However, the ability to couple in and out of this system would offer a challenge to make even further utilization of this part of the human machine just as we rely on tools to multiply the capacity of our manual effort.

THERAPY AND ANESTHESIA There are experiments in progress that promise exciting developments in electrical anesthesia. Electrical therapy for other purposes has been used for years, but mainly in very primitive forms and practices. In this field, electronics offers the ability to sense phenomena that may be used to control the treatment, and to generate peculiar electrical stimuli that may be the most effective treatment. In the most advanced form, the latter is responsive to the former.

A major problem is the stagnation type of fatigue, which develops when a patient is unconscious (perhaps asleep) or is immobilized (perhaps under traction). In a lesser degree, the same problem confronts a healthy person when he is "restless," or has difficulty getting the rest he needs. Electronic sensing of this condition may be possible, or the condition may be recognizable by ordinary senses. A future treatment may be the use of electronically generated stimuli of certain patterns to which the nerves and muscles will respond with little or no disturbance to the person.

One problem is how to couple electrically between the exterior and the well protected nerves and muscles; that is beyond the scope of this discussion.

This approach may be termed "exercising," but the objective goes much further. One desired result is some form of exercising, or a substitute, which will accomplish relaxation without effort or motion or any objectionable sensations. They may also be helpful for building up muscle tone in a normal person while he is engaged in a sedentary occupation.

Diagnostics

URNER LIDDEL FELLOW IRE

It has been the custom of the medical profession for years to measure the temperature of a patient as an indication of whether a general fever exists. This is a lengthy (at least 3 minutes) operation and does not identify any area of illness. Advanced electronic techniques should assist the diagnostician to identify local areas of fever which might be identified instantly by such devices as infrared "heat scopes." In a more erudite fashion, it is conceivable that some "resonance technique" could be used in determining disfunction of the various critical organs of the body. Old time members associated with the medical profession will remember the various "quack" devices foisted off on unsuspecting medical doctors as "scientific" diagnostic machines. But we are learning more about proton magnetic resonance now and how it might measure the characteristics of molecules undergoing metabolic action, so that a true diagnostic machine is not inconceivable.

The most difficult area of research is that of the study of electroencephalography, without the use of attached or inserted electrodes. Should it be possible to measure the movement of ions in the brain without penetrating the epithelial tissue, information concerning cerebral action would be available which would open up an entirely new field of research. The ramifications of such studies could have profound effects on our civilization taking extra-sensory preception from the realm of the conjectural into the area of practical activity.

Another intriguing use of electronics is the possibility of using this kind of device for remote surgery, *i.e.*, destruction of tissue within the body without surgical incision. It is at least worth thinking about that one could destroy the tissue in such a fashion that it could be removed by standard metabolical processes without the use of metal devices, and without the dangers of external infection of tissues.

U. Liddel is Assistant Director, Hughes Research Laboratory, 3011 Malibu Canyon Rd., Malibu, Calif. *(Received June 15, 1961.)*

Bio-Medical Electronics— 2012 A.D.

LEE B. LUSTED FELLOW IRE

ORGANS Nearly all of the body organs can now be replaced by compact artificial organs with built-in control systems. This includes the heart, kidneys, stomach and even liver. A great variety of plastic materials enable the organs to be moulded to size and still operate efficiently. The micro-micro electronic components are embedded in the wall of the organ, and the micro-control system for the organ are connected directly to the proper nerves. The ability to find the correct nerve circuits and to trace the circuit like a telephone repairman was developed in the early 1990's and it is this technique, made possible by microelectrodes and very compact flexible computers, which permits the quick hook-up of artificial organs. Incidentally, these same trace techniques made it possible to find faulty nerve connections in the brain or the spinal cord where nerves had been damaged by tumor or trauma. Very fine biological conducting materials are inserted and quickly connected after the damaged nerve tissue has been removed.

If the artificial organ required for some reason an external power supply, a pickup circuit is either tattooed in the skin or the natural pigment of the skin is rearranged to form a pickup circuit. Of course, the connections between the artificial organ and the skin pickup circuit are made very readily.

Inspection and Repair. Long flexible fiber optic bundles can be inserted into any body orifice and also into arteries and veins so that the physician can look around inside any organ in the body. Coherent light sources provide ample intensity to inspect any surface. Small instruments at the end of the long shaft can be manipulated by the physician. This permits him to cut off a small piece of tissue from the bowel wall, and hold it before the end of the shaft for inspection. If he is in doubt whether the tissue is malignant (the cause of some types of cancer have been discovered by 2012 A.D., but some are still an enigma) he feeds the image on the fiber bundle eyepiece to the video input of his diagnostic aide computer which immediately gives him an expert opinion on whether the cells are malignant. If the cells are malignant he can switch on high light intensity and cauterize the malignant cells. However, before he does this he has taken some sample malignant cells for analysis. He intends to find the chemical system or virus causing the cancer, and by feeding the patient the proper "homing" enzymes and chemicals to eradicate any cancer cells which may have spread to other parts of the body.

With the fiber optic system and small tools on the end of the shaft, the surgeon is easily able to repair faulty heart valves by going inside the heart itself through the aorta and vena cava. Of course, he has a large variety of excellent synthetic materials to use in the repair. He is also able to find and remove blood clots in blood vessels before harm is done. The old unsolved problem of the 1960's of trying to find a small bleeding vessel in the small intestine is easily solved now because the surgeon can look around inside the intestine, with one fiber scope and simultaneously inspect the blood vessels to the intestine with a fiber scope which he has inserted into the blood vessel system.

The repair of blood vessels in the brain or for the heart has now been made quite easy and brain or heart damage from strokes or heart attacks are no longer to be feared.

GENETIC STUDIES Micro chemical analysis, high speed computers, the extensive use of information theory and recently developed new mathematical techniques now help identify the exact composition of the nucleic acids and a host of enzyme systems. The position of the chemical constituents such as amino acids in all of the complex chemical systems is now known. It is also possible with about a 90 per cent probability of success to replace any defective amino acid within a nucleic acid. This technical development has made it possible to develop a superior system of genetics. Genetic defects which were unavoidable in the twentieth century can now be avoided, or if they do occur they can be repaired.

L. B. Lusted, M.D., is Professor of Bio-Medical Engineering and Associate Professor of Radiology at the University of Rochester, Rochester, N.Y. *(Received June 29, 1961.)*

It should be pointed out that parents can now choose to have a boy or a girl (with about 90 per cent probability of success). Some countries have used this knowledge of genetic material to control the birth rate.

ELECTRONIC DIAGNOSIS, TREATMENT AND PROGNOSIS Beginning in the 1960's, rapid progress on electronic systems of diagnosis was made. At first the most attention was given to diagnosis of disease, but as various diseases were conquered the electronic systems were turned to preventive medicine and problems of checking people's health regularly. The early bio-medical computing systems and electronic diagnosis systems developed in the United States were sent to other countries, such as India, Africa and South America where they were very helpful to the few available doctors who were trying to look after the health of large populations.

By 2000 A.D. people have become accustomed to the regular health checkup most of which they can have in their own homes. On a given day at an appointed time the person goes to some room in the house where equipment has been installed and connected to the local medical center. There is a two-way closed TV system to the center. Sensors for temperature, weight, etc., are connected to the automatic data processor at the medical center. Temperature and weight are recorded, the physician can see and question the person—all data being recorded automatically at the medical center. A small cuff placed about the forefinger does a blood analysis. The blood is analyzed for all the most necessary constituents to tell whether the person is doing well and is healthy. The electrical properties of blood have been well worked out. The values of the blood tests are compared with the values from previous years by the medical center computer. If any abnormality is noted, or if a trend is noted which could be serious, the person is advised to go to see the physician for further tests and necessary treatment. The blood tests have been worked out so that they can characterize an individual much more completely than fingerprints.

Electronic Spectro-Analysis of Chemical Compounds

HENRI BUSIGNIES FELLOW IRE

On a previous occasion,[1] I disclosed a few thoughts about the prediction of events at very low levels. My comments were as follows:

"The urge of educated man to see, feel, and analyze smaller and smaller quantities, as well as to see and communicate farther, has been satisfied with no greater brilliance than in the production of the microscope, telescope, and electronic amplifier. Actually, amplification has already reached in some fields that extreme limit where distinguishable patterns of intelligence or orderliness have disappeared in the apparent chaos of random effects.

"It is at this low level of random effect that all events and all things that later become perceptible to man are born and grow. This applies to thoughts that, in the originating or some other brain, are developed intentionally or otherwise to produce a masterpiece of art or literature, an airplane crash, the tallest building, or a concept of human progress. It includes such things as the molecules that assemble into the seed that may precede by even a decade or two a mature living creature, and also the agitation of the still air that produces the tornado.

"History shows that at some particular place and time and in a favorable enviroment a very minute pattern emerges from chaos and, supported by other forces, feedback, and correlation, attains a growth that finally produces a significant phenomenon of distinguishable proportions. Then through some such effect as saturation, growth stops and stabilization occurs, only to be succeeded by decay. Growth is not linear in most cases; it

[1] H. Busignies, "Growth and amplification," PROC. IRE, vol. 38, p. 379; September, 1950.

H. Busignies is Vice President and General Technical Director of International Telephone and Telegraph Corporation, New York, N.Y. *(Received June 30, 1961.)*

proceeds through thresholds and intermediate steps, and consumes from microseconds to years to produce an end result. At its origin and for a fraction of its early life, the magnitude of the emerging pattern is so minute that it could be influenced, shifted, modified, or destroyed by a force of the same order of magnitude. One is led to believe that, however small they may be, brain waves could influence the outcome of physical events if they were available in suitable form at just the right time, and thus by the proper application of man-made patterns of very small magnitude, beneficial control of large physical effects could be achieved.

"In the superregenerative circuit of Professor Armstrong, a small electromotive force corresponding to a signal pattern just above the thermal-agitation level builds up in a fraction of a millisecond to become perceptible to the human senses: but it could have been influenced by a dissimilar, although just as minute, force applied with a timing accuracy of some microseconds.

"An effect of very small relative magnitude is used by Doctor Langmuir in seeding clouds to throw out of balance a low-level threshold controlling atmospheric conditions over a large area. Similar thresholds exist in many forms and at varying levels; they represent stages of development where potential growth has momentarily stopped.

"In the fields of education and propaganda, thought seeds that are properly timed produce opinions and prejudices that become strongly entrenched. Although they may encounter many thresholds and, particularly, saturations, they are very difficult to modify once they have reached a high level of acceptance.

"From an engineering point of view, one marvels at the possibilities of low-level control of the growth of events and things favorable to humanity. The electronic engineer, with his knowledge and experience in amplification and control, together with his accurate notion of timing, may soon explore and harvest in this promising field of growth control."

Prediction of events before they become strongly apparent is an undertaking in which millions of people are engaged but generally at a level of growth where signs become available directly to our senses or through some of the technical aids available. Prediction at very low levels of energy has only been dealt with in a limited number of occasions.

Let us take an example in the medical field. A cancer probably starts in one cell among the billions constituting the human body and within the cell it probably shows up at its birth in one or a group of molecules as a chemical compound different from normalcy. In order of increasing difficulty in the detection of the disease, we can list the following:

1) Visible symptoms
2) Detection by X-ray examination
3) Detection by an analysis of the blood stream carrying by-products of the cancer
4) Detection when the cancer is limited to a few cells, invisible by X-ray examination
5) Detection when the cancer is within one or a group of cells in the form of chemical compounds different from normalcy.

Detection at the fourth and fifth level, for instance a small group of cells at the fourth level, would represent an enormous progress if it could be achieved.

This is obviously very difficult at present. The solution might consist in scanning rapidly cells or groups of cells with the radiations to which they are responsive and observing their response to identify the changes taking place. The greatest problems lie, first in the access to the deep layers of cells, second in a rapid system of scanning and evaluation of the signals in order to limit the time necessary for the exploration of billions of cells, and third in obtaining a sufficient resolution. These are all major problems, but every one will agree to the importance of their solution to humanity, in providing a machine which would detect harmlessly very early cancers and obviously many other diseases of lesser gravity. A community of 50,000 people could probably afford to possess one in its hospital. As there are only 525,600 minutes a year on a twenty-four hour a day basis, a scanning should not last more than approximately ten minutes in order to permit one check-up a year per person, which would soon be found insufficient. Such machines, if at all subject to invention and design, would probably still be of cost which would allow for more installations, for more checkups or increased scanning time. Aside from the humanitarian viewpoint, the field may be a fruitful one for the proper industry. Few discuss the price of medical services and drugs that they can afford when a human life is in the balance.

I feel very strongly that the applications of electronics to medical research and treatment will develop into an important field and industry in which we can contribute. I hope that the following will incite others to join up in an effort in this direction.

Most biological natural elements, chemicals resulting from disease as well as all the hormones and other chemical compounds are carried by the bloodstream from gland to gland and organs. The last two decades of research in medicine and chemistry has shown the great importance of hormones and vitamins in the control of the entire operation of the human body including the mental faculties. A proper balance of the hormones results in an euphoric state; but when such balance is altered this state changes into a state of mental disability, disease or death. The bloodstream carries also many chemical compounds, natural or by-products of degenerative or infectious diseases or indicative of some unsatisfactory physical condition. Some of these compounds can normally be present in very minute quantities.

It would be extremely useful to make available an instrument of analysis which could in a short period of time report the contents of a small sample of blood in such chemical compounds as the physician wants to investigate. When "tuning" this instrument of analysis on any of the chemical compounds known that the blood can carry, the proportion of this chemical would be known in a matter of seconds. This instrument could be used for the detection of unknown or known by-products of disease and contribute to the discovery of disease conditions, perhaps prior to the outbreak of such disease. It must be sensitive to very small quantities—1 part in 100,000 would be useful in many applications—1 part in a million would furnish very valuable information. Such an instrument would normally have its place in each hospital and clinic, and would also find very wide applications in the chemical and connected industries. Application of elementary information theory and of the present knowledge of electronic systems indicates that such an instrument is subject to design and that work in such a field should be fruitful.

An amplification of the methods mentioned would lead to the detection of a relatively small number of molecules and atoms by observing their emission in a change of energy level. Present observation techniques would not permit noticing this emission, but further development of maser techniques or what we have at times designated molecular radar allowing measurements in very low levels of noise would lead to this result with many possible applications:

Medical diagnosis and treatment
Physical and chemical analysis
Prediction of life of given components
Understanding of life principles
Quality inspection.

Magnetic Recording and Reproduction—2012 A.D.

Marvin Camras fellow IRE

Fifty years ago, in 1962, magnetic recording had already fulfilled earlier prophecies as to sound recording, and had entered into other fields of information storage, serving as a comprehensive memory for computers, and for recording television signals. It became apparent that information storage was essential to every automatic or intelligent system, whether a machine, a living being, or a government. Magnetic recording was an ideal storage medium from a standpoint of compactness, stability, and convenient transduction to or from electric waves.

The first magnetic memory devices resembled tape recorders, just as the first automobiles resembled horse carriages. Eventually tape recorders evolved into the standard memory pack which is presently manufactured in large quantities by specialists. The memory pack of today is a sealed box about the size of a package of playing cards. It holds upwards of 10^{20} bits of information, and has no mechanically moving parts; the recording, readout, and scanning are all electronic. The storage density is so high that the information of entire libraries is condensed into a few cubic feet. Additional modules are added to extend the capacity to any required level.

An important use of the memory pack is for our home viewing consoles. Here we view a recording in stereoscopic color on a large curved mural display, with lifelike sound coming from the imaged performers. Our home movie cameras use a similar record. Immediately after taking a sequence, we can transfer the memory pack to our display console to observe and hear the action that took place a short time previously. Recorded entertainment and information is no longer sold in stores, but is purchased via telephone connections. A person merely requests a recording that he desires, whereupon the recording at the central office is transmitted at high speed over the telephone, and recorded on the home unit. Entertainment is only one aspect of the display system; for example, shopping is largely done at home from memory packs which serve as catalogs. Merchandise is displayed and described very realistically in the home viewer. In the same way we have an animated encyclopedia which gives us the facts on any subject in a more interesting manner than the old books. Our catalogs and encyclopedias are constantly revised by subaudible signals that come over our telephone lines. Many people subscribe to a tele-newspaper service of this kind.

In medicine the memory pack and tele-revision provides every physician with the latest information on diagnosis, treatment, and techniques. He has a magnetic

M. Camras is a Senior Engineer at the Armour Research Foundation, Technology Center, Chicago, Ill., and is Editor of IRE Transactions on Audio. *(Received July 20, 1961.)*

card for each patient giving the complete medical history. He inserts this into a processor and adds current symptoms. The readout gives the present status, further symptoms that should be observed, the most probable treatment, and expected reactions.

Industrial activity now depends largely on three factors: raw materials delivered by pipelines, energy delivered by electric wires, and information received over coaxial cables. The raw materials and the energy are used continuously and there is a steady flow, but the information for directing the processes is stored on the site in memory packs which are revised by remote control according to changes of style or product.

Consumable everyday items such as foods, drugs, motor fuel, and towels are piped into each home in elementary form of fluid or suspension, and reconstituted by a machine which converts or separates the suspension into the desired products. The complex chemical-electric-mechanical processing is controlled by a memory pack.

Money is no longer a medium of exchange. All purchases are now charged directly to one's bank account whenever one presents his magnetic credit card. At the same time the latest balance and a record of what has been purchased is placed on the card for one's personal record. Wages and earnings are also credited continuously to the account. For every transaction the appropriate tax is deducted so that the government is always up to date on collections.

Extending Man's Intellect by Electronics

SIMON RAMO FELLOW IRE

We are all aware of the enormous impact of science and technology in shaping our future lives. This new and growing influence is usually illustrated by pointing to H-bombs and more recently, to space achievements. In the long run, there is an area of technological endeavor of even greater significance. It is the extending of man's intellect, the source itself of all future scientific and social progress. Viewed as a technical field, it is the extension of human intellect by electronics—what we call for convenience "Intellectronics"—into every aspect of the running of our businesses, our educational systems, our government—wherever intellectual processes are involved.

LAW In the practice of law—or at least law as it might be practiced if intellectronics is used to the fullest —every practicing attorney might have in his office a decade or two from now means for convenient electronic connection to a huge national central repository of all of the facts, rules, procedures, and precedents that he needs. For the routine filing of papers, records, petitions, he or his assistant will introduce his data into the intellectronic system by operating a properly designed input device looking a little like a typewriter. Almost immediately, there will be displayed to him on a special viewing screen any conflict, omission, inconsistency, or other shortcoming of his statement—any problems with the law, the existing records, or the claims of another. And this display will not cover just a few possibilities that an unaided, though trained, human brain might have produced in a few days. Instead, the intellectronics system will scan, select, reject, and present the equivalent results of thousands of trained searchers covering many decades of records over the entire nation in a split second.

So the intellectronic system—and we know how to design such systems today—will elevate the lawyer's intellect to the more complex intellectual tasks, give him better tools with which to work. It will alter a substantial fraction of legal practice. Even on the nonroutine, the more intellectual legal processes, the attorney will be able to consult with the equivalent of a host of informed fellow attorneys. His request to the system for similar cases will yield an immediate response from the central store, together with questions and advice filed by other attorneys on those cases, even as he will later add his experience and ideas into the system for future use by all.

The national intellectronics network to serve the legal profession will serve government as well. It will become a vital tool for orderly law enforcement. Moreover, its authoritative and broadly based statistical data will disclose inadequacies in the nation's laws and be of prime aid to our law-making bodies. Speculation will often be replaced by hard, cold, up-to-date facts dis-

S. Ramo is Vice Chairman of the Board of Thompson Ramo Wooldridge, Inc., Los Angeles, Calif. *(Received June 12, 1961.)*

closing clearly where confusion exists and, consequently, where new rules are needed. As a matter of fact, without intellectronics, we might build up to an impossible condition of saturated legal processes and hopelessly jammed and confused procedures. The world is becoming ever more complicated, every action interacts with more, other actions; so the rules of our society are correspondingly becoming highly detailed and the procedures increasingly lengthy. Something new, like an intellectronics system, seems imperative, not just desirable.

MEDICINE How many physicians can your doctor consult with? A practical answer would most often be: maybe two or three. The physician in the future technological age, towards which we are in such rapid transition, will also routinely introduce his data on a patient to this network of "consultative wisdom." The patient's entire past history, the results of all tests, the symptoms and complaints, and even a statement of the family background of inherited tendencies—all this would be efficiently introduced to the medical intellectronics system. The system will quickly react to give the doctor not all, but key portions of what would have been the result of many consultations with other physicians. It will call out questions and possibilities our physician may not have asked himself. It will give statistical probabilities (because it will have the amassed data to do so) of the relative effectiveness of various treatments—with numerous variations and warnings to account for corollary possibilities and complications, all automatically turned up within the machine, triggered by some of the detailed data that the physician introduced about the patient.

With diseases nationally monitored, the statistical approach to medical practice will take on an entirely new stature. Cause and effect relationships will be studied on a large but rapid scale, tying ailment to treatment. And again it will raise the intellectual effort of our physicians.

ENGINEERING The engineering field has already been revolutionized by intellectronics. The intercontinental ballistic missile would today be many years away if we had not extended the brains of the designer by electronics. Thousands of flights would have been needed in a clumsy trial and error approach. Instead, mere dozens of actual flights have been sufficient to finalize the engineering and prove out the design. The thousands of flights took place in the simulation laboratories, and the selection of the right combination of design parameters out of the myriad of possibilities—too huge a task for human comparison and sorting alone—was made by a man-machine partnership.

LIBRARY One of the things that we can do with electronics is create a library of information that is mammoth and yet accessible with electronics speed. But the memory extension is far from the total effect. The properly designed artificial intelligence works on the library's information. It has built into it a set of logical rules chosen by the human partner. With this logic, and with the information, and with its electronic competence, the machine can handle the low intellectual tasks of the first sorting, the categorizing, the comparing, the selecting, and the presenting. These tasks, done well, require typically that tremendous volumes of data be processed quickly. The machine member of the partnership does this high quantity, high rate part of the intellectual job, allowing the higher intellect of the human partner to concentrate on more subtle, less predictable, less routine —the conclusion-drawing, decision-making, judgment—aspects of the intellectual task.

MONEY AND BANKING In money and banking—and the whole process of keeping track of who owns what, where it is, and who owes whom—it is absurd that for this kind of work we still have millions engaged all day in putting little marks on pieces of paper, reading them off, introducing similar ones—without much need for deliberation in the process. This is mostly as unsuitable for the human intellect as pulling huge boulders to build the pyramids was to human muscles. Someday currency and coins will be for the rural areas. Even checks and most other of today's forms of human record originations may become extinct. If you buy a necktie or a house, your thumb before an electronic scanner will identify you, and the network will debit your account and credit the seller. The system will automatically do the routine accounting, will call out the violations of rules, or any problems in the transaction, and list alternatives. (Occasionally a transistor burning out in Waco may accidentally wipe out someone's fortune in Boston. There will be some continued dangers and risks in life in the period ahead, although they will likely be new ones.)

LOGISTICS—RESERVATIONS How about the making of reservations—airlines, hotels—and the waste in human intellectual capacity which this increasingly important aspect of keeping the operations of the world going entails? In the future, we should step up to a telephone-like device and, after consulting something much like a telephone book, dial a reservation request. The instrument will respond by giving a red signal, meaning "no," or displaying alternatives that are available, or indicating an acceptance of our request. It will print out the ticket, automatically charging our account, avoiding errors, confusion, disappointment, uncertainty, and poor use of available accommodations.

So far as technology is concerned, there is nothing about the reservation problem (or money or banking) that requires any new invention in pure science for the application of intellectronics. It takes a great deal of engineering to achieve any one of the specific advances we might cite, and we can apparently do without any given one today. However, in each instance the increased volume and rates, and, again, the interactions and complexities of our growing information-controlled world, press for new systems for handling these intellectual processes.

TRAFFIC CONTROL It is hard to find a better example of the urgent need for man-machine, intellectronics partnerships to handle thinking tasks than the control of things moving in the sky and even on the ground. It is clear that a human brain unaided—whether that of a pilot, an airport controller, or a Los Angeles automobile driver—cannot integrate all of the changing, dynamic, split-second facts to make decisions leading to the smoothest, safest, maximum use of the artery. Intellectronics measurements are needed instantaneously on the nature, quantity, rate of change, spatial spread, and interconnection of traffic. Automatic predictions then need to be made as to consequences, and directions given to control the flow. Not only will the handling of airplanes and the role of the pilot be drastically changed with time, but it is not completely ridiculous to imagine automobiles of the future which go on electronic control, if directed onto a crowded, speedy freeway—the driver limited to push-buttoning his chosen exit.

MANAGEMENT The average business can shift between a profit or a loss position easily—by just raising or degrading the available information for decision-making by management. We are working on management control systems in which information as to what is happening will be electronically compared with the plan—deviations will automatically produce directives for changes in plan in accordance with "stored" logic set in by management to cover many possible situations—or the system will call out warnings, when the unexpected happens, to the higher intellect of the human supervisors who are kept free for the unpredictable, more difficult, nonrecurring situations. This is the same as having more, smarter, and more knowledgeable managers, and it helps to put management on a scientific basis.

LANGUAGE TRANSLATION Most of the physical operations of the world—production, transportation, communication—are candidates for passing under intellectronic systems control, but also much of the operations involve world-wide integration and interconnection. In the coming years, many millions of human minds—and their extensions in the form of signals and data and information collection—will be connected together, often crossing national and language barriers. No wonder one of the most interesting intellectronic areas in which we are engaged is in the translation of natural language. Again, the machine member of the team provides the crude first cut. It also identifies double meaning possibilities and weighs alternative meanings based on what preceded. It assists and sets up for the more brainy human partner, and the combination turns out a far better performance than either alone.

The concept of what language is will probably change drastically in the years ahead. The machine partners in the universal electronic systems of the future will want facts and rules in the most efficient form possible. They will create pressure for a common universal, purely informational, completely logical and consistent kind of language. Natural language, originated in an earlier nontechnical age, does not meet all these conditions. The technological, intellectronics period of the future may force on the world a new kind of language reform.

EDUCATION The most truly intellectual activity of all must be the education of the human brain. Now, we are approaching a crisis in education because, while the needs of the more complicated, more populous world are increasing rapidly, our ability to place human resources behind the educational system is apparently decreasing. But an intellectronics system could make a hundredfold change in the effectiveness of education. The human educator can have tools analogous to the physician's X rays and electrocardiographs. The routine material can be machine-presented, leaving the more difficult concepts for the higher intellects of the human educator. Programmed machines can stimulate thinking of the student. The electronic presenter can speed up or slow down, add more explanations, skip steps, as it makes the presentation—all automatically as a result of continually noting student's push-button response to questions.

An intellectronic system can remember the progress of millions of students. It can compare their tested learning with the estimate, with the plan. It can measure and report deviations. Yet, that same system can immediately recognize an individual student and can give him an accelerated or special presentation or test—all by a virtually instantaneous scan of his record and a following of rules that have been set in by the wiser human educator.

Such a future educational system would involve new large industries employing experts in the subjects to be taught, in the design of the programs, and in the design and production of the devices and the systems. There will be new professional groups within an augmented educational profession to provide for statistical study and planning, diagnosis, and generally for the matching of the synthetic intelligence of the machine with the human brain to achieve the fullest utilization of both human and machine resources in an educational system really suited to the coming technological age.

DEFENSE In assuring the defense of this nation, it is some years now since the first priority was to create the force, exemplified by the bomb, so as to provide the necessary military might for defense purposes. We are already now passing through even the next phase, where we provide the capability of delivering the force wherever it is needed on any part of the earth's surface in a very short time. The next big, urgent, highest priority area in the application of technology to the military is in the intellectual aspect of the military problem—in the control and command, observation and communications. In a world in which one nation can inflict decisive damage on another in a few minutes, even though they are separated by half the earth, sound military action requires knowing what is going on everywhere, and putting together a tremendous amount of diverse, confusing, and continually shifting warnings and other information into sound conclusions ready for major all-out decisions and, indeed, to make minor ones automatically and rapidly in accordance with some prearranged plan. A man-machine, intellectronics partnership is the only answer in sight to this problem.

ELECTRONIC VOTING Now, all of this description, these examples of facets of the coming technological society, can easily create one gloriously wrong impression, namely: The world of the future will be an automated, robot-like, regimented world, with man (even though the controller and originator) nevertheless only a part, a cog. The world will be one in which freedom, democracy, the creative spirit—all these concepts—will be pushed into the background as the mad complex of men, machines, cables, and vehicles interact to produce, move, and keep track of everything that makes up the material aspects of life.

This is not only unnecessary, but unlikely. Suppose we want to assure the greatest of participation by the people of the world in determining plans and policies. We want greater democracy. We want to insure that we are all occupied with thinking about the issues that control our lives. Technology makes possible a higher democracy, a greater assurance of this, if properly employed. Suppose that the deliberative bodies, the congresses of the world, are exposed to every home continuously by a special television-like unit. Suppose also that it is our practice in that technological society of the future for a large fraction of the issues discussed to come up for a vote, not just by the congress, but by the people. To emphasize this point of view, let us imagine an extreme, which would doubtless be beyond either practicality or desirability, in which it is our custom that the registered adult voters would expect several times a day to identify themselves to the machine (with their scanned thumbprints) and to put in a "yes" or a "no" or a vote for "A," "B," or "C." The system would automatically register, check, add up, determine, and announce the majority viewpoint. Clearly, the highly technological society of the future can be one in which communication and action can be so wide-spread, efficient, easy, and commonplace that participation is virtually guaranteed, interest is heightened, and the apathy and ignorance that can destroy democracy is virtually eliminated.

Communication as an Alternative to Travel

J. R. PIERCE FELLOW IRE

While dictators have striven for self-sufficiency and prophets have foretold huger and more concentrated cities, we are actually becoming ever more interdependent, and our industries and ourselves ever more dispersed. Surely one of the great and persisting problems of the future will be that of transportation, of bringing people and industries what they need and of sending forth what they produce.

Goods may spoil through delay, but they do not suffer; men do. They suffer from home or office to airport and they waste time on planes. A day, two days, are spent for the sake of a few hours of face-to-face communication. Perhaps medicine can make up for the lost time by increasing our life span. But, in fifty years, electronics could eliminate the loss altogether by substituting communication for travel.

Surely, one need not move human bodies in order to achieve a meeting of minds. The ingenious gadgets, the broad-band circuits my fellow Fellows have written about, can provide face-to-face or group-to-group sound and vision, safely encrypted to preclude eavesdropping. Documents can be transmitted by rapid facsimile or by data devices. Some control or manipulation at a distance will be possible, equivalent to operating a computer, turning the pages of a catalog, or moving from department to department in a store while shopping. With cheap enough communication, students could attend classes by television.

One may object that he would not care to visit his wife by television. The clear answer is that he could stay home and communicate to work.

J. R. Pierce is Executive Director of Research Communications Principles Division, Bell Telephone Laboratories, Murray Hill, N.J. *(Received November 24, 1961.)*

Controlling Man's Environment

HAROLD A. WHEELER FELLOW IRE

In the control of environment, electronic instrumentation enables sensitive detection and versatility of sensing. A responsive action is likely to be effected by a source of power that need not be electronic.

STORMS The prevention of any catastrophe (a world war or just a tumble) requires sensitive detection of the cause at an early stage of development, so that corrective measures can be applied before it is too late. In political troubles, the causes are human and are difficult to interpret. On the other hand, a major storm is usually an orderly development from a localized cause. This being in the realm of physical science, we have hope for steady progress. The results to be obtained are so spectacular as to justify a major effort, comparable with that being expended on flood control and soil conservation. However, the problem of the storm is different in that the causes and the corrective measures are not spread over a period of months or years.

The first requirement is early detection. Let us consider the tornado or hurricane, which is born in a "calm," an area of stagnation of air currents. If there is also rising air in some location, the resulting vortex is accumulated rather than being dissipated by turbulence. All of these contributing factors are susceptible of measurement, and the rancher or the sea captain may be able to sense them by experience and association. Unfortunately, a large area has to be covered continuously in order to sense the origin of a storm, presumably by interpretation of air currents.

The second requirement is an immediate countermeasure commensurate with the progress of the storm before detection. Here a "clean" nuclear explosion may offer the first workable solution. It is conjectured that one or several such explosions would be able to confuse a vortex that would otherwise develop into a hurricane. There may be found a preferable approach, such as a random pattern of many smaller explosions. But the success of the plan is still dependent on sensing the vortex at a very early stage.

A program of weather control would naturally be directed also to other problems, notably the fog in the vicinity of an airport or seaport.

AIR-CONDITIONING, MULTIDIMENSIONAL An advanced air-conditioning plant controls temperature, humidity, and rate of air replacement. None of these requires any unusual type of sensor.

We are beginning to appreciate some benefits from other characteristics of the air we breathe. For example, a predominance of negative ions over positive ions has been found helpful to some individuals, or for some complaints. (A principal investigator in this field is C. W. Hansell of RCA Laboratories, one of our colleagues in the IRE.) Doubtless there will be discovered other elusive properties of the air that will be correlated with specific benefits to be sought by control.

As an example, we should soon be utilizing electronic instrumentation for sensing the ion content in the air in a room, and then applying this information for controlling the ion content at a favorable level.

H. A. Wheeler is President of Wheeler Laboratories Inc., Great Neck, N.Y. (A subsidiary of Hazeltine Corporation.)
(Received September 5, 1961.)

Public Functions and Standing of the Engineer

C. F. Horne FELLOW IRE

Few authors of our era have been able to match George Orwell in his perspicacious attitude toward what life may be like in the next two generations. In his novel "1984," Orwell paints a wretched scene of a society where a man "can be made to believe that black is white, two plus two equals five, and evil is good."

Orwell's warning against the loss of individual freedom is a suitable backdrop against which we must consider the role of the engineer in 2012. Fifty years from now we can have a society in which all of our scientific and technological advancements are used for the betterment of man. Or we can become the slave of our creations and we can witness the breaking of our individual spirit in "Room 101."

Hopefully, we will choose the former course. In 2012 our society will have progressed to the point that our discoveries in atomic energy and electronics should be used intelligently for the betterment of man. In this society the engineer will be at the pinnacle.

During these next fifty years the engineer and scientist will be forced to accept more and more responsibility outside of his classroom or laboratory. He will be called upon to accept responsibility for basic policy decisions in government, business, and education because he, the scientist or engineer, will have the "tools" and the experience necessary to solve problems which have a high technical or scientific "input." It is already apparent that, with each passing year, more and more national, local and business problems require both training and experience usually found only in engineers and scientists.

Thus our scientist and engineer in the next fifty years must be broadly educated. He must understand the world of people who surround him. He must appreciate the cultural heritage of many civilizations. He must have a spiritual resource which will help provide a sense of balance. Above all our engineer-scientist must maintain his own individuality. He must resist the temptation to become the slave of his machines or of his society.

I know that many words have been written about the need for broadly-trained engineers. I am equally aware that too often our acknowledgment has been little more than "lip service." But the time has come when we must face the fact that in the next fifty years our engineers and scientists must become thoroughly conversant with this new dimension in their careers.

Can we honestly say that our profession will be ready to accept this responsibility? Last year Dr. David A. Lockmiller, the President of Ohio Wesleyan University, noted in a speech at Detroit, Mich., the dilemma of a retiring corporation executive: I have 50 younger executives, but not one is ready to succeed me. Each is well qualified in such fields as *engineering*, sales, accounting, and advertising, but none of them knows enough about public affairs." "That is not too serious," he continued, "because they can learn. More difficult is that none of them has a framework or scheme of values against which to cast and evaluate the needed knowledge. Last, and worst of all, none of them seems to realize that he lacks anything."

In his last years H. G. Wells, one of the most observant writers of our age in scientific fiction, reflected on the fate of mankind. In his provocative book, "Mind at the End of Its Tether," Wells finally concluded that humanity was dead—that indeed a force hostile to man's survival ruled supreme. Wells dismissed once and forever all of mankind. Critics have pondered the question: "Was Wells an embittered and ill man, or was he truly a prophet?"

In the next fifty years we must provide the answer to these critics. 1962 is the time when we must prepare to see that our free society will survive. Now is the time to ensure that high-quality, well-educated scientists and engineers are available not only in our laboratories but also in other important positions. Engineers will never "take over" government and management. But much that is good will go by default if we do not participate adequately in decision making. By 2012 it will be too late to reflect philosophically on whether our profession is creating for the benefit of man and not for his destruction.

C. F. Horne is President of General Dynamics/Pomona, Pomona, Calif. *(Received July 31, 1961.)*

Electrons and Elections

Sir Robert Watson-Watt FELLOW IRE

Despite the convenience of having substantial freedom of retrospective choice, for a review to culminate with our present state in May, 2012, there is some justification, beyond mere chronological conventionalism, for making the review period start in 1912. That year may well have been the last year of an unselfconscious happiness which had carried, comparatively lightly, the burdens of political and military conflict in a civilization permanently reoriented by the French Revolution. The same year is as good as any other single approximation to the burgeoning of an exciting new era in the physical sciences, and to the incipient decline of the engineer's "handbook," as it made way for the *ad hoc* and direct application of a sound mathematical, physical and chemical armoury in the mind of the individual technologist.

1912 saw the introduction of a comprehensive measure to safeguard human life on the high seas, by the statutory requirement of radio-communication facilities in U. S. passenger and cargo ships, and by the holding of an International Radio Telegraph Conference in London, to be followed in the succeeding year by a Conference on Safety at Sea, also in London. 1913 saw the initiation of a cooperative research project, between the United States and France, on the travel of radio waves across the Atlantic. These events, of international importance, were outward and visible signs of a new upsurge of technological innovation, consequent on the renaissance of physics already mentioned, and strongly influencing national and international policies, both civil and military.

The renaissance can in fact be dated, with some precision, by the lingering decline of that "universal ether" whose decease was postponed by an array of decreasingly aesthetic and increasingly clumsy (because increasingly complicated) mechanical models, such as those to which the great physicist-technologist, Lord Kelvin, clung even in the early years of the twentieth century. Michaelson—later joined by Morley, and, so to speak, "translated" by Fitzgerald and more searchingly by Lorentz—made, in 1881, the *experimentum crucis* which had been suggested by Clerk Maxwell in the last year of his life—the year, 1879, in which Einstein was born. This symbolic continuity ties the *pli cacheté* (as the French Academy would have called it) which Faraday deposited with the Royal Society on March 12, 1832, saying "I am inclined to think the vibratory theory will apply to these phenomena ["magnetic action" and "electric induction"] as it does to sound and most probably to light," with the Einstein doctrines of relativity, the "special theory" having been enunciated in 1905, the "general theory" completed in 1915-17.

Clerk Maxwell's conclusions of 1867, that there must exist electromagnetic waves of frequencies lower than those of visible light and infrared, were experimentally confirmed by David Hughes in London, 1879, and more exhaustively studied by Heinrich Hertz in Karlsruhe in 1886. In 1894-95 Lodge in Great Britain, Marconi in Italy, Popov in Russia and Jackson in Great Britain established the potential utility of radio communications. These pioneering steps led to the growth of great networks, still indispensable to our twenty-first century "way of life," for the carrying of "intelligence" by radiotelegraph, radiotelephone and television (in black and white or in color) over a range of distances transcending merely terrestrial distances.

In parallel with these scientific, technological and sociological innovations—and indeed at their roots—lay, a wealth of theoretical and experimental developments, in the physical and biological sciences, which were alike an enrichment of the human intellect, the raw material of technological advances, and (less generally recognized) aesthetic and artistic triumphs not inferior to those of the other "Great Masters" to whom this title is still commonly restricted.

Sir Robert is Chairman of the Scientific Advisory Board of Axe Science and Electronics Corp., New York, N.Y., a Consultant to the Center for the Study of Democratic Institutions, Santa Barbara, Calif., and an Advisor at Sterling Forest Research Center, Tuxedo, N.Y. *(Received September 25, 1961.)*

That the "parity principle" just enunciated was not recognized until the century 1912-2012 (the arbitrary period of our present review) was far advanced, is explainable only by a longer and more searching survey of the historical conflict between "words" and "works" than can be attempted here. It may suffice to recall, by way of example, the injunction of St. Augustine, c. 400 A.D., "Go not out of doors; return into thyself. In the inner man dwells truth." St. Augustine is still with us; agoraphobia is still a social disease—as is the claustrophilia of the second-hand bookshop. But it moves, nevertheless, and from both sides. At the end of the first quarter of the century under review, 1937, it was still permissible, within normal good manners, for a "humanities" man to express, without the wrappings of any elegant choice of words, the deepest contempt for, and the loftiest superiority to, the world of mechanics, of "stinks" and even of math. Before the half-century had run out, in 1962, the *avant garde* of the old school had discovered that physics was now, indubitably, what it had formerly been to all its devotees, "natural philosophy." The "pure mathematicians," who had sometimes been more inhumanist than the humanist, saw their most exquisitely refined constructions becoming worthily intertwined with those of the scientist, being, in fact, used as treasured tools of the trade, in the weaving of gossamer threads of conceptual thought and acute observation into an abstract mind-painting of the universe. Abstract it must be, for the natural philosopher had, by then, learned the modesty which says "It is as if" rather than "It is."

It was still too early, in 1962, for the desegregation movement to be completed. The schools and colleges were just emerging from the flatly sodden swamps of the "social adjustment" curriculum, onto the steep and stony, but vastly satisfying, tracks of intellectual mountaineering. The graduates of the third quarter-century were nearer to being well equipped for the running of an always new world than any who had gone before. Neither the unfulfilled threats of the inadequately educated politicians of 1961 nor the forecasts of the author of "1984" had matured into the universal wastelands of body and mind that had been the bogey of the 1950's to the 1970's.

A vitally important element in progress towards this brave new world was the relatively early improvement in the understanding between the scientist and the technologist on the one hand and the statesman, politician, administrator and citizen on the other. The history of this politico-social problem has not yet been written; to the historical researcher it appears "as if" the problem had been intermittently forgotten, seldom more than superficially examined, and never, but never, within range of being so much as half-solved.

It would be idle to call Archimedes in evidence. For long after his day there was little to be said about the scientist, as scientist, in the service of government. He was called in only as technologist, to provide "hardware" in a hurry because the need for it—or for him—had not been foreseen sufficiently early. Thus, indeed, was Archimedes "persuaded" says the historian "to turn to military technology as a mere holiday sport for a geometer"—and to die untimely in the Roman assault on Syracuse. Two millennia later, the father of scientific meteorology, Napier Shaw, consulted the British War Office on how he could best personally contribute towards victory in World War I, by applications of the science which he had nurtured from the first shoots as a dilettante hobby to its bud as an economic and social service. His mild assault was bloodily repulsed by a red-tabbed general who said "Sir, do you suppose the British Army goes into battle carrying umbrellas?" At about the same time, an obscure neophyte in physics put a somewhat similar question to the War Office, but had the good fortune to be advised by a staff colonel—a distinguished lawyer in civil life—that his smattering of radiophysics might be as well applied to war-winning in Napier Shaw's Meteorological Office as anywhere else. Had he met the general and not the colonel, radar would almost certainly have been available within World War II—but it would almost certainly have been too late for the crucial Battle of Britain.

In this brief and necessarily patchy study of science in the service of government, British history is a more tractable source of illustrative fragments, than is U. S. history—though no account, however fragmentary, can overlook Thomas Alva Edison's historic advice to the President of the United States about the Naval Consulting Board which Edison had envisaged in 1910 and which he headed on its foundation in 1915, "We might have one mathematical fellow, in case we have to calculate something out." It was to this modest ambition that the Board owed the presence of its one lonely physicist member!

There can be no single date marking an explicit recognition of science as essential to the satisfaction of national (and thus governmental) needs in war and in peace. The mariner had for long relied on sun, moon, planets and stars as aids to navigation; the warrior had long forgotten that the technologies which served him were rooted in earlier science, and that new science emerged from these technologies. Perhaps Charles the Second, aided by the administrative genius of Samuel Pepys (successively Clerk of the Acts and Secretary of the Admiralty) and by the galaxy of natural philosophers who founded a Society for promoting physico-mathematical experimental learning (under the patronage of Charles II) should be specially honored as a Chief of State who appreciated the pursuit of science, alike in its intellectual and its utilitarian aspects. Great as was the importance of the informal advice tendered by the Council of this Royal Society, and by its individual Fellows, to the policy makers and the administration in the governmental agencies of Great Britain, such service was not sufficiently continuous, not sufficiently widely and deeply informed of the specific needs, to ensure adequate use of scientific knowledge and judgment in service to the nation.

There was not, even at the midpoint, 1962, of our chosen century, any adequately researched, documented and constructively analyzed account of the slow growth of a still inadequate body of scientific advice to government in the affairs of any one of the Great Powers. Without this, the interpretation of the progress achieved in the succeeding fifty years is devoid of a scale with which to measure its success or to predict its future growth and utility. It is, however, possible to trace here the main reason for the inadequacies of 1937, the awakening which was just sufficient to buy time in the early forties, and the forward surge in the late eighties and nineties of the twentieth century.

C. P. Snow, in his Rede Lecture on "The Two Cultures," Cambridge University, 1959, dealt with the great gulf which he saw as yawning between "The Intellectuals" on the one hand (the literary intellectuals in particular)—and "The Scientists" (the physical scientists in particular) on the other. Politicians, and even statesmen, were almost all nearer to the intellectuals than to the scientists, though, of course, many called down "a plague on both your houses." The highest echelons of the Civil Service, "The First Division" as it used to call itself, was recruited almost exclusively from the camp of the Humanities, with a sprinkling of Mathematicians and a sparse handful from the Sciences.

It was very rarely indeed that there emerged from either of the two great universities, Oxford and Cambridge, a strange multiple personality who combined the high distinctions of being at once Senior Wrangler, Senior Classic and Smith's Prizeman. That the third Marquis of Salisbury married the daughter of one such rarity, and that she had inherited the paternal abilities "across the board," goes some considerable way towards explaining how Lord Salisbury as Prime Minister had the combined wisdom and good fortune to have Lord Rayleigh (who had succeeded Clerk Maxwell in the Cavendish Chair of Experimental Physics) as an unofficial adviser, and the philosopher-statesman, A. J. Balfour, later Lord Balfour, as a nephew! Balfour, in turn, depended heavily on the judgment and administrative genius of R. B. Haldane, later Viscount Haldane. Balfour and Haldane were Scots, born eight years and 25 miles apart, widely different in their abstract philosophy, widely different in their political affiliations, neither of them well-suited to the atmosphere of the House of Commons, but both passionately and fruitfully united in their determination to achieve more effective machinery of government, for military and civil strength alike.

For our present purposes we must be content with a mere mention of isolated samples of the British story—which, it must be admitted, is so characteristically British as to be nonrepresentative of the global picture, even were there a single global picture achievable by the diligent needlework of the historian. There are, of course, admirable books on The Social Relations of Science, but they do not deal with the social relations of the supreme statesman and the supreme scientist, for example, the third Marquis of Salisbury and the third Baron Rayleigh. Nor do they explain the highly atypical collaboration between the politically Conservative nephew of the aforesaid third Marquis (the nephew who was in time to be the first Earl Balfour) and his politically Liberal and left-wing friend, who was in time to be the first Viscount Haldane. Yet this collaboration was, in all human probability, necessary to the just-timely advent of British radar, and to much else in the field of science in the service of government. The full story belongs elsewhere; here it suffices to say that but for Balfour's proposal for the founding of a new governmental Department of Scientific and Industrial Research, and for Haldane's care of the new Department, the conditions for the inception and rapid development of Radar in Great Britain might not have existed at the critical dates.

Nor, indeed might the Royal Air Force itself, with the Royal Aircraft Establishment which served it, have been in the state of just-adequacy which, albeit precariously, prevailed in 1939. Rayleigh, Balfour, Haldane all were, very directly, essential to the survival and growth of that decisive arm.

In a still wider field, the military posture of the British armed forces as a whole might well have been hopelessly weak but for the much earlier work of Balfour in making the strong and flexible organ of military policy called The Committee of Imperial Defence, the work of Haldane in guiding that Committee for many critical years, and the remarkable work of Maurice Hankey in establishing and refining its organizational efficiency so that its machinery was ready to be converted into the War Cabinet Secretariat when the need came. When Winston Churchill, as First Lord of Admiralty, decided that a Naval General Staff was needed, he found its prototype in the Imperial General Staff which Haldane had set up for the Army. The philosophers had served their country well.

Balfour and Haldane together set up, and made into another valuable instrument of national importance, the civil Department of Scientific and Industrial Research. The Department benefited greatly by the advice of Sir J. J. Thomson and Lord Rutherford, each serving as Chairman of its Advisory Council. This civil department of government was the parent of a number of Research Boards. One of these, the Radio Research Board, was presided over by Admiral of the Fleet Sir Henry Jackson, one of the four pioneers of radiotelegraphy—in alphabetic order Jackson, Lodge, Marconi and Popov. The functions of this Board were thus described by Sir Henry Tizard, Permanent Secretary of the DSIR 1927-29, "The functions of the Radio Research Board were as described in its first report, namely 'To provide for research work of a fundamental character in directions where it was lacking and where it would be likely to lead to useful applications' . . . It is no part of the normal responsibilities of the D.S.I.R. to undertake defence research, nor to provide new devices for the use of the Fighting Services."

This orientation, towards general and nonmilitary developments, did not prevent the work done within the Board's program from making possible timely development of radar in 1935-39 by Watson Watt and his colleagues, and also of an earlier device which is described by U. S. and Canadian naval authorities as having been, at the lowest, no less important than radar itself to victory in the Battle of the Atlantic. This device, the High Frequency Visual Radio Direction Finder, depended on the Instantaneous Visual Radio Direction Finder proposed, for purely meteorological research, by Watson Watt in 1916, demonstrated by him in 1923, and introduced into the Royal Navy in 1940. It may well be that neither of these vital tools of victory in World War II would have been available at the crucial moment but for the wisdom and foresight of Balfour and Haldane in the years before 1916.

Despite these enlightened contributions from the philosopher-statesmen, and the consequent brilliant work of Hankey in improving the administrative machinery of government at the highest level, the lack of any continuing and growing understanding of the nature of science and of its subsidiary but profoundly important role as the "mother of invention," through technological utilization of scientific knowledge, continued. Intermittent crises evoked temporary *ad hoc* devices for obtaining scientific advice in limited areas of urgent need. Such devices included the Aeronautical Research Committee, and later the Committee for the Scientific Study of Air Defence (the Tizard Committee), the Air Defence Research Committee of the Committee of Imperial Defence, the "Tube Alloy" Committee on the development of the atomic bomb, and the Radio Board of the War Cabinet of which the Chairman was Sir Stafford Cripps, who had at one time been a research student under Sir William Ramsay and thus a *rara avis,* scientist-statesman.

Corresponding problems in the U.S.A. and other nations produced corresponding efforts, corresponding institutions which did not wholly bridge the gap between the two cultures in the service of government, and *ad hoc* committees whose useful life usually ended with the passing of the crisis which engendered them.

The radical solution has not yet become fully effective in this current year of 2012 A.D., but great advances have been achieved. They have not been made by the old method of "staggering from crisis to crisis," nor have they been forced marriages between "classical" and "scientific" minds still ill-acquainted one with the other. Somewhere in the 1960's the "educators" awoke to an already dazzling glimpse of the already obvious. Thumbing through the "Who's Who"s and the "Who Was Who"s of the time, they discovered that there was a whole galaxy of scientists, theoretical and experimental, who had, in the preceding eighty years or so, demonstrated a level of imaginative artistry at least as high as any ever attained in the literary, musical and visual arts, together with a philosophic penetration far surpassing that achieved by the philosophers-without-adjective.

Without attempting any comprehensive catalogue of these artist-philosopher-scientists, and without assessing priority or relative importance, a representative list must have a preface including Henry, Faraday, Mendeleyev and Clerk Maxwell, immediate forerunners of the revolution. Then come—in a list that must be incomplete and in some measure arbitrary—Michaelson, Fitzgerald, Lorentz, Ramsay, Rayleigh, Hertz, the Thomsons, Millikan, Rutherford, Bohr, Urey, Einstein, Dirac, Anderson, Yukawa, Fermi, Pauli, Planck, de Broglie, Davisson, Schroedinger, Roentgen, Moseley, the Curies, Leonard, Compton, Born, Heisenberg, Wigner, Weizsäcker, Hahn, Meitner, Frisch, *et al.* The inclusions are not in any rigid order; the omissions are due to a happy *embarras de richesse*.

Physics as a representational art died with Lord Kelvin; physics as an abstract art was born with Michaelson and Fitzgerald. Unlike the abstractions of abstract graphic art, however, the abstractions of the new physics were disciplined by the rules of a new arithmetic which clarified but did not confine, which was strictly quantitative but subtly elastic. In its modest abandoning of any claim to have penetrated to the definite mechanism which the schoolboy Clerk Maxwell sought in has passion for "the *real* go of things," in its substitution, for the arrogant "it is," of the modest "it is as if," it released conceptual thought from the shackles of the individual senses while keeping it on a well-posted highway towards the infinite.

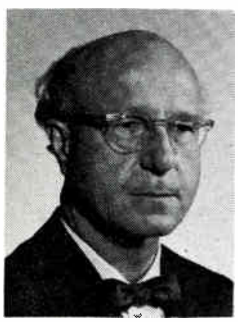
A century of increasingly awe-struck coexistence with a "Culture" which had banished "infinite" and "infinitesimal," "eternal" and "transient." "real" and "imaginary" from their professional lexicons, forced even the obstinately laggard rear-guards of the "humanities"—so often deeply concerned with inhumanities—to admit that the new scientists must be accepted, on at least a comparable level, into a united and cooperative (the two adjectives are far from identical!) House of Intellect. They had traveled a long way, from the opprobrious epithet "stinks" to a combined operation of reply to the Riddle of the Sphinx.

Now, in 2012 A.D. (somewhere around the millionth year of man the tool-maker?), we have almost the first generation of mature minds educated into seeing a single picture extending from the elementary particle to that galaxy which at the beginning of this sentence was five-billion light-years distant from us—and at the still remote end of the same sentence will be several million miles further from us, into visualizing the whiff of ammonia which switched from the inorganic to the organic state, into recognizing that although the number of quasi-independent variables in the statesman's equations is vastly greater than in those of the scienceman; yet each "political variable" is, in practice, *individually* much easier to evaluate than is each of the fewer "scientific variables." This is the generation which is now engaged in steering the world towards the state in which the still young Siegfried is indeed, as Wotan foresaw, "Stronger than I, the God."

So it can at last be said that the scientist is no longer sentenced to be "on tap but not on top," and this the more happily because he is both at the top and yet on tap. The essential blending of the cultures, all of them aesthetic and all of them important to the fuller life, has brought us an electorate which chooses, as the steersmen of the state, men who are both natural philosophers and philosophers *tout court*.

Electronic Nirvana

DANIEL E. NOBLE FELLOW IRE



The identification of trends is not difficult. For the generation of perspective, it is often a useful exercise to extrapolate trends into the nether regions of their unqualified terminal configurations with all modifying and inhibiting influences held in abeyance. Two presently discernible trends are first, the exponentially accumulating techniques for the extension of brainpower, or more properly, the extension of informational processing, and second, the very marked trend toward greater and greater degrees of vicarious living. The two trends are not only related, the first trend accentuates the second, for the rise of informational processing will accelerate the drift toward higher and higher levels of static and innocuous existence. Of course, it is quite true that some degree of vicarious living has been a characteristic of the search for an innocuous state of bliss since the telling of the first story, but it took the invention of the printing press, and more recently, the introduction of moving pictures, radio, and TV, to carry the living-by-proxy technique to such a high state of perfection that we can predict for the future that "personal participation" will be regarded as an eccentricity. Even in the field of sports, we have hired professional baseball players to play baseball for us, professional football players to play football for us, and professional golfers to play golf for us. We have now taken the ultimate step to limit physical action by making it unnecessary to participate personally as a spectator; we need not bother to go out to the field to watch; we sit at home and view the progress of the contest on the ubiquitous TV eye.

D. E. Noble is Executive Vice President of Motorola Inc., Phoenix, Ariz. *(Received July 31, 1961.)*

Since children are conditioned to watch TV from early childhood, the educators are reaching for TV as a desperate means for re-establishing communications. Children will increase the number of hours they spend staring at the glassy eye, and the total viewing hours will become an incredibly high percentage of their total waking hours. While there will be many steps in between, the terminal pattern for the extrapolation of this trend is easy to project.

Today's parents use the TV as a pacifier for temporarily immobilizing children of all ages, but the parent of the future will resort to a much more scientific solution. At birth, the infant will be clamped in front of the TV eye by means of a suitable supporting structure, and two sections of tubing will be connected to provide nourishment and to carry away the waste materials. From this time on, the subject will live an ideal vicarious life, scientifically selected for compatibility with the fixed influences of the inherited genes and chromosomes. For the early years, the automatically programmed TV will process the storing of the necessary bits in the memory brain cells of the subject for use in the initial "thinking" responses. As the accumulation of working material in the memory progresses, the "teaching" will be alternated with programmed vicarious living experiences. As the subject grows older, the victim—I mean the subject—will be trained in the necessary mental processes for selecting, storing, rejecting, correlating, disciplining, arranging and re-arranging new combinations of bits in the memory cells to enhance the potential for this more advanced thinking capability. It is obvious, of course, that various essential sensors will be connected to master points in the subject's nerve responsive system, and as the intellectual and emotional patterns emerge, the subject will be exposed to graded influences which will be selected automatically and will be totally controlled by the feedback from these sensors. A permanent recording of all sensor reactions properly coded to the influencing subject matter will be made for analysis and for use later in the continuing programmed activation of the telememory transducers which will directly stimulate a selected retrieval from the memory brain cells. As a result of the transducer stimulation, the subject will relive past vicarious experiences without recourse to sight or sound. Other transducers, properly controlled by the sensors, will be used to create and store additional sensations associated with the senses of touch, smell and taste. A library of qualified selections of vicarious living experiences will accumulate in solid-state memory form and will be automatically selected by the coded sensors for maintaining the "happy" state of the subject. In time, the memory cells of the subject will become saturated, or at least will be supplied to an adequate level with vicarious living experiences so that listening, and viewing the TV eye will no longer be necessary. From this point on, the sensors will automatically select the "living" experience which is to be *relived*, and the coded memory transducer will provide the controlled brain stimulus which will release the stored memory for the "reliving" of the vicarious living experience. In this way, the subject will be maintained in a perpetual state of static, innocuous bliss.

As the practice of "electronic nirvana" becomes universal, the natural result will be race suicide. Before this happens the development of machines which will reproduce themselves will have been completed, and the machines, drawing their raw materials from the ocean, will continue the redundant pattern of automatic machine reproduction and the recreation of simulated human living experiences until finally all of the human subjects will leave the infinity of innocuous existence for the infinity of nonexistence. And so, dear friends, as the sun sets slowly in the east, we leave the beautiful earth with its millennium of technological achievement, with no strife no mess, and no men, while the machines continue to grind out their fruitless patterns of human vicarious living with no humans present to respond to the stimuli.

Now that we have gained perspective by following the trends to their directly extrapolated limits, we may return to an examination of the interim patterns of informational processing which will flower within the fifty-year period of our present projection. Certainly, the demands for recording, disciplining and programming information will tend to move us away from the random approach which has been characterized by the development of isolated and unrelated solutions to related problems. Our whole approach to the necessary informational processing of the future will become more orderly, with an increase in the integration of related solutions. While we are busy pacing the trend in applied science away from the macroscopic and toward the microscopic in the processing of materials for electronic subsystems, we shall also change from a gross averaging of political, social and business statistics to refined selection, to individual characterization and to a more enlightened and integrated control of the environmental forces.

At birth, each individual will be given a serial number, which he shall retain for life. This number will be the key to the constantly accumulating storage of solid-state-memory bits, and it will serve as a telephone number, a social security number, an automobile license number, a passport number, a credit card number, and it will be the means of retrieving all other numbers and bits of information which will be important to the individual and to his political, business and social life over the span of his lifetime. Perhaps a nine digit number will be sufficient in the U.S.A. for a couple of hundred years. As an example, the number attached to one of our descendants in the year 2012 A.D. may be 345-462-179. For the child, the number would provide the key to the recording of all scholastic records generated by the teaching machines. A continuously accumulated health record would be stored in a master computer memory and whenever the individual became ill, as a child or as an adult, the coded latest symptoms would be fed into the computer, which would respond with the pertinent history, a diagnosis, and treatment information based not only upon this individual record, but also upon the information automatically retrieved by the computer from a search of all similar health patterns.

All mail would utilize the serial number automatically, whether it was used for the routing of written material or for the transmission of facsimile messages. For mailing a letter, the addressor and the addressee numbers would appear on the face of the envelope. At the post office, the numbers would be read and checked by rapid search of the central solid-state computer memory, and the letter would then be automatically coded and routed to the proper destination. Telephone calls would be routed automatically in a somewhat similar fashion.

Additional applications for this personal serial number system would be for continuous up-to-date census information, for automatic banking procedures covering deposits, payments, billing, supermarket charges, and for the handling of loans and time payments. The FBI and police would also make use of the number. Whether the problem was one of law enforcement or automatic mail delivery, or the routing of TV and entertainment programs, or the use of the telephone, or phono-vision, or facsimile, or newspaper delivery, or even one of supermarketing, or automatic cab dispatching, the number would serve as the key identification for scanning the solid-state memories, and for automatically processing the transaction.

Instant-delivery communications capability will be greatly enhanced by the use of coaxial cable service to all homes and business establishments. With the broadband laser coaxial service in operation, two-way visual as well as audio and facsimile communication service will become routine. One very important extension in the field of mass communication as a result of home coaxial cables will be the response communications potential of the system. No longer need the observer be a passive victim of least-common-denominator programming. With the coaxial cable in use, the observer can applaud or he can originate an electronically coded degree of disapproval. By suitable coding procedures, the individual response will be properly weighted in terms of intellectual, as well as social, political and financial background. The availability of many programs, which by this time will be stored in monolithic solid-state memory systems, will offer a wide choice to the individual seeking entertainment, information or instruction. Voting on national questions will be carried out with no significant lapse of time, and the vote will be automatically analyzed in terms of any selected categories of coding or weighting. The very important difference between the present system and the future system will be the element of choice, which is now theoretical rather than real. With the development of solid-state molecular memory capability, the storage of programs will require little space, and will become low in cost and routine in procedure. By selection, the potential customer can call out a Shakespearean play, a musical comedy, a modern drama, a lecture on space exploration, or a lecture on child care. The instant selection and freedom of choice will forever silence the Madison Avenue voices which tend to seek the least common denominator of program mediocrity to capture the mass audience.

With the wide-band two-way home communications service available in the year 2012 A.D., the potential customer may make use of the coded retrieval system to arrange for a display and for full information about any desired products. Each dealer will store in a solid-state memory system a carefully prepared selection of presentations for each product. The programs will vary from a simple picture of the product and information about the essential specifications, to elaborate demonstrations of the product in use, emphasizing its features and competitive characteristics. At home, the potential customer will take a quick look at the products in the selected category of interest, using the home three-dimensional color screen, and when a product is found with special appeal, the full program may be requested to aid in making the final decision.

Since there will be no newspapers and magazines in the usual form, and since there will be detailed product information always at command, the generalized propaganda advertising will be a thing of the past. Three-dimensional color video display with sound, home facsimile publication of pictures and reading material (available also with sound, which may be selected), and the talking books available for dialing directly into the home from the central library solid-state memory storage, will minimize the use of many of the traditional forms of mass communications.

By means of satellite broad-band communications, the informational processing from a home in the United States to a home anywhere in the world will be as easy and as routine as communication between the homes in the same city or in the same area. The language barrier will be eliminated automatically by processing communications between two different language groups through the automatic solid-state translator. The extension of this communication to universal world coverage will gradually bring about the development of mutual confidence and understanding among the peoples of the world, so that in three or four hundred years, war will be outlawed and the people can then devote all their time and energy to the more important task of keeping the matrix of worldwide informational processing properly functioning and forever expanding.

This perfection of informational processing will not, as might appear from first examination, isolate, restrict and control the individual, nor will it condemn intellectually qualified individuals to vicarious living or to "electronic nirvana." Quite the contrary, the qualified intellects will have at command the almost instant retrieval of any selected part of the total store of human knowledge accumulated in the archives of the solid-state memory systems of the world. From almost any location in the world an individual may call up, in both visual and audio form, information from any other part of the world and accomplish the whole operation of query and response in a matter of minutes, if not a matter of seconds. By the use of effective teaching machines and with instant retrieval of information available to support the expanding intellect, the newly developing human mind will start out on a foundation of knowledge exceeding in depth and breadth anything which has been possible in the past, with the result that during the normal lifetime, which will be extended to 100 years by 2012 A.D., the human intellect will rise to levels of knowledge and to the total use of knowledge never before achieved by man. Certainly, as is the case today, there will be all degrees and levels of intellectual competence, but for every level from the moron to the scholar, the rise of informational processing will contribute to the interest and minimize the cruel and painful elements of life and living. The art of living will be enhanced in all of its multifold relationships and the potential for experiencing and sharing experiences will be mushroomed by the development of effective informational processing in all forms. While the chronological age span will be increased, far more important than this, the real life of experiencing and sharing experiences will be increased for everyone. Since the brain is where we live, the enhancement of communications in all forms will speed up life and living. We will live faster, not in the degrading sense, but in the sense of receiving a far greater quantity and quality of stimuli and of reacting, and sharing reaction to such stimuli. This irresistible and irrevocable rise of informational processing will extend man's reach for comprehension and accelerate his march toward maturity.

Electronics and Evolution

JEROME B. WIESNER FELLOW IRE

Technology can be thought of as a man-made substitute for natural evolution in the process of adjusting to our environment and in which many of the steps of natural evolution are being repeated. The modern human being is the culmination of an eon-long evolutionary process that began with the existence of individual simple cells and culminated in the complex organism of today. In the process a great variety of specialized cells—such as muscle cells, nerve cells, blood cells, and many others—evolved to carry out the functions of energy conversion, communication and transportation which are necessary for the existence of a complicated organism. The increasingly complex specimens which developed are the survivors of a natural selection which made dominant those species best equipped to deal with the hardships encountered in the environment. It may be that the evolutionary development of man has not ceased—though there is some reason to believe that he does represent about as much organized complexity as can be satisfactorily achieved with components having the properties of biological materials—but even if it hasn't stopped, it is a very slow process.

Fortunately, man has found a substitute for continued biological evolution to aid him in his struggle against hardships of the environment. He has learned to exploit scientific knowledge to his advantage. Nearly all of the fields of science, biology, mathematics, physics, chemistry and astronomy, have made their important contributions to the technical evolution. By co-operation involving communication, the control of energy and the manipulation of materials, man has greatly increased his mastery over nature and his accommodation to the environment in which he must live. Our own fields of electronics and communications, which are making so large a contribution to this effort, depend primarily upon knowledge from the fields of physics and chemistry for their devices—electron tubes, transistors, magnetic cores and a multitude of other components—and upon mathematics—field theory, circuit theory, switching theory, etc.—for their underlying logic. In the last one-hundred years we learned how to harness almost limitless amounts of energy, particularly electrical energy, and direct its use for our purposes. Simultaneously there were developed electrical means for communicating between individuals—the telegraph, the telephone, radio, television and teletype—in which the high speed of electrical signals was exploited.

In these early developments man was learning to enhance his own physical capabilities, to replace menial human effort with machines, and to extend the distance over which he could be effective.

More recently, feedback control and automatic computing systems have been developed to eliminate the need to perform personally a variety of routine mental tasks such as steering aircraft, guiding a milling machine tool, reading a check or regulating a steel mill.

The digital computer which makes these feats possible is undoubtedly the most important single electronics development of the last two decades and ultimately will have as widespread consequences in the development of our civilization as did the steam engine. Digital computers have already been employed to perform a variety of computation, information processing and storage functions in which they greatly enhance the capacity of man's unaided brain, and yet their exploitation has hardly begun.

J. B. Wiesner is Special Assistant to the President for Science and Technology, The White House, Washington, D.C. *(Received October 24, 1961.)*

The most important attribute of the modern computing machine is its very great speed compared to biological computing systems, an advantage that is easy to understand. Signals on nerve fibers travel at velocities up to 300 meters per second. Furthermore, neurons (neural relays) have maximum response rates of the order of one hundred impulses per second, while electronic switching rates of greater than one billion per second have been achieved, and this is hardly the limit.

Ultimately, computers should have advantages other than speed. As electronic components develop especially satisfactory molecular-sized elements, a computer's memory should become better than man's in at least two respects: better retentivity and more rapid access to stored information. Conceivably it could also have greater storage capacity, though information storage capacity of the human mind is so extremely large that need for greater amounts of active memory may never occur. These considerations would justify the conclusion that one should ultimately be able to create "thinking machines" much brighter than the smartest human being, if presently unforeseen limitations on the logical design or construction of computers do not appear.

At the present time our ability to use machines instead of human brainpower in many tasks, particularly those involving insight or intuition, is prevented by our lack of understanding of the logical processes involved, but there is a vast research effort in this challenging field and understanding is advancing very rapidly. Machines are already assisting in their own programming and development; as this technique develops it is certain to extend the scope of computer capabilities.

How will these developments be exploited in the next half-century? It is probably impossible to be visionary enough to match reality, though the opportunities are obvious. With endless amounts of energy available where needed, and with automata to control machines and to substitute for or augment most mental labor and to extend greatly the human mind's best creative efforts, man can truly become master of his physical environment. One can even hope that the "thinking machines" will make it possible to understand and cope with the enormous and vexing social problems that the technological revolution is leaving in its wake. Possibly the most important aspect of these scientific and technical developments will be the assistance they will provide man in making adjustments to the environment which he has created—economic, political and social.

Index to 2012 A.D.

SUBJECTS

AUTHORS

AUTHORS

SUBJECTS

L

M

N

O

P

Q

R

S

T

U

V

W

X

Communications and Electronics 1912–1962

Organized with the assistance of the IRE Professional Groups

Section 1

AEROSPACE AND NAVIGATION

Organized with the assistance of the IRE Professional Group on Aerospace and Navigational Electronics

Fifty Years in Aeronautical Navigational Electronics*

VERNON I. WEIHE†, FELLOW, IRE

***Summary*—The state-of-the-art in aircraft operations and in radio technology for the very early period of aeronautical navigation is given, and the technical evolution during the first twenty-five years of IRE is covered. Policy and organizational problems of the transition to the second twenty-five-year period are mentioned. A discussion is given of the World War II system developments, the introduction of pulse techniques and the race toward the microwave portion of the spectrum. This is followed by a discussion of self-contained systems. Illustrations are provided covering some of the more fundamental spatial characteristics of electronic navigation systems.**

Navigation and Power

Prince Henry the Navigator of Portugal (1394–1460), patron of voyagers and organizer of many serious projects to improve the science and practice of navigation, is credited with doing more than any other single person to make the 15th, 16th and 17th centuries the Great Age of Discovery. Except for this period the rate of progress in navigation during all prior centuries, and even to the 20th, was extremely slow. This probably can be attributed to lack of urgent need, since man could travel no faster than sail or animal power could propel him.

Even in the early twentieth century, during the fledgling days of aviation, designers concentrated on lift and stability while disdaining the less esoteric power approach as the "brute force" method. Yet the dynamic progress of aviation has been paced by the constantly increasing amount of power made available. Each new improved source of thrust spawned a new generation of higher performance aircraft and each, in turn, brought new requirements for navigational services.

But, lest we pass those early days too hastily, a quote from Talbot[1] on the status of navigation in World War I naval aviation may help us start from the beginning:

> When once the flying machine had indicated its possibilities in connection with land operations, it was only natural that endeavors should be made to adapt it to the more rigorous requirements of the naval service. But the conditions are so vastly dissimilar that only a meager measure of success has been recorded. Bomb-throwing from aloft upon the decks of battleships appeals vividly to the popular imagination, and the widespread destruction which may be caused by dropping such an agent down the funnel of a vessel into the boiler room is a favorite theme among writers of fiction and artists. But hitting such an objective while it is tearing at high speed through the water, from a height of several thousand feet is a vastly different task from throwing sticks and balls at an Aunt Sally on *terra firma;* the target is so small and elusive.
>
> Practically, it is impossible to employ the flying machine, whether it be a dirigible or an aeroplane, in this field.
>
> In operations over water the airman is confronted with one serious danger—the risk of losing his bearings and his way. . . . Many people suppose that because an airman is equipped with a compass he must be able to find his way, but this is a fallacy. . . .
>
> Unless the airman has some means of determining his position, such as landmarks, he fails to realize the fact that he is drifting, or even if he becomes aware of this fact, it is by no means a simple straightforward matter for him to make adequate allow-

* Received by the IRE, October 2, 1961; revised manuscript received, October 30, 1961.

† Advanced Systems Planning, General Precision, Inc., Washington, D. C.

[1] F. A. Talbot, "Aeroplanes and Dirigibles of War," William Heinemann Co., London, Eng., pp. 243–246; 1915.

ance for the factor. Side-drift is the aviator's greatest enemy. It cannot be determined with any degree of accuracy. . . . More than one bold and skillful aviator has assayed the crossing of the English Channel and being overtaken by fog, has failed to make the opposite coast. His compass has given him the proper direction, but the side-drift has proved his undoing, with the result that he has missed his objective.[2]

Aircraft of the wood, cloth and "bailing wire" period were not all-weather machines, and the number of them did not, of course, present a serious air traffic control problem. In addition, closing speeds were until very recently such that "see and be seen" was normally adequate to prevent air-to-air collisions over land. The earliest basic requirement was to find intermediate navigational checkpoints and terminal landing places or airports, while today's most urgent need is that of adhering continuously to a flight plan track for air traffic control, rather than for basic navigational purposes. A major challenge of our profession is to devise and introduce new electronic systems which can help maintain the flexibility and efficiency of air operations while drastically reducing the incidence of mid-air collision.

[2] Talbot, *op. cit.*, pp. 244, 245, and 246.

Early Radio Technology

During the first twenty-five years of the IRE, only continuous wave techniques were useful for navigation, and only in the relatively long-wave portion of the radio spectrum. Hence, antenna structures of practical dimensions provided low aperture, usually much less than one-half wavelength. This lack of aperture made high spatial sensitivity difficult to achieve, while the inability to differentiate between direct and reflected signals made acceptably accurate and stable navigational lines of position at times impossible to accomplish. In Figs. 1–3, navigation systems are categorized and chronologically related to emphasize first of all, the changes brought about by wavelength/aperture considerations; secondly, to point out the impact of the introduction of pulse techniques and most importantly to dramatize the explosive impact on navigational development produced by high performance aircraft requirements.

Early systems such as the aural-null airborne homing device or direction finder, and the LF/MF four-course radio range required the pilot to listen for signal nulls or to compare signal amplitudes. Atmospheric noise and the monotony of keyed tone signals produced consider-

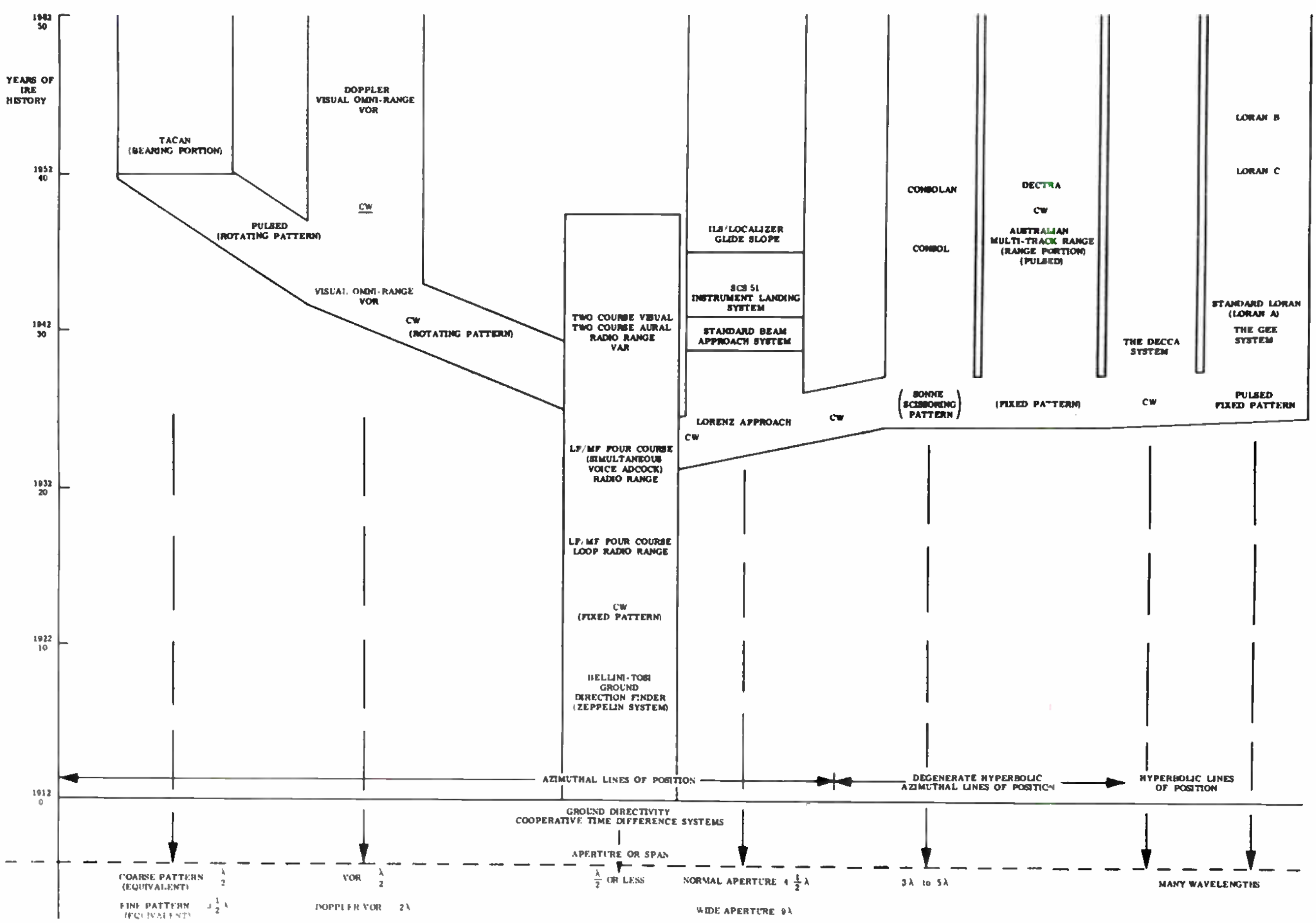


Fig. 1—Ground directivity—cooperative time difference systems.

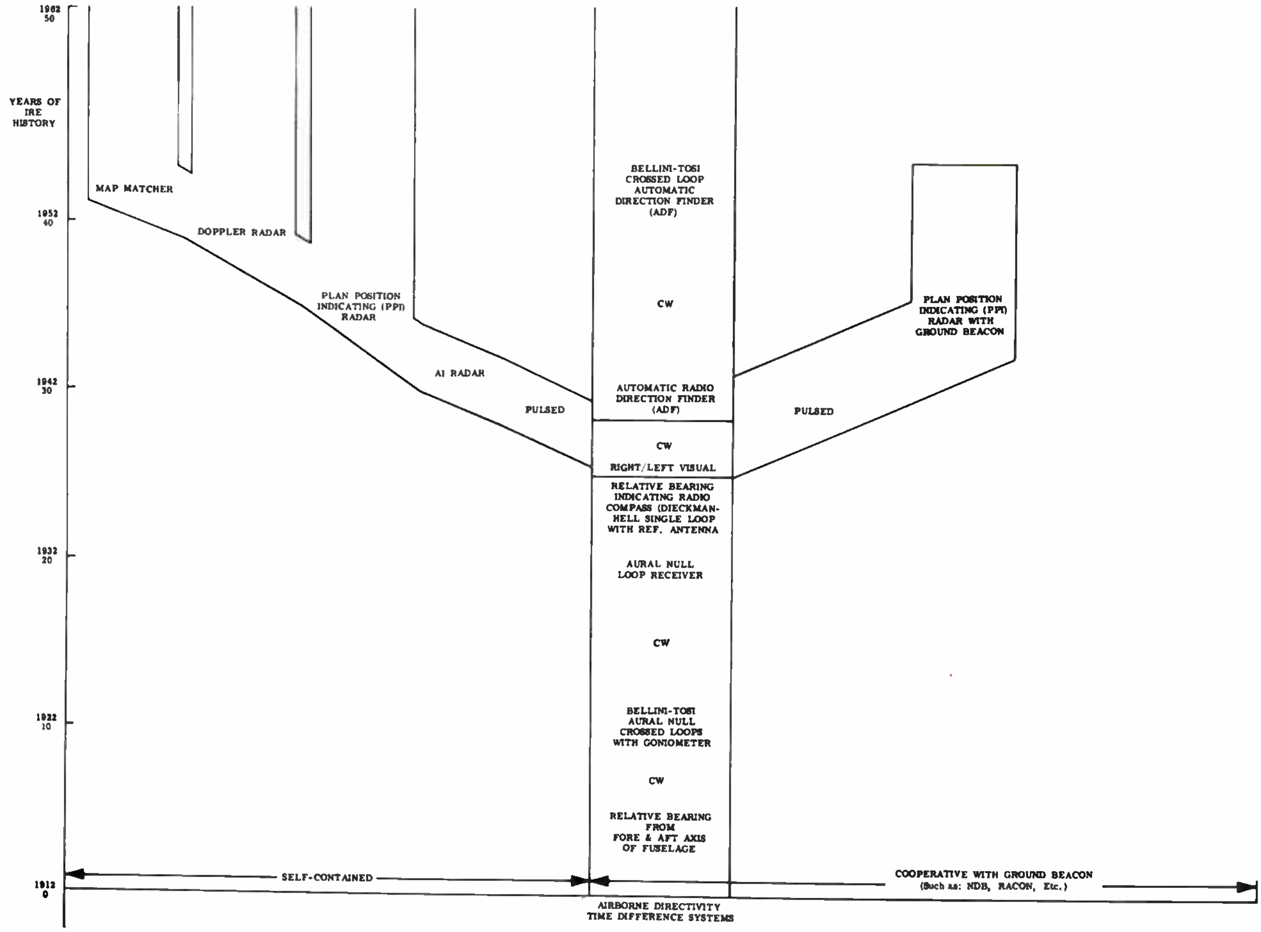


Fig. 2—Airborne directivity—time difference systems.

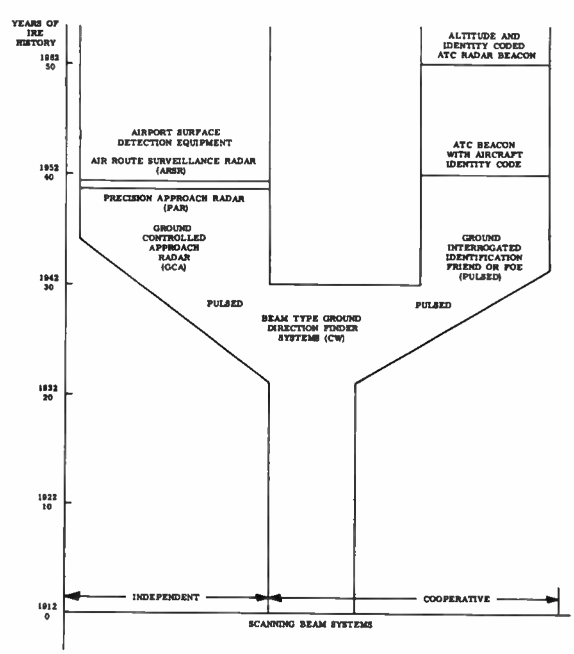


Fig. 3—Scanning beam systems.

able pilot fatigue, thereby emphasizing a requirement for the presentation of visual information in the cockpit.

This need caused development of zero-center meter systems. Such meters may have the needle at center, either when on-course or when the equipment is inoperative, hence are not considered to "fail safely." Later, zero-center meters were provided with flag alarm systems to indicate to the pilot the status of system operation, and in some cases the navigational signals were employed as servo inputs to provide "an active needle automatic follow-up" for presentation on an azimuth scale.

The need for cooperative lock and key relationships between airborne and ground portions of the system far in access of those for conventional radio communications soon became apparent. It became essential that transmitting and receiving units not emit or receive radio-frequency energy through extraneous paths, such as power leads, ground leads, chasses, etc. Only those antenna system relationships which were spatially sensitive were employed. For example, radio altimeters have been difficult to calibrate on more than altitude scale because of an inability to compare the reflected signal against a single reference. Leakage between transmitter and receiver within the equipment or within the installation wiring produced one reference, while direct signals between the transmitting aerial and the receiving aerial on the same aircraft provided another reference. The dynamic range, automatic gain control and phase characteristics of receivers now, for the first time, were required to remain uniform throughout the radio spectrum of the system and between severe environmental limits.

Cooperative Navigation Creates Organization and Policy Problems

National standards covering signal characteristics, tolerances, operational utilization, minimum acceptable performance criteria, and test standards had to be evolved. In the United States, many Presidential boards and committees were formed and the Radio Technical Commission for Aeronautics was created. This Commission, founded in 1935, still proveds a forum for all aeronautical interests—government, industry, and private—to exchange views and to recommend procedures and standards for the improvement of radio technical services to aviation.

The introduction of new navigational systems into aircraft must be accompanied by scores of prior agreements, since the number and type of aids used in the aeronautical service must be limited to a few broadly implemented standard types. Aircraft cannot change airborne equipment when flying from one country to another and, of course, "dead-heading" equipment during periods of non-use is detrimental to payload. Cooperating systems which require a ground counterpart, for practical reasons, must be more limited than other types, since administrations of many countries must install, operate and maintain the ground facilities. For airborne self-contained systems which radiate (*e.g.*, airborne radar and Doppler autonavigators), the problem is much simpler since radio-frequency channel licensing and airborne equipment certification only are involved. Much of the ponderous policy and administrative machinery necessary for ground equipment implementation and maintenance fortunately can be by-passed.

Continuous Wave vs Pulse Techniques

The start of the second twenty-five years of IRE coincided with the military build-up for World War II. Early emphasis on detection and early warning via radar produced an almost immediate employment of pulse techniques in electronic navigation. Real competition quickly developed between those who contended that navigation had a low information rate requirement and therefore should utilize very low bandwidth per channel and those who supported the full employment of pulse techniques in order to achieve multipath discrimination regardless of spectrum economy. This latter view was held regardless of the fact that in some instances techniques utilizing coding or repetition rates also had to be used as a means for providing navigational information.

Pulse techniques have two fundamental uses in electronic navigation:

1) to facilitate the measurement or display of distance through round-trip transit time of the direct wave when a multiplicity of targets or a multiplicity of sources is involved; and
2) to increase the accuracy of line-of-position determination through the reduction of multipath effects.

Feature 1) above is fundamental to radar and radar beacon range measurement and to distance measuring systems, while 2) is employed in hyperbolic systems (such as GEE and LORAN).

In continuous wave systems, the geographic line of position data at any point at any time is produced by the integrated total of all of the energy received, regardless of whether it is propagated over the direct path, indirect paths, or both. These effects are also very much influenced by the plane of polarization of the radio waves employed. For instance, the horizontally polarized component of the propagated wave is particularly detrimental to the performance of the loop radio range, causing multiple courses in regions of rough terrain and being particularly susceptible to course shifts resulting from ionospheric changes. The Adcock range was developed to assure improved course stability through minimization of these effects.

The navigational accuracy of CW aids can also be improved by transmitting on a number of wavelengths simultaneously or by increasing the aperture or span of the antenna array. The former method, although tried successfully, was not adopted because of poor spectrum and equipment economics. The latter method generally produced ambiguities which had to be resolved by including a less accurate subsystem to provide ambiguity resolution (*i.e.*, DECCA lane identification, CONSOL sector identification, and the 15-cps subsystem of TACAN).

Doppler systems of the type which systematically move a real or quasi-antenna through a span of several wavelengths obtain an increase in aperture without coincident development of ambiguous lines of position. The increase in aperture materially diminishes derogating multipath and related site sensitiveness (*e.g.*, Doppler direction finder, Doppler VOR). Systems of this type overcome serious disadvantages of CW systems without incurring some of the basic limitations of pulse systems.

In all but a few cases, equal performance can be had from either pulse or CW techniques, if equal cleverness is shown in system conception and development. Many of the better systems today utilize a combination of the two techniques either in the basic concept (LORAN B and LORAN C), in the measuring circuits, or in the safety monitoring subsystem. Certain other systems would be more useful if modified to include both CW and pulse techniques. For example, if TACAN had a CW feature like the audio channel of television, voice identification could supersede the somewhat archaic slow code identification method utilized; other uses also could be made of the available channel.

All accepted systems have relatively good sampling, integration, or correlation circuitry, evolved either through trial and error, or through theoretical analysis and development, yet little firm agreement exists concerning the superiority of one technique over another. CW signals are often squared and converted to pulses for measurement in airborne receivers or ground monitors, while pulse signals are often used to synchronize local sine wave generators to facilitate measurement of CW phase.

Full utilization of either seems to create unwarranted compromise, while full utilization of both seems, in some instances, to bring about considerable equipment sophistication. Sophistication can increase the cost and, in turn, decrease the number of potential users of a system. The unending competition between the CW and pulse systems proponents will no doubt improve both types of system; however, general agreement on the superiority of one over the other is not to be expected in the near future.

The Race Toward Microwaves

In addition to the introduction of pulse techniques, the impact of World War II also set in motion a vigorously supported development race to shorter and shorter wavelengths. By the end of World War II, microwave systems were in general use. Employment of shorter and shorter wavelengths made it feasible to obtain wider apertures with airborne and ground antennas of practical size. In the general case, wider aperture means improved geographic or spatial sensitivity and produces improved navigational accuracy.

In the line-of-sight bands, 100 Mc and above, means were also found for reducing or overcoming atmospheric and precipitation static difficulties, while increasing not only the number of radio frequency channels available, but the flexibility of their geographic employment. In some aeronautical services, the reduced coverage of line-of-sight limited systems (below line of sight and in the overhead cone) was preferred to the vulnerability to atmospheric and precipitation static and the susceptibility to long-range co-channel interference and multipath effects (VHF vs LF/MF aids).

Detection and Measurement of Spatially Significant Information

Earlier systems relied primarily on amplitude, in null detection and amplitude comparison forms, for sensing spatially configured signals, while later systems placed emphasis on phase comparison and pulse time difference measurements.

Unlike aural systems, visual systems require the establishment of measuring references, the formation of measuring scales, the application of these to the sensed signals, and visual presentation of the information to the user. In most pulse systems, reference signals are placed in time sequence with spatially varying signals, while in the CW systems the reference and variable signals are generally transmitted simultaneously, either on separate radio frequency channels or on different sub-carriers on the same radio channel.

Round-trip transit time of radio waves is extensively employed in primary and secondary radar (radar beacons) for the measurement and/or display of distance information.

Distance measuring systems are, in fact, multichannel, nondirectional radar beacon systems, wherein a multiplicity of aircraft are accommodated through the process of time variant or random airborne interrogation, rather than through azimuth scanning of a beam antenna.

Airborne Self-Contained Navigation—Doppler and Inertial

Janus, an ancient Italian deity, was regarded by the Romans as presiding over doors and gates, and over the beginnings and endings, and was commonly represented with two faces in opposite directions. The Janus airborne Doppler radar is so called because it points beams fore and aft. The reflected signals in the forward semicircle return with increased radio frequency, while those to the rear have lower frequency. Hyperbolic loci of equal Doppler frequency return are sensed to provide the pilot with indications of ground speed and drift angle. Through electronic integration and automatic computation, using as reference the heading subsystem of the aircraft, a very accurate self-contained automatic navigation service is made available. Airborne Doppler navigation was borne in the immediate post World War II period and is now common equipment on most long-range jet aircraft.

In the immediate future, inertial navigation will be employed on a broader scale. Doppler/inertial systems, where each subsystem improves the accuracy of the other, are apt to come into common use. Long-range scheduled air carrier flights have in the past twenty to thirty years primarily relied on the services of a professional navigator, and flights have been predicated on celestial rather than electronic aids to naviagtion. Now that Doppler autonavigator equipment with direct display to the pilot is coming into general use there is a decreasing need for the services of the professional navigator. As the heading reference subsystems (magnetic, inertial, or both) are improved, even more accurate autonavigator performance can be expected.

In contrast to cooperative electronic navigation systems, autonavigators can be employed anywhere in the world over land or sea and are not degraded by atmospheric, precipitation or auroral disturbances.

Airborne Astro Navigation

From the early Polynesian times, the surface navigator has employed the stars and planets for fix determination and heading reference with many types of instruments. The Polynesians knew which stars and planets passed over each island of interest and used a gourd with fluid as a horizontal reference to form a crude zenith sextant. On early long-distance flights, astronavigators soon learned to use a bubble rather than a sea horizon for a reference in determining star altitude. In World War II, astro domes were installed on most long-range multi-engine aircraft. In one high-altitude, high-performance aircraft, explosive failure of the astro dome in flight occurred with fatal results. The pressure differential propelled the navigator through the astro dome opening to his death in mid-Atlantic, thousands of feet below. The periscopic sextant was developed and placed into service, and until that time, a mere decade ago, electronics was not a predominant factor in classical astro navigation. The next step in the development was automatic astro-tracking, so that we now have automatically controlled and computed astro navigation. More recently, we have even violated the classical self-contained principle and are beginning to "hang out our own stars." In the TRANSIT system, the TRANSIT satellite is tracked by ground-based (Doppler) tracking stations. The using vehicle observes the satellite while receiving orbit-correcting signals from it.

As we move further into the space age, optical band radio and radar navigation systems, using navigational principles not unlike those of their predecessors chronicled here, will be developed and come into practical service. Only the time scale will be different; in today's development environment, one expects to finish in months tasks which formerly required years to complete.

The Air Traffic Control Equipment Subsystem—Present and Future*

P. C. SANDRETTO†, FELLOW, IRE

Summary—The air traffic control problem consists of operational, technical, and political elements. This paper, however, is confined to the technical (equipment) aspects of the problem. The equipment subsystem consists of five essential elements. These are: Type I—means whereby pilots of aircraft may know their positions and means whereby they may be able to conduct their aircraft over any arbitrarily chosen routes from terminal to terminal; Type II—means whereby a central decision-making entity may be able to know the position of all aircraft within its jurisdiction; Type III—means whereby the decision-making entity can determine instantly safe, nonconflicting flight procedures for all aircraft within its purview; Type IV—means whereby instantaneous safe-flight instructions may be transmitted unambiguously to all aircrafts concerned; Type V—means whereby action may be coordinated among the several central flight-instruction, decision-making entities. The paper lists the significant actions which have affected the air traffic control equipment subsystem. It gives the history of the various units which form the present equipment subsystem and forecasts the equipment that may constitute the future traffic control subsystem.

* Received by the IRE, February 16, 1961; revised manuscript received, April 17, 1961.

† U. S. Defense Group, International Telephone and Telegraph Corp., Nutley, N. J.

I. Introduction

THE air traffic control problem is a triad [1] for it consists of three problems that must be solved individually, otherwise there can be no solution to the whole. First, it is necessary to obtain an agreement among all air operating entities as to the rules, conditions, and methods by which they will operate and control (or order to be operated and controlled) the air traffic under their jurisdiction. Second, it is necessary to solve technical problems so that there may be constructed practical equipment which will perform within the framework of the air operations system. The solution of problems one and two, however, can result in a practical solution only when the political organizations which have been endowed with the appropriate statutory powers take steps to obtain the necessary agreements and implementing action to bring a system into being. It is evident, therefore, that within the space allotted to this paper, the air traffic control problem cannot be covered in all its details. For this reason, this paper is confined narrowly to the equipment aspect alone.

The air traffic control system is the means for controlling traffic which consists of a set of air operations procedures aided in their application by an equipment subsystem which is mainly (but not wholly) electronic. The air traffic control equipment subsystem is a collection of suitable equipment by which aircraft positions and movements are determined and through which the movements of the aircraft are controlled. Since this paper is written for engineers, it is directed at equipment, which is their principal domain, but it is necessary to emphasize that their best inventions will come to naught if the operational and political environment attending this equipment is ignored. Significant actions which have played a part in determining the character of the present air traffic control equipment subsystem are listed in Table I.

II. The Present Air Traffic Control Equipment Subsystem

To meet its requirements, the air traffic control equipment subsystem must consist of the following five sets of equipment:

Type I: Means whereby pilots of aircraft may know their positions (in three dimensions) and means whereby they may be able to conduct their aircraft over any arbitrarily chosen routes from terminal to terminal.

Type II: Means whereby a central decision-making entity may be able to know the position of all aircraft within its jurisdiction.

Type III: Means whereby the decision-making entity can determine instantly safe, nonconflicting flight procedures for all aircraft within its purview.

Type IV: Means whereby instantaneous safe-flight instructions may be transmitted unambiguously to all aircraft concerned.

Type V: Means whereby action may be coordinated among the several central flight-instruction, decision-making entities.

A. Type I Equipment

This equipment is navigation equipment which must serve in the en route long-distance zone, the en route short-distance zone, the approach and landing zone, and the airport zone. A discussion of the present equipment for these various zones follows, but reference to newer developments has been placed in Section III.

1) The En Route Long-Distance Zone: This zone is that portion of the air space in which aircraft are primarily served by radio aids spaced at intervals of 200 miles or more. There was little need for navigation equipment for the en route long-distance zone prior to 1935, for it was not until November of that year that the first transpacific airmail flight was made and not until 1939 that the first transatlantic airmail flight was inaugurated. As early as 1927, however, four-course, low-frequency radio ranges were installed [29] in the Hawaiian Islands and in San Francisco for the purpose of guiding flight over the Pacific. The station at each terminal could be heard at the opposite terminal, so the four-course radio range became the principal Type I equipment and was not supplemented in this zone with another aid until 1937 when airborne direction finding was installed on American aircraft. Pan American Airways developed a ground-based HF direction finder about 1936 and placed it in use in many parts of the world [30]. During the war, loran equipment was developed [31] and installed in many places throughout the world. Some of these loran chains were decommissioned immediately after the war, but a few years later new equipment [32] was developed and the coverage was extended [33] after about 1950. During the war, the Germans invented a LF azimuthal system known as "Sonne" [16], and they made several installations on the European mainland. This equipment was refurbished under the name of "Consol." In 1954, ICAO approved Consol for long distance use together with loran and nondirectional radio beacons. By 1960, the CAA had installed and commissioned Consol stations at Nantucket, R. I., and San Francisco.

Since the inception of transoceanic flying, the bubble octant has been the most important of the devices for navigation over transoceanic routes; however, in 1945 it was joined by the high-altitude radio altimeter [34] used in conjunction with the aneroid altimeter. In 1960, a few installations of the third self-contained aid (Doppler navigator) were made on aircraft in transoceanic service.

2) The En Route Short-Distance Zone: To fulfill the navigational requirements of the en route short-distance

TABLE 1 - SIGNIFICANT ACTIONS WHICH HAVE AFFECTED THE AIR TRAFFIC CONTROL EQUIPMENT SUBSYSTEMS

Date	Organization	Meeting Place	Chairman	Purpose	Remarks	Bibliographic Ref.
Nov., 1944	52 Nations	Chicago		To set up an organization of a technical and advisory nature.	Upon ratification of the Chicago convention, the International Civil Aviation Organization was formed and later became a branch of the United Nations.	2
Aug. 7-20, 1945	Third Commonwealth and Empire Conference on Radio for Civil Aviation	London	Viscount Swinton (Over-all) Sir Frederick Tymms (1) Sir Robert Watson Watt(2) (3)	(1) To establish communications procedures, regulations, preferred standards for radio equipment, (2) to select from military radio and radar systems those which should be installed on international routes, (3) to suggest a program of radio development to be undertaken in the Commonwealth and Empire and in the United States.	The program developed includes practically all of the items which today are being implemented. A pulse type omnirange with distance measuring equipment was recommended for development.	51
Feb. 4-8, 1946	U.S. Air Force	Washington, D.C.	General Curtis E. LeMay	To discuss all known systems for navigation and control of aircraft with a view towards (subject to national and international commitments and agreements) concentrating our development of those systems which appear most promising.	Several manufacturers presented proposals for coordinated air traffic control systems based on the technology which had developed during the war. Contracts were placed later for many of these systems. The CAA presented a "long range program" consisting of ILS, CW omnidirectional radio range, automatic communications and flight data posting (said to have been under development). Also proposed direct radio communications via VHF from aircraft to tower.	3, 5, 4, 6, 7, 8, 9, 10, 17
Aug. 20-27, 1946	Radio Tech. Commission for Aviation	Washington, D.C.	Howard K. Morgan	To develop recommendations for the United States policy on aeronautical telecommunications.	The program laid down was almost identical with that of Third CERCA but endorsed the VHF, CW phase comparison omnirange. Indicator signaling (ground-air radio link) also recommended.	11
Oct. 28 Nov. 11, 1946	COT Division of ICAO	Montreal	G/C Campbell	To recommend standardization of certain radio aids to navigation and communications.	Accepted ILS for installation not later than Jan. 1, 1951, at all international airports. GCA considered a supplemental standard where warranted. Major discussion centered around the VHF, CW omnidirectional radio range and DME. Recommended that standard Loran and Consol be retained. No single aid to air navigation for long range found to justify universal adoption.	12, 13, 14, 31, 16, 17
Jan. 22 - Apr. 4, 1947	Committee on Interstate and Foreign Commerce House of Representatives	Washington, D.C.	Mr. Wolverton	To study the problems involved in increasing the safety of air transportation.	The following quote is from the report of this group: "When we take the long-range view, it becomes apparent that individual devices do not constitute a coordinated system for safe air navigation. Sometime in the early 1950's there will be an urgent need for substantial and comprehensive improvement." This report lead to the establishment of Committee SC-31.	21
June, 1947 to Feb., 1948	RTCA Special Committee 31	Washington, D.C.	Phase 1 - Col. J. B. Duckworth Phase 2 - Capt. A. S. Born	To undertake a study of air traffic control for the purpose of developing recommendations for the safe control of expanding air traffic. Phase 1 - Development of basic air traffic control principles. Phase 2 - Development of recommendations regarding equipments and procedures required to implement the air traffic control principles.	Developed a plan for coordinated systems with bearing, landing, situation display, and airport surface navigation functions to be multiplexed on DME. Introduced the concept of digital computers with data links to constitute the automatic air traffic control equipment. Accepted an interim system which was essentially the system previously described by RTCA. The coordinated system was to require 15 years for complete development, installation, and training of operators, and was designated the ultimate system in contrast to the revised older RTCA plan designated the interim system.	22
1947 to Mar., 1948	Ad Hoc Committee on Navigation of the Joint Research and Development Board	Washington, D.C.	R. S. Damon	a) Determine the basic problems of military navigation of all types with regard to electronic aids. b) Determine what efforts have been made by the armed services to solve these problems in terms of research and development programs and the provision of experimental facilities. c) Ascertain the relation of the present military effort to solve the problems of navigation and control to that of the various civilian agencies active in this field. d) Recommend to the Joint Research and Development Board through the Committee on Electronics action that should be taken to promote safe and militarily effective navigation. e) Advise the Committee on Electronics on the feasibility of consolidating a national system of air navigation and traffic control with the air warning network.	Recommended the establishment of an organization for research, development, and standardization of the system described by RTCA Committee SC-31, the organization to be "An authoritative full-time planning and steering group consisting of a single representative from each of the departments of Commerce, the Navy, and the Air Force." This organization known as the Air Navigation and Development Board was established by a joint charter by the Secretaries of Defense and Commerce issued on November 2, 1948. It remained in existence until the Airways Modernization Board was established.	23
Jan. 11 to Feb. 26, 1949	Com. and Radio Aids Div. 3rd Session ICAO	Montreal	H. Costa	Foster further international standardization of certain navigation and communications facilities.	Specification for "The purpose of guiding development of UHF (1000 mc) distance measuring equipment accepted but not as an ICAO standard or recommended practice." Developed a GCA specification and refined further the ILS specification.	15, 19
Jan. 19 to Dec., 1950	Special Working Group 5, Air Traffic Control and Navigation Panel, Operational Policy Group - Air Coordinating Committee	Washington, D.C. Boston Wright-Paterson AFB Indianapolis USS Midway at Sea	A. W. Wuerker	To review and evaluate the SC-31 report, all air navigation and traffic control and navigation panel papers that deal with the common system, review and evaluate all available material pertaining to such items as the volume of present and anticipated air traffic and the characteristics of aircraft in use before 1963, and to establish operational policy and procedures for transition system implementation.	More than 200 persons participated in demonstrations and discussions. Work of this committee had the effect of bringing surveillance radar into use in daily air traffic control operations.	25
Apr. 24 to June 2, 1951	Comm. Div. of ICAO 4th Session	Montreal	R. Lecomte	Foster further international standardization of certain navigation and communications facilities.	Recommended where international air service required that 1000 mc DME be installed. Recommended extending standard Loran, Consol, and non-directional beacons for the long distance zone as required.	18
Mar. 9 to Apr. 9, 1954	Do 5th Session ICAO	Montreal	R. Lecomte	As above.	Recommended reviewing the operational objectives for standard long distance radio aids as established by COT Division. Recommended consideration of ground-air radio links known as air traffic control signaling system.	20
Early 1955 to Oct., 17, 1955	Special Working Group 13 Air Coordinating Committee	Washington, D.C.	Capt. Sam P. Saint	To find the way by which the advancement of U.S. operations can be brought about through the most modern operational and technological means.	Final report was published in March of 1957 and recommends use of defense radar in traffic control.	76
May 4 to Dec. 31, 1955	Aviation Facilities Study Group of the Director of the Bureau of the Budget	Washington, D.C.	Wm. Barclay Harding	Study U.S. Aviation facilities requirements for the next decade or two.	Recommended that a study should be made under the direction of an individual of national reputation with a broad understanding of civil and military aviation. President Eisenhower acted on this recommendation by announcing the appointment of General Edward P. Curtis on February 11, 1956.	77
May, 1957	System Engineering Team Office of Avia. Facilities Planning (Curtis Committee)	Washington, D.C.	Preston R. Bassett	To plan for meeting aviation facility needs in the U.S. for the next two decades.	Report of this committee described a new on the spot invention and was known as the Curtis report. In May of 1957, General Curtis wrote to the President recommending that the plan set forth by this committee be used as a basis for airways modernization and further that an independent agency known as "The Airways Modernization Board" be established. This Board was established on August 14, 1957, by public law 85-133. This Board remained in existence until it was absorbed as part of the Federal Aviation Agency which came into being on August 23, 1958, under public law 85-726.	27
Mar., 1957	Air Coordinating Comm.	Washington, D.C.	Louis S. Rothschild	To review the air traffic control systems and procedures in light of the introduction of jet aircraft and the existing defense situation.	Included the findings of Special Working Group 13 in a complete report which recognized Vortac and is the basis for the present air navigation and traffic control equipment program in the U.S. The Air Coordinating Committee was abolished on August 11, 1960, by executive order 10833.	26
1957	Working Group of European Air Space Coordination Committee of NATO	Paris		To study the problems of air traffic control in the upper air space over the six common market countries.	The six common market countries joined by the Benelux countries and Federal Republic of Germany joined by France and Italy drafted a convention to establish a permanent commission for air navigation security. The purpose of this organization is to standardize air traffic regulations, formulate common policy on radio aids, study technical progress on air navigation services, form an international agency to be known as Eurocontrol to provide air navigation services.	28
Feb. 10 to Mar. 2, 1959	Special Com/Ops/Rac Meeting of ICAO	Montreal	J. W. Stone	(Among others) to develop recommendations for standards and recommended practices and guidance material for short distance navaids.	Recommended that distance measuring equipment corresponding in principle to that specified by the U.S. (for its precision polar coordinate navigation system) should be adopted as the ICAO Standard for application as a complement to VOR.	78

zone, four-course radio ranges were installed [36] beginning in 1929. This equipment operated in the band of 200–400 kc. Since the number of equipments grew from 9 installations in 1929 to 378 in 1949 [37], it was necessary to reduce the power (and coverage) to allow for the use of the same frequencies simultaneously. Some LF nondirectional stations were installed in 1930 to mark the intersections of range legs but, beginning in 1939, these stations were replaced with markers [38] operating at 75 Mc. In 1938, however, it was required that aircraft carry direction finders, and as time went on, LF nondirectional stations were installed purely for use with these airborne direction finders. In Europe during the period extending to 1939, ground direction finders were the most important navigational facility. In 1937 there were seven direction-finder nets in Europe [39], each employing three stations for giving positions to aircraft from ground observation. At that time, however, eleven additional ground direction-finder nets were planned for installation. Upon outbreak of the war, about 100 four-course radio ranges were installed throughout the world. Fig. 1 shows the facilities installed for navigation in this short-range zone. While the RTCA 1945 meeting [11] had considered that the four-course radio ranges would be obsolescent in 1950, only 56 of these stations had been decommissioned by 1959.

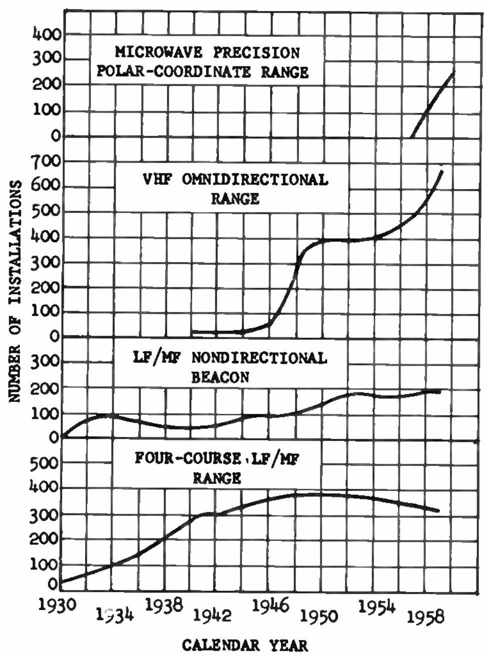


Fig. 1—Navigation equipment of the short-distance en route zone.

VHF radio ranges employing two visual and two aural courses [40] were installed beginning in 1940, but were later replaced with omnidirectional radio ranges. The VHF omnidirectional radio range (VOR) [41] program became an important factor in navigation for the United States short-distance en route zone in 1948, at which time over 300 units were installed. By 1959 the number had increased to 661 [37]. By 1960 the number of VOR's in other parts of the world had reached 279 stations. The mainstay of international navigation is the nondirectional beacon used with airborne direction finders. Over 200 stations are in operation at the time of this writing.

The RTCA program of 1945 envisioned the installation of a precision polar-coordinate system in 1955. Some installations were made by the military prior to that time. However, the civil program [24] began in 1957. By the end of 1960 this program totaled 343 stations. These stations (when commissioned) provide distance measurement to VOR equipped aircraft and precision bearing and distance to aircraft equipped with the complete microwave (Tacan) system [42]. Largely, due to installation by the military, about 170 precision microwave polar-coordinate ranges were in operation in other parts of the world (mainly in Europe).

3) The Approach and Landing Zone: Development of a so-called "blind landing system" was begun at the Bureau of Standards in the U. S. as early as 1928 [43]. A contract for a new development was placed by the Civil Aeronautics Administration in 1938 [44], but by 1941 the U. S. did not have ready for use a satisfactory low-approach system. In Europe in 1932, the Lorenz Company began development of a low-approach system [45] based on that outlined by the Bureau of Standards, and by 1939 this system (termed Standard Beam Approach System) had been installed throughout Europe and in some locations in South America. With U. S. entry in the war, the CAA development was completed by the U. S. Army Air Corps [13], and 500 equipments were manufactured in the period from 1942 to 1945. Installation of this equipment for civil use was begun in 1945. By 1959 a total of 174 stations [37] had been commissioned in the U. S. ILS was one of the first systems to be adopted as a world standard by ICAO in 1946 and more than 154 stations are installed external to the U. S. The world program is scheduled at 208 stations. A radar-beam approach system [46] was developed by the Radiation Laboratory of the Massachusetts Institute of Technology during the war and approximately 200 units were constructed prior to 1946. Installation of these devices for civil use began in the U. S. in 1947, and by 1959 a total of 11 were in operation [37]. The number in use by the military is far greater. Approximately 40 installations are in operation external to the U. S.

4) The Airport Zone: Little work has been done to install an effective guidance and control system for the airport zone. At the time of this writing, 10 radars operating at wavelengths of less than one cm have been installed in the U. S. for the purpose of detecting the movement of aircraft located on the surface of civil airports.

B. Type II Equipment

This equipment constitutes the means whereby a central decision-making agency may be able to know the position of all aircraft within its jurisdiction. From the time of the inception of air traffic control in 1935 to the war years, the only means provided for this purpose were reports by the pilots via HF radio transmitters. An airline pilot would report his position to a ground radio station owned by his company. The company personnel would, in turn, through the use of land-line telephony (often via a direct wire), pass the information to the local traffic control center. A similar procedure was followed for some of the military. Private flyers could, through the use of a frequency of 3105 kc, report their positions to LF radio range stations and could receive replies via the stations' normal navigational frequency. Communications were established from the aircraft to airport towers via the HF transmitter, and instructions from the tower were received on a frequency of 278 kc. About 1940, the airlines of the U. S. placed orders for VHF airborne equipment. However, the airline developments were taken over by the military prior to the U. S. entry into World War II. The RTCA had on February 8, 1937, discussed the desirability of employing VHF frequencies to replace the 278-kc communications used in airport control work [47]. Shortly thereafter, the U. S. Bureau of Air Commerce (the predecessor of the CAA) undertook tests at 125 Mc.

Aircraft positions could have been determined on the ground by networks of HF direction finders (as in Europe), and some work toward developing and installing such systems was conducted in the U. S. prior to World War II; however, these installations never became fully effective.

In the post-war era, RTCA Committee SC-31 recommended the utilization of primary radar [22] for the purposes of equipment of Type II as part of the interim system. For the ultimate system, it recommended the development of an elaborate secondary radar system. Special Working Group 5 [25] of ACC stated, "of all the aids now becoming available, airport surveillance radar is capable of making the most significant contribution toward increased safety and efficiency of operation." As a consequence of this recommendation, the CAA augmented its ground radar program. By 1959, 48 radars [37] were in operation in the U. S. serving the purpose of Type II equipment. The Standards and Recommended Practices of ICAO [19] first mentioned performance of surveillance radar in 1949, and this equipment is installed at civil airports throughout the world where heavy traffic is experienced. In the U. S. in 1957 Special Working Group 13 recommended [26] the inauguration of a program whereby information from military radars would be employed for civil and military air traffic control. This program was implemented, thereby greatly increasing the total radar coverage available for traffic control purposes.

The most significant change in means of Type II was the introduction of VHF communications between aircraft and control centers. The installation of such a system was begun in the latter part of 1950. Of course, transoceanic air traffic must continue to utilize HF communications although tests with directional arrays on the ground to provide communications with jet aircraft have shown ranges [48] over the North Atlantic up to 600 nautical miles. In Europe, some VHF direction-finding networks are employed to determine aircraft positions. In the U. S. (based on SC-31 recommendations), VHF and UHF direction finders [49] are used as a means for assisting in the identification of targets appearing on the radar scopes.

C. Type III Equipment

This equipment constitutes the means whereby the decision-making entity can determine instantly safe, nonconflicting flight procedures for all aircraft within its purview. As this is written, these means [50] are nonelectronic. By means of manual time and distance computers, the traffic controller determines when an aircraft will be over some specific fix. Another device is the flight progress strip. These strips list the number of the flight, altitude, and other pertinent information and fit into racks in time sequence of arrival over a check point represented by a rack.

D. Type IV Equipment

The means whereby safe, nonconflicting instructions are transmitted to the aircraft are the same means that are used for the pilot to report his position. That is, until the war years, communications from the ground to the aircraft took place via "company" or military stations via HF and to itinerant flyers via the LF radio range stations or LF tower stations. After the war, instructions were transmitted to pilots largely via VHF voice transmission except for transoceanic aircraft. Voice is also available on some VHF navigational aids. During the war the British developed an indicator signaling system [51] operating at VHF frequencies. This system was known as "Beechnut" and in the United States there was a parallel development called "Voflag." By means of this system, a series of pulses transmitted to the aircraft would cause visual indicators in the cockpits to appear with clearance messages inscribed on them (such as hold, proceed, climb, etc.). A system of this type was recommended by the RTCA meeting of 1946 for installation in 1948 but was never implemented.

E. Type V Equipment

The means whereby action may be coordinated among the several central flight-instruction, decision-making entities is of two different types. One of these is by direct-wire telephony (referred to as interphone). The second means is the teleprinter. In 1937 only 57 miles of interphone system were employed in the U. S. connecting 57 stations. By 1960, 167,821 miles of interphone were in use serving 4564 drops. The first interphone for control tower purposes was installed between 114 stations in 1939. By 1960, 404 stations were connected over a network of 49,906 miles. In addition to the wire service in the U. S., under the auspices of ICAO, radio teleprinter is being employed between many countries and a cable system is in process of installation between Canada and the United Kingdom to serve the needs of international air travel.

F. Conclusion

From the foregoing description of the equipment which constitutes the *air traffic control equipment subsystem*, it is evident that rather than to speak of it as a "system" it would be more accurate to say that it is an agglomeration of units with man providing the necessary flexible interconnections. The coordinated system mentioned in the Congressional Report of the 80th Congress and described in detail by RTCA Committee SC-31 was never implemented. SC-31 found that a factor of merit no greater than 67 per cent would result if the procedure employed was that of adding units to existing units, but its findings have not been heeded.

III. Forecast

Since events like those listed in Table I[1] doubtlessly will dictate the future systems, it is not possible to make a forecast with any degree of confidence. Therefore, the following is based on known developments and prejudices with which the reader may freely disagree.

A. Type I Equipment

This equipment is the navigation means for the four zones as follows:

1) The En Route Long-Distance Zone: The COT Division meeting in 1947 [12] found no system that could justify universal adoption, and looked to new developments to fill the requirements. In the years that have ensued, several systems [35], [53], [54], [55] have been developed which could meet most but not all of the requirements, but these newer systems have not been adopted. The lack of action in providing these aids leads to the conclusion that they have not really been needed. Yet air traffic control over the North Atlantic suffers from the necessity of each aircraft being spaced by departure times alone and, therefore, each carries with it thousands of cubic miles of air space. The increase in traffic will force the adoption of some aid. At the moment, tests are in progress attempting to determine the success of a VORTAC installation aboard weather ships. A VHF aid is advantageous for air traffic control use because it is unaffected by atmospheric and propagation anomalies such as those associated with the aurora borealis. It would be readily possible to install radio ranges capable of reception on existing equipment and operating in the VOR frequency band. It is predicted that transoceanic aids, at least for the North Atlantic, will utilize a technique similar to that of the present short-distance aids. Some loran C [35] installations are being made in various parts of the world. This device has been shown to have a ground-wave range over water of 1700 miles in the daytime and 1200 nautical miles at night, with probable errors of the order of 0.1 to 0.2 μsec. Sky-wave ranges are greater with correspondingly greater errors [56]. With such outstanding performance, then, what is the probability of loran C being accepted as the international standard? One can readily picture an international conference at which the proponents of the system would outline its range, accuracy, and the existence of certain networks. The opponents of the system would bring out 1) airborne equipment complexity and weight, 2) susceptibility of the low frequencies employed to propagation anomalies, 3) the necessity for locating cooperating stations in different countries (therefore, not under the control of a single responsible agency), and 4) the lack of automaticity and difficulty of adaption to computers and displays. The outcome of these discussions would lie solely in the laps of the techno-political gods.

In recent years there have been installations [57] of Doppler navigators [59] on transoceanic aircraft and also the development of inertial systems. There have been some forecasts [58] that these self-contained aids will replace the ground-based aids. While the Doppler navigator often ceases to function over calm water and its accuracy is dependent on the accuracy of the heading device employed [60], its ability to give ground speed and track makes it a valuable aid even though it is not associated with a positional computer. Inertial systems, although corrected by Schuler tuning so that their average errors [61] may be small, may at times have large instantaneous errors. The accuracy of the heading information which inertial systems provides has been sorely needed for flying ever since instruments were put on aircraft. Self-contained aids are not new to long-distance flying. The octant and the high-altitude altimeter have produced navigational performance of comparable accuracy to that provided by the Doppler and inertial systems, although not as conveniently. ICAO in the 6th Session of its Communications Division [79] said that it is extremely doubtful whether these aids alone could meet the needs of air traffic control. It is predicted, therefore, that both Doppler and inertial systems will become standard equipment on all aircraft that traverse the long-distance zone, but the requirement for ground-based aids for traffic control will not diminish.

2) En Route Short-Distance Zone: Experience of many years has shown the desirability of employing in this zone a precise navigation system and a homing device. With time, it is believed that the VORTAC system will perform the same function as that performed by the two older systems. As pictorial and course-line com-

[1] In March, 1961, President Kennedy issued a directive calling for a scientific, engineering review of aviation facilities and related research and development. This effort, known as "Project Beacon," was under the chairmanship of R. R. Hough. As of November 1, 1961, the report of this committee had not been made public.

puters are installed in aircraft, the desirability of obtaining both bearing and distance from the Tacan portions of VORTAC [62] with its higher accuracy [63] and lack of course scalloping will become evident, while the VOR will serve as a homer and for communications. It is forecast, therefore, that VORTAC will be the navigational aid for the short-distance en route zone for the next thirty years. Should it become necessary, it is possible to increase the accuracy of Tacan by one order of magnitude without too great a modification of the equipment.

3) The Approach and Landing Zone: To permit fully automatic landings in the fourthcoming years, there will be additions of equipment (such as the radio altimeter) and other improvements to the present low-approach systems. A large directional array has been developed [64] for use with the ILS (fixed-beam low-approach system) and has been in use by the U. S. Air Force for some time. The employment of this array so improves the localizer course that the use of automatic pilots for landings will force its adoption. Again, however, the necessity for using the localizer frequencies of 108 to 112 Mc for communication will force the replacement of this localizer with a similar device operating at a higher-frequency band.

The use of the frequency of 329.3 to 335 Mc for the glide slope equipment of the ILS has long been a detriment to optimum performance [65]. In the future, then, it can be expected that this equipment will be moved to a higher frequency (of the order of 5000 Mc). At the same time, techniques will be employed which will allow for selective glide slopes suitable for VTOL aircraft.

With the adoption of fully automatic landings [66], the radar low-approach system takes on greater significance. It will be mandatory that each automatic landing be monitored from the ground by a radar low-approach system. The radar low-approach systems currently in use have a data rate that is too low, and the complex wartime development has been made still more complex. It is predicted that the future will see the development of a simpler automatic landing monitoring ground radar with improved performance over anything heretofore employed for this purpose.

4) The Airport Zone: With improvements in en route flying and automatic landings, the bottleneck to the movement of aircraft will be at the airport. This bottleneck will so reduce the return from the high airport investment that a revolution will literally be caused in the development of navigation and control systems for the airport zone. A special signaling and detection system will be developed for operation on the airport surface so that aircraft may be guided to and from the loading ramps at top speed.

B. Type II Equipment

It has been shown that the means which has been enjoying the greatest increase in popularity in recent years and whereby the central decision-making agency may know the position of all aircraft within its purview is the surveillance radar located on the ground. At the same time the inherent weaknesses of the ground surveillance radar (even though it may be improved to indicate altitude) are very well known. Its data rate is too low; it is unable to determine the identity of its targets; targets are frequently obscured by precipitation and ground clutter, yet if MTI (moving target indicator circuit) is used, then orbiting aircraft disappear [67]. If the radar output is to be connected to a computer, it must be equipped with complex track-while-scan circuits. After all, radar is a development of twenty years ago and should be obsolescent. To remedy some of these defects, an airborne transponding beacon [68] was recommended by SC-31 and the combination of radar and beacon was to be obsolescent in 1961. The development of the beacon has been in process since at least 1948 and it appears only now that it will become standard for air-carrier aircraft. The primary radar program, even with beacon, is an example of a program that is "too little too late." There is in the U. S. a large network of defense radar that is now being used to furnish traffic control information. If the air-breathing threat to the nation's defense is determined to have disappeared, there will be no use for this radar [69] system except for traffic control. It might be argued that the high cost of telephone line rental from these radars, together with the high cost of maintenance and operation, would render this defense radar network uneconomical for traffic control use alone.

The above arguments tend to forecast the disappearance of primary radar in the air traffic control system, but quite on the contrary, radar is the only device that will be available for some time for this valuable purpose; consequently, it appears very certain that ground surveillance radar will be in the traffic system for at least thirty years hence.

The use of VHF reporting of position via voice will not disappear. Quite on the contrary, this system employed with high-accuracy ground direction finders [70] attached to digital computers will form a valuable aid, particularly to track those aircraft which will not be equipped with radar beacons. The importance of VHF for voice will render the proposed air-to-ground automatic reporting link [71] on these frequencies obsolete before the system ever comes into use. If, for no other reason than the fact that the airborne distance-measuring equipment interrogator has frequencies which are lightly used, air-to-ground automatic reporting will eventually find its place as a simple adjunct to this equipment [72]. As the interrogator performs its normal function in cooperation with already installed ground stations, the identity, altitude, and position of the aircraft will be conveyed to the ground where, over already existing telephone lines, this information will be relayed to traffic control centers.

C. Type III Equipment

An accurate forecast of the means by which the central decision-making entity of the future will determine nonconflicting flight procedures for all aircraft within its purview is that this equipment will consist of an electronic digital computer associated with suitable displays for which the man-machine relationship has been adequately solved. Such a forecast, although accurate, has little meaning unless the philosophy governing the procedures by which the aircraft would be operated is also forecast.

Perhaps the basic conditions underlying the air traffic control problem can best be understood by contrasting them with those for railroad operation. At each railroad terminal there are many sidings where the trains may be loaded and unloaded, but once out of a railroad terminal trackage is expensive and many miles of railroad right-of-way consist of only a single pair of rails. The condition just described is diametrically opposite to that for air operations. Airports require vast areas and very expensive runways. Only very few runways are present at the largest of our airports. Once away from the terminal area, however, the aircraft has unlimited space. If one observes the total number of aircraft in a flight between New York and Washington (over the world's heaviest density area), only very few aircraft will be seen during the entire flight. Why then is there an airways traffic control problem? It has been shown that if aircraft are allowed to fly at random that the probability of collision is very small. This probability is greatly increased by compacting aircraft on airways. A few freethinkers [73] in the air age have been pointing out that if the aircraft can be equipped with an effective collision avoidance system, they could fly where they will and in any number desired with a lower probability of collision than exists with the present system which is subject to error. It is not forecast that airborne true collision-avoidance devices will become practical within the next twenty to thirty years, although proximity and collision warning devices may come sooner. Experience at sea has shown that a mere proximity detector only increases the danger of collision. However, it is readily possible to develop and install a highly effective ground-based collision-avoidance system [74]. With the position reports relayed to a digital computer on the ground via the data links previously described, it is possible to send the pilots a proceed signal or a signal indicating a collision-avoidance maneuver. It is believed, then, that the future will see the elimination of the airways and the installation of ground digital computers which take their information from air-to-ground links that are automatically relaying the position of aircraft.

The above forecasts one type of digital computer but it has been pointed out that the difficult air traffic control problem lies not on the airways but at the terminals, so it is necessary to suggest the character of the digital computers that will solve the terminal problem. The popular solution to the terminal problem which has been proposed by many electronic engineers is to control (restrict) the traffic along the airways in order to solve the terminal problem. This control procedure is completely at variance with the characteristics of the aircraft and would force the aircraft to operate in such a manner as to solve the problem of the electronic engineer or the controller rather than solving the problem for the aircraft. It is conceivable that the aircraft at the terminals could be controlled entirely by automatic means. However, at terminals many instances occur which are the sources of potential accidents, and to allow for all of the possible contingencies that might occur, it appears that the present manual procedures which have been laboriously worked out through the years in actual practice will form the best basis for future operations. There is no reason why a very large amount of automation should not be applied. The communications systems between controllers are crude and the display systems have had no appreciable development to adapt them to the tasks for which they are used. It is forecast, therefore, that there will be a substantial automation assistance to the terminal air traffic controller in the future. Digital computers will update flight progress and compute release times for various terminal points and in other ways will relieve the controller of much of the labor, thus giving him a far greater degree of accuracy of control than he now has time to exercise. The machine will be used to monitor the actions of the controller [75] rather than vice versa.

D. Type IV Equipment

It is forecast that VHF communications will continue to be an important means whereby the instructions will be transmitted to the aircraft. There has been standardization among the military services of an indicator signaling ground-to-air link operating at UHF frequencies. Aside from the frequency space that is occupied by this type of communications, the weight and space penalty paid in the aircraft for its adoption is not great. At the same time, the amount of information to be transmitted from ground-to-air (in contrast to that required for transmission from air-to-ground) is not required to be too extensive. Data link from ground-to-air at VHF frequencies, therefore, will probably be adopted. Data link at 960 and 1215 Mc from ground-to-air has been demonstrated to be satisfactory and could serve for the collision-warning link, which has been mentioned in the previous section, to bring the output of the collision-avoidance computers to the aircraft.

E. Type V Equipment

If aircraft on the airways are to be allowed to fly as they wish only under the surveillance of ground-installed collision-warning systems, it will be necessary that these systems be capable of passing positional information from one to the other. The art of machine talking to machine is being developed rapidly, and such a system would be installed with the collision-avoidance system previously described. By relegating airways traffic control to the fully automatic means, the amount of information to be passed between terminals would be greatly reduced. Thus, there would be no need for an extensive augmentation of the teleprinter network which now connects terminal to terminal. By the use of indicator-type signaling directly to the face of the controller's console, much of the present voice communications via interphone would be eliminated.

F. Conclusion

The foregoing forecast is not based on any new and startling technological breakthrough. Having read in the public press for these many years of the seriousness of the traffic control problem, the reader might tend to conclude that 1) important technological breakthroughs have for years been awaited in order to solve this problem, and 2) that the writer is forecasting that the problem is unlikely to be solved. Neither of the above conclusions is correct. Quite on the contrary, the technology necessary to solve the problem has, since the war years, existed in a sufficiently advanced state to permit solution. The traffic control problem has not been solved because it consists in establishing agreements and actions among groups with divergent ways of thinking and divergent aims. It is hoped that in the reasonably near future these groups can agree on a common ***modus operandi***, for only this action is required to solve the traffic control problem very quickly.

Bibliography

[1] P. C. Sandretto, "The air traffic control paradox," IRE Trans. on Aeronautical and Navigational Electronics, vol. ANE-5, pp. 80–85; June, 1958.
[2] "What is ICAO?" Internatl. Civil Aviation Org., Montreal, Canada; May, 1947.
[3] Electronic Subdivision Advisory Group on Air Navigat on, Wright Field, Dayton, Ohio, Rept. No. TSELC-SP2; June, 1946.
[4] "They call it Tricon, General Electric's triple coincidence electronic method of air navigation," *Aero Digest*; April–May, 1948.
[5] "The Block System for Airways Traffic Control," General Railways Signal Co., Rochester, N. Y., Pamphlet 630; October, 1946.
[6] R. I. Colin, *et al.*, "Aerial navigation and traffic control with Navaglobe, Navar, Navaglide, and Navascreen," *Elec. Commun.*, vol. 23, pp. 113–143; March, 1946.
[7] K. McIlvain, "Hazeltine Lanac system of navigation and collision prevention," *Proc. Radio Club of Am.*, vol. 26; February, 1949.
[8] L. F. Jones, *et al.*, "Pictorial display in aircraft navigation and landing," Proc. IRE, vol. 38, pp. 391–400; April, 1950.
[9] J. Lyman and G. Litchford, "A precision omnidirectional radio range for the terminal area," *Aeronaut. Engrg. Rev.*, vol. 9, p. 16; June, 1950.
[10] "Traffic Control Without Pilot Effort," Teleregister Corp., New York, N. Y.; 1945.
[11] "Recommended United States Policy," Air Navigation-Commun.-Traffic Control, RTCA, Washington, D. C.; August 28, 1946.
[12] ICAO, 1st Session, Special Radio Tech. Div., Montreal, Canada, Final Rept., pp. 25–26; January, 1947.
[13] S. Pickles, "Army Air Force portable instrument landing system," *Elec. Commun.*, vol. 22, pp. 262–294; 1945.
[14] John S. Hall, "Radar Aids to Navigation," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 237–251; 1947.
[15] ICAO, 3rd Session, Commun. and Radio Aids to Air Navigation Div., Montreal, Canada, Doc. 6625-COM/521; March, 1949.
[16] P. C. Sandretto, "Electronic Avigation Engineering," IT&T, New York, N. Y., pp. 116–144; 1958.
[17] H. C. Hurley, *et al.*, "CAA VHF Omnirange," CAA Tech. Dev. and Evaluation Ctr., Indianapolis, Ind., Tech. Dev. Rept. 113; June, 1950.
[18] ICAO, 4th Session, Commun. Div., Montreal, Canada, Doc. 7171-COM/544; 1951.
[19] "Standards and recommended practices," *Aeronaut. Telecommun.*, ICAO, Montreal, Canada, pp. 45–55; June, 1949.
[20] ICAO, 5th Session, Commun. Div. Montreal, Canada, Rept., Doc. 7480-COM/548, March 9–April 9, 1954.
[21] "Safety in Air Navigation," 80th Congress, Washington, D. C., p. 16, House Rept. No. 1065; 1947.
[22] "Air Traffic Control," RTCA, Washington, D. C., Paper No. 27-48/DO 12; May 12, 1948.
[23] "Navigation and National Security," The Natl. Military Establ., Res. and Dev. Board, Washington, D. C., p. XIV; March 5, 1948.
[24] "ANDB recognizes Tacan," *Am. Aviation Daily*; February 8, 1955.
[25] "Air Traffic Control and the National Security," Air Coordinating Comm., Oper. Policy Group, Supt. of Doc., U. S. Govt. Printing Office, Washington, D. C.; December, 1950.
[26] "Accelerated Modernization of the U. S. Air Traffic and Navigation System," Air Coordinating Comm., U. S. Dept. of Commerce, Washington, D. C.; March, 1957.
[27] "Modernizing the National System of Aviation Facilities, Supt. of Doc., U. S. Govt. Printing Office, Washington, D. C.; 1957.
[28] "Eurocontrol," *Civil Aviation Radio News*, CERCA Ctl. Office, Berkeley Sq. House, London, England; March, 1960.
[29] C. C. Shangraw, "Radio beacons for transpacific flights," Proc. IRE, vol. 16, pp. 1203–1235; September, 1928.
[30] A. E. Harrison, Washington, D. C., CAA Tech. Dev. Confidential Note 15; January, 1939.
[31] J. A. Pierce, *et al.*, "Loran," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 97–102; 1948.
[32] R. H. Meyers, "Design of a loran transmitter," *Elec. Commun.*, vol. 28, pp. 31–45; March, 1951.
[33] "Ocean Electronic Navigation Aids," U. S. Govt. Printing Office, Washington, D. C., Rept. No. CG157-1; pp. 26–27, 1949.
[34] H. F. Langdon, "Navigating by altimeter," *Air Transport*, pp. 53–55; May, 1946, and pp. 47–49; July, 1946. (Also Sandretto [16], *op. cit.*, pp. 341–351.)
[35] W. P. Frantz, *et al.*, "A precision multi-purpose radio navigation system," 1957 IRE National Convention Record, pt. 8, pp. 79–98.
[36] H. J. Walls, "The civil airways and their radio facilities," Proc. IRE, vol. 17, pp. 2141–2157; December, 1929.
[37] "FAA Statistical Handbook of Aviation," Supt. of Doc., U. S. Govt. Printing Office, Washington, D. C., p. 23; 1960.
[38] Sandretto [16], *op. cit.*, pp. 105–113.
[39] H. W. Roberts, "Radio in Aviation," William Morrow and Co., New York, N. Y., p. 12; 1945.
[40] "Development of the Ultra-High Frequency Radio Range," CAA, U. S. Dept. of Commerce, Washington, D. C., pt. 3, p. 2. Tech. Dev. Rept. No. 49; June, 1945.
[41] Sandretto [16], *op. cit.*, pp. 373–427.
[42] *Ibid.*, pp. 521–557.
[43] W. E. Jackson, "The Status of Instrument Landing Systems," U. S. Govt. Printing Office, Washington, D. C., pp. 1–3, Tech. Dev. Rept. No. 1; October, 1937.
[44] W. E. Jackson, *et al.*, "Development of the CAA instrument landing system at Indianapolis," *Elec. Commun.*, vol. 18, pp. 285–302; April, 1940.
[45] E. Kramar, "Present state in the art of blind landing of airplanes using ultra-short waves in Europe," Proc. IRE, vol. 23, pp. 1171–1182; October, 1935.
[46] Sandretto [16], *op. cit.*, pp. 683–710.
[47] W. E. Jackson and J. C. Hromada, "125-Megacycles Airport Traffic Control Tests at Indianapolis," U. S. Govt. Printing Office, Washington, D. C., CAA Tech. Dev. Rept. No. 2; January, 1938.
[48] "Report on Extended Range VHF Communications at Shannon, Ireland," Pan American Airways, Atlantic Div., New York, N. Y.; September 15, 1959.
[49] Sandretto [16], *op. cit.*, pp. 4502–509.
[50] "Operation of the Air Traffic Control System," U. S. Govt. Printing Office, Washington, D. C., pp. 8–23; September, 1953.
[51] "Report on the Third Commonwealth and Empire Conference on Radio for Civil Aviation," H. M. Stationery Office, London, England, pp. 152–159; Summer, 1945.
[52] "Statistical Handbook of Civil Aviation," CAA, Washington, D. C., p. 17; December, 1945.
[53] J. A. Fusca, "Navarho," *Aviation Week*, August 5, 1957.
[54] J. A. Fusca, "Omega navigation aid could serve the world," *Aviation Week*, November 11, 1957.
[55] T. D. Johnson, "Status of Dectra," *J. Inst Navigation*, vol. 5, pp. 303–310; Summer, 1957.
[56] W. T. Dickinson, "Engineering Evaluation of the Loran C Navigation System," U. S. Coast Guard, Washington, D. C., Contract No. TCG-40547 (CG-40502-A); September 30, 1959.
[57] "Airline tests Doppler radar," *Electronics*, Business Issue, p. 15; October 10, 1957.

[58] I. W. Boggiss, "Inertial Doppler systems seen key to automatic navigation," *Electronic News*, December, 1959.
[59] W. J. Tull, "Doppler navigation," *Aeronaut. Eng. Rev.*, pp. 58–61; May, 1957.
[60] W. R. Fried, "Principles and performance analysis of Doppler navigation systems," IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS, vol. ANE-4, pp. 176–196; December, 1957.
[61] F. Stevens, "Aiding the inertial navigation system," *J. Inst. Navigation*, vol. 6, p. 168; Autumn, 1958.
[62] See [26], *op. cit.*, pp. 74–76.
[63] D. F. Latimer, "Azimuth errors of the TACAN systen," IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS, vol. ANE-3, pp. 150–156; December, 1956.
[64] R. A. Hampshire, "A new fixed-beam aircraft approach system," 1954 IRE NATIONAL CONVENTION RECORD, pt. 5, pp. 77–83.
[65] C. B. Watts, Jr., "Theoretical Consideration of an Improved Glide Path Antenna System," CAA, Tech. Dev. Rept. No. 81, Washington, D. C.; March, 1959.
[66] "Automatic Landing," *Flight*, vol. 00, pp. 233–246; February 19, 1960.
[67] L. N. Ridenour, "Radar System Engineering," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 646–655; 1947.
[68] "The art of radar beaconry," *Skyways*, pp. 21–27; November, 1959.
[69] "FAA Research and Development Program and Progress Report," U. S. Govt. Printing Office, Washington, D. C., p. 38; September, 1959.
[70] *Ibid.*, pp. 89–90.
[71] R. E. Davis and D. T. Gross, "Automatic communication system for air traffic control," *Proc. NAEC*; pp. 522–530; 1959.
[72] B. Alexander and R. C. Renick, "Background and Principles of Tacan Data Link," *Elec. Commun.*, vol. 34, pp. 171–174; September, 1957.
[73] H. W. Boggess, "What's ahead for business aviation?" *Air Force Mag.*, vol. 39, pp. 64–67; March, 1956.
[74] P. G. Sandretto and M. Rogoff, "A novel concept for application to the control of airways traffic," *J. Inst. Navigation*, vol. 6, pp. 102–107; Summer, 1958.
[75] P. M. Fitts, "Human Engineering for an Effective Air Naviga-t on and Traffic Control System," Natl. Res. Council, Washington, D. C., pp. 5–11; March, 1951.
[76] "The Continued Implementation and Improvement of the U. S. Common Civil/Military System of Air Traffic Control and Navigation," ACC Air Traffic Control and Navigation Panel, 2nd Oper. Policy Group, SWG 13, Washington, D. C.; October 17, 1955.
[77] "Aviation Facilities," Rept. of Aviation Facilities Study Group to the Director, Bureau of the Budget, Washington, D. C.; December 31, 1955.
[78] ICAO, Special COM/OPS/RAC Meeting, Montreal, Canada, pp. III–15, par. 1.8, Doc. 7973, COR/1; February 10–March 2, 1959.
[79] ICAO, 6th Session, Commun. Div., Montreal, Canada, Rept., pp. VII-103, Doc. 7831-COM/551-2; September 10–October 14, 1957.

Space Navigation*

ALTON B. MOODY†

Summary—**In space a successfully navigated craft must not only reach its intended destination, but must arrive there at the right time for rendezvous, and with a suitable velocity vector. Power limitations impose severe restrictions upon the amount of maneuvering that can be performed. Navigation techniques differ somewhat during the escape phase, in the vicinity of a planet and its natural satellites, during the midcourse phase of an interplanetary flight, and during the terminal phase. The most useful coordinate system is radius from a celestial body and some form of "latitude" and "longitude" on the sphere thus defined. Quantities measured are speed, direction, distance, acceleration, angles, and time. Of these, speed and direction of travel cannot now be measured to navigational accuracy. Position in space can be determined electronically, optically, and by inertial navigation, each under suitable conditions. Use of physical phenomena is not promising. A general-purpsoe digital computer will be needed. No order of magnitude improvement in state of the art is needed to produce a first-generation fully automatic space-navigation system, but considerable development work is required.**

INTRODUCTION

The Problem

NAVIGATION on or near the surface of the earth is usually considered successful if the craft arrives at the desired destination by a reasonably expeditious route. In contrast, space navigation is successful only if the space craft arrives at the desired destination at the right time to effect rendezvous with its target, and with a suitable velocity vector. The probability of recovering from a major mistake is remote.

The lack of adequate power to launch large payloads and provide the thrust needed to control their motions in space and return them to earth constitutes the principal limitation of space travel. It also dictates the requirement that the accuracy and reliability of any navigation system be very high. With a more satisfactory solution of this problem will come greater flexibility in space travel.

Space craft have been guided during powered flight extending over the first few minutes after liftoff. Following initial cutoff, positions near the earth have been determined by means of tracking facilities on the earth, and limited changes in orbits have been made by command from the earth. This is the extent of space navigation to date.

For the foreseeable future, space navigation can be considered in several phases or areas: 1) during escape, 2) near a planet and its natural satellites, 3) midcourse portion of a flight to another planet, and 4) terminal phase to an orbit or soft landing. Purposeful interstellar

* Received by the IRE, June 28, 1961. The opinions or assertions in this paper are the private ones of the author, and are not to be construed as official, or as reflecting the views of the Dept. of Defense, the Navy Dept., or the Naval service at large.
† Capt., USNR, U. S. Navy Hydrographic Office, Washington, D. C.

and intergalactic flights, and beyond, will have to await the development of more efficient fuels and the solution of other formidable problems.

Kinds of Navigation

Terrestrial navigation techniques are often conveniently classed as piloting, dead reckoning, or celestial. The same classifications can be useful in space.

Piloting: Piloting is position determination relative to one's environment. Traditionally, it has involved the use of bearings, distances, map-matching in some form, etc. A prior survey to locate accurately the various landmarks used is essential. The application of electronics to navigation during the present century has largely eliminated the visibility limitation and has extended the range to provide coverage over much of the ocean areas.

Dead reckoning: Dead reckoning involves the extrapolation of a known position in accordance with assumed or anticipated motion during a given time interval. The basic elements of dead reckoning are velocity and elapsed time, or direction and distance. The fact that the velocity vector in space travel may be varying continuously does not alter the basic concept of dead reckoning.

Celestial navigation: Celestial navigation is usually defined as navigation utilizing celestial bodies. It is similar to piloting in that position is determined relative to points on the surface of the earth. But in celestial navigation these points, each vertically below a celestial body, are constantly in motion as the earth rotates. Thus, real time is involved even if the observer is stationary.

In space travel the use of any celestial body for navigation is usually considered celestial navigation. However, when celestial bodies are used as lighthouses and position is determined relative to them, the technique might more properly be considered a form of piloting.

Navigation Techniques

Escape Phase

Whether the escape phase starts at the launching pad, at a mobile launching site, from an orbit around a planet, or from a natural satellite, the position at the start of a space flight is known accurately. During the initial stage of such a flight, the desired trajectory is determined in advance. Position may be determined from measurements made either from the planet or aboard the space craft, but any detected deviations from the precomputed flight path can be corrected. When the desired velocity vector has been achieved, power is cut off and the space craft is on its way. The accuracy with which navigation is performed during this phase of the flight establishes, in large measure, the requirement for corrections later. For instance, it has been determined that, for a flight from earth to Mars, no subsequent correction will be needed to pass within 10,000 miles of Mars if the cutoff speed is accurate to 16 ft/sec and initial direction of travel is correct to 0°.01.

In the Vicinity of a Celestial Body

Here navigation consists primarily of determining position relative to that body or some part of it. In its usual form it consists of determining the distance from the body and the point on the surface of the body over which the craft is located. Position may be determined either by tracking from the surface of the celestial body, or aboard the space craft. Power is applied as needed to produce desired changes in the orbit.

Midcourse Phase

Normally, the time and place of rendezvous with the destination celestial body is determined in advance. With present chemical rockets of relatively low specific impuse, the planned trajectory is invariably a minimum power ellipse. In flight the position of the craft may be monitored by determination of position from time to time and comparison of this with the scheduled position. This is similar to commonly used techniques on the earth with one significant difference. As long as a terrestrial craft remains on the desired track and progresses toward the destination, it will arrive in due time. If the craft is ahead or behind schedule, the only effect is a revision of the estimated time of arrival (ETA). The original ETA can be maintained if the speed is decreased or increased as needed. In space navigation time is important in another way because the destination is not a fixed point and will be at the rendezvous point only for an instant. A change of speed of the space craft will change its orbit and may increase the discrepancy.

When discrepancies between the precalculated [dead-reckoning (DR)] position and the actual position are established, at least three possible correction techniques are available. The first is to return to the original schedule. To minimize the amount of fuel required, the return might be scheduled over a period of time. Several weeks is not unreasonable during an early portion of a flight of several months, but, if the period of time is too long, the delay in determining the effectiveness of the maneuver might result in a still larger correction later. If small continuous thrust were available, correction of DR positions might be essentially a continuous process. This technique requires a minimum of computation, but is somewhat wasteful of fuel. Fuel expenditure might be minimized by planned undercorrection to avoid possible overcorrection due to errors in position determination.

A second technique of correcting a discrepancy is to increase the correction period until it occupies the entire time remaining until rendezvous. To do this accurately involves the determination of a new orbit from the present position to the originally planned time and position of rendezvous. This is less wasteful of fuel but imposes a more severe computation problem. The thrust vector needed to change the present orbit to the new one requires computation of the present orbit. This may be determined, in increasing order of difficulty, by 1) present heliocentric position and present velocity vector, 2) three positions of the craft, or 3) two positions of the craft and the elapsed time between them. Although the first involves the simplest computation, the determination of present velocity vector can probably not be determined with adequate accuracy.

A third technique of correcting a discrepancy is to compute the new orbit to a new rendezvous point so as to involve minimum fuel requirements. If this computation proves to be beyond the capability of the computer aboard the space craft, solution might be obtained by a trial-and-error process, selecting successive rendezvous points (and corresponding times), until a minimum fuel expenditure is found to change the orbit from the present one.

A completely different technique for the midcourse phase involves the homing principle. Instead of following a precomputed ellipse from the origin to the destination, the craft travels first to some identifiable line through the destination, and then remains on this line until it reaches the destination.

Of various possible homing techniques, one that appears attractive for an interplanetary flight involves the radial line from the sun through the destination planet. The craft first follows an elliptical orbit until it reaches some plane through the radial sun-planet line. For minimum fuel consumption, the space craft should arrive in this plane with the same angular velocity as the destination planet, and with just enough radial velocity to reach its orbit. A properly directed thrust within the plane then moves the craft to the radial line between the sun and destination planet. The craft then applies thrust to eliminate motion across the radial homing line. Short periodic bursts of a chemical rocket may be adequate to keep the space craft approximately on the radial homing line, but the method will become much more attractive when low-thrust, high specific-impulse power plants, such as ion or plasma-jet propulsion, become available. It has been calculated that a continuous thrust of as low as 6×10^{-5} g is adequate to utilize this technique on a flight from earth to Mars. If the plane through the radial sun-destination planet line is also through the origin of the flight, one step might be eliminated.

One advantage of the homing technique is the simplicity of navigation involved. Once on the homing line, the craft need determine only that it is on the required line and moving in the right direction along this line. Position on the line might be determined by any device capable of accurately measuring angles near 0° or 180°. Crude successive measurements of distance or position would be adequate to establish the direction of motion if any reasonable doubt existed. Such position or distance measurement would be useful for establishing ETA at destination, but time would not be significant in terms of making or missing rendezvous.

Another advantage of this technique is that the destination planet might be approached at any desired speed, resulting in as much time as needed for making decisions and with relatively simple navigation during the terminal phase. In fact, along the radial approach line there exists a point where the centrifugal force and the gravitational attraction of the sun and destination planet exactly balance. A space craft could remain at this point indefinitely, with very small thrusts from time to time to correct for slight drift from this position. In approaching Mars from the earth's orbit, this point is some 600,000 nautical miles from Mars. This might well be used as a staging point for reconnaissance flights closer to the planet.

Terminal Phase

As the destination is approached, the leisurely tempo of the midcourse phase changes. The coordinate system is shifted from heliocentric to one centered on the destination object. Higher accuracy is needed, and this at a time when the astronaut may be subjected to relatively high g forces after having been weightless, or essentially so, for a long period. If the approach is from an elliptical path, computation of the space craft's orbit is made to determine the closest point of approach. As the path of the craft changes from a heliocentric ellipse to a hyperbola around the destination planet, the terminal phase is entered. Unless the closest point of approach is dangerously close, no thrust is applied before this point is reached, but the axis of the retro rockets is aligned with the velocity vector at the closest point. At this point sufficient thrust is applied to bring the space craft to the desired distance at the opposite side of the planet, and at this point additional thrust is applied as needed to produce the desired orbit. From this orbit a landing is made, if desired, by the technique selected in advance.

An alternative technique is to correct the approach orbit at some distance from the destination planet so as to arrive at the closest point of approach at the desired distance, and then apply the required thrust at this point to place the space craft in the desired orbit.

Measurements in Space

Quantities Measured

Space navigation involves determination of speed, direction, distance, acceleration, angles, and time.

Speed: Speed is a basic element of dead reckoning. In space navigation a knowledge of the velocity vector and position of the craft provides the information needed for a relatively simple determination of the orbit of the craft. Unfortunately, the determination of speed at any time during a space flight is not easily accomplished. Average speed between two established positions can be computed by means of the distance and elapsed times between the positions.

Rate of change of range can be used to compute speeds. In the vicinity of a celestial body, rate of change of range might be established by means of radar, providing radial speed determination with an error of the order of 1 per cent. Relatively close to a celestial body, the Doppler effect of a reflected CW signal can also be used to determine radial speed to perhaps 0.5 per cent, but with error introduced by unevenness in the surface of the celestial body and pointing error of the antenna if the beamwidth is less than the angle subtended by the body.

During periods of acceleration, either positive or negative, the speed increment can be determined by integration of acceleration.

Direction: Direction of travel, together with speed, constitutes the velocity vector which, with time, forms the basis of dead reckoning. Celestial bodies are used as references to indicate directions. Simple angle measurement related to such bodies can be used to determine the orientation of the space craft, but no technique is available for measuring directly the direction of motion of the craft to navigational accuracy. Measurement of the Doppler effect of radiation, such as visible light, from stars in different directions, or the observation of starlight, might provide a rough approximation. The most satisfactory method is probably the determination of direction from successive positions of the craft.

Distance: Distance traveled is a product of speed and elapsed time. Without knowledge of speed, the astronaut can resort to determination of distance by means of successive positions of his craft and elapsed time. When the space craft is moving toward or away from a celestial body, distance traveled relative to that body can be determined by successive measurement of distance from the body and elapsed time.

Distance from an object, particularly the celestial body forming the origin of the coordinate system in use, is a useful quantity in determining position. In the vicinity of a celestial body, distance from that body can be determined by radar, either primary or secondary, mounted either in the craft or on the celestial body. If mounted in the space craft, power requirements might be prohibitive, but could be reduced somewhat by use of relatively long pulse length. Slant range of a stable oscillator, either aboard the space craft or on the surface of the celestial body, can be determined by means of the frequency shift due to the Doppler effect, if the space craft is orbiting the celestial body within a few hundred miles of its surface. From a known point on the surface of a celestial body of known radius, the position of a space craft relative to the background of stars can be used with elevation angle to determine distance. Several indirect methods based upon observations of bodies of the solar system are discussed later, under position determination. A method useful over a wider range of distances is to measure the apparent angular diameter of a celestial body of known linear diameter. If the apparent diameter can be determined to an accuracy of one second of arc, the uncertainty of range from the sun at the distance of the earth will be about 42,000 nautical miles. At 100,000 miles from the earth, the error for the same accuracy of measurement is about 7 miles. An error as small as $0''.2$ may be possible in the measurement of diameter by optical means, and even smaller errors might be attained by the use of interferometer techniques. However, some error would be introduced by uncertainty or irregularities in the edge of the celestial body measured, and for accurate results some care might be needed in measuring the correct diameter in the case of a highly oblate body. A large error might be introduced by optical measurement of the apparent diameter of a body illuminated by reflected light if any part of surface toward the observer were in darkness, but this error might be reduced by use of infrared instead of visible light. The error by this technique increases with distance from the body.

Acceleration: Acceleration measurement during periods of thrust application is needed to establish the cutoff time. It is best determined by direct measurement, using suitable accelerometers. This is the primary function of an inertial guidance system during space flight.

Angle: Angle measurement is needed for determination of postion and for determination of the orientation of the space craft.

If the arc to be measured is small enough to be entirely within the field of view of the optical instrument, it might be scaled directly, using the reticle of the instrument. Another method would be to expose a photographic plate, as is frequently done at optical observatories on the surface of the earth, and use an optical micrometer to measure the arc on the developed plate. This would have the advantage of providing a permanent record of the measurement.

A somewhat different technique is required for measurement of angles between widely separated bodies. This might be done by pointing one telescope or star tracker at each body and measuring the angle between the axes of the two instruments. A single instrument could be used if some means were provided for marking the first line of sight. The space craft itself, if not rotating, might provide a sufficiently stable reference. A gyroscope is another possible means of marking the first direction while the optical axis is shifted to the second body. Mirrors or prisms might be used to bring the two bodies into optical coincidence, as is done in a sextant. Methods applicable at particular times might be used when one body of the solar system passes in front of or across the face of another body, or a photograph might be used to determine the position of a body relative to its background of stars.

Direct optical measures are perhaps the most accurate, errors as small as $0''.2$ or less probably being practical. Such measurements in space are free from the troublesome effects of atmospheric refraction. However, they are subject to several errors, as follows.

1) Optical center: Observation, to be accurate, must be made on the center of a celestial body. When only part of the surface is illuminated, or if the albedo differs over the exposed surface, an error is introduced. The complete elimination of this error may not be practical,

but the use of infrared instead of visible light might increase the accuracy. A somewhat similar error may be introduced when a double star is observed.

2) Motions of bodies of solar system: An appreciable time is needed for light to travel from the celestial body to the observer, during which the body moves a short distance. This is easily corrected by allowing for the time of travel of the light.

3) Aberration: A celestial body observed perpendicular to the direction of travel appears to be displaced in the direction of motion of the observer. This can amount to an error of several seconds of arc, but if the speed is known to 0.1 nautical mile per second, and the direction of travel is known to within a few degrees, the correction can be computed. One advantage of photographing a body against the background of stars is that no correction need be made for aberration because all bodies in the line of sight are displaced together.

Time: Time is needed for determining position by measurement of relative positions of bodies of the solar systems, for comparing a precomputed dead-reckoning position with a measured fix to determine the extent of any discrepancy, for determining future positions of objects when an orbit to a rendezvous point is to be computed, and for determining speed by means of distance traveled between two positions with a known elapsed time between them.

An ordinary high-grade chronometer should provide time to an accuracy of one part in 10^5. This is adequate for many purposes of space navigation. For greater accuracy, a quartz-crystal clock might provide time with an error as small as one part in 10^{10}, or slightly smaller. For even greater accuracy, of perhaps one order of magnitude or better, an atomic clock might be practical. Fast moving bodies of the solar system, such as the inner natural satellites of some of the planets, provide a means of determining time if the position of the space craft is known. Relative positions of various bodies of the solar system can be used by any of several techniques discussed under Position Determination.

Position Determination

The Situation

Position determination in space is analogous to position determination on the surface of the earth, with two significant differences:

1) *Third dimension:* On earth navigational position is essentially two-dimensional. The third dimension, the radius vector of the earth, is established, and minor variations due to location of the craft above or below this surface are too small to be significant. Because of the need for determining the third dimension in space, three quantities are needed to define a position. Where a single observation on earth provides a line of position, in space it defines a surface. On earth the intersection of two lines of position define a point, and, where lines intersect in more than one point, a third line may be used to eliminate the ambiguity. In space the intersection of two surfaces of position provides a line of position, and a third surface may be needed to resolve an ambiguity. An additional surface with a different origin is generally needed to establish a position in space.

2) *Time:* On earth time is needed for celestial navigation because of the rotation of the earth. It is not needed for piloting, except as it affects the motion of the craft. In space the opposite is generally true. Unless a position is desired relative to a particular place on a celestial body, time is not needed for celestial navigation relative to the celestial body itself. In space piloting, where position is determined by measurements made on bodies of the solar system to determine position within the solar system, time is needed because of the orbital movements of the bodies.

Coordinate Systems

The coordinate system generally considered most suitable for space navigation is similar to the equatorial system on earth, with the addition of the radius vector. Thus, in three-dimensional space a position report consists of distance from the center (sometimes the surface) of a celestial body, some form of "latitude," and some form of "longitude" on the sphere established by the radius vector.

The celestial body used as the origin of such a system varies with the situation, and will probably change during different phases of an interplanetary flight. During escape and for all navigation near a planet and its natural satellites, the planet is usually most suitable, although a natural satellite might be preferable under some circumstances. During the midcourse phase of an interplanetary flight, the sun is the logical origin. During the terminal phase, the destination celestial body becomes the origin. If flights are eventually made beyond the solar system, some other origin might prove useful.

Near the earth the equatorial system of coordinates (latitude and longitude) is used invariably to locate position on the sphere defined by distance from the earth. A similar set of coordinates with respect to some selected reference meridian on another celestial body will prove useful. However, when position relative to a celestial body itself is desired, rather than to its surface features, an external reference direction such as the sun or a star might be selected as an origin for "longitude."

This is generally done when heliocentric coordinates are used. The ecliptic is a suitable reference "equator." The vernal equinon may seem the obvious choice for origin of the "longitude" measurement in this system, but it has two serious disadvantages. No bright visible body defines it, and it moves at the rate of more than 50 seconds of arc per year. The perihelion or aphelion of the earth's orbit is similarly disadvantageous, although the rate of motion is slower. The most satisfactory point is probably a bright celestial body near the ecliptic, with a small component of proper motion parallel to the ecliptic. Spica, with a component of proper motion along the ecliptic of some 0″.02 per year, has been suggested as suitable for this purpose.

Optical Methods

Optical methods provide the most versatile and generally the most reliable techniques for determining position, particularly during the midcourse phase of an interplanetary flight.

On the earth measurement of the angular altitude of a celestial body above the horizon provides a circle of position on the surface of the earth. Two such circles provide a fix, and a third resolves any ambiguity. In space, measurement of the angle, at the observer, between the lines of sight to a body of the solar system and a star provides a conical surface of position with apex at the body of the solar system. Two such cones with apex at the same body provide a radial line of position, and a third removes any reasonable doubt as to which of the two lines of intersection of the two cones constitutes the line of position. If the origin of the coordinate system in use is at the apex of the cones, the line of position establishes the "latitude" and "longitude" of the space craft.

A special case of the usual celestial line of position on earth occurs when the circle of position established by observation of the altitude of a celestial body shrinks to a single point with the celestial body in the zenith. In space the determination of the position of a body of the solar system against the background of stars establishes a line of position radial to that body. If the body is the origin of the coordinate system, "latitude" and "longitude" are established. The position of a body relative to the background of stars can be obtained by photography or by measurement of the arcs from the center of the body to several stars near the line of sight to that body. If the body has an appreciable apparent angular diameter, the mean of measurements to opposite sides establishes the measure to the center. Close to a celestial body an infrared scanning system might be useful in establishing the center of the body.

The position of the space craft on the line of position, established by measurement of angles between stars and a body of the solar system, or by locating the zenith among the stars, can be determined by any of the methods of finding distance from the origin body, discussed earlier, or by measuring the angle between a star and a different body of the solar system, thus establishing another cone with a different origin, intersecting the line of position at two points. If reasonable doubt exists regarding which intersection represents the position of the craft, it can be resolved by another observation. Accurate time would be needed for this technique, unless two observations were made in which case the method might be used to establish time. Radial lines of position from two bodies, however established, fix the position of the craft. Three such position lines provide a means for determining time.

These are celestial navigation techniques. Several piloting techniques are available. Distances from three bodies of the solar system provide three spheres intersecting at the position of the space craft. A fourth sphere would remove any possible ambiguity and provide a measure of time. The angle between two points in a plane defines an arc of a circle with the straight line through the two points constituting a chord of the circle. If the two points are bodies of the solar system, in three-dimensional space, the measurement of the angular separation of the two bodies, as observed at the space craft, defines a surface of position formed by rotating the arc around the chord as an axis of rotation. The intersection of this surface with another formed in any manner constitutes a line of position.

With any of the methods discussed, allowance is needed for proper motion of stars. This is simply a matter of updating the ephemeris. Parallax of some of the nearer stars, too, might introduce a small error. Although the ephemerides provided by the Naval Observatory are sufficiently accurate for terrestrial navigation, some improvement in positions of bodies of the solar system may be needed for accurate space navigation, Also, because the size of the astronomical unit in terms of linear units in use on the earth is known only to an accuracy of some two parts in 10^5, a serious systematic error will be present until this accuracy can be improved. In general, "latitude" and "longitude" can be determined to somewhat greater accuracy than radius, particularly when distance is determined by measurement of the apparent angular diameter of a body.

Electronic Methods

Various electronic methods of position determination might be available in the vicinity of a celestial body. Measurement of distance from a body has been discussed. An electronic system providing space-oriented directional information (perhaps of the VORTAC or consol type) or hyperbolic surfaces of position (such as loran or Decca) is theoretically possible, but does not appear to be practical except very close to the celestial body on which the system is located. The short base line would be a severe limitation. If synchronized transmitters were placed on different celestial bodies, such as the earth and the moon, the rotational and orbital motions of the bodies would introduce serious problems.

Further, any method relying upon electronic or other equipment on the celestial body would severely limit the regions in which such techniques would be available, at least during the early periods of space travel. Any method requiring measurements made on the celestial body and transmitted to the space craft would be dependent upon maintenance of communications with the celestial body.

Thus, electronic systems of the type familiar to earth navigators can be expected to be of limited application, near celestial bodies, particularly the earth. However, because the volume of space travel near the earth will probably be comparatively heavy compared to traffic well beyond the moon, the possibility of utilizing earth-based electronic systems should not be overlooked. A precedent has already been established in the various forms of electronic tracking that have been utilized.

Inertial Navigation

As discussed earlier, inertial navigation can be expected to be used only during periods of thrust, particularly during escape from the earth, orbit corrections in flight, and during approach and landing.

When the approach is to some body other than the earth, knowledge of the gravitational force of the body will be needed. The drift of gyros during a long space flight might limit the effectiveness of an inertial system, but stellar supervision could be provided. Another technique would be to mount the inertial components directly on the space craft and replace the stable platform with a computer. Large gyro torquing rates, instrument fabrication, and computer requirements would be problems to be overcome in such a system.

Physical Phenomena

Physical phenomena such as magnetism, gravity, cosmic radiation, ion streams, or solar pressure offer possibilities for measurement of lines or surfaces of position, direction, speed, or time. These have been investigated, but with the knowledge available none are considered useful at this time. The most promising is the application of the Doppler effect to natural radiation in the visible spectrum for determining the velocity vector of the space craft. Although much development work remains to be done before this technique can be evaluated, the development of the laser offers some promise of success. Somewhat less probable is the measurement of the aberration of light or other radiation from stars to determine speed perpendicular to the direction of observation.

Data Processing

The mass of data to be processed during a space flight constitutes a problem of considerable magnitude. A study of this question indicates that the optimum solution is probably a general-purpose digital computer that can be programmed for the various computations anticipated during the flight. Rapid advances in computer technology indicate that within a few years the size and power requirements might be reduced to values compatible with space-craft limitations.

Conclusions

Although a considerable amount of research and development work is needed to provide adequate, reliable space navigation systems, it appears that no order of magnitude advances over present state of the art are needed to produce a workable first-generation space-navigation system. Such a system could probably be fully automatic, although some functions might be performed manually if this is determined to be desirable. In any event, human monitoring and a manual back-up system appear desirable.

No one technique is best for all situations. Therefore, the technique or combination of techniques selected for any one space flight should be chosen to suit the circumstances.

Section 2

ANTENNAS AND PROPAGATION

Organized with the assistance of the IRE Professional Group on Antennas and Propagation

Early History of the Antennas and Propagation Field Until the End of World War I, Part I—Antennas*

P. S. CARTER†, MEMBER, IRE, AND H. H. BEVERAGE‡, FELLOW, IRE

Summary—The prediction by Maxwell in 1865 that electric oscillations in a circuit produce electric waves in the surrounding space stimulated scientists to devise experiments to detect the presence of these waves. Following the classical experiments of Hertz in 1888, many attempts were made to communicate at a distance by electric waves. Marconi was the first to demonstrate a complete workable system. His success was largely due to his clear understanding that high antennas with top loading were essential to transmitting signals over considerable distances. Since the early transmitters utilized spark gaps directly between the antenna and ground, the wavelength of the energy radiated increased as the dimensions of the antenna increased. This naturally led to the idea that long wavelengths were required for operation over long distances, particularly during daylight as indicated by Marconi.[5] Thus, the usefulness of short waves for long-distance communication during daylight was destined to remain unknown until some 30 years later when Marconi himself pioneered the development of long-distance communication with short waves. The development of various types of long-wave antennas is described to the end of World War I.

IN 1865 James Clark Maxwell, by mathematical reasoning, predicted that electric oscillations in a circuit produce electric waves in surrounding space, and that these waves travel with the velocity of light.

Professor Amos Dolbear of Tufts College, Medford, Mass., was granted a patent in 1882 for a wireless system which consisted of an induction coil with one terminal of the secondary grounded, and the other terminal con-

* Received by the IRE, July 10, 1961; revised manuscript received, October 24, 1961. Manuscript prepared by H. H. Beverage, based on discussion with P. S. Carter prior to his confinement in New York Hospital.

† RCA Laboratories, Rocky Point, N. Y. Deceased, April 20, 1961.

‡ Consultant, RCA Laboratories, Stony Brook, N. Y.

nected to a condenser. A microphone and battery were connected to the primary of the induction coil. A similar arrangement was proposed for reception with a telephone receiver and battery connected between a condenser and ground. While Dolbear's arrangement had some of the elements later used with success by Marconi, there were no electrical oscillations of high frequency. Consequently, any transmission would be due to induction rather than to radiation.[1]

It remained for Heinrich Hertz to prove the existence of electric waves in space as predicted by Maxwell. The first true antenna appears to have been used by Hertz in his classical experiments at Karlsruhe in 1887. His antenna consisted of two flat metallic plates, 40 cm square, each attached to a rod 30 cm long. The two rods were placed in the same straight line, and were provided at their nearer ends with balls separated by a spark gap about 7-mm long. The spark gap was energized by a Ruhmkorff coil. In order to detect the radiated waves, Hertz employed a receiving circuit consisting of a circular loop of wire broken by a microscopic gap. The radius of the loop was 35 cm, which was found by experiment to be the proper size to be in resonance with the oscillator.[2]

After Hertz had succeeded in proving that the action of an electric oscillation spreads out as a wave into space, he planned experiments with the object of concentrating this action and making it perceptible to greater distances by putting the oscillator in the focal plane of a large cylindrical mirror. With this arrangement, he performed his classical experiments with polarization, refraction, and reflection of electromagnetic waves.[1]

Sir Oliver Lodge, on June 1, 1894, delivered before the Royal Institution a lecture entitled, "The Work of Hertz." In this remarkable lecture he described, among other things, the filings coherer with automatic tapper, the connection of the coherer to a grounded conductor, and the method of detecting distant thunderstorms by connecting the coherer to a grounded gas-pipe line. Lodge estimated that the apparatus would respond to signals at a distance of half a mile.[3]

In 1895 Professor Popoff of Kronstadt set up apparatus designed for the study of atmospheric electricity. He used the metallic filings coherer developed by Professor Branly of Paris in 1890. One terminal of the coherer was connected to a metallic rod extending above the housetop, and the other terminal was connected to earth, so that electric currents produced by atmospheric electricity were conducted to earth through the coherer. He included a bell which sounded when an atmospheric discharge was received. The bell was so placed that its hammer, while in vibration, also struck the coherer, causing it to decohere. He also included a telegraphic registering apparatus in shunt to the bell, so as to obtain a written record of the duration of each electric disturbance. In a note dated December, 1895, Popoff said, "I entertain the hope that when my apparatus is perfected it will be applicable to the transmission of signals to a distance by means of rapid electric vibrations—when in fact a sufficiently powerful generator of these vibrations is discovered."[3,4]

The practical application to wireless telegraphy of the principles discovered by Maxwell, Hertz and others awaited the genius, skill, and forceful initiative of Signor Guglielmo Marconi. In 1896 Marconi filed a provisional specification, and in 1897 complete specifications for a system of wireless telegraphy, including a spark transmitter and a receiver utilizing an improved coherer. One terminal of his transmitter and corresponding receiver was connected to an elevated wire, and the other terminal was connected to earth. Marconi's success was to a great extent based on his understanding of the importance of the elevated antenna. In a lecture before the Royal Institute of Great Britain in 1908,[5] Marconi stated:

> My early tests on wireless transmission by means of the elevated capacity method had convinced me that when endeavoring to extend the distance of communication it was of little utility merely to increase the power of the electrical energy applied to the transmitting circuits, but that it was necessary to increase the area or height of the transmitting and receiving elevated conductors. As it was economically impractical to use vertical wires of very great height, the alternative was to increase their size or capacity, which, in view of the facts I had first noted in 1895 seemed likely to make possible the efficient utilization of electrical energy.

Obviously, as the length and capacity of the elevated antenna was increased, the wavelength was also increased. This led to the idea that, for long-distance operation, long waves (low frequencies) should be used, particularly during daylight. Marconi stated in his lecture, "Apparently the amplitude of the electrical oscillations and the length of the waves radiated have much to do with the interesting phenomena, small amplitude and long waves being subject to the effect of daylight to a less degree than large amplitudes and short waves."[5] Thus the usefulness of short waves for long-distance communication was destined to be unknown until some 30 years later, when Marconi himself pioneered in the development of long distance daylight communication with short waves.[5,6]

In 1901 Marconi erected a large cone-type antenna at Poldhu in Cornwall for an attempt at signaling across the Atlantic. This antenna was wrecked by a storm in September, 1901. The transmitting antenna actually used at Poldhu during the experiments in transmitting to Newfoundland consisted of 50 vertical copper wires supported at the top by a horizontal wire stretched be-

[1] G. W. Pierce, "Principles of Wireless Telegraphy," McGraw-Hill Book Co., Inc., New York, N. Y., p. 77 and chs. 9 and 12; 1910.

[2] *Ibid.*, p. 44.

[3] R. A. Fessenden, "Wireless telephony," *Trans. AIEE*, vol. 27, no. 1, pp. 558–629; June, 1908.

[4] Pierce, *op. cit.*, p. 82.

[5] G. Marconi, "Transatlantic Wireless Telegraphy," presented before the Royal Inst. of Great Britain; March 13, 1908.

[6] E. H. Armstrong, "The spirit of discovery—an appreciation of the work of Marconi," *Elec. Engrg.*, vol. 72, pp. 670–676; August, 1953.

tween two masts 157 ft (48 m) high and about 200 ft (61 m) apart. These wires converged together at the lower end in the shape of a large fan. The wavelength was 1200 ft (366 m) and the actual power for the production of the waves was about 15 kw. In Newfoundland Marconi used an antenna supported by kites. He stated:

It was soon discovered that an ordinary syntonic receiver was not suitable, although at one time a number of doubtful signals were recorded. I therefore tried various microphonic self-restoring coherers placed either directly in the aerial or included in the secondary circuit of the oscillation transformer, the signals being read by telephone.

On December 12, 1901, the signals transmitted from Cornwall were clearly received at the prearranged times. In many cases a succession of S's being heard distinctly, although in consequence of the weakness of the signals and the constant variations in the height of the receiving aerial, no actual message could be deciphered. The following day, we were able to confirm the result. The signals were actually read by myself and by my assistant, Mr. G. S. Kemp.[5]

In a note in the *Electrical World*, December 21, 1901, it was stated that the letter S was sent from 3 to 6 P.M. GMT or 11:30 A.M. to 2:30 P.M., St. John's time. Considering the time of day, the stated wavelength of 1200 ft (366 m) and the insensitivity of the coherer, this was a remarkable achievement. In 1901 means for measuring the wavelength of radiation were very crude. It is quite possible that the wavelength actually used at Poldhu was considerably longer than 1200 ft.

The same Poldhu transmitter was used in February, 1902, to transmit to the SS *Philadelphia* which had the receiving antenna fixed to the main mast, the top of which was 200 ft (61 m) above sea level. Good results were obtained on a syntonic receiver, and the signals were all recorded on tape by the ordinary Morse recorder. Readable messages were received from Poldhu to a distance of 1551 miles, S's and other test letters as far as 2099 miles at night.[5,6]

In 1902 antennas were constructed at Poldhu and Cape Cod, Mass., consisting of four wooden lattice towers, each 210 ft (64 m) high erected at the corners of a square of 200 ft (61 m) on a side. The towers carried insulated triatic stays from which a conical arrangement of 400 copper wires was suspended. The wavelength used was 3600 ft (1100 m). The antenna at Poldhu was shortly afterwards extended by the addition of wires sloping downwards, umbrella fashion. This arrangement was a considerable improvement. By the adoption of much longer waves than had been hitherto employed, namely a wavelength of 14,000 ft (4270 m or 70 kc), it was possible to telegraph over a distance of 550 miles with an expenditure of about 1 kw.[5]

The first commercial service across the Atlantic was established in October, 1907, between Clifden, Ireland, and Glace Bay, Newfoundland. The antennas consisted of a flat top with multiple wires fed at one end. The wavelength was 12,000 ft (3650 m).[5]

By 1907 the advantage of using top-loaded antennas was widely recognized. In 1910 the Telefunken Company constructed an antenna at Nauen, Germany, in the form of an umbrella. The central tower was 100 m high, supported by a carefully insulated ball at the base. The tower served partly to support the entire antenna and partly as a current carrier in conjunction with a bundle of wires with which it was connected.

In 1911 this tower was increased to a height of 200 m, but in April, 1912, during a severe storm, this tower collapsed. Shortly thereafter, the construction of an entirely new antenna and tower was undertaken. Another umbrella-type antenna was erected by C. Lorenz at Eberswalde, Germany.[7]

In the United States an umbrella-type antenna was erected at Brant Rock, Mass., by the National Electric Signaling Company (tower height—130 m).[7] During 1913 and 1914, several long-wave stations were erected in the United States for overseas radio communication. The Telefunken Company of Germany erected a station at Sayville, N. Y., with an antenna of the umbrella type supported by a central insulated tower about 500 ft (152 m) in height. A counterpoise consisting of 56 wires 16 ft (5 m) above ground was provided.[8] Another antenna of the umbrella type was erected by German engineers at Tuckerton, N. J., utilizing a central tower, insulated at the base, having a height of 825 ft (250 m).[9] The American Marconi Company erected stations at New Brunswick, N. J., Marion, Mass., Bolinas, Calif., and Kahuku, Hawaii. The Marconi transmitting antennas were of the so-called Marconi-bent type, consisting of a flat top with the horizontal portion containing a number of wires 5000 ft (1530 m) long supported by triatics 600 ft (183 m) long suspended between adjacent masts. Thirteen tubular guyed masts 400 ft (122 m) high were used.[10]

By 1914 the U. S. Navy had over 50 shore stations and 250 ship stations in operation.[11] The first high-power station to be put into operation by the Navy was the famous "NAA" at Arlington, Va., in 1916. Other high-power stations for the Navy were in the process of construction at several other points, including San Diego, Calif., Pearl Harbor, Hawaii, Cavite, P. I., and Guam. Most of the Navy high-power stations used antennas suspended between three self-supporting towers 600 ft (183 m) high spaced in the form of a triangle 1100 ft (336 m) along the base, each of the other sides being 1000 ft (304 m) between towers.[12]

Prior to the start of World War I, the techniques were available for establishing transoceanic radio communication over distances of several thousand miles. Several of the stations described above were on the air making tests, some with spark transmitters and others with continuous waves utilizing Poulson arcs or alternators.

[7] J. Zenneck, "Wireless Telegraphy," McGraw-Hill Book Co., Inc., New York, N. Y., p. 152; 1915.

[8] A. E. Seelig and F. Van Der Woude, "The Sayville station of the Atlantic Communication Company," PROC. IRE, vol. 1, pt. 3, pp. 23–37; July, 1913.

[9] E. E. Mayer, "The Goldschmidt system of radio telegraphy," PROC. IRE, vol. 2, pp. 69–108; March, 1914.

[10] E. F. W. Alexanderson, "Transoceanic radio communication," PROC. IRE, vol. 8, pp. 263–285; February, 1920.

[11] W. H. G. Bullard, "The naval radio service," PROC. IRE, vol. 3, pp. 7–28; 1915.

[12] W. G. H. Bullard, "Arlington radio station," PROC. IRE, vol. 4, pp. 421–447; 1916.

The start of the war in August, 1914, soon interrupted all attempts at establishing an international commercial radio service. During the war years, the development of high-power radio facilities continued, primarily for military purposes.

Also during the war, in addition to the stations mentioned above, the U. S. Navy constructed high-power stations at Darien (Panama Canal Zone),[13] Annapolis, Md., and elsewhere. The Navy also installed arc transmitters at the former German stations located at Sayville, N. Y.,[8] and Tuckerton, N. J.[9]

In cooperation with the U. S. Navy the General Electric Company installed Alexanderson alternators at the Marconi station located near New Brunswick, N. J. An important feature of the New Brunswick station was the multiple tuning applied to the Marconi antenna, which increased the efficiency from 1.85 per cent to about 14 per cent. This development is described elsewhere in this issue.[14]

In Europe the power and efficiency of the station at Nauen, Germany, was considerably increased. The principal long-wave stations which were in operation before the end of the war were as follows:

Germany—Nauen (POZ) and Hanover (OUI) using alternators.

England—Carnarvon (MUU) using Marconi timed-spark transmitters.

Norway—Stavanger (LCM) using Marconi timed-spark transmitters.

France—Lyon (YN) and Lafayette (LY) using arc transmitters.

After the war, several of the European stations mentioned above were used for commercial service, mostly to the United States.

[13] R. S. Crenshaw, "The Darien radio station of the U. S. Navy (Panama Canal Zone)," Proc. IRE, vol. 4, pp. 35–40; 1916.

[14] H. H. Beverage, "Antennas and transmission lines," this issue, p. 879.

The History of Radio Wave Propagation Up to the End of World War I*

CHARLES R. BURROWS†, FELLOW, IRE

Summary—**Hertz in the 1880's demonstrated electromagnetic wave propagation predicted by Maxwell from his equations in 1864. Heaviside and Kennelly postulated the ionosphere to explain Marconi's historical transatlantic reception of radio waves in 1901. Austin derived the first formula for radio propagation in 1911 from experimental data in the kilometer wavelength range taken in the daytime.**

Much theoretical effort was expended on the effect of the electrical properties of the ground but the problem was not resolved until a later date. Watson, however, cleared up the problem of diffraction around a perfectly conducting sphere in 1919.

Up to the end of World War I, it was generally believed that radio transmission improved with increase in wavelength so the experimental data is concentrated in this region.

THE HISTORY of radio-wave propagation began in 1864, when Clerk Maxwell [1] introduced the concept of dielectric current to satisfy the equation of continuity. This added symmetry to the equations that now bear his name. From these "Maxwell Equations," he deduced the existence of electromagnetic waves that would be propagated with the velocity of light.

In the 1880's, Heinrich Hertz [2], experimenting with the spark discharged from a Leyden jar, demonstrated that the effect was propagated at a distance so as to cause a spark to occur across the terminals of a receiving antenna. In these experiments he produced the electromagnetic waves predicted by Maxwell. In describing his experiments, Hertz used the concepts of rectilinear propagation, polarization, reflection and refraction.

On the theoretical side, Hertz [3] calculated the field from a rectilinear current element, and Abraham [4] calculated the field from a half-wave dipole. Hack [5] plotted the fields around rectilinear antennas half wavelength, full wavelength and three-half wavelengths long, based on the formulas of Abraham [4]. Rudenberg [6] calculated the radiation resistance of the dipole which made possible the expression of attenuation as the ratio of the power radiated from one antenna to the maximum power that could be extracted from the receiving antenna [7].

Blondel [8] pointed out in 1898 that the effect of the earth on a vertical antenna was to produce an image of the antenna, an approximation good for long waves. Hence, for moderate distances the field was found to be inversely proportional to the distance. Experimental evidence for this was obtained by Tissot [9] in 1906.

In the 1890's, Heaviside interpreted the propagation of electromagnetic energy along a pair of wires as a guiding of the wave energy, a concept also used by Hertz during this period. Hertz discovered that by

* Received by the IRE, July 5, 1961; revised manuscript received, October 23, 1961.

† Datronics Engineers, Inc., Bethesda, Md.

forming the inducing circuits, so as to become what is now known as a Hertzian dipole, the radiation was greatly increased. In January, 1897, Marconi [10] was able to greatly extend the range of radio wave propagation by employing a vertical grounded antenna, hence demonstrating the superiority of vertical polarization over horizontal polarization for propagation up to several hundred miles on the longer waves.

The successful reception of radio signals across the Atlantic by Marconi [11] on December 12, 1901, introduced much speculation as to the mechanism of radiowave propagation around the curvature of the earth. Three mechanisms were proposed and seriously considered for several decades: 1) ionospheric propagation, 2) guided propagation by means of a surface wave and 3) diffraction into the earth's shadow [12]. To explain the first mechanism, Kennelly [13] in America and Heaviside [14] in England independently postulated [15] the existence of an ionized region in the upper atmosphere that would reflect the radio waves back to earth. Eccles [16] showed mathematically how this ionized region could bend the waves back to earth. In view of our present knowledge, this was the mechanism responsible for Marconi's transatlantic reception.

Zenneck [17] gave credit to Blondel [8] and Lecher [18] for the concept that radio waves were of the nature of surface waves. In 1907 Zenneck [19] showed that the interface between two dielectrics, such as earth and air, could support a surface wave. In 1909 Sommerfeld [20] studied the problem of the propagation of electromagnetic waves from a Hertzian dipole in the interface separating media of two different dielectric properties. In his formulation he had a component which corresponded to a cylindrical surface wave which at great distances over dielectric ground predominated and was the same as Zenneck's surface wave. From this he concluded that radio waves would be guided around the curvature of the earth. He believed that the effect of the imperfect conductivity of the earth (which introduced cylindrical divergence described as a surface wave) would probably be intensified by the curvature of the surface and hence might counteract attenuation of the waves by the curvature. Ten years later Weyl [21] investigated the same problem, but his answer did not explicitly include the surface wave component of Sommerfeld. He, however, believed that quantitatively his results were the same as Sommerfeld's. Many years later Burrows [22] discovered that these two formulations, in fact, differed by exactly the surface wave of Sommerfeld, and, in collaboration with his associates, L. E. Hunt and A. Decino, conducted the crucial experiment which showed that Weyl's formulation was correct and that the surface wave of Sommerfeld did not exist. Later Rice [23] and Niessen [24] independently found the source of the error in Sommerfeld's work—the incorrect choice of the square root of a complex quantity in an intricate mathematical derivation.

The problem of the diffraction of radio waves around the earth's curvature occupied the efforts of several of the leading mathematicians for many years before Watson [25], in 1919, pointed out the errors of the earlier investigators and presented values for the diffraction of radio waves around a perfectly conducting sphere. Much later Van der Pol and Bremmer [26] generalized Watson's solution to take into consideration arbitrary ground constants, and Gray [27] extended the work to include horizontally polarized antennas.

In 1902 Marconi observed that the same transmitter had a longer range at night, sometimes having a range two and a half times the daytime range. This was the first observation of atmospheric absorption by the ionosphere. Austin's [28] measurements confirmed this and showed that the nighttime propagation was irregular. Marconi [29] observed that when part of the propagation path was in darkness and part in daylight, the signal was weaker than the steady daytime signal. This was followed by stronger signals when it was sunset over the entire path and preceded by stronger signals just before sunrise at the eastern terminal. When it was nighttime over the entire path, the signals were stronger and variable. Signals were stronger in the daytime on 7000 meters than on 4000 meters, but the sunrise and sunset maximums were stronger on the shorter wavelength.

Austin [28] gave us our first formula for radio wave propagation [30]. It was an empirical formula representing the summary of experimental data for the propagation on wavelengths in the kilometer region over the sea during the daytime. It is

$$I_r = 4.25 \frac{I_s h_1 h_2}{\lambda d} \exp\left(\frac{-0.015d}{\sqrt{\lambda}}\right)$$

where the currents are in amperes and lengths in kilometers. I_r is the current in the vertical antenna of height h_2 resulting from a current I_s in an antenna of height h_1, separated by a distance d on a wavelength λ. Hogan's [31] experiments confirmed Austin's formula and showed that it was valid over a greater range of distances.

Austin's formula corresponds to the usual formula

$$E = \frac{120\pi h_1 I_s}{\lambda d}$$

for the field propagated along a perfectly conducting plane [7] multiplied by an exponential absorption factor. Austin's formula showed that the exponential attenuation factor was proportional to the distance and inversely proportional to the square root of the wavelength. In Watson's [25] theoretical formulation for the diffraction of radio waves around the curvature of the earth, the exponential attenuation factor is inversely proportional to the cube root of the wavelength. In a later paper [32], he considered the effect of the ionosphere, which then changed the dependence on the wavelength from the cube root to the square root.

On the longer waves (in the kilometer range) the propagation was studied as a function of wavelength, ground conditions, time of day and season. The strength of signals on a wavelength of 1000 meters at a distance of some 200 km was found to be maximum in winter and minimum in summer with a variation of a factor of about six [33]. Daytime transmission over land on 3500 meters was more attenuated than over sea and depended on weather and time of year [34]. Shorter waves (2500 meters) were more attenuated than longer waves (3500 meters). Taylor [34] attributed the increase in attenuation in spring and summer to poorer conductivity which he attributed to the breaking up of the ground by frost and cultivation. Kennelly [35] discussed the influence of solar radiation on radio propagation. The dip both at midsunrise and at midsunset with reinforcement both before and after each dip were explained by reflection effects at the shadow wall.

The first report of selective fading was made by Dr. de Forest on the observations of the federal telegraph operators on their transmissions between Los Angeles and San Francisco, a distance of 560 km on long waves [36]. They observed that the mark wave would fade on 9.7 kc while the space wave some 5 kc lower in frequency remained normal.

In 1915 the AT & T and Western Electric Companies, in cooperation with the United States Navy, demonstrated voice radio communication between Arlington, Virginia, and Hawaii and Paris. On November 29, 1915, signals from Nauen, Prussia, were copied in Honolulu, a distance of 9000 miles [37].

Many high-power radio transmitters were put in operation during the war. Little is available in regard to the radio-wave propagation resulting from them, except that transmissions were successful.

References

[1] J. C. Maxwell, "A dynamical theory of the electromagnetic field," *Proc. Roy. Soc. (London)*, vol. 13, pp. 531–536; December 8, 1864.

[2] H. Hertz, "Ueber sehr schnelle elektrische Schwingungen," *Ann. Physik und Chemie* (*Wiedeman*) *N. F.*, vol. 31, pp. 421–448; May 15, 1887.

[3] H. Hertz, "Die Kräft electrischer Schwingungen, behandelt nach der Maxwell'schen Theorie," *Ann. Physik und Chemie* (*Wiedeman*) *N. F.*, vol. 36, pp. 1–22; December 15, 1888.

[4] M. Abraham, "Die electrischen Schwingungen um einen stabformigen Leiter, behandelt nach der Maxwell'schen Theorie," *Ann. Physik und Chemie* (*Wiedeman*) *N. F.*, vol. 66, pp. 435–472; October 20, 1898.

[5] F. Hack, "Das elektromagnetische Feld in der Umgebung eines linearen Oszillators," *Ann. Physik.*, ser. 4, vol. 14, pp. 539–546; July 12, 1904.

[6] R. Rüdenberg, "Der Empfang elektrischer Wellen in der drahtlosen Teltgraphie," *Ann. Physik*, ser. 4, vol. 25, pp. 446–466; March 3, 1908.

[7] C. R. Burrows, "Radio Propagation over plane earth—field strength curves," *Bell Sys. Tech. J.*, vol. 16, pp. 45–75 and 547–577; January and October, 1937. See p. 48.

[8] A. Blondel, Compt. rend. Assoc. franc. Avancement des sciences, Congres de Nantes, France, p. 212; 1898.

[9] C. Tissot, "Note on the use of the bolometer as a detector of electric waves," *Electrician*, vol. 56, pp. 848–849; March 9, 1906.

[10] E. H. Armstrong, "The spirit of discovery," *Electronic Engr.*, vol. 72, pp. 670–676; August, 1953.

[11] G. Marconi, "Signals across the atlantic," *Elec. World*, vol. 38, pp. 1023–1025; December 21, 1901.
—— "Notes," *Electrician*, vol. 48, p. 329; December 20, 1901.

[12] C. R. Burrows, "Mechanisms of Propagation," presented at Atti del Congresso Internazionale per il Cinquantenario della scoperta Marconiana della Radio, Roma, Italy, September 28–October 5, 1947. Published by Consiglio Nazionale della Ricerche, Rome, 1948, pp. 43–51; 1948. (In English.)

[13] A. E. Kennelly, "On the elevation of the electrically-conducting strata of the earth's atmosphere," *Elec. World and Engr.*, vol. 39, p. 473; March 15, 1902.

[14] O. Heaviside, Encyclopedia Britannica, 10th ed.; 1902.

[15] This ionized region had been previously postulated by Balfour Stewart to explain the variations in terrestrial magnetism. See B. Stewart, "Terrestrial Magnetism," *Encyclopedia Britannica*, 9th ed., vol. 16, pp. 159–184; 1878.

[16] W. H. Eccles, "On the diurnal variations of the electric waves occurring in nature, and on the propagation of electric waves round the bend of the earth," *Proc. Roy. Soc.* (*London*) *A*, vol. 87, pp. 79–99; August 13, 1912.

[17] J. Zenneck, "Wireless Telegraphy," McGraw-Hill Book Co., New York, N. Y.; 1908. Translated by A. E. Seelig, 1915.

[18] E. Lecher, "Ueber drahtlosen Telegraphie," *Physik. Z.*, vol. 3, p. 274; April, 1902.

[19] J. Zenneck, "Ueber die Fortpflanzung ebener elektromagnetischer Wellen längs einer ebenen Leiterfläche und ihre Beziehung zur drahtlosen Telegraphie," *Ann. Physik*, ser. 4, vol. 23, pp. 846–866; September 24, 1907.

[20] A. Sommerfeld, "Ueber die Ausbreitung der Wellen der drahtlosen Telegraphie," *Ann. Physik*, ser. 4, vol. 28, pp. 665–756; March, 1909.

[21] H. Weyl, "Ausbreitung elektromagnetischer Wellen über einem ebener Leiter," *Ann. Physik.*, ser. 4, vol. 60, pp. 481–500; November 20, 1919.

[22] C. R. Burrows, "The surface wave in radio propagation over plane earth," Proc. IRE, vol. 25, pp. 219–229; February, 1937.

[23] S. O. Rice, unpublished memorandum, privately communicated to the writer.

[24] K. F. Neissen, "Zur Entscheidung zwischen den beiden Sommerfeldschen Formeln für Fortpflanzung von drahtlosen Wellen," *Ann. Physik.*, vol. 29, pp. 585–596; April, 1937.

[25] G. N. Watson, "The diffraction of electric waves by the eatth," *Proc. Roy. Soc.* (*London*) *A*, vol. 95, pp. 83–99; October 7, 1918.

[26] B. Van der Pol and H. Bremmer, "The diffraction of electromagnetic waves from an electrical point source round a finitely conducting sphere, with applications to radiotelegraphy and the theory of the rainbow," *Phil. Mag.*, ser. 7, vol. 25, pp. 141–175, July, 1937; vol. 24, pp. 825–864 (suppl.), November, 1937; also vol. 25, pp. 817–834 (suppl.), June, 1938; and vol. 27, pp. 261–275; March, 1939.

[27] M. C. Gray, "The diffraction and refraction of a horizontally polarised electromagnetic wave over a spherical earth," *Phil. Mag.*, ser. 7, vol. 27, pp. 421–436; April, 1939.

[28] L. W. Austin, "Some quantitative experiments in long-distance radio telegraphy," *J. Res. NBS*, vol. 7, no. 3, pp. 315–366; October, 1911.

[29] G. Marconi, "Radio-Telegraphy," *Electrician*, vol. 67, pp. 532–534; July 14, 1911.

[30] Much effort was spent in calculating the field for propagation over the earth, but failure of subsequent authors to critically compare their results with earlier workers left the situation far from satisfactory. Sommerfeld's [20] equations for good conducting ground (conduction current large compared with dielectric current in the ground) are sensibly the same as Weyl's [21], but this was not known until a later period. The situation for propagation around a spherical earth was very confused until Watson [25] cleared it up. The effect of imperfect conductivity of the ground and the effect of the earth's curvature was finally resolved by C. R. Burrows and M. C. Gray, "The effect of the earth's curvature on ground wave propagation," Proc. IRE, vol. 29, pp. 16–24; January, 1941.

[31] J. L. Hogan, Jr., "Quantitative results of recent radio-telegraph tests between Arlington, Va., and U. S. S. "Salem," *Electrician*, vol. 71, pp. 720–725; August 8, 1913.

[32] G. N. Watson, "The transmission of electric waves around the earth," *Proc. Roy. Soc.* (*London*) *A*, vol. 95, pp. 546–563; July 15, 1919.

[33] L. W. Austin, "Seasonal variations in the strength of radiotelegraphic signals," Proc. IRE vol. 3, pp. 103–105; June, 1915.

[34] A. H. Taylor, "Diurnal and annual variations in overland radio transmission," *Phys. Rev.*, vol. 4, pp. 435–439; November, 1914.

[35] A. E. Kennelly, "The daylight effect in radio telegraphy," Proc. IRE, vol. 1, pt. 3, pp. 39–52; July, 1913.

[36] L. DeForest, "Recent developments in the work of the federal telegraph company," Proc. IRE, vol. 1, pt. 1, pp. 37–51; January, 1913. (See p. 42.)

[37] "A Honolulu station gets war dispatch from Nauen, Prussia," QST, vol. X, p. 23; February, 1916.

Growth of the Antennas and Propagation Field Between World War I and World War II

Part I—Antennas*

L. J. CHU†, FELLOW, IRE

Summary—**The advances in antenna theory concepts and designs between World War I and World War II are summarized. Rapid growth of world-wide and regional broadcasting and point-to-point radio communication stimulated the design of high-power low-frequency transmitting antennas of reasonable efficiency, directive-wave antennas for signal-to-noise improvement in low-frequency reception, rhombic and V-shape wave antennas and arrays of various types for directive short-wave transmission and reception.**

Introduction

The development of antenna techniques and theory during the period between World Wars I and II was motivated by two major events. The first was the rapid growth of radio communication in the form of world-wide and regional broadcasting and point-to-point communication. The second was the development of power-generation techniques toward higher and higher frequencies. In this paper the major antenna advances during this period are outlined. A number of other advances are regretfully omitted. Among these are the engineering aspects of antenna development, the improvment of direction-finding aircraft and marine antennas and the "scattering" and "diffraction" theories, which were not treated as antenna problems during that period and exerted great influence on microwave antenna development during World War II.

Fundamental Theorems

The foundation of antenna theory was solidly laid out before World War I. Maxwell's equations were valid for the purpose of antenna analysis. Hertz-vector and magnetic-vector potentials were used extensively to calculate the radiation from electric currents. Poynting's theorem was applied to determine total radiated power from antennas and their radiation resistances.

Carson [1], [2] and others extended Lorentz's reciprocity theorem for electromagnetic field to antenna terminals. Their contribution to antenna theory was immeasurable. Aside from other applications, it allowed us to specify the receiving characteristics of an antenna in terms of its transmitting characteristics. Before Carson's work, receiving antennas were sometimes analyzed as such by questionable procedures. After Carson's work, the analysis of receiving antennas as distinguished from transmitting antennas practically disappeared from the literature of the subject.

At the turn of the century, A. E. H. Love and H. M. MacDonald formulated the mathematical procedure for replacing a given current distribution by equivalent surface currents over an arbitrary surface, enclosing the given current distribution to obtain identical field outside the surface. This "equivalence theorem" was rediscovered by Schelkunoff [3], and applied to the problems of radiation from open ends of a coaxial line and waveguide. The engineering significance of the equivalence theorem lies not so much in its formal exactness. It provides a self-consistent procedure for calculating the external field from the aperture type of antenna, such as the horn, reflector, and lens, for which a reasonable approximation of electromagnetic field can be made over a closed surface that includes the "aperture."

VLF Transmitting Antennas

To achieve world-wide reliable communication, a number of VLF antenna systems were installed and in operation in the 1920's. Following Alexanderson's design procedures, these antennas consisted of top-loaded vertical down leads connected to huge coils over elaborate ground systems. The top loading serves not only to provide a greater effective height for the down leads, but also allows a large current in the down leads to radiate up to 250 kw at 15 kc without voltage breakdown. The measured radiation efficiency at this frequency for some of the systems is as high as 26 per cent [4].

Wave Antennas

The most significant discovery of the capability of the wave antenna was made by Beverage [5]. With a no. 14 rubber-covered wire, six miles long, laid in the scrub oak and sand of Long Island, he observed a significant increase in signal-to-noise ratio from Europe when the

* Received by the IRE, July 31, 1961; revised manuscript received, November 1, 1961. This work was supported in part by the U. S. Army Signal Corps, the Air Force Office of Scientific Research, and the Office of Naval Research.

† Department of Electrical Engineering, and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass.

receiver termination was at the remote end and ground termination at the other end nearer to Europe. Concurrently, C. W. Rice and E. W. Kellogg concluded, on purely theoretical grounds, that a long horizontal antenna pointed in the direction of the sending station would constitute a unidirectional receiver, provided that the receiver were placed at the end of an antenna remote from the sending station, while the other end was grounded through a noninductive resistance equal to the surge impedance of the antenna. These discoveries were followed by an exhaustive experimental and theoretical study to establish the "Beverage antenna" as an effective and practical directional antenna for low-frequency reception.

The analysis of the Beverage antenna by E. W. Kellogg was an analysis of it as a receiver. As a transmitting antenna, one would find traveling-wave solutions along the antenna. The natural continuation of the development of wave antennas was carried out by RCA, Bell Laboratories, Inc., and the International Telephone and Telegraph Company, in the forms of V and rhombic antennas [6], [7]. The peak of the radiation pattern of a traveling-wave current along a straight conductor lies along a cone that is coaxial with the conductor. The angle of cone is a function of the conductor length measured in wavelength. By arranging the conductor in the form of an inverted V above ground, a rhombic in free space, with proper termination impedance at the end opposite the input terminal to minimize reflection, the cones of maximum radiation intersect with equal phase in a direction along the axis of the antenna, when the angles of the antennas are properly chosen.

While traveling-wave antennas inherently possess a broad-band characteristic both in input impedance and directivity, the other type of wave antenna developed during this period made use of standing waves to achieve the broadside effect. The Marconi-Franklin antenna is a typical example. Quarter-wave shorted transmission lines or lumped high-impedance circuits were inserted in series every half-wavelength along a straight conductor to set up a standing wave along the conductor, with phase reversals at the inserts. This method results in a narrow-band broadside array.

Antenna Arrays

For a decade after World War I, investigations of multiple-element antenna systems were carried out by many workers in the United States, Great Britain, France, Germany and Japan, both experimentally and theoretically. This development took two distinct directions. The first recognized the coupling of field from the excited element to the parasitic elements that are not connected circuitwise to the transmitter or receiver, and the resonant phenomena of a linear conductor. By choosing a proper length for the parasitic elements, the phase of the current induced relative to that of the excited element could be controlled to form a directive beam along the axis of the array [8]. The well-known "Yagi" is one example. The mutual coupling among elements was reduced to a multiterminal circuit problem by England and Crawford [9]. Calculations of self-impedance and mutual impedance were carried out by Carter [10], who started from the field viewpoint. Carter made ingenious use of Carson's reciprocity theorem.

The second significant development of linear-array theory was based upon assumed relative current distributions of a uniformly spaced array of identical elements. In 1928, Bohm [11] published experimental and theoretical polar diagrams of broadside and end-fire arrays. For an array of many elements, it was found that the mutual coupling does not play an important role in determining the relative current distribution of elements that are connected together through a transmission-line system. Southworth [12], before his fundamental work on waveguides, made an exhaustive theoretical study of three-dimensional arrays of uniformly spaced elements of constant current amplitude and constant phase difference between adjacent elements in any one of the orthogonal directions. With R. M. Foster's help, he evaluated the gain of the array and applied it to many special cases of importance.

Not all arrays developed before World War II were uniformly spaced. To conform to the area coverage assigned by the FCC to a broadcasting station, the technique of antenna synthesis or beam shaping was developed after 1930. To obtain specified patterns, the problem was to determine the coordinates of the towers, which were used as radiating elements, and their complex currents and then to design the proper matching and transmission network to achieve the desired relative currents [13].

One of many significant developments during this time was MUSA (multiple unit steerable antenna) by Friis and Feldman [14]. Their work laid the foundation for a new technique of antenna design, which now is known as electronic scanning. This antenna system consists of six identical rhombic antennas uniformly spaced colinearly. Each rhombic is connected to a detector and then to three IF variable phase shifters of RC type in parallel. The outputs of the three sets of phase shifters from the six rhombics are combined to feed three receivers. Each set of phase shifters is mechanically arranged to produce progressive phase shifts to give equal phase at the receiver for signals coming in at a given direction. The system was designed for point-to-point high-frequency communication between New York and London, with a view toward achieving signal-to-noise improvement and eliminating fading resulting from multipath transmission.

A significant contribution was made by Hansen and Woodyard [15] in the design of end-fire arrays, which led to the investigation of super-gain antennas after 1940. By allowing 2.94 radians of additional phase shift along a conventional end-fire array of linear phase variation, they found a decrease in beamwidth and an improvement in gain by a factor of 1.82. The theory was subsequently proved experimentally. A conventional end-fire array has a total phase difference between end elements of $2\pi/\lambda$ times the length of the array.

Linear Antenna Theory

Most of the antennas that were in use at the beginning of World War II were designed with wires and towers. Their designers used quasi-static theory, transmission-line theory, Hertz-vector and experimental data. Many theoretical investigations of impedance and other characteristics of linear antennas had been carried out between 1930 and 1940. The work of Carter [10] and Labus [16] was based upon assumed current distribution and the EMF method. From the assumed current distribution, the component of electric field tangential to conductor surface was computed. The integral of the component of the complex Poynting vector normal to the surface of the conductor was considered to be the complex power output at the input terminals. This method yielded rather accurate results, except when the input terminals were near the current minimum. It was then recognized as an electromagnetic boundary-value problem. Stratton and Chu [18], extending the work of Page and Adams [17], used a prolate spheroidal model for a linear antenna and expanded the wave function into the orthogonal series of spheroidal wave functions. King [19] and Hallen [20] formulated the boundary-value problem of a linear antenna as an integral equation of the current distribution itself. Perhaps, the most illuminating approach to the problem was made by Schelkunoff [21]. He used, initially, a biconical horn model, which supports, among others, a TEM mode of transmission in the space enclosed by the cones, and matched the field to the spherical wave functions outside the biconical horn. It demonstrated that conductors in a linear antenna serve to guide the electromagnetic waves, and expelled the illusion created by the EMF method that the radiated power was emitted from the conductor.

Electromagnetic Horns

The development of waveguides in the middle of the third decade led, naturally, to the development of electromagnetic horns [22]–[24]. By tapering cross-section dimensions of waveguides, one obtained sectoral, pyramidal, and conical horns. The dominant mode of propagation in the waveguide is transformed into the corresponding mode in the horn with cylindrical or spherical wave front. At the aperture, the wave leaves the horn and diffracts to form a directive beam. For a given length of horn, the beamwidth reaches a minimum at a critical flare angle. The introduction of electromagnetic horns paved the way for the engineering use of the aperture type of antenna, which, today, dominates the microwave field.

References

[1] J. R. Carson, "Generalization of the reciprocity theorem," *Bell Sys. Tech. J.*, vol. 3, pp. 393–399; July, 1924.
[2] J. R. Carson, "Reciprocal energy theorem," *Bell. Sys. Tech. J.*, vol. 9, pp. 325–331; month, 1930.
[3] S. A. Schelkunoff, "Some equivalence theorems of electromagnetics and their application to radiation problems," *Bell. Sys. Tech. J.*, vol. 15, pp. 92–112; January, 1936.
[4] S. C. Hooper, "The Lafayette radio station," *Elec. J.*, vol. 18, pp. 112–113; April, 1921.
W. A. Graham, "The multiple tuned antenna," RCA Eng. Dept. Circular No. 425, October, 1924.
E. H. Shaughnessy, "The Rugby radio station of the British Post Office," *J. IEE* (London), vol. 64, pp. 683–713; June 1926.
[5] H. H. Beverage, *et al.*, "The wave antenna," *J. AIEE*, vol. 42, pp. 215–266; March, 1923.
[6] E. Bruce, "Developments in short-wave directive antennas," Proc. IRE, vol. 19, pp. 1406–1433; August, 1931.
[7] P. S. Carter, *et al.*, "Development of directive transmitting antennas by R.C.A. Communications, Inc.," Proc. IRE, vol. 19, pp. 1773–1842; October, 1931.
[8] S. Uda, "Wave Reflector and Wave Director," Tokyo University Tokyo, Japan, Tech. Repts. 8.4, pp. 63–73; 1929.
[9] C. R. England and A. B. Crawford, "The mutual impedance between adjacent antennas," Proc. IRE, vol. 17, pp. 1277–1295; August, 1929.
[10] P. S. Carter, "Circuit relationship in radiating systems and applications to antenna problems," Proc. IRE, vol. 20, pp. 1004–1041; June, 1932.
[11] O. Bohm, "Concentration of energy of short electric waves," *Elek. Nachr.-Tech.*, vol. 5, pp. 413–421; November, 1928.
[12] G. C. Southworth, "Factors affecting gain of directive antennas," Proc. IRE, vol. 18, pp. 1502–1536; September, 1930.
[13] G. H. Brown, "Directional antennas," Proc. IRE, vol. 25, pp. 78–145; January, 1937.
[14] H. T. Friis and C. B. Feldman, "A multiple unit steerable antenna for short wave reception," Proc. IRE, vol. 25, pp. 841–917; July, 1937.
[15] W. W. Hansen and J. R. Woodyard, "New principle in directive antenna design," Proc. IRE, vol. 26, pp. 333–344; March, 1938.
[16] J. Labus, "Mathematical determination of antenna impedance," *Hoch. frequenz und Elektroak.*, vol. 41, p. 17; January, 1933.
[17] L. Page and N. J. Adams, Jr., "Electric oscillations of a prolate spheroid," *Phys. Rev.* vol. 53, pp. 819–831; May, 1938.
[18] J. A. Stratton and L. J. Chu, "Steady state solution of electromagnetic field problems, "*J. Appl. Phys.*, vol. 12, pp. 230–248; March, 1941.
[19] L. V. King, "On the radiation field of a perfectly conducting cylindrical antenna over a perfectly conducting plane earth and the calculation of radiation resistance and reactance," *Trans. Roy. Soc. (London)*, vol. 236, pp. 381–422; November, 1937.
[20] E. Hallen, "Theoretical investigation into the transmitting and receiving qualities of antennas," *Nova Acta Uppsala*, vol. 11, pp. 1–44; 1938.
[21] S. A. Schelkunoff, "Theory of antennas of arbitrary size and shape," Proc. IRE, vol. 29, pp. 493–521; September, 1941.
[22] W. L. Barrow and F. D. Lewis, "The sectoral electromagnetic horn," Proc. IRE, vol. 27, pp. 41–50; January, 1939.
[23] W. L. Barrow and L. J. Chu, "Theory of electromagnetic horn," Proc. IRE, vol. 27, pp. 51–64; January, 1939.
[24] G. C. Southworth, and A. P. King, "Metal horns as directive receivers of ultra-short waves," Proc. IRE, vol. 27, pp. 95–102; February, 1939.

Radio-Wave Propagation Between World Wars I and II*

STEPHEN S. ATTWOOD†, FELLOW, IRE

Summary—**The period between World Wars I and II, 1918 to 1941, is particularly noteworthy for the significant developments that led to a comprehensive understanding of the factors that control the propagation of radio waves through the atmosphere. In slightly over two decades it became possible to identify the basic physical mechanisms with the aid and correlation of theory and experiment, so that by the end of this period it was possible to prepare predictions of future propagation conditions covering a wide variety of frequencies, transmission paths, for day and night, and seasonal variations.**

During the entire period, numbers of scientists were working toward the calculation of field strengths of radio waves propagated by both the ground wave and the ionosphere. Watson's paper in 1918 initiated the ground wave calculations by considering the propagation over a perfectly conducting sphere. It was soon recognized that the ground could not be treated as perfectly conducting, and the theory of propagation for an imperfectly conducting ground was not produced in usable form until the 1937–1941 period. Watson's predictions for field strengths in the shadow zone proved to be much lower than experimental values. This led to increased interest in the ionosphere as a possible explanation.

Effective work on ionospheric problems began in 1924 with the development of the Eccles-Larmor theory of electronic behavior. The effect of the earth's magnetic field extended this theory to include doubly refracting layers and ionospheric layer anistropy. There soon were developed experimental techniques that determined the existence of layers and their electron density variations. The rate of accumulation of data was greatly increased by the development of automatic ionosondes. From the data accumulated came increased understanding of the effects of solar flares, physical processes, diurnal and seasonal variations of propagation.

Tropospheric propagation is discussed briefly since most of the developments were made during World War II and lie outside the scope of this paper.

DURING this period the study of the propagation of radio waves led to a marked increase in the understanding of the physical mechanisms and the evaluation of signal strengths over a wide range of frequencies, with varying ground and atmospheric conditions. The study fairly naturally divides into consideration of the ground wave at all frequencies, ionospheric propagation at frequencies of about 30 Mc and below, and tropospheric propagation above 30 Mc and up into the microwave range of thousands of megacycles.

The Ground Wave

The ground wave is defined to include those factors affecting the propagation of radio-wave energy through the atmosphere and along the surface of the earth, excluding any ionospheric effects, or tropospheric effects other than standard refraction. Among the factors of prime importance are: 1) diffraction of the wave energy around the curve of the earth, 2) refraction or bending of the waves in passing through the atmosphere, 3) reflection from the earth's surface, with consequent phenomena of interference between direct and ground-reflected rays, and 4) absorption of energy by the earth and in the atmosphere. The latter proves to be significant only in the microwave range.

At all frequencies, electromagnetic energy is diffracted into the geometric shadow of an obstacle, the amount increasing with decreasing frequency. In the radio-frequency range, diffraction yields significant and usable signal strengths in the shadow zone beyond the geometric horizon of the earth. The mathematical problem of calculating the diffraction field strength in the shadow zone of the earth was attacked by many mathematicians over 15 years with resulting divergent answers. The first significant *break-through* was produced by Watson[1] in 1918, who showed that waves radiated by an antenna on the surface of a *perfectly conducting sphere* would be attenuated exponentially at great distances. The numerical values of field strength predicted by Watson under these conditions proved to be far lower than known experimental values. This discrepancy promoted increased interest in the ionosphere as a mechanism that might explain the wide divergence in the mathematical and experimental values of field strength, particularly in the kilocycle range of frequencies.

Meanwhile Austin[2] and others, over a period of years, had collected a considerable amount of quantitative field strength data in the LF and VLF frequency ranges. The results were summarized in the now-classical Austin-Cohen formula. This formula proved quite satisfactory as a basis for the engineering design of long dis tance LF and VLF circuits.

Interest in the behavior of the ground wave per se tended to diminish until the advent of MF broadcasting in the 1920's, when the strength of the wave in both the interference region and the shadow zone over an *imperfectly conducting earth* became of importance in the

* Received by the IRE, July 10, 1961.

† College of Engrg., The University of Michigan, Ann Arbor, Mich.

[1] G. N. Watson, "The diffraction of electric waves by the earth," *Proc. Roy. Soc.* (*London*) *A*, vol. 95, pp. 83–89, October, 1918; and vol. 96, pp. 546–563, July, 1919.

[2] L. W. Austin, "Some quantitative experiments in long distance radiotelegraphy," *Bull. Bur. Standards*, vol. 7, pp. 315–363; October, 1911.

—— "Long distance radio receiving measurements and atmospheric disturbances at the Bureau of Standards in 1925," PROC. IRE, vol. 14, pp. 663–674; October, 1926.

allocation of frequencies to a rapidly increasing number of broadcasting stations.

This problem was solved completely during the years 1937 to 1941, covering pertinent factors such as heights, gains and spacing of the antennas, both vertical and horizontal polarization, arbitrary ground constants, and *standard* refraction of the *average* atmosphere. With standard refraction the radio-wave rays are curved downward with a radius of curvature equal to four times the actual earth radius. Passing over important mathematical contributions, it is possible to list only those authors and publications that contributed directly to the production of universal field strength curves. These are Eckersley and Millington,[3] van der Pol and Bremmer,[4] Burrows,[5] Gray,[6] Burrows and Gray,[7] and Norton.[8]

Ionospheric Propagation

Following Marconi's success in 1901 in transmitting signals across the Atlantic, Kennelly and Heaviside postulated the existence of a conducting (ionized) layer in the earth's upper atmosphere and suggested that such a layer might cause the waves to follow the curvature of the earth. After it became clear that diffraction could not explain the substantial field strengths actually received at great distances, increased attention was directed to this proposal of an ionized region.

The theory of radio-wave propagation through the ionosphere is based on work by Eccles[9] in 1912, on the ionizing effect of solar radiation, and on the effective refractive index of an ionized medium. Larmor[10] in 1924 re-examined the work of Eccles and others and ascribed the major part of the refractive effect to the presence of free electrons in large numbers. The Eccles-Larmor theory, as later extended and developed by Appleton,[11] Hartree,[12] and others to include the effect of anisotropy due to the earth's magnetic field, is now considered the basic theory of radio-wave propagation in the ionosphere. This work was later extended by Booker[13] and others to cover oblique propagation in a nonhomogeneous ionosphere.

The Eccles-Larmor theory leads to the concept of an effective refractive index n less than unity in an ionized medium; in the simplest theory $n=\sqrt{1-(Ne^2/m\epsilon_0\omega^2)}$, where ω is the radian wave frequency, ϵ_0 is the mks free space dielectric constant, N is the number of electrons per cubic meter and e and m are the electronic charge and mass. In terms of the wave frequency f, $n=\sqrt{1-(81N/f^2)}$. 81 N is the so-called critical frequency, now known as the *plasma frequency*. Since $n\leq 1$, a radio wave going up into the ionosphere encounters a decreasing refractive index and, therefore, is bent back toward the earth according to the usual refraction law. The maximum height it reaches is that corresponding to N_{max} given by

$$N_{max} = \frac{1}{81} f^2 \cos^2 \theta$$

where θ is the angle of incidence on the bottom of the ionosphere.

Absorption in the ionosphere is accounted for by including ν, the number of times per second each electron collides on the average with neutral molecules. This modifies the relative dielectric constant slightly, but adds a conductivity term that results in absorption of energy and consequent attenuation of the wave.

Again using the simplest theory and making suitable approximations, the absorption coefficient α is proportional to ν and to N but is inversely proportional to the refractive index n. Thus maximum absorption can occur at two different levels: one where the product νN is a maximum with $n\approx 1$, the other where n is a minimum. The first case (nondeviative absorption) occurs in the lower ionosphere, the second (deviative absorption) occurs in the upper ionosphere for near-vertical inci-

[3] T. L. Eckersley and G. Millington, "Application of the phase integral method to the analysis of the diffraction and refraction of wireless waves round the earth," *Phil. Trans. Roy. Soc. (London)*, vol. 237, pp. 273–309; June, 1938.

—— "The diffraction of wireless waves round the earth," *Phil. Mag.*, vol. 27, pp. 517–542; May, 1939.

—— "The experimental verification of the diffraction analysis of the relation between height and gain for radio waves of medium length," *Proc. Phys. Soc. (London)*, vol. 51, pp. 805–809; September, 1939.

[4] B. van der Pol and H. Bremmer, "The diffraction of electromagnetic waves from an electrical point source round a finitely conducting sphere, with applications to radio telegraphy and the theory of the rainbow," *Phil. Mag.*, vol. 24, pp. 141–176; July, 1937; suppl. 825–864; November, 1937; vol. 25, pp. 817–834; June, 1938; vol. 27, pp. 261–275; March, 1939. (For vertical polarization.)

[5] C. R. Burrows, "Radio propagation over plane earth—field strength curves," *Bell Sys. Tech. J.*, vol. 16, pp. 45–75, 574–577; January and October, 1937.

[6] M. C. Gray, "The diffraction and refraction of a horizontally polarized electromagnetic wave over a spherical earth," *Phil. Mag.*, vol. 27, pp. 421–436; April, 1939.

[7] C. R. Burrows and M. C. Gray, "The effect of the earth's curvature on ground wave propagation," Proc. IRE, vol. 29, pp. 16–24; January 1941.

[8] K. A. Norton, "The propagation of radio waves over the surface of the earth in the upper atmosphere," Proc. IRE, vol. 24, pp. 1367–1387, October, 1936; vol. 25, pp. 1203–1236, September, 1937.

—— "The physical reality of space and surface waves in the radiation field of radio antennas," Proc. IRE, vol. 25, pp. 1192–1202; September, 1937.

—— "The calculation of ground-wave field intensities over a finitely conducting spherical earth," Proc. IRE, vol. 29, pp. 623–639; December, 1941.

[9] W. H. Eccles, "On the diurnal variations of the electric waves occurring in nature and on the propagation of electric waves round the bend of the earth," *Proc. Roy. Soc. (London) A*, vol. 87, pp. 79–99; June, 1912.

[10] J. Larmor, "Why wireless electric waves can bend round the earth," *Phil. Mag.*, vol. 48, pp. 1025–1036; December, 1924.

[11] E. V. Appleton, "Geophysical influences on the transmission of wireless waves," *Proc. Phys. Soc.*, vol. 37, pp. 16–22; February, 1925. URSI Reports (Washington, D. C.); vol. 1, 1927.

—— "Wireless studies on the ionosphere," *J. IEE (London)*, vol. 71, pp. 642–650; May, 1932.

[12] D. R. Hartree, "The propagation of electromagnetic waves in an isotropic stratified medium," *Proc. Camb. Phil. Soc.*, vol. 25, pp. 97–120; January, 1929.

—— "The propagation of electromagnetic waves in a refracting medium in a magnetic field," *Proc. Camb. Phil. Soc.*, vol. 27, pp. 143–162; January, 1931.

[13] H. G. Booker, "Propagation of wave packets incident obliquely upon a stratified doubly refracting ionosphere," *Phil. Trans. Roy. Soc. (London) A*, vol. 237, pp. 411–451; September, 1938.

dence propagation, where $n \approx 0$ at the level of reflection.

Experimental evidence for the existence and the properties of the ionosphere were soon forthcoming. In 1925 Appleton and Barnett[14] proved the existence of a reflecting layer and calculated its height by measuring the angle the down-coming ray makes with the earth. With the transmitter and receiver separated an appreciable distance, the problem becomes one of simple triangulation.

The most significant and useful experimental technique is that of Breit and Tuve.[15] These authors in 1926 used transmitting and receiving antennas practically side-by-side, and sent a succession of HF pulses of short duration (30 to 100 μsec) at regular intervals (1/30 to 1/120 sec) up to the ionosphere and measured the time delay of the returning pulses. Assuming velocity c, the virtual height of the ionospheric layer (actual heights are somewhat less) returning the pulse of a given frequency is readily calculated. Continued experiments using different frequencies have enabled experimenters, using in conjunction the Eccles-Larmor relations, to measure the variation of electron densities with height in the various regions of the ionosphere, namely E, F_1 and F_2, ranging in height from about 100 to 400 km, and electron density maxima from less than 10^{10} to 4 times 10^{12} electrons per cubic meter.

Theory and experiment have shown that the ionosphere layers are doubly refracting as a result of interaction with the earth's magnetic field. Waves moving horizontally as well as vertically are split into ordinary and extraordinary groups that move in different paths with different velocities, and return to earth at distances beyond the so-called skip distance that exceeds the reception distance by the ground wave.

A big step forward resulted from the development by Gilliland[16] and others of sweep frequency ionosphere sounding equipment. These were the prototypes of the present day *ionosondes* that consist of transmitting and receiving sets in which the frequency is continuously varied from 500 kc or so to 15 or 20 Mc at an appropriate rate, generally proportional to the logarithm of the frequency. Modern ionosondes can sweep the MF/HF spectrum in a few seconds. The virtual heights are recorded as a function of frequency photographically by an oscillographic recorder. Fig. 1 is a typical example of the records that may be obtained for the E, F_1 and F_2 layers. The use of automatic recorders led at once, and ever since, to an enormous accumulation of factual information showing normal and frequently abnormal ionospheric behavior. Centers for the collection and study of this material were established in England, the United States, Canada and Australia.

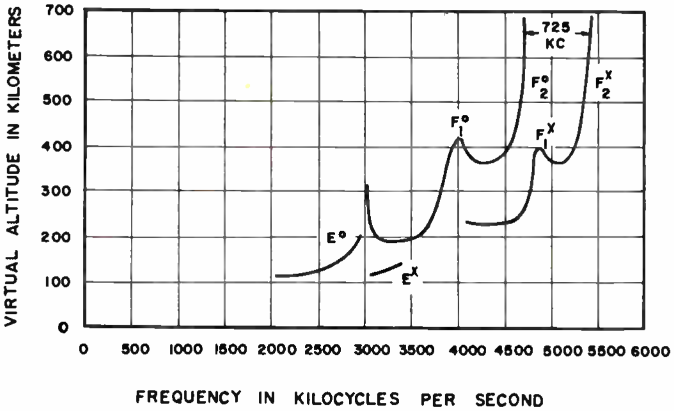


Fig. 1—Typical ionosonde record. Superscripts o and x correspond to the ordinary and extraordinary rays resulting from interaction of the wave with the earth's magnetic field. The sudden changes in height correspond to the critical frequencies of the layers.

As a result of the extensive theoretical and experimental work during the 1930–40 decade an extensive body of knowledge was built up regarding the structure of the ionosphere and its effect on the propagation of radio waves. This involved a number of disciplines: solar physics and ionizing radiations, atomic and molecular processes in the upper atmosphere of the earth, terrestrial magnetism and auroral phenomena, as well as terrestrial-radio communication over both long and short paths. The International Polar Year in 1932–33 represented a great world-wide cooperative effort among nations to obtain geophysical data, a large part of which directly added to the understanding of the ionosphere. Direct solar-ionospheric relations were discovered, such as ionosphere storms and solar-flare disturbances characterized by radio blackouts. Observations during solar eclipses led to the determination of recombination coefficients in the ionosphere. The important correlation between the sunspot cycle and electron densities in the ionosphere was discovered.

Especially significant, from the standpoint of radio communications, was the work directed toward calculation and prediction of radio-propagation conditions from the ionospheric observation. Seasonal, diurnal, and sunspot-cycle variations were established, and a beginning made on the delineation of the geographical distributions of the ionosphere. This would later be greatly extended by the network of observatories established during World War II. Ionospheric absorption measurements were begun, as well as the study of atmospheric radio noise, which generally sets the lower limit on usa-

[14] E. V. Appleton and M. A. F. Barnett, "Local reflection of wireless waves arising from the upper atmosphere," *Nature*, vol. 115, pp. 333–334; March, 1925.
—— "Some direct evidence for downward atmospheric reflection of electric rays," *Proc. Roy. Soc. (London) A*, vol. 109, pp. 621–641; October, 1925.
—— "Wireless signal variations," *Electrician*, vol. 95, p. 678, 1925.

[15] G. Breit and M. A. Tuve, "A test of the existence of the conducting layer," *Phys. Rev.*, vol. 28, pp. 554–575; September, 1926.

[16] T. R. Gilliland, "Note on a multifrequency automatic recorder of Kennelly-Heaviside layer height," Proc. IRE, vol. 21, pp. 759–760; June, 1933.
—— "Note on a multifrequency automatic recorder of ionosphere heights," Proc. IRE, vol. 22, pp. 236–246; February, 1934.
—— "Multifrequency ionosphere recording and its significance," *J. Res. NBS*, vol. 14, pp. 283–303; March, 1935.

ble high frequencies for long distance communication.

The general picture resulting from this work indicated an ionosphere that was highly variable, turbulent and patchy, and had unpredictable variations, but nonetheless had sufficiently regular characteristics so that, within certain limits of apparently random variations, reasonably reliable radio propagation predictions could be made.

It became possible, for oblique angle incidence, to calculate and disseminate information regularly about the maximum usable frequencies (MUF) and the lowest usable high frequencies (LUHF) for given paths and to predict the probable occurrence of ionospheric storms. In 1943 the Interservice Radio Propagation Laboratory of the National Bureau of Standards published the IRPL Radio Propagation Handbook under the guidance of the director, Dellinger, and the assistant director, Smith. The Handbook reappeared in expanded form in 1948 as "Ionospheric Radio Propagation," National Bureau of Standards Circular 462.

Excellent review articles for the period between Wars I and II were written by Mitra,[17] Mimno,[18] Dellinger,[19] and Dellinger and Smith.[20]

Tropospheric Propagation

The term Tropospheric Propagation refers to the effects of the earth's lower atmosphere on the field strengths of radio waves, with particular attention being given to the VHF, UHF and SHF frequencies, and also to various atmospheric conditions that produce nonstandard types of refraction. The atmospheric effects, due to temperature and moisture vapor gradients, are not particularly noticeable at frequencies below about 30 Mc because of the predominance of the ground-wave and ionospheric propagation modes. In fact, frequencies above the MUF were commonly called line-of-sight frequencies because they were useful only out to the radio horizon.

In the middle of the 1930–40 decade, with the advent of increasing use of VHF, attention was called to the effect of the troposphere on radio propagation[21] by several kinds of observations, chief among which were fading of VHF signals in and beyond the radio horizon, and extension of VHF ranges to several horizon distances, as evidenced by amateur radio contacts and unexpected interference from distant stations.

There appeared to be a noticeable relation between these transmissions and the weather, although little was accomplished in quantitatively explaining or predicting the effects before World War II, when it became extremely important to understand the vagaries of long-range radar performance at VHF and UHF, and later, SHF. Most of the work up to 1940 was concerned with studying the types, rapidity, and depth of fading and attempting, with only moderate success, to correlate the results with meteorological data.

Measurements of the temperature and moisture vapor gradients, made prior to and during the early period of World War II, showed the existence of atmospheric ducts. These ducts were found to exist at times upward a few hundred meters above ground, and at other times might exist as elevated layers. The ducts appeared to act somewhat like leaky waveguides. With sufficiently high frequencies, and with ducts of adequate thickness and strength to cause nonstandard refraction, the wave energy could be propagated largely inside the ducts and follow the curvature of the earth far beyond the radio horizon.

The full appreciation of this type of HF propagation was not developed until the war period itself. This period is outside the scope of this paper. Publications summarizing the work on atmospheric ducts and high frequencies done during the war are available.[22]

Acknowledgment

The author wishes to acknowledge the very valuable assistance of Prof. N. Smith in the preparation of this paper.

[17] S. K. Mitra, "The present state of our knowledge of the ionosphere," *Proc. Nat. Inst. Sci. India*, vol. 1, no. 3, p. 131; 1935.

[18] H. R. Mimno, "The physics of the ionosphere," *Revs. Mod. Phys.*, vol. 9, pp. 1–43; January, 1937.

[19] J. H. Dellinger, "The role of the ionosphere in radio wave propagation," *Trans. AIEE*, vol. 58 (*Commun. and Electronics*, pp. 803–822; January, 1939.

[20] J. H. Dellinger and N. Smith, "Developments in radio sky-wave propagation research and applications during the war," Proc. IRE, vol. 36, pp. 258–266; February, 1948.

[21] R. Jouaust, "Some details relative to propagation of very short waves," Proc. IRE, vol. 19, pp. 479–488; March, 1931.

C. R. Burrows, *et al.*, "Ultra-short-wave propagation over land," Proc. IRE, vol. 23, pp. 1507–1535; December, 1935.

[22] D. E. Kerr, "Propagation of Short Radio Waves," *M.I.T. Rad. Lab. Ser.*, vol. 13; 1951.

"Summary Technical Report of the Committee on Propagation," NDRC, vol. 1, "Historical and Technical Survey"; vol. 2, "Wave Propagation Experiments"; vol. 3, "The Propagation of Waves Through the Standard Atmosphere"; 1946.

"Meteorological Factors in Radio-Wave Propagation," Report of a Conference held by The Physical Society and The Royal Meteorological Society, April 8, 1946, Royal Institution, London, Eng.

Contributions to the Antenna Field During World War II*

L. C. VAN ATTA†, FELLOW, IRE, AND S. SILVER‡, FELLOW, IRE

Summary—**During World War II intensive efforts of radio engineers and physicists resulted in the invention of many new types of antennas and in advances in fundamental antenna theory. The paper presents the results of this work in relation to the principal problem areas that were recognized during that period: 1) sidelobe suppression; 2) beam-shaping techniques; 3) beam-scanning techniques; 4) broadbanding; 5) antenna siting. In each of these areas major advances were made, both in operating hardware and in theoretical understanding.**

I. Introduction

The developments in the antenna field during World War II, the result of intensive and combined efforts of radio engineers and physicists, were marked by the invention of many new types of antennas, and also by advances in fundamental antenna theory. Indeed, the whole subject of microwave optics was born and grew to maturity during the war years. A review of the field leads us into many different directions. In choosing the subjects for this short résumé we have arranged the material, including theoretical considerations, according to certain operational developments. Because the achievements were so largely the result of the combined efforts of many people in the United Kingdom, Canada, and the United States, we have avoided ascribing any development to a particular person, even though in a few instances this would have been possible.

The review is divided into five sections: 1) sidelobe suppression, which relates to developments based on the concept of flat phase fronts; 2) beam-shaping techniques, which utilized new developments in microwave optics; 3) beam-scanning techniques; 4) broadbanding; 5) antenna siting, a subject of great importance in realizing the optimum performance of the systems and in measuring the characteristics of antennas.

II. Pencil Beams and Sidelobe Suppression

The first operational concept in the design of antennas for radar applications was that of providing a narrow beam whereby targets could be located with high precision in azimuth and elevation. The pulse technique of radar, of course, provides the range information, and the combination of beam and pulse techniques yields the three coordinates of the target in space. The operational requirement, based on the searchlight concept of optics, required not only that the main beam be narrow, but also that the sidelobe structure be kept low to avoid the ambiguity of targets detected on the sidelobes. Two main approaches were taken to achieve a narrow beam: the use of optical devices, lenses and reflectors, having the property of transforming a family of rays from a point source into a family of parallel rays; and the use of discrete arrays of elements based on the principle of interference.

The directive antennas to produce a main beam circularly symmetric about the axis utilized a paraboloid of revolution with a primary feed at the focus. In order to achieve maximum forward gain out of a given aperture the distribution in the field over the aperture should be uniform in both amplitude and phase. However, the sidelobe structure is determined by the distribution of amplitude within the uniform phase constraint, and a major aspect of the theoretical work in this field was the study of the interrelation between the aperture distribution and the entire radiation pattern of the system. This greatly enlarged the understanding of diffraction theory over what had been available prior to the war. It also led to a rediscovery of long forgotten monumental contributions made to the field of optics over a hundred years ago.

Much effort was devoted to the realization of a feed system having, on the one hand, the characteristic of a point source, that is, generating a spherical phase front, and, on the other hand, a power pattern which would properly illuminate the paraboloid. When the coaxial line was the preferred form of transmission line, the natural choice of primary radiator was the dipole and, in particular, the reflector-backed dipole. Dipole feeds did only a fair job of illuminating the reflector since they produced unequal illumination in the orthogonal planes, and hence asymmetry in the mainlobe. In addition, their considerable backlobe radiation had deleterious effects on gain and near-in sidelobe characteristics of the final pattern.

As was indicated, it is necessary to trade sharpness of the main beam for sidelobe reduction. The general result established was that the aperture illumination should be tapered by some 15 db between the center of the aperture and the edge. The half-power width of the main beam is given by $\theta = k\lambda/D$, where λ = free-space wavelength, D = diameter of the aperture; designs were developed for circular apertures which yield values of k in the range $1.2 \leq k \leq 1.5$ with near-in sidelobes down 20 to 25 db from the peak intensity of the mainlobe.

* Received by the IRE, June 30, 1961.
† Hughes Research Labs., Malibu, Calif.
‡ Dept. of Elec. Engrg. and the Space Sciences Lab., University of California, Berkeley, Calif.

The availability of higher power, the move toward higher frequency and the toleration of heavier antenna structures, especially airborne, led to the use of rectangular uniconductor waveguides and compatible feed systems. The development of these waveguides and horns is certainly an outstanding feature of antenna work during World War II. Sectoral horns derived from rectangular waveguides by flaring in one or the other of the two principal planes of the guide, composite sectoral horns derived by flaring first in one plane and then in the orthogonal plane, and pyramidal horns derived by flaring simultaneously in both principal planes were developed for various applications. By virtue of the flexibility available in the design of a horn feed, it was possible to illuminate a reflector more effectively and to achieve improved performance with respect to backlobe and sidewise radiation. Three horn-feed types are shown in Fig. 1.

A major development in reflector design that accompanied the use of horn feeds was that of directing the axis of the horn pattern onto the reflector at an angle to the reflector axis. The reflector is then cut along an equi-intensity contour, again at an edge illumination some 15 db below the peak value over the aperture. Since the phase center of the horn remains at the focus, the phase distribution over the aperture remains plane yielding a directive beam. The importance of the technique is its flexibility in controlling sidelobe levels while retaining the required mainlobe characteristics. Such an antenna is shown in Fig. 2.

Operational needs arose for beams having mainlobes not circularly symmetric but of different beamwidths in two orthogonal planes, one very narrow to retain high resolution in the plane of scan, and the other relatively large to give extended coverage in the orthogonal direction. Such beams are known as *fanned* beams, and basic diffraction theory shows that they can be obtained by using rectangular or elliptical apertures with a corresponding ratio of their principal dimensions. The aperture illumination problem remains that of providing a uniform phase and a tapered amplitude distribution to control the sidelobe level; the horn feed solved this problem. Fig. 3 shows a paraboloid of revolution cut into an elongated elliptical shape and fed by a flared horn.

Another family of antennas designed to produce fanned beams consists of a parabolic cylindrical reflector between parallel plates illuminated by a sectoral horn feed at the focus. The exit pupil of the system is then a narrow rectangular aperture. Such antennas, known variously as cheese or pillbox antennas, served as ends in themselves or as line sources for illuminating larger parabolic cylinders, as shown in Fig. 4.

Microwave lenses were used in Germany prior to the war and were investigated extensively in the United States during the latter years of the war. In most military applications, however, lenses were not competitive with reflectors because of such factors as reduced gain, higher sidelobes, frequency sensitivity, greater weight and unfavorable shape. One development of note, however, must be mentioned in this survey. It was recognized that the dispersive property of a waveguide could be used in creating a medium having an effective index of refraction less than unity. The waveguide or metal-plate lens thus came into being but did not find actual application during the war. Fig. 5 shows an early type of waveguide lens.

Linear arrays also received a great deal of attention. The designs which were built around the coaxial line utilized dipole radiators, and there were many ingenious configurations devised for beacon antennas and related operational systems. Basically, however, these antennas were adaptations of ideas and developments already in use in radio-communications and direction-finder systems before the war. The distinct war-period contribution to the array system was the slotted waveguide array. The theory of slot radiators was developed extensively, and here it is appropriate to state that the basic work was largely provided by groups in the United Kingdom and Canada.

The excitation of a slot in the wall of a rectangular waveguide can be controlled by its position on the wall and the orientation of the axis of the slot with respect to the axis of the waveguide. It is this flexibility that makes possible the relatively easy control of excitation along the array. Of the many developments that were made in this field, the so-called Dolph-Tchebycheff array deserves special recognition. In this type of array the amplitude distribution along the array is related to the coefficients of a Tchebycheff polynomial. The result is a beam having all sidelobes of equal amplitude and therefore the narrowest beamwidth consistent with a prescribed sidelobe level. Sidelobe levels 30 db below the mainlobe peak were obtained, and in later developments following the war even lower sidelobe levels were achieved.

III. Beam Shaping

The operational requirements that motivated the design of fanned-beam antennas became rather quickly more demanding with respect to the more efficient utilization of the available power in air-search systems and more uniform ground illumination in airborne navigational and bombing antennas. One solution to the problem was obtained by the use of an extended feed in the focal plane of a paraboloid reflector. Each element of the feed produces a beam displaced from the axis by an angle proportional to the displacement of the element from the focal point. The resultant of the overlapping beams is a flared beam. High resolution in the transverse aspect is preserved until the coma aberration overrides the collimating property of the reflector. The theory of aberrations was advanced markedly in the course of this developmental program. The antenna on the left in Fig. 7 achieves a shaped beam by means of a three-horn distributed feed.

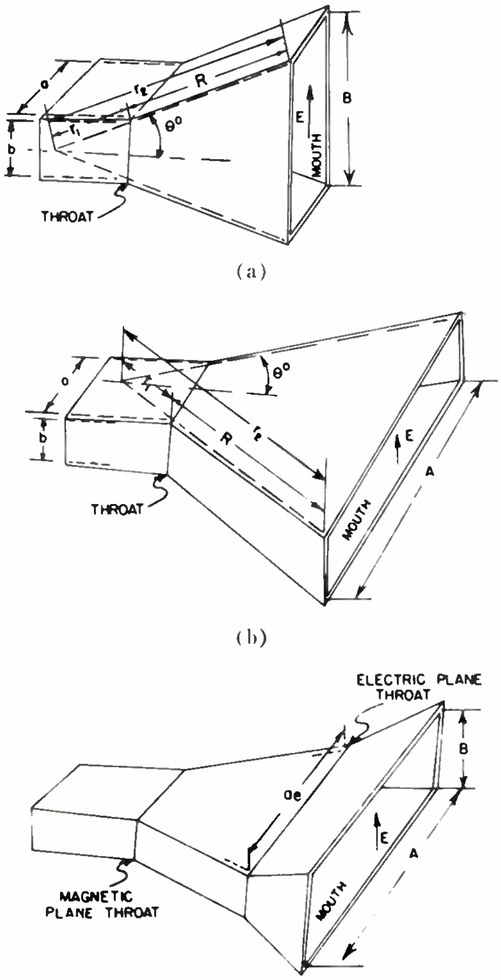


Fig. 1—Horn-feed types. (a) Electric-plane horn. (b) Magnetic-plane horn. (c) Compound horn.

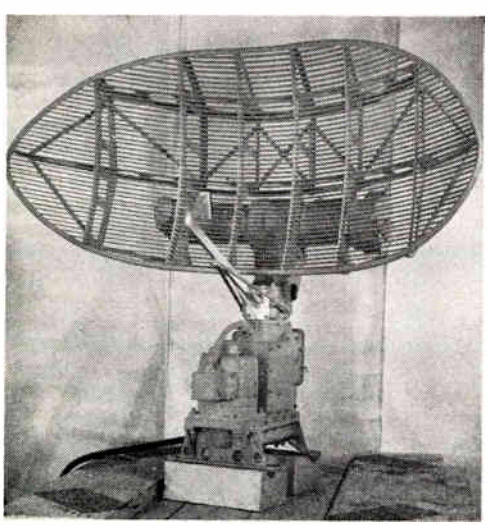


Fig. 2—Antenna with asymmetrically cut reflector.

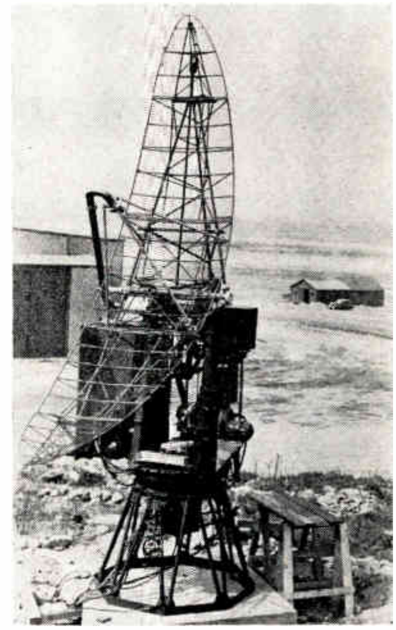


Fig. 3—Fanned-beam antenna.

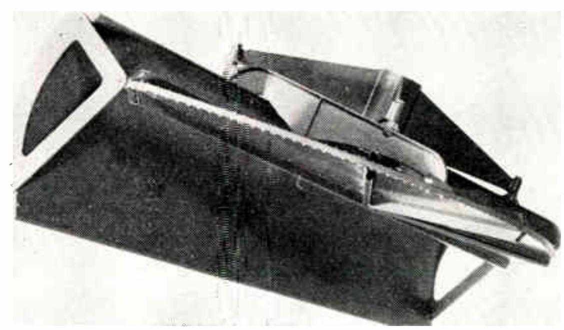


Fig. 4—Pillbox antenna feeding a cylindrical reflector.

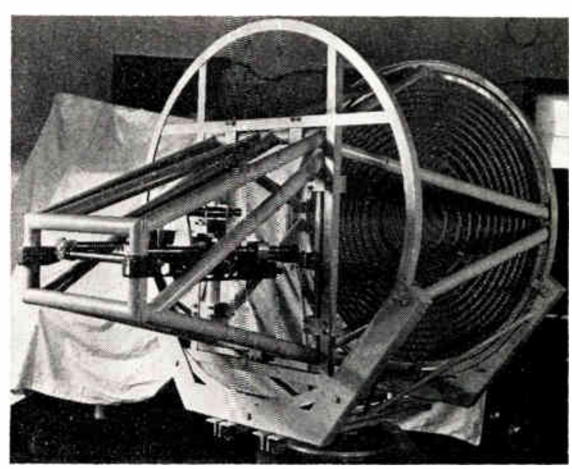


Fig. 5—Waveguide lens antenna.

The outstanding development, however, was the production of cylindrically curved phase fronts over the aperture of the antenna. By developing a phase front which is a section of a generalized cylinder, it is possible to obtain a flared beam which, in the transverse aspect, retains sharpness over large flare angles. The coma effect is eliminated since the rays are necessarily perpendicular to the generator of the cylinder. Both the underlying concepts and the realization of the system were major advances in microwave optics.

The requisite phase front was realized in several ways. One system utilized a line source feeding a cylindrical reflector whose cross section was designed to produce the dispersion of the ray system required for the flaring of the beam. The line source was aligned to be parallel to the generator of the reflecting cylinder. The sidelobe level in the transverse planes is determined entirely by the amplitude distribution along the line source. Both pillboxes and arrays were used as line sources, and all of the techniques for controlling sidelobe levels derived from flat phase fronts can be used in the design of the feed. (Fig. 4 represents this type of antenna.)

A second type of system utilized a point source feed with a modified paraboloidal reflector. The earliest form was made up of a split paraboloid of revolution with one section displaced relative to the other, so that the resulting phase front had a large amount of third-degree phase error. Advances in theory, however, led to a new type of reflector whose curvature varied from point to point so as to fit the ray requirements completely over the entire range of the beam. The complexity of the resulting surface was more than balanced by the simplicity inherent in the point source feed. This type of reflector, referred to as the doubly curved reflector, was used increasingly toward the end of the war. Fig. 6 shows an operational antenna of this class.

IV. Beam Scanning

A radar system must be able to scan its directive beam at a rate compatible with information rate called for by operational requirements. This must be accomplished while preserving resolving power and an effective SNR. Among the outstanding contributions made to the antenna art during World War II, were the developments in scanning techniques.

The obvious method of scanning by moving the entire antenna had soon to be superseded by other techniques as antennas became larger, and as higher scanning rates and complex scanning patterns became needed. In essence, the scanning problem is one of changing the orientation of the phase front at the aperture of the antenna. This must be done with minimum distortion of the phase front to preserve the structure of the beam. The various rapid-scanning techniques which were invented can be designated by the war-time categories as optical scanning, phase-shift scanning, and frequency-shift scanning.

In optical scanning only a part of the antenna is

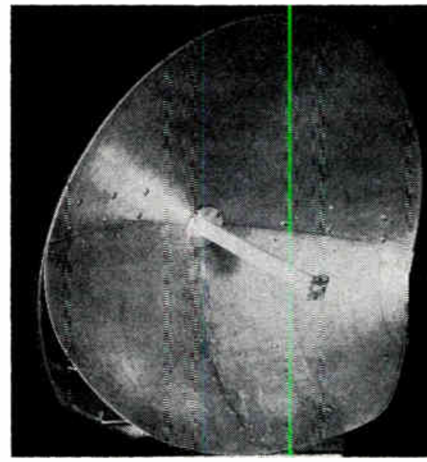


Fig. 6—Shaped-beam antenna employing a modified paraboloidal reflector.

moved. For example, the feed of a parabolic reflector can be moved off axis in the focal plane to obtain a limited angle of scan. In particular, the feed may be displaced slightly and then spun about the axis to achieve a conical scan. The differential signal from either side of the cross-over point yields a high accuracy of pointing and a convenient tracking signal.

The Robinson Roll antenna (right-hand antenna in Fig. 7) was the most ambitious application of the displaced-feed technique. The reflector had a long focal length in one plane and a relatively short focal length in the other. The feed was located at the far focal point, but was confined in one plane between parallel plates to meet the short-focus requirement. Moving the feed between the plates and parallel to the aperture then provided a scan in one plane. The final invention was to fold the plates in such a way that the oscillatory feed motion was replaced by a circular motion.

One system which was developed to effect the same purpose as conical scanning is essentially a data-processing technique. It uses a stationary feed system comprised of four feeds clustered in a square about the axis. These, when fed all in-phase and in out-of-phase pairs, form a central or sum lobe and two split or difference lobes in orthogonal planes, respectively. By using the sum lobe in transmission and the difference lobes and the sum lobe simultaneously on reception and comparing signals, one obtains a superior tracking antenna.

Three different examples of phase-shift scanning deserve mention. The Navy Mark 8 fire-control antenna (Fig. 8) scanned in an azimuth sector the beam formed by a two-dimensional array of "polyrod" endfire radiators, by separately controlling the phase to each vertical bank of radiators. This control was accomplished by rotating impedance elements in the circularly polarized field of cylindrical waveguide sections. In the Foster scanner a linear variation in phase was accomplished by control of physical path length; a wave between parallel plates was conducted around a variable portion of the circumference of a cone. The Eagle antenna was a

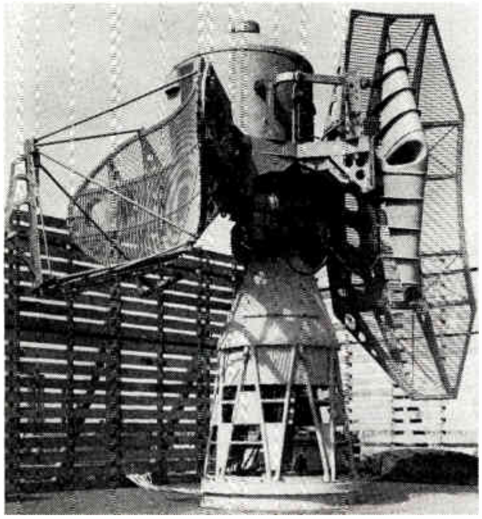


Fig. 7—Shaped-beam antenna (on the left) and Robinson Roll scanning antenna (on the right).

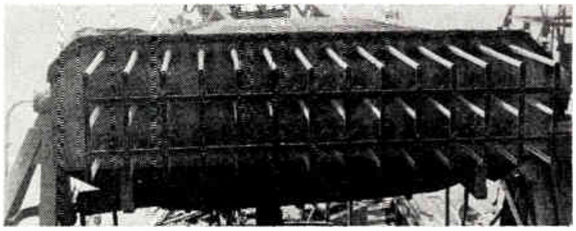


Fig. 8—Mark 8 "polyrod" scanning antenna.

linear array of dipole elements fed from rectangular waveguide. A linear-phase variation was obtained in this case by varying the major dimension of the waveguide and, hence, the optical path length between radiating elements.

Frequency-shift scanning makes use of the fact that the wavelength, and therefore the phase, along an array of radiators fed from a waveguide depends upon the frequency of the radiation. The scanning effect can be amplified by using a frequency-sensitive waveguide and by wrapping or folding the guide between radiators. This technique was explored during the war but did not find extensive application until later because of lack of frequency flexibility in available radio-frequency power sources.

V. Broadbanding

In the early days microwave power sources, even though designed for a spot frequency, came off the line with a spread in frequency. Later an even greater spread in frequency was required to reduce friendly jamming and to make enemy jamming more difficult. For these reasons the RF system, including the antenna, had to perform satisfactorily over a band of frequencies. Studies of a large variety of systems led to a clearer understanding of the factors limiting impedance bandwidth and of the methods of combining mismatches to cancel one another.

In the case of the antenna, satisfactory performance required that any of the essential antenna characteristics be maintained within certain limits over the specified band. These characteristics included power gain, beam-width, sidelobe level and, in the conical scan pattern, the cross-over level. The characteristic of greatest concern, however, in those days of the sensitive magnetron was the impedance of the antenna.

A constant impedance mismatch, even if large, could be corrected by a transformer at the magnetron, but a mismatch that varied widely and rapidly with frequency had to be eliminated or compensated at the point of origin. Frequency-sensitive mismatches resulted primarily from the combined effect of a series of discontinuities distributed along the RF line. Antenna mismatch was most serious since the antenna was farthest from the RF source.

Feeds for paraboloid reflectors were matched first independently of the reflector. Dipole feeds from coaxial lines, with their associated chokes and directive dipoles or plates, were relatively frequency sensitive with many critical dimensions requiring elaborate adjustments. Horn feeds from waveguides were basically better matched with fewer critical dimensions and were more susceptible to calculation.

The art of horn design reached an advanced state during the war years. If the waveguide was flared in both dimensions, the flare angles and their positions in the guide were chosen so that their discontinuities tended to compensate each other and also that caused by the mouth of the horn (see Fig. 1). Final correction was accomplished by capacitive or inductive strips at or near the mouth of the horn, or better, in some cases, by the thickness and placement of a plastic cover over the horn.

When the feed was placed in the reflector, an additional discontinuity resulted from reflection back into the feed. In some cases this discontinuity was corrected by a small plate placed a fraction of a wavelength in front of the vertex of the reflector. The size and placement of the vertex plate could be calculated from the geometry and the feed pattern. The vertex plate saw limited use, however, because it disturbed the aperture illumination and increased sidelobes. The ideal solution was found in the asymmetrically cut reflector (see Section II) which took the feed out, or almost out, of the reflected beam. In this case the feed was located at the focal point but directed off-axis approximately toward the center of the reflector area.

The reflector mismatch was much more serious in cylindrical reflectors, such as pillboxes fed by horns or long cylindrical reflectors fed by linear arrays. For these antennas it was essential that the feed be removed from the reflected beam. This was accomplished for the pillbox by the use of the folded pillbox or by the "hoghorn." In the folded pillbox the feed at one level was connected to the linear aperture at another level by a parabolic bend. In the hoghorn the feed, still at the focus, illuminated half a pillbox with the aid of a guiding horn. Both of these pillbox modifications, however, introduced additional weight and a less satisfactory form factor. In the case of a cylindrical reflector illuminated by a long linear array, the reflected wave striking the array seriously disturbed both the impedance and the radiation characteristics of the array. Therefore the array had to be kept well out of the path of the reflected wave.

The long linear array, in which only a small fraction of the incident power reached the far end of the array, was usually terminated in a load. This so-called nonresonant array provided a relatively good impedance match since the reflections from successive elements could be made to curl up into a small resultant mismatch at the input to the array. Short arrays, operated resonantly with a short-circuit termination, were more frequency sensitive in their impedance. In both cases, of course, well-matched or individually compensated radiating elements greatly improved the over-all performance.

VI. Antenna Siting

An antenna tested under relatively "free" conditions was found to perform quite differently under operating conditions because of reflection of its radiation by the surroundings. Reflection back into the antenna affected its impedance match; forward reflection affected the mainlobe or sidelobe characteristics of its radiation pattern. Toward the end of the war great strides were made in understanding and solving the complex problems created by the installation of many antennas on a single ship or aircraft.

A shipborne antenna with a horizontal beam had its elevation pattern sharpened and split into multiple lobes, and its gain increased by reflection of half of the mainlobe by the sea surface. The over-all effect was not simple since the reflectivity of the sea surface depends on wavelength and polarization of the radiation, angle of incidence, and sea state, and the combined pattern depends on height and orientation of the antenna.

A more complex installation problem was created aboard ship because many antennas were competing with each other in the presence of stacks, masts and superstructure for some semblance of free-space conditions. This so-called antenna-system problem was particularly severe for omnidirectional antennas. In the case of HF communication antennas, intercoupling between transmitting and receiving antennas and the sharply lobed patterns of individual antennas were known only through operational experience, so that maintaining satisfactory HF communications required black art of a high order. Radars were frequently installed in pairs with their antennas fore and aft or port and starboard in order to achieve complete azimuthal coverage.

Military aircraft also presented a serious antenna-system problem. Antennas for navigation, short- and long-range communication, and radar competed for favorable sites in the face of increasingly severe aerodynamic restrictions as aircraft speeds increased. The radar antennas, aircraft intercept or bombing and navigation, had the additional problem of operating through a radome, a plastic housing necessary for aerodynamic reasons.

The radome problem first presented itself dramatically when a naval aircraft search radar was found to be blanked out in certain sectors. This was found to be due to reflections from the radome back into the RF system pulling the frequency of the magnetron out of the pass band of the receiver. This impedance effect continued to be the most serious consideration. Under certain conditions, however, refractions and forward reflections could distort the mainlobe and introduce objectionable sidelobes.

Radome wall designs included the thin wall (zero-thickness approximation), the sandwich wall (two thin skins spaced a quarter wavelength apart by plastic foam), and the half-wave wall (a solid, high-density wall, one-half wavelength thick). During the war much attention was given to the sandwich wall, but shorter radar wavelengths, higher temperature requirements, and more severe microwave optical requirements eventually favored the half-wave wall.

Antenna siting problems were first encountered in connection with measurements made in the course of antenna development and test. With directive microwave antennas, measurements of the pattern and impedance were normally handled separately. For impedance measurements, satisfactory results could be obtained merely by pointing the antenna out through a large open window. For pattern measurements, especially since there was considerable interest in measuring low-level sidelobes, it was necessary to have transmitting and receiving antennas facing each other across a considerable distance from elevated vantage points, so that reflecting surfaces were well out of both beams. As antenna apertures increased, the minimum pattern range also had to be increased according to the relation $R_{min} = 2D^2/\lambda$, where D was the maximum dimension of either transmitting or receiving antenna.

Radio-Wave Propagation During World War II*

KENNETH A. NORTON†, FELLOW, IRE

Summary—**Many publications have appeared which describe the very extensive wartime radio-wave propagation research. References to some of these summaries are given which cover some aspects of this research as carried out in the United States, England, the U.S.S.R., Japan and Germany. These summaries are most complete for the work done in the United States, England and Japan.**

A summary is then given of some of the wartime research in the United States on direction finders and the polarization of downcoming ionospheric radio waves. This is followed by some heretofore unpublished material on Japanese ionospheric research which, by way of example, clearly indicates that an April, 1942, paper by Maeda, Uyeda and Shinkawa contains the first definite identification and interpretation of the F2-layer longitude effect. Finally a brief summary is given of a few selected topics in propagation research which arose out of the wartime development of radar.

I. Introduction

BECAUSE of the importance of radio to the efficient conduct of many military operations, an intensive effort was made during World War II to understand and to predict the nature of radio-wave propagation. Thus great advances were made in the application of such information to radio communication, radio direction finding, and radar. For the improvement of communications and direction finding, these advances were primarily in the prediction of ionospheric propagation and atmospheric noise levels. As regards radar, our knowledge of propagation characteristics was extended initially from about 100 Mc in an abrupt jump to 3000 Mc (*S* band) and later to 1000 Mc (*L* band), 10,000 Mc (*X* band) and 25,000 Mc (*K* band). This wartime research, begun in these higher frequency bands in England and later conducted vigorously in the United States, Canada, New Zealand and Australia, increased the usable portion of the radio spectrum by a factor of 250. Thus, immediately after the war it was possible to make extensive use of this new knowledge in the establishment of such peacetime services as FM and TV broadcasting, microwave radio relays and aircraft communication and navigation services.

Prior to World War II, the study of radio-wave propagation was confined to a few small laboratories scattered throughout the world. During and subsequent to the war, these activities were expanded many fold and now represent a continuing activity of relatively enormous extent. In a brief article of this kind, it will not be possible to mention more than a few of the highlights of this extensive research. Fortunately, several good summaries are available in the literature.

* Received by the IRE, January 2, 1961.
† Radio Propagation Engineering Division, National Bureau of Standards, Boulder, Colo.

A large part of the research in the United States on radio propagation was conducted under the auspices of the National Defense Research Committee [1], and summaries of this research were issued by the Columbia University Press for distribution by the War and Navy Departments. The reports in this series of interest in connection with radio-wave propagation are contained in the "Summary Technical Reports" of Division 13, Electrical Communication, Division 14, Radar, and the Committee on Propagation. The Division 13 report is in four volumes, of which vols. 1, 2A and B contain some discussion of propagation. Volume 1 includes a discussion of the direction of arrival of downcoming ionospheric radio waves, and the bibliography lists several reports dealing with polarization errors in high-frequency direction finders and studies of the polarization of downcoming ionospheric radio waves which have not subsequently been published elsewhere; a summary is given in the following section of this paper. Volume 2A includes a discussion of the effects of hills and trees on radio propagation and a brief discussion of ionospheric propagation. Volume 2B deals with electronic navigation systems. The Division 14 report on radar contains a bibliography dealing with radio propagation. The NDRC Committee on Propagation published a three-volume report which gives extensive coverage of the work conducted during the war; in particular, volume 3, entitled "Propagation of Radio Waves," should be in the library of all serious students of this subject. In addition to these summary reports, the NDRC also sponsored the well-known Radiation Laboratory Series [2], and the following volumes of this series contain information on radio-wave propagation: vol. 2, "Radar Aids to Navigation," edited by J. S. Hall; vol. 4, "Loran," edited by J. A. Pierce, A. A. McKenzie, and R. H. Woodward; vol. 13, "Propagation of Short Radio Waves," edited by D. E. Kerr; vol. 24, "Threshold Signals," edited by J. L. Lawson and G. E. Uhlenbeck. The NDRC also sponsored the publication in two volumes of "Very High Frequency Techniques" [3], which were compiled by the staff of the Radio Research Laboratory of Harvard University.

Dellinger and Smith [4] published an excellent review of the developments in radio sky-wave propagation research and applications during the war, and Gladden [5] published a history of vertical-incidence ionosphere sounding at the National Bureau of Standards.

In England good summaries of wartime propagation research were published by Appleton [6], Tremellen and Cox [7], and Smith-Rose [8]. A most valuable source of war research results obtained in England is "Meteorological Factors in Radio-Wave Propagation," a Report of a Conference held on April 8, 1946, at the Royal Institution in London by the Physical Society and the Royal Meteorological Society.

Bailey [10] has prepared a thorough report on the very extensive Japanese research on radio-wave propagation before and during World War II; the ionospheric research portion of this report is summarized in Section III of this paper.

As regards Germany, a good review of their scientific activities during the war is given in the "FIAT Review of German Science" 1939–1946 [11]. However, there is very little discussion of radio propagation in this review; pertinent books in this series are: "Electronics Including Fundamental Emission Phenomena," vols. 1 and 2, by G. Goubau and J. Zenneck; "Geophysics and Geodesy," vols. 1 and 2 by J. Bartels; "Meteorology and Physics of the Atmosphere," by R. Mügge.

In an effort to obtain a balanced account of the research on propagation in the U.S.S.R. during this period, I solicited the assistance of Vladimir A. Kotelnikov who responded as follows:

> I received your letter with the request for information about the investigation on the propagation of radio waves carried out in the Soviet Union during the period of the Second World War. I am sending a list of fundamental work on the ionospheric and tropospheric propagation of radio waves belonging to this period. The questions which interest you in greater detail are discussed in the review by B. A. Vvedenskii and A. V. Sokolov, "Investigations of Tropospheric Propagation of Meter, Decimeter and Centimeter Radio Waves in USSR," published in *Radio Technika and Electronica*, vol. 2, no. 11, pp. 1375–1389, November, 1957; and likewise in the articles "Investigations of Propagation of the Radio Waves Along the Surface of the Earth in U.S.S.R.," by M. P. Doluckanov, pp.1344–1359, and in A. N. Kazantzev, "Investigation of the Ionospheric Propagation of Radio Waves in USSR," published in the same issue on pp. 1360–1375. These articles are accompanied by extensive bibliographic lists.

Enclosed with his letter were references [12]–[29] which, together with the above, appear to cover the more significant research carried out in the U.S.S.R. during the war.

II. Studies of Polarization Errors in Direction Finders and the Polarization of Downcoming Ionospheric Radio Waves

Work in long-distance radio location and direction finding at the National Bureau of Standards began in the Radio Section in Washington, D. C., which during World War II became the Interservice Radio Propagation Laboratory (IRPL), and afterwards the Central Radio Propagation Laboratory (CRPL), now located at Boulder, Colorado.

A direction finder may be responsive to vertical or horizontal polarization or to both polarizations phased in various ways, but one polarization may produce an undesired response which causes an error in the reading. This undesired response is a function of the geometrical properties of the instrument, its mountings and various accessories, such as RF cables. A combination of modes in the downcoming ionospheric wave results in more or less random polarization. From instant to instant the relative amplitudes and phases of the wanted and unwanted components of the arriving wave keep changing, causing the polarization error to change.

It is difficult to compare direction finders and evaluate their polarization errors by observing sky waves, for which bearings depart from their true great circle values for a variety of causes. Thus some simple test technique was desired which would offer an estimate of the polarization-error propensities of any particular direction-finder design. The first test of this type had been evolved by Barfield in England [30]. A balloon with a target transmitter and antenna producing a circularly-polarized wave at a downcoming angle of 45° was placed near the equipment under test. The polarization error observed was called the *standard-wave error*.

At NBS a group under the late H. Diamond [31]–[39] devised techniques whereby pickup factors for the desired and undesired polarizations were determined from fixed target transmitters. These pickup factors were then treated analytically in accordance with a procedure designed for each type of direction finder and polarization errors determined for any set of conditions.

In the case of the simple Adcock antenna, the desired polarization is parallel to the plane of incidence and the undesired polarization is normal to the plane of incidence. The plane of incidence is the plane normal to the ground which contains the incident downcoming ionospheric ray.

The pickup factor for a wave polarized parallel to the plane of incidence was called h. That for a wave polarized normal to the plane of incidence was called k. If h and k were known then the polarization angular error ϵ could be determined (in the notation used by the Diamond group) from

$$\tan \epsilon = \frac{k}{h} \frac{E_n}{E_{pz} \cos \psi},$$

where E_n is the resultant total vector electric field component normal to the plane of incidence (*i.e.*, horizontal polarization) after combination of the incident and ground reflected wave, and E_{pz} is the resultant total vector-electric-field component parallel to the plane of incidence and normal to the ground after the combination of the incident and ground-reflected wave components. ψ is the angle of elevation at the ground of the downcoming ray.

A figure of merit ϵ_0 was proposed for a direction finder which may be determined from $\tan \epsilon_0 = (k/h)$; it is clear that a direction finder may be less subject to polarization errors the smaller the value of ϵ_0.

The advantage of the method was that h and k could be measured on the ground. One type of error called *radiator parallax* occurred in measuring k in direction finders like the balanced H Adcock. This was due to a net unwanted vertical pickup in the system when a horizontal dipole was used as a source of perpendicular polarization. In the case of spaced vertical loops another type of error called *controlled parallax* was studied. The forward tilt component of the target transmitter's magnetic vector produced an unbalanced voltage in the system. Both of these errors in measuring h and k could be minimized by increasing the distance to the target transmitter.

In support of this experimental work, a comprehensive theoretical study of the polarization of downcoming ionospheric radio waves was made by the author [40]. Extracts from the conclusions to this report are given herewith. Its principal purpose was to provide a picture of the nature of the polarization of downcoming ionospheric radio waves with emphasis on those characteristics which affect the operation of a direction finder. It was found convenient to separate the study into two parts: 1) a study of the polarization of the downcoming waves before they reach the surface of the earth, and 2) the modification in the polarization caused by the fact that the wave collectors of the direction finder are placed near the ground so that the resulting wave at the direction finder consists of both a direct wave and a wave reflected from the ground.

For a study of direction finding, the polarization of a downcoming wave is adequately described when we can specify the ratio of the intensities of the components of the electric and magnetic fields polarized parallel and normal to the plane of incidence and the relative phase between these components. It was found that, at very high frequencies, the magnitude of the ratio (E_{nd}/E_{pd}) in the downcoming wave varies from a value much less than one to a value much greater than one, the median value of this ratio being one; for one per cent of the time it is greater than 10, and for one per cent of the time it is less than 0.1. These results follow from the assumption, which is reasonable at very high frequencies, that the normal and parallel components of the downcoming wave are each random variables, independently Rayleigh distributed in time about the same median value; on this assumption it follows that the magnitude of the ratio (E_{nd}/E_{pd}) will exceed q with a probability $p=1/(1+q^2)$. At very low frequencies (E_{nd}/E_{pd}) is distributed over a narrower range of values; its median value will be much less than one when the transmitted wave is polarized parallel to the plane of incidence at the ionosphere and will be much greater than one when the transmitted wave is polarized normal to the plane of incidence at the ionosphere. Near the gyro-frequency the extraordinary wave is absorbed in the ionosphere and the ratio (E_{nd}/E_{pd}) becomes equal to the constant value $R_d \exp(-iB)$ which is the limiting value of the ratio (E_{nd}/E_{pd}) for the ordinary downcoming wave as it emerges from the ionosphere. Just below the maximum usable frequency only the extraordinary wave returns to the earth and the ratio (E_{nd}/E_{pd}) becomes equal to the constant value $(1/R_d) \exp\{-i(B-\pi)\}$. At very high frequencies the relative phase between E_{nd} and E_{pd} is random, whereas at very low frequencies values of the relative phase near B are more probable when the transmitted wave is polarized parallel to the plane of incidence, while values near $(B-\pi)$ are more probable when the transmitted wave is polarized normal to the plane of incidence. Near the gyro-frequency the relative phase becomes equal to the constant value B, and just below the maximum usable frequency the relative phase becomes equal to the constant value $(B-\pi)$.

For wave collectors near the ground, the effect of the ground is to suppress the normal component in favor of the parallel component of the resulting wave under most circumstances. When the elements of the direction finder responsive to the undesired normal components of the resulting wave are buried underground as in the shielded U direction finder, this suppression will be very large indeed so that a direction finder should be comparatively free from polarization error. On the other hand, when a direction finder is to be operated primarily at the shorter distances so that the elevation angles of the downcoming ionospheric waves are large, the parallel components of the downcoming waves will be suppressed in favor of the horizontal components for wave collectors above the ground and a direction finder designed to operate on these latter components, such as one with spaced vertical magnetic-dipole wave collectors, would be expected to have some advantages. When, as in most direction finders, the parallel component of the electric field (or the normal component of the magnetic field) is the desired component while the normal component of the electric field is the undesired component, there will be less polarization error when the direction finder is located over soil with the largest possible value of the index of refraction, n, *i.e.*, over soil with the highest possible dielectric constant and conductivity. In Part 4 of Norton [40] a simple method for measuring these ground constants is described.

The problems arising when testing a direction finder by means of fields transmitted from a source at a comparatively short distance away were also discussed in Part 4. Due to the great complexity of the fields at distances less than one wavelength from a transmitting antenna and the lack of any reasonably simple, yet accurate, relations for the field at these short distances, it is considered desirable to keep the source for testing the direction finder at a distance greater than one wavelength. It was shown that, owing to the presence of a surface wave from the nearby transmitter which is absent in the case of the downcoming ionospheric waves, tests for determining the magnitude of the polarization error of a direction finder can only be made for high elevation angles, unless the source is removed to a very great distance, or unless appropriate correction is made for the effect of the surface wave. In general, it would appear that the National Bureau of Standards' method of testing a direction finder for polarization error is to be preferred; in this method the polarization of the test wave is measured at the direction finder, and the direction finder performance is then calculated using the relations given in Part 3.

The method of calculating the electric field corresponding to the magnetic field as measured by a loop antenna is also given in Part 4.

A direction-finder site should be on very level land with uniform ground conductivity. It was proposed that information of value concerning a proposed site could be obtained very easily and rapidly by transmitting from a vertical magnetic dipole and measuring the ground constants at various points on the site using the method described in Part 4 [40].

Since the electric and magnetic fields at a depth Δ below the surface of the earth are of interest in connection with the operation of some types of direction finders, expressions were given for these fields in this paper. These results anticipated the treatment of this subject and the problem of transmission from underground antennas given in a 1951 thesis by Moore, a revision of which has recently been published [41], in reports by Banös and Wesley [42] and in papers by Wait [43] and others.

The above rather extensive summary of this work has been given since it has not been published subsequent to declassification.

There is continuing interest in the problem of ionospheric radio-wave polarization. In 1950 Aden, de Bettencourt and Waterman [44] pointed out an error in the above report by the author; they showed that the major axes of the ordinary and extraordinary ellipses are not in general mutually perpendicular, but instead have the property that the sum of the angles which they make with the direction (normal to the plane containing the direction of propagation and the earth's magnetic field) is an odd multiple of $\pi/2$. A still more recent summary of this question is given by Murty and Khastgir [45].

The behavior of reflected downcoming ionospheric radio waves described above is to be expected when the ordinary and extraordinary wave components fade independently after traversing different paths through the ionosphere. This case is essentially different from the polarization fading of satellite signals which involve ordinary and extraordinary waves which follow essentially the same path in traversing the ionosphere; this latter phenomenon has recently been used by Little and Lawrence [46] to measure the electron content along the path through the ionosphere traversed by satellite signals.

III. Japanese Ionospheric Observations During the War

Shortly after the end of the war in October, 1946, a brief summary of Japanese ionospheric research was prepared by D. K. Bailey and N. Smith under the title appearing in the heading of this section. It was submitted to an English journal for publication, but no word was ever received from that journal relative to its acceptability and, as a consequence, this account was never published. It seems appropriate to include it in this historical summary and it is given here in full:

A substantial contribution to the scientific knowledge of the world is to be found in the work of Japanese scientists, during the war, on radio wave propagation.

This work is summarized in a recent report distributed by the Chief Signal Officer, United States Army Forces, Pacific, in Tokyo [10]. Besides abstracts of the research papers on radio propagation, the ionosphere, atmospheric noise, radiation, reflection, tropospheric propagation, antennas, direction finding, and diversity reception, this report gives also the results of vertical-incidence pulse ionosphere measurements of the usual *h'f* type made at many locations in Asia and the Western Pacific. These data and reports are now on file at the Central Radio Propagation Laboratory of the National Bureau of Standards and may be consulted for further details.

Particularly interesting are the ionospheric data. Observations of the *h'f* type were made between 1934 and 1945 at nineteen locations scattered throughout Eastern Asia and the Western Pacific, over periods ranging in length from ten months to ten years. It is of interest to note that only the developments of the Pacific War prevented the operation by the Japanese of eight additional stations in 1944 and 1945.

The stations actually operated by the Japanese are listed in Table I. In addition to these, the following stations were planned or under construction, but never actually operated:

Naha, Okinawa
Tainan, Formosa
Seishin, Korea
Hanoi, French Indo-China
Saigon, French Indo-China
Palembang, Sumatra
Kwajalein Atoll
Rabaul, New Britain.

The geographical disposition of the stations included in the Japanese program extended from Paramushiro in the north to Bandoeng in the south, and from Rangoon in the west to Kwajalein Atoll in the east. At the height of the program, in the spring of 1944, there were fourteen different stations in simultaneous operation—nearly one-third of the total number of stations in operation throughout the world at that time.

Nearly all the Japanese ionosphere data have fortunately been preserved. By far the most interesting feature of the data is the ample confirmation they provide for our previous ideas about the longitude effect in the *F2* layer. The data from the southern stations, at Hankow, San-Ya (Hainan), Rangoon, Manila, Palau, Penang, Singapore, Makassar, and Bandoeng are of particular interest, since they supply for the first time a nearly complete definition of the remarkable trough in daytime critical frequencies centered near the magnetic equator, and concerning which observations elsewhere, particularly in the Western Hemisphere, had provided little more than a suggestion of existence.

Particularly noteworthy is the fact that the Japanese first recognized the possibility that the *F2* layer of the ionosphere might exhibit different characteristics at the same local time and latitude in different longitudes, a behavior now generally termed the longitude effect, when they attempted to reconcile the data from Washington (lat. 39°N) with their data from Tokyo (lat. 36°N) and their Hokkaido observations (lat. 42.5 and 43.5°N) taken in June 1936 in con-

TABLE I

LIST OF JAPANESE IONOSPHERE STATIONS

Place	Operating Agency	Latitude	Longitude	Period of Operation	Last Available Data
1) Paramushiro, Kurile Islands	Navy	50.1°N	155.3°E	Nov, 1943–Sep, 1944	Sep, 1944
2) Shikuka, Karafuto	Navy	49.3	143.0	Oct, 1940–Dec, 1941	Dec, 1941
3) Tsitsihar, Manchuria	Army	47.3	123.9	Apr, 1940–Aug, 1945	Jul, 1945
4) Toyohara, Karafuto	Ministry of Education	46.9	142.8	Sep, 1942–Jul, 1944	Jul, 1944
5) Dairen, Kwantung	South Manchurian Railway	39.0	121.6	Dec, 1944–Unknown	Feb, 1945
6) Tokyo and vicinity					
a) Hiraiso	Ministry of Communications	36.4	140.6	Jun, 1936–Jul, 1945	Jul, 1945
b) Kokubunji	Ministry of Education	35.7	139.5	Aug, 1943–Dec, 1944	Dec, 1944
c) Meguro	Navy	35.6	139.7	Jun, 1934–Nov, 1940	Nov, 1940
d) Kaminoge	Ministry of Education	35.6	139.6	Jan, 1945–Aug, 1945	Aug, 1945
e) Hiratsuka	Navy	35.3	139.3	Dec, 1940–Apr, 1945	Apr, 1945
7) Hankow, China	Ministry of Education	30.7	114.3	Sep, 1941–Dec, 1944 (No observations, 1942)	Dec, 1944
8) San-Ya, Hainan Island	Navy	18.3	109.3	Jun, 1942–Jun, 1943	Jun, 1943
9) Rangoon, Burma	Army	16.8	96.5	Jun, 1943–Feb, 1945	Dec, 1944
10) Manila, P. I.	Army	14.6	121.0	Mar, 1944–Dec, 1944	Nov, 1944
11) Palau Islands	Navy	7.3	134.5	Jul, 1939–May, 1944	May, 1944
12) Penang, Malay States	Navy	5.5	100.4	Mar, 1944–Aug, 1945	Jul, 1945
13) Singapore, S. S.	Army	1.4	103.7	Feb, 1943–Aug, 1945	Aug, 1945
14) Makassar, Celebes	Navy	5.2°S	119.5	Jan, 1944–Oct, 1944	Oct, 1944
15) Bandoeng, Java	Army	6.9	107.6	May, 1943–Aug, 1945	Aug, 1945

nection with the solar eclipse of that month. In 1940 new stations were established at Tsitsihar, Manchuria (lat. 47.3°N) and Shikuka, Karafuto (lat. 49.3°N) which soon gave a sufficient body of data to permit a more detailed study of the longitude effect.

In April, 1942, the Japanese published a paper [47] giving the results of, so far as is known, the earliest quantitative as well as qualitative study of the longitude effect and attributing it to a geomagnetic influence. The freely translated title of the paper is "Differences of Ionospheric Conditions in the F2 Region between Two Points Lying in the Same Latitude but Having Different Longitudes." The paper, which is fully summarized in the Tokyo report [10], gives much attention to the preparation of contour charts of the critical frequencies of the F2 layer as a function of local time and both geographical and geomagnetic latitude. The writers found that the hitherto unreconcilable Washington observations could be satisfactorily used in Eastern Asia for the preparation of contour charts if the data were used at geographical latitude 60°N, the latitude having, in Japanese longitudes, the same geomagnetic latitude as Washington.

The Tokyo report [10] has the following to say in connection with failure of scientists outside of Japan definitely to recognize the longitude effect before 1943: "In fairness to the work on this subject done on the Allied side, it must be pointed out that the Japanese had nearly complete results of ionosphere observations made elsewhere in the world up to the end of 1941, whereas practically none of the Japanese work has been available before the present time. Nevertheless it is clearly necessary to assign the credit for the discovery of the longitude effect to the Japanese."

The Japanese ionospheric data now available go far toward filling in one of the three major gaps remaining, at the end of the war, in coverage of the land areas of the world by ionospheric observing stations. It is to be hoped that the future will see ionospheric stations in the other two gaps, which include temperate South America and Africa.

Somewhat later Bailey [48] published a more complete discussion of the geomagnetic nature of the F2-layer longitude effect. The existence of this longitude effect has always constituted one of the principal complications to the compact prediction of F2-layer maximum usable frequencies on a world-wide basis. Currently this is accomplished by publishing world maps [49], [50] for each even hour of the day and month of the year. In the near future a still more accurate method of prediction will be available; a first paper on this new method has been published by Jones and Gallet [51], and further papers are expected in the near future.

IV. Topics from the Propagation-Related Wartime Radar Research

The development of radar introduced a very large number of wave-propagation problems and at the same time provided a very useful tool for exploring some of these problems. These problems were for the most part associated with the matter of determining the maximum range of the radar for a target with a given effective echoing area and discriminating between the desired target and reflections from land masses or other objects [1], [2], [3], [8], [9]. A summary of the influence of some of these radar-propagation phenomena was given shortly after the war by Norton and Omberg [52].

The maximum range expected for a standard atmosphere was often shown in the form of a lobe diagram giving contours of maximum range in the vertical plane with distance as the abscissa and height above the curved surface of the earth as the ordinate. Such a plot indicated the gaps in coverage caused by interference between the direct and ground-reflected waves. When these two waves are in phase, the range reaches its maximum value, this maximum being determined by the transmitter power, antenna gain, a visibility factor and the operating noise factor. The concept of an operating noise factor was originated by North [53], [54], who showed that it depended not only on the noise factor of the radio receiver but also on the effective noise temperature of the radiation impedance of the receiving antenna. These concepts opened up a whole new field of research. Friis [55] introduced the concept of available power, thus obtaining a precise definition of the receiver noise factor which he called a *noise figure*. The author [56] has generalized these results to allow for image responses and for differing ambient temperatures of the various lossy components in the receiving system, thus obtaining a very general expression for North's operating noise factor and a general expression for the effective input noise temperature for a receiving system; this latter temperature has also been called an *operating* or *system* noise temperature. This operating noise temperature t_e is expressed in degrees Kelvin and provides a convenient measure, which is currently coming into widespread use, of the sensitivity of a receiving system. The operating noise temperature is related to the operating noise factor f and to an absolute arbitrarily-fixed reference temperature t_0 by $t_e \equiv f t_0$. Since f and thus t_e depend in general on losses or gains and image responses in various components of the receiving system and may also depend on the direction in which the receiving antenna is pointed, t_e can be identified with an actual temperature only by virtue of the fact that it has the dimensions of a temperature. This follows from the fact that the operating noise factor f is a dimensionless positive factor which is usually much larger than one but which may be very much less than one for microwave receiving systems employing low-noise receiving antennas and masers. The concept and method of calculation of an effective temperature of the receiving antenna have been described by Slater [57] and by Lawson and Uhlenbeck [58]. At frequencies of the order of 100 Mc and lower, cosmic and atmospheric noise determine the effective antenna temperature, whereas at higher frequencies the temperatures of the antenna foreground and of the absorbing constituents of the atmosphere usually control the antenna temperature; the effective temperature of an antenna pointed directly at the sun may be determined from its surface temperature of 6000° K and the antenna directivity as discussed by Southworth [59].

The most important gap in the radar coverage occurs near the radio horizon as a result of the weakening by diffraction of the radio waves in propagation over the curved surface of the earth. Under certain circumstances, however, radio waves at sufficiently high frequencies may be trapped in an atmospheric duct. This type of propagation was studied intensively during the war because of its great operational importance. A paper by Booker and Walkinshaw [9] provides a theoretical discussion of the main features of this interesting phenomenon. Duct propagation was found to be the normal mode throughout the day at sufficiently high frequencies over the sea, but at lower frequencies and for propagation over land trapping was a sporadic phenomenon.

Prior to and during the war, radio waves scattered from the troposphere were observed; a good summary of these early observations is given in a paper by Bullington [60]. Shortly after the war, these beyond-the-horizon fields were sometimes confused with the waves propagated via ducts, but research at NBS and elsewhere [60] soon established that these scattered waves were present even under standard atmospheric conditions when no ducting was expected. This led to the development of the Pekeris and Booker-Gordon theories of scatter propagation which will no doubt be discussed in a companion paper [61].

Finally, it is of interest to point out that radar reflections from the moon were first observed on January 10, 1946, by DeWitt and Stodola [62] using a war-developed radar set. Here again these observations opened up a new field of research which has recently had a revival of interest in connection with the current intensive investigation of outer space.

References

[1] Natl. Defense Res. Committee, "Summary Technical Reports," Columbia University Press, New York, N. Y.; 1947. (Distributed by the War and Navy Departments.)

[2] "Radiation Laboratory Series," McGraw-Hill Book Co., Inc., New York, N. Y., 28 vols.; 1948.

[3] "Very High Frequency Techniques," McGraw-Hill Book Co., Inc., New York, N. Y., 2 vols.; 1947.

[4] J. H. Dell nger and N. Smith, "Developments in radio sky-wave propagation research and applications during the war," PROC. IRE, vol. 36, pp. 258–266; February, 1948.

[5] S. C. Gladden, "A history of vertical-incidence ionosphere sounding at the National Bureau of Standards," NBS Tech. Note No. 28; September, 1959. (Available from Office of Technical Services, U. S. Dept. of Commerce, Washington 25, D. C.)

[6] E. Appleton, "The investigation and forecasting of ionospheric conditions," *Proc. IEE*, vol. 94, pt. 3 A (*Radiocommun.*, no. 11, pp. 186–199; 1947.

[7] K. W. Tremellen and J. W. Cox, "The influence of wave propagation on the planning of short-wave communication," *Proc. IEE*, vol. 94, pt. 3 A (*Radiocommun.*), no. 11, pp. 200–219; 1947.

[8] R. L. Smith-Rose, "Radio-wave propagation research in the Department of Scientific and Industrial Research during the years 1937–46," *Proc. IEE*, vol. 94, pt. 3A (*Radiocommun.*), no. 16, pp. 879–892; 1947.

[9] "Meteorological Factors in Radio-Wave Propagation," a report of a conference held at the Royal Institution, London, Eng., by The Physical Soc. and The Royal Meteorological Soc., April 8, 1946; published by The Physical Society, London, Eng., 1946.

[10] D. K. Ba ley, "Report on Japanese Research on Radio Wave Propagation," Office of the Chief Signal Officer, GHQ U. S. Army Forces Pacific, Tokyo, Japan, vol. 1, 201 pp., and vol. 2, 343 pp.; May, 1956.

[11] "FIAT Review of German Science," Office of Military Gov. for Germany, Field Information Agencies Tech.; 1948.

[12] B. A. Vvedenskii, "Diffraction charts for ultra short radio waves within line-of-sight for vertical polarization," *Izvest. Akad. Nauk SSSR, Otdel Tekh. Nauk*, no. 1/2, p. 3; January–February, 1942.

[13] B. A. Vvedenskii, "Graphs for determining the diffracted field for horizontally polarized ultrashort radio waves," *Izvest. Akad. Nauk SSSR, Otdel Tekh. Nauk*, no. 9, p. 60; September, 1942.

[14] B. A. Vvedenskii and M. I. Ponomarev, "Reception method of geometrical optics for determining the path of ultrashort radio waves in a nonuniform atmosphere," *Izvest. Akad. Nauk SSSR, Otdel Tekh. Nauk*, no. 9, pp. 1201–1210; September, 1946.

[15] M. I. Ponomarev, "Influence of refraction on the propagation of radio waves near the earth," *Izvest. Akad. Nauk SSSR, Otdel Tekh. Nauk*, no. 9, pp. 1219–1233; September, 1946.

[16] V. A. Fok, "Diffraction of radiowaves around the earth's surface," *Izvest. Akzd. Nauk SSSR, Ser. Fiz.*, vol. 10, no. 2, pp. 187–188; March–April, 1946.

[17] V. A. Fok, "Freznel reflection laws and diffraction laws for ultrashort waves," *Uspekhi Fiz. Nauk*, vol. 36, no. 3, pp. 308–327; November, 1948.

[18] V. A. Fok, "Field from a vertical and horizontal dipole elevated over the earth," *Zhur. eksp. teor. fiz.*, vol. 19, no. 10, pp. 916–929; October, 1949.

[19] P. E. Krasnooshkin, "The Application of the Normal Wave Method to the Problem of Long Distance Radio Communication," Izd. MGU (Moscow University Press) U.S.S.R.; 1947.
[20] S. Ya. Braude and I. E. Ostrovskii, "On the question of propagation of ultrashort waves from a dipole in the line-of-sight zone," *Izvest. Akad. Nauk SSSR, Ser. Fiz.*, vol. 10, no. 2, pp. 225–234; October, 1946.
[21] A. G. Arenberg and R. I. Perets, "Calculation of ultrashort radio field strength in the directly illuminated region," *Izvest. Akad. Nauk SSSR, Otdel Tekh. Nauk*, no. 1/2; January–February, 1942.
[22] M. A. Leontovitch, "On the approximate boundary conditions for electromagnetic waves over a surfa e of good conductivity." "Research in radiowave propagation," *Izdaniya Akad. Nauk SSSR*, collection 21; 1948.
[23] M. A. Leontovitch and V. A. Fok, "A solution to the problem of the propagation of electromagnetic waves over the surface of the earth by the method of smooth parabolas," *J. Exptl. and Theoret. Phys.*, vol. 16, no. 7, p. 557; 1946.
[24] L. I. Mandelshtam, N. D. Papaleksi and E. Ya Shegolev, "Measurement of the velocity of propagation of radio waves over the earth's surface," and "Latest measurements of radiowave propagation," Noveishie issledovania rasprostranenüa radiovoln sbornik 1, 1945.
[25] A. N. Kolmogorov, "Local structure of turbulence in a nonstationary and viscous fluid for very large Reynolds numbers," *Dokl. Akad. Nauk SSSR*, vol. 30, no. 4, pp. 299–303; 1941.
[26] Ya. L. Alpert and V. L. Ginzberg, "On the absorption of radiowaves in the ionosphere," *Izvest. Akad. Nauk SSSR, Ser. Fiz.*, vol. 8, pp. 42–65, 1944.
[27] V.L. Ginzberg, "On the Luxemborg theory—Interaction effect," *Izvest. Akad. Nauk SSSR, Ser. Fiz.*, vol. 12, no. 3, pp. 293–309; May–June, 1948.
[28] A. N. Kazantzev, "A theoretical account of the absorption of radiowaves in the ionosphere using different laws of the variation of ionization with height," *Izvest. Akad. Nuak SSSR, Otdel. Tech. Nauk*, no. 9, p. 1261; 1946.
[29] Ya L. Alpert, "Propagation of radiowaves in the ionosphere," *Gostekhizdat*, 1947.
[30] R. H. Barfield, "Some principles underlying the design of spaced aerial direction finders," *J. IEE*, vol. 76, pp. 424–443; April, 1935.

References [31]–[40] have been declassified since they were issued originally. Although the contents of these reports are likely to be of continuing interest to those concerned with direction finding, they have not been published subsequently. Copies of these reports are available in the library of the Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

[31] Natl. Bur. Standards, "High-Frequency Direction Finder Apparatus Research," November 15, 1941.
[32] ——, "Preliminary Measurements of a Navy Model DY Direction Finder," December 9, 1941.
[33] ——, "Correlations for Polarization Errors of Bearings with the DY Direction Finders," December 11, 1941.
[34] —— and Collins Radio Co., "Polarization Error Test of the CXAL Direction Finder," March 10, 1942.
[35] ——, "Polarization Errors of the SCR-551 Direction Finder," March 12, 1942.
[36] ——, "Polarization Errors of the W.E.-C.A.A. Direction Finder," April 20, 1942.
[37] ——, "Polarization Errors of a Direction Finder With Spaced Loop Antennas," June 25, 1942.
[38] H. M. Diamond, H. Lifschutz, and L. M. Poast, "High Frequency Direction Finder Apparatus Research," Res. Project C-18, Natl. Bur. Standards, Washington, D. C., July 1, 1942; Div. 13-101-M3 1a, *Ibid.*, pp. 61–62.
[39] L. M. Poast, "Coordinated Study of Correlation of High-Frequency Direction Finder Errors with Ionospheric Conditions," Natl. Bureau of Standards, Washington, D. C.; August 31, 1944.
[40] K. A. Norton, "The Polarization of Downcoming Ionospheric Radio Waves," FCC Rept. 60047, Natl. Bur. Standards, Boulder, Colo.; May, 1942.
[41] R. K. Moore and W. E. Blair, "Dipole radiation in a conducting half space," *J. Res. NBS*, vol. 65D, no. 6, pp. 547–563; November–December, 1961.
[42] A. Banös and J. P. Wesley, "The Horizontal Dipole in a Conducting Half Space," University of Calif. Marine Physical Lab., Repts. No. 53-33, 1953; and No. 54-33, 1954.
[43] J. R. Wait, "The electromagnetic fields of a horizontal dipole in the presence of a conducting half space," *Canad. J. Phys.*, vol. 39, pp. 1017–1028; July, 1961.
[44] A. L. Aden, J. T. de Bettencourt, and A. T. Waterman, Jr., "A note on ionospheric radio wave polarization," *J. Geophys. Res.*, vol. 55, pp. 53–56; March, 1950.
[45] Y. S. N. Murty and S. R. Khastgir, "Polarization parameters of the downcoming radio wave," *J. Geophys. Res.*, vol. 65, pp. 1449–1457; May, 1960.
[46] C. G. Little and R. S. Lawrence, "The use of polarization fading of satellite signals to study the electron content and irregularities in the ionosphere," *J. Res. NBS*, vol. 64D (*Radio Propagation*), pp. 335–346; July–August, 1960.
[47] K. Maeda, H. Uyeda and H. Shinkawa, "Reports of the Physical Institute for Radio Waves," no. 2, Tokyo, Japan; April, 1942.
[48] D. K. Bailey, "The geomagnetic nature of the F2-layer longitude-effect," *Terrestrial Magnetism and Atmospheric Elec.*, vol. 53, pp. 35–39; March, 1948.
[49] D. H. Zacharisen, "World Maps of F2 Critical Frequencies and Maximum Usable Frequency Factors," Natl. Bur. Standards Tech. Note No. 2, April, 1959; and "Supplementary World Maps of F2 Critical Frequencies and Maximum Usable Frequency Factors," Natl. Bur. Standards Tech. Note No. 2-2, October 1960. (Available from the Office of Technical Services, U. S. Dept. of Commerce, Washington 25, D. C.)
[50] "World Maps of F2 Critical Frequencies and Maximum Usable Frequencies for 4000 KM," Radio Res. Labs., Ministry of Posts and Telecommunications, Tokyo, Japan; August, 1958.
[51] W. B. Jones and R. M. Gallet, "The Representation of Diurnal and Geographical Variations of Ionospheric Data by Numerical Methods," in press.
[52] K. A. Norton and A. C. Omberg, "The maximum range of a radar set," PROC. IRE, vol. 35, pp. 4–24; January, 1957.
[53] D. O. North, "The absolute sensitivity of radio receivers," *RCA Rev.*, vol. 6, pp. 332–343; January, 1942.
[54] D. O. North, "Discussion on 'Noise figures of radio receivers,'" PROC. IRE, vol. 33, pp. 125–127; February, 1945.
[55] H. T. Friis, "Noise figures of radio receivers," PROC. IRE, vol. 32, pp. 419–422; July, 1944. "Correction," PROC. IRE, vol. 32, p. 729; December, 1944.
[56] A. P. Barsis, K. A. Norton, P. L. Rice, and P. H. Elder, "Performance predictions for single tropospheric communication links and for several links in tandem," Natl. Bur. Standards Tech. Note No. 102; August, 1961. (Available from the Office of Technical Services, U. S. Dept. of Commerce, Washington 25, D. C.)
[57] J. C. Slater, "Microwave Transmission," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 252–255; 1942.
[58] J. L. Lawson and G. E. Uhlenbeck, "Threshold Signals," M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 24, pp. 103–108; 1948.
[59] G. C. Southworth, "Microwave radiation from the sun," *J. Franklin Inst.*, vol. 239, pp. 285–297; April, 1945.
[60] K. Bullington, "Characteristics of beyond-the-horizon radio transmission," PROC. IRE, vol. 43, pp. 1175–1180; October, 1955.
[61] L. A. Manning, "Radio propagation following World War II," this issue, p. 709–711.
[62] J. H. DeWitt and E. K. Stodola, "Detection of radio signals reflected from the moon," PROC. IRE, vol. 37, pp. 229–242; March, 1949.

Advances in the Field of Antennas and Propagation Since World War II: Part I—Antennas*

E. C. JORDAN†, FELLOW, IRE, AND R. W. P. KING‡, FELLOW, IRE

Summary—**Progress in the quantitative understanding of antennas as circuit elements, transmitters, receivers and scatterers of electromagnetic radiation is reviewed briefly for the period 1945–1961. Advances in the design of selected radiating systems with special properties are indicated.**

Specific reference is made to the impedance, current distribution, and pattern characteristic of cylindrical dipoles, singly and in arrays. Particular developments touched upon include slot and surface wave antennas, microwave antennas and microwave lenses, super-gain antennas, and very large antennas and arrays for radio astronomy and satellite communication. Frequency-independent "angle" antennas and log-periodic structures are reviewed briefly.

Introduction

The WIRELESS transmission of signals depends critically on the radiating and receiving properties of antennas. In the period 1945–1961 significant advances were made in the quantitative understanding of the fundamental properties of antennas and many new and useful types of radiating systems were developed. In this brief review references can be made to only a small cross section of this rapidly growing field.

The Dipole Antenna

The basic radiating element is the dipole in the broad sense of a symmetrical center-driven conductor, half of such a conductor with its image in a metal screen as shown in Fig. 1 (or, by complementarity, a slot in a metal plane). The groundwork in the quantitative study of cylindrical, biconical, and spheroidal dipoles as boundary-value problems in electromagnetic theory instead of as sources with arbitrarily assumed distributions of current was laid in the years just preceding 1945. Research since that time has been extensive and so fruitful that the behavior of these antennas is now well understood in a detailed quantitative sense.

The transmitting and receiving properties of the cylindrical dipole have been determined theoretically in overlapping ranges by iterative, variational, Fourier-series, and Wiener-Hopf methods. The impedance is now known for antennas of any length over a broad range of ratios of radius to wavelength. Typical curves are shown in Fig. 2. A simple and quite accurate formula for the current when $kh=2\pi h/\lambda \leq 3\pi/2$ is

$$I(z) = \frac{j2\pi V}{\zeta\Psi \cos kh} \cdot [\sin k(h - |z|) + T(h)(\cos kh - \cos kz)], \quad (1)$$

where $\zeta = 120\pi$ ohms and where Ψ and $T(h)$ are complex coefficients that depend on the length and the radius of the conductor. Excellent correlation between theory and extensive series of measurements was achieved as soon as the interesting problem of end and coupling effects for various types of transmission lines had been solved and the properties of the idealized generators used in the theory were understood. Extensions of both the theoretical and the experimental techniques have been made

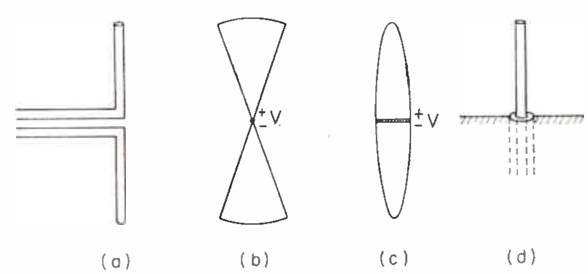


Fig. 1—(a) Cylindrical dipole driven from two-wire line. (b) Biconical antenna driven by idealized point generator. (c) Spheroidal antenna driven by idealized slice generator. (d) Cylindrical half dipole over conducting image plane.

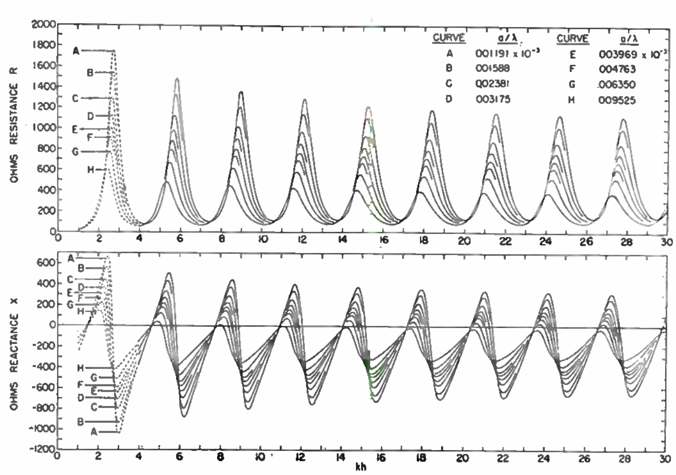


Fig. 2—Resistance and reactance of cylindrical antenna. Solid line = Wu. Broken line = King-Middleton 2nd-order theory.

* Received by the IRE, July 21, 1961; revised manuscript received, November 1, 1961.

† College of Engineering, University of Illinois, Urbana, Ill.
‡ Division of Engineering and Applied Physics, Harvard University, Cambridge, Mass.

to cylindrical antennas in dissipative media including their application as probes to study ionized regions in space. The response of dipoles to voltage pulses has also been studied theoretically and experimentally.

Corresponding progress has been made in the study of asymmetrical, folded and special antennas, as well as of the closely related loop antenna both with and without a ferrite core (which behaves like a magnetic dipole when electrically small).

The power transferred to the load of a cylindrical receiving antenna in an arbitrarily polarized plane-wave field incident from any direction has been obtained from an analysis of the distribution of current (which differs significantly from that in a transmitting antenna). The directional properties and the EMF in an equivalent circuit are expressed in terms of a complex effective length. The scattering of electromagnetic waves from variously loaded cylindrical antennas has been analyzed and measured; in particular, the back-scattering cross section has been determined for a thin dipole of any length.

The complete electromagnetic field of the driven dipole with an assumed sinusoidally distributed current was well known in 1945. More correct fields have since been obtained with the current given by (1) and from the application of the reciprocal theorem to the quite accurately known effective length of the receiving antenna. For very long antennas a Wiener-Hopf solution has yielded good results. Measured field patterns are in satisfactory agreement with theory.

Many of the properties obtained for thin cylindrical antennas have been translated to apply approximately to narrow slot antennas in a conducting plane with the help of the principle of complementarity. Advances similar to those summarized for the cylindrical dipole have also been made for biconical and spheroidal antennas.

Dipole Arrays

An unlimited variety of directional patterns may be achieved with suitably designed arrays of dipole antennas in air or slots in metal surfaces. In 1945 a very substantial knowledge existed about the properties of arrays of elements that are geometrically alike and are *assumed to have identically distributed* currents. Important advances in arrays of this type since 1945 include the proper choice of currents in the elements in order to achieve optimum relationships especially between the beamwidth and the minor lobe level, the use of elements that are of unequal length and unequally spaced in order to obtain broad-band properties, the development of methods for the synthesis and for the scanning of patterns, and new mathematical techniques that involve the application of potential theory and the treatment of the currents in the elements as sampled values of continuous functions.

Significant progress has also been made in the determination of the actual current distributions and the self- and mutual impedances of coupled elements in collinear, circular, and curtain arrays including Yagi, corner-reflector and special types of antennas. Except in arrays of half-wave dipoles and regular phase-sequences in circular arrays, the distributions of current along the several elements are far from alike, so that the actual field patterns with specified input currents may differ greatly from those predicted by conventional theory, especially with reference to the nulls and minor lobes. Distributions of current on a couplet of two full-wave elements with specified input currents, $I_{20}=jI_{10}$, are shown in Fig. 3 and the resulting field pattern in Fig. 4, together with the ideal pattern for identically distributed currents.

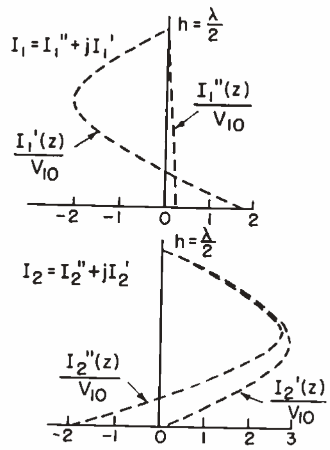


Fig. 3—Currents in the upper halves of the two elements of a full wave couplet; $I_2(0)=jI_1(0)$; $h/a=75$.

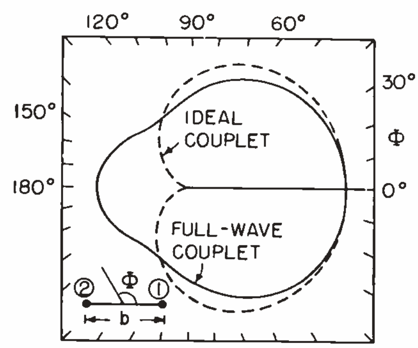


Fig. 4—Horizontal field pattern of couplet wjen $b=\lambda/4$; for ideal couplet with $I_2(z)=jI_1(z)$ and for full-wave couplet with currents shown in Fig. 3.

Surface Wave Antennas

Surface waves are associated with an interface between two media, their most important characteristics being the exponential decay of their field components in an equiphase surface perpendicular to the direction of propagation. The surface wave does not transmit power away from a lossless supporting interface unless the supporting structure is terminated or a discontinuity of some kind is introduced. The radiation pattern may be analyzed in essentially the same way as that of other aperture antennas from the field distribution in the terminal plane. For a nonuniform or modulated supporting structure there are, in addition, the radiation fields generated by the nonuniformities.

Although many of the characteristics of surface waves have been known for some time, not much attention was paid to practical uses in transmission lines and radiating systems until after the development of high-powered microwave oscillators during and after World War II. It was found that suitable interfaces that support surface waves generally consist of plane or cylindrical dielectric materials, such as sheets, rods or tubes as well as plane or cylindrical transversely corrugated metallic interfaces.

The dielectric rod antenna has been extensively used both singly and in arrays (see Fig. 5). It may have a square or circular cross section and may be excited in one of several possible surface-wave modes. The maximum of the radiation pattern is along the axis of the rod.

Corrugated end-fire antennas were developed where a higher degree of rigidity and heat resistance was required (see Fig. 6). In many cases it has been found possible to support the surface wave by means of corrugations in the aerodynamic structure itself. The corrugations are filled with a suitable dielectric material so that streamlining of the aerodynamic body is undisturbed.

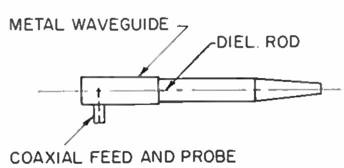


Fig. 5—Dielectric rod antenna.

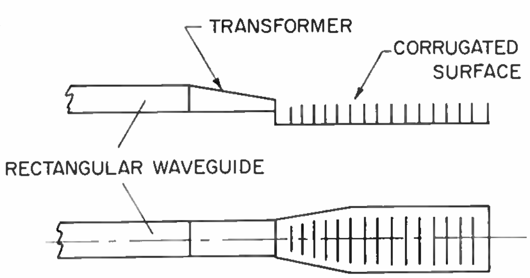


Fig. 6—Corrugated surface antenna.

Other antennas whose behavior may be quite well understood in terms of surface-wave modes are the Yagi and helical antennas which support surface waves with a phase velocity that is less than the velocity of light. Radiation occurs only from the terminal plane or from other discontinuities such as tapers. Antennas of this type are useful where end-fire operation over a wide band of frequencies and low wind resistance are required.

Corrugated or dielectric-coated surfaces operating in the circularly-symmetrical cylindrical surface-wave mode have been designed for use as omnidirectional beacon antennas.

Microwave Antennas

Advances in microwave antennas have been associated with increased size, more precise beam shaping, and improved scanning capabilities. Conventional paraboloidal, spherical, and parabolic cylinder reflectors with apertures of several hundred wavelengths have been constructed. Primarily this development represents a mechanical achievement in maintaining close tolerances although automatic servomechanism control of the surface alignment is also in use. For scanning or multiple-beam operation numerous microwave systems have been developed from their optical counterparts. Microwave lenses have been constructed with artificial dielectrics of many forms, and recent advances in the polymer industry are making available a wide range of natural dielectrics suitable for this purpose. For beam scanning, the Luneberg lens is a particularly interesting type that has been developed in many variations. In its simplest form it is a dielectric sphere with the useful property that rays from a point source, located at any point on its spherical surface, emerge as parallel rays from the other side of the sphere.

Broad-Band Antennas

A highly significant advance in the antenna field since 1945 is the successful development of frequency-independent antennas. Before 1945, a bandwidth of 2 to 1 was considered large, and few antenna engineers would then have predicted the possibility of designing antennas with impedance and pattern characteristics that are *independent* of frequency over as wide a frequency band as the designer may specify (and, of course, is willing to pay for in terms of dimensional requirements).

The concept of frequency-independent antennas has been realized in two separate but related practical developments: equiangular structures and log-periodic structures. The first of these uses the "angle concept" which is based upon the observation that an antenna whose geometry can be specified entirely by angles should have characteristics which are independent of frequency. Of course, all such structures extend to infinity, so the problem is to determine which retain their frequency-independence when truncated to a finite length. (The infinitely-long biconical structure is an example which does *not* remain frequency-independent when truncated to form a practical antenna.) The equiangular spiral antenna of Fig. 7 was the first of a class of antennas which are frequency independent when truncated. The feed voltage is applied between the two arms at the center, and the decay of current is sufficiently rapid that the current is negligible at a distance of a wavelength measured along the arms. Hence, the antenna may be truncated at this length, and its characteristics will be frequency independent for all higher frequencies (up to that for which the diameter of the feed region becomes an appreciable fraction of a wavelength).

In a more practical version of this antenna the spiral arms are developed on the surface of a cone as shown in Fig. 8. This structure is frequency independent, and yields a desirable unidirectional, circularly-polarized beam, having its maximum off the apex of the cone. Antennas of this type have been constructed to cover a frequency band of more than 20 to 1.

Log-periodic structures have radiation characteristics that repeat periodically with the logarithm of the frequency and in addition, that do not vary greatly over a period. Such antennas are essentially frequency independent. One of the simplest of many types is the log-periodic dipole array. In this array adjacent elements have a fixed length-ratio τ, and adjacent spacings bear the same ratio. From the principle of scaling, whatever radiation characteristics result at a frequency f, must also result at a frequency τf, and indeed at all frequencies $\tau^n f$, where n is an integer. When appropriately excited (with phase reversals introduced between adjacent elements) a truncated version of this array radiates a single-lobe linearly-polarized beam directed toward the apex. In operation, the active region of the antenna remains near the element of length $\lambda/2$. As the frequency is increased the active region moves forward through the array.

If the dipole arms are slanted forward, the log-periodic resonant-V array of Fig. 9 results. With care such an array can be designed so that when the active portion runs off the front end with increasing frequency, it returns to the rear end in the $3\lambda/2$ mode, moves through the array and again returns to the rear to advance through the array in a still higher-order mode. By this means very large bandwidths can be achieved from a relatively short array.

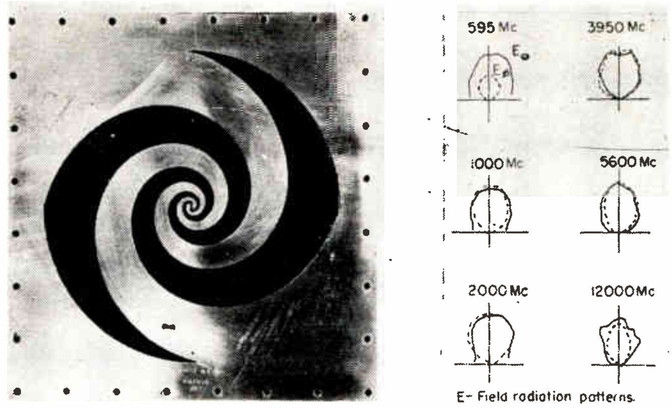


Fig. 7—Planar balanced equiangular spiral slot antenna.

Fig. 8—Conical balanced equiangular spiral antenna.

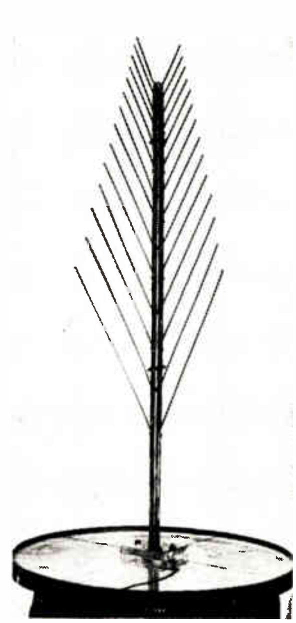


Fig. 9—Log-periodic resonant-V array.

Large Antennas and Related Problems

Applications in radio astronomy and in satellite tracking and communication have spurred the development of very large antennas and arrays. In radio astronomy where resolution can be more important than gain, large interferometer-type antennas, Mill's cross antennas, and synthetic arrays are in use. The latter, instead of using many elements simultaneously, use a fixed and a movable element to measure the relative phase and amplitude over the aperture of the equivalent array being synthesized. The data are stored and operated upon by a computer to yield information equivalent to that given by an ordinary array. In some large arrays utilizing many elements data processing techniques are being applied to combine the data from individual elements to produce a maximum of information.

During the 1940's it was demonstrated that there is no limit, theoretically, to the narrowness of the beamwidth obtainable from a given-sized aperture, or array. Although several excellent analyses showed conclusively the impracticality of these "super-gain" antennas, considerable effort was expended in some quarters to obtain a small degree of supergaining.

The advent of extremely low-noise receivers, using masers and parametric amplifiers, has introduced the antenna engineer to the notion of antenna temperature as a unit of measurement for the noise picked up by the antenna. The extremely low-noise temperatures seen by a microwave antenna looking at a "cold" sky has required that additional effort be devoted to reducing side and backlobes which "see" the warm earth. By careful design, antenna temperatures of less than 5°K have been achieved for an upward-looking antenna.

Acknowledgment

The authors acknowledge with thanks the assistance given by their colleagues, particularly J. D. Dyson, Y. T. Lo, and Y. Shefer.

Radio Propagation Following World War II*

LAURENCE A. MANNING†, SENIOR MEMBER, IRE

Summary—Since the early 1940's the development of new propagation modes has greatly increased man's ability to communicate. Some of the outstanding advances are briefly reviewed.

Introduction

A REMARKABLE REVOLUTION has taken place since World War II in the field of radio propagation. In the early 1940's, the recognized modes of signal transmission useful for communications and broadcasting were principally 1) ground waves at MF and below, 2) ionospheric layer reflections, and 3) line of sight at VHF and above. These modes are still important, and many new developments have led to their more efficient use; however, a really surprising number of exciting new propagation modes have now arisen that enormously extend the communications capabilities that are open to us.

Orbiting Relays and Reflectors

Potentially the most significant development to radio propagation has been the growth of man's capability to launch artificial earth satellites. Using satellite relay stations, wide-band VHF and UHF communications become possible on a line-of-sight basis. The presence of signals depends only on the operation of man-made devices, and not upon ionospheric critical frequencies, or tropospheric conditions. The possible bandwidths are limited only by system sensitivity, so that such services as world-wide television are within reach.

Four principal forms of orbiting relays may be distinguished.[1] First is the active satellite repeater. This approach minimizes the power and sensitivity required by the ground transmitters and receivers and leads to the highest signal-to-noise ratios. If three such repeaters were placed in 24-hour orbits, communications would be possible between non-polar regions at all times. However, since a 24-hour orbit has an altitude of about 22,000 miles, a transmission delay of about one-half second would occur; such a delay would be objectionable for some applications. If an active satellite is placed in a lower and more easily achieved orbit, the time delay can be greatly reduced, but at the expense of reduced geographical coverage. However, a network of perhaps 100 active satellites in random polar orbits would make communication possible 99.9 per cent of the time and appears likely to be economically feasible.

A second form of satellite repeater has been achieved by orbiting a large, spherical, conducting balloon. Echo I, launched in August, 1960 by the National Aeronautics and Space Administration, was the first test of this technique. Because passive balloon reflectors of this type require high ground-system sensitivity, they lead to relatively small rates of information transmission. An alternative to the spherical balloon reflector is the use of resonant dipole elements. Several hundred million short resonant wires placed in orbit with slight differential velocities soon spread into an extended belt. The high reflectivity of the resulting belt of reflectors can be used to return signals near the resonant dipole frequency.

Although not man-made, the moon is an earth satellite which has been tested for communication purposes. Despite its great distance, its size makes usable signals possible. The moon has the merit of unique durability. However, the bandwidths are relatively limited, and reflected signals are delayed by nearly two-and-one-half seconds. When the moon is not in view, communication is impossible. Lunar reflections have been of importance principally in studies of propagation through the ionosphere, and in studies of the lunar surface.

Beyond-the-Horizon and Scatter Propagation

The applications of satellites as relays are just beginning to be felt. Tropospheric and ionospheric scatter propagation, on the other hand, have already assumed an important role in the communication scene.[2,3] Systems of meteor-burst propagation have been developed which promise certain advantages for some applications.[4]

Diffraction theory predicts that, at VHF and UHF, beyond-the-horizon signal strengths should attenuate very rapidly so that propagation over distances of several hundred miles should not be possible. During the early 1940's, as high power transmitters became available in these frequency ranges, it became apparent that beyond-the-horizon signals were well above the values predicted by diffraction. In some cases, the signals could be explained by trapping in an atmospheric duct or waveguide caused by certain meteorological conditions. However, in absence of ducting, a weak signal remained that has been attributed to scattering from ir-

* Received by the IRE, November 24, 1961.

† Radioscience Lab. and Dept. of Elect. Engrg., Stanford University, Stanford, Calif.

[1] J. R. Pierce, "Communication satellites," *Sci. Am.*, vol. 205, pp. 90–102; October, 1961.

[2] Scatter Propagation Issue, Proc. IRE, vol. 43; October, 1955.

[3] Joint Technical Advisory Committee, "Radio transmission by ionospheric and tropospheric scatter," Proc. IRE, vol. 48, pp. 4–44; January, 1960.

[4] Meteor-Burst Communication Papers, Proc. IRE, vol. 45, pp. 1642–1733; December, 1957.

regularities in the tropospheric dielectric constant. Tropospheric scatter communication has been found to be useful at distances up to about 600 miles, and with bandwidths of as much as several megacycles.

At distances of roughly 600 to 2300 km, VHF signals can also be scattered in useful intensities from irregularities in electron density in the lower E region. Discovery of ionospheric forward-scatter signals was a result of a conscious search for a new communications mode, and was envisioned as an extension of the theory of tropospheric scatter to ionospheric irregularities.[5] Unlike tropospheric scatter, the irregularities in refractive index causing ionospheric scatter are strongly frequency dependent, so that frequencies from 25 to 60 Mc are preferred. Bandwidths of 40 kc are quite feasible. Ionospheric scatter circuits are relatively immune to the blackout and storm effects that plague HF ionospheric circuits. High power transmitters, large antenna gains, and diversity reception are the rule.

Ionospheric scatter signals are made up in part of reflections from a continuous background of ionized meteor trails. Also present are less frequent bursts of signal strength caused by relatively large meteors. Intermittent communication has been shown to be feasible by transmitting condensed bursts of information whenever the signal level exceeds an arbitrary high threshold level. Closed-loop systems are used to determine when transmission is possible. Since signals are sent only when the path loss is low, relatively low power is needed. Because of the aspect sensitivity of meteoric echoes, the antennas are not directed along the great circle. Meteor-burst communication is useful for distances out to about 2000 km, and at frequencies up to about 200 Mc. A transmission delay of a fraction of a minute is associated with the intermittent rate of meteoric incidence.

Higher-Order Scatter Systems

Tropospheric and ionospheric scatter systems depend on a second-order effect, the fluctuations in refractive index in the troposphere or ionosphere. Recently a third-order scatter mechanism has been experimentally demonstrated in the ionosphere,[6] and calculations show that communications possibilities exist.[7] It was shown by Gordon[8] that a neutral plasma containing electrons will scatter weakly because of the randomness in the positions of the individual electrons. More careful studies have shown that the electrons follow the positive ions, and the number of positive ions within successive regions separated by a fraction of a wavelength is unequal because of their random motion. Energy is scattered in proportion to the amplitude of the component of irregularity of scale equal to the effective space wavelength.[9] These "incoherent scatter" signals are very weak, and were first detected from the F layer using 5-Mw pulses and an antenna array of 1024 dipoles. Calculation shows that, at frequencies of 100 to 1000 Mc, communication would be possible at distances of over 4000 km. Using a million-watt transmitter and 1000 m^2 antenna apertures, an average of 100 teletype channels could be supported with an error rate of 1 in 10^3. Propagation would be relatively unaffected by ionospheric blackouts.

Developments in HF Communications

Great progress has been made in use and prediction of the properties of ordinary HF ionospheric layer propagation. Of primary significance has been the gradual evolution of a world-wide ionospheric sounding network. Since the war, the Central Radio Propagation Laboratory (CRPL) of the National Bureau of Standards has made predictions of radio propagation conditions based on results from the sweep-frequency ionosondes in the system. The development of the "f chart" has greatly improved reporting of data from individual stations.[10] CRPL is now adapting the prediction process to computers, thus reducing the subjective element. Predictions are made not only of the normal propagation frequencies but also of the probable occurrence of storm conditions and fade-outs.

Progress has also been made in measurement of propagation conditions as they occur. The scatter-sounding technique is important in this regard. Pulse signals are transmitted to a distant point by way of the ionosphere. Upon striking the ground, energy is scattered back along the original path to the transmitter site. Presence of echoes at given frequency and range thus indicates the possibility of communication on a given circuit.[11]

VLF Propagation

Despite the venerable place of VLF and LF propagation in the history of radio, important developments have occurred in the field. Greater theoretical understanding has arisen because of experimental and theoretical work.[12] The mode theory of propagation between earth and ionosphere has been carefully developed and

[5] D. K. Bailey *et al.*, "A new kind of radio propagation at very high frequencies observable over long distances," *Phys. Rev.*, vol. 86, pp. 141–145; April 15, 1952.

[6] K. L. Bowles, "Observations at vertical incidence scatter from the ionosphere at 41 Mc/sec," *Phys. Rev.* (letters), vol. 1, pp. 454–455; December 15, 1958.

[7] A. M. Peterson, "Free Electron Scatter as a Communication Mode," Paper no. 31/3, presented at WESCON, San Francisco, Calif.; August 22–25, 1961.

[8] W. E. Gordon, "Incoherent scattering of radio waves by free electrons with applications to space exploration by radar," Proc. IRE, vol. 46, pp. 1824–1829; November, 1958.

[9] T. Hagfors, "Density fluctuations in a plasma in a magnetic field with application to the ionosphere," *J. Geophys. Res.*, vol. 66, pp. 1699–1712; June, 1961.

[10] W. R. Piggott and K. Rawer, "URSI Handbook of Ionogram Interpretation and Reduction," Elsevier Publishing Co., distributed by D. Van Nostrand Co., Inc., Princeton, N. J.; 1961.

[11] O. G. Villard, Jr. and A. M. Peterson, "Scatter-sounding: A technique for study of the ionosphere at a distance," IRE Trans. on Antennas and Propagation, vol. AP-3, pp. 186–201; August, 1952.

[12] K. G. Budden, "Radio Waves in the Ionosphere," Cambridge University Press, London, England; 1961.

used to explain the experimental results. Pulse soundings of the ionosphere were made at LF for the first time soon after the war,[13] and quite recently the resonant properties of the concentric earth-ionosphere cavity has been demonstrated for the low order modes.[14] It is of interest that the fundamental frequency is about 7.8 cps, and the Q is about 4.

Study of the waveform of signals from lightning discharges led to detailed analysis of "whistlers," a highly dispersed form of atmospheric with energy return spread over a second or more. These signals were shown to propagate between northern and southern hemispheres along a dispersive guided path controlled by the geomagnetic field. Thus the signals extend several earth radii into space.[15] Man-made signals have also been transmitted using this mode.

Underlying Research

Back of the salient developments in propagation discussed above has been a tremendous and growing research effort, directed in large part to the physics of the troposphere and ionosphere.[16] One might draw fleeting attention to the development of refractometer techniques of tropospheric study, the enormous amount of work devoted to study of auroral ionization and polar propagation, and the study of ionospheric storms and polor blackouts. One should not ignore the development of techniques to measure ionospheric winds using drifting meteor trails, the fading properties of radio signals, or the drift and expansion of rocket-borne chemical releases. Nuclear detonations both below and above the ionosphere, as in the ARGUS experiment, have had far-reaching effects. Detailed study has been made of ionospheric layer formation, ionospheric composition, electron concentration, and the incident solar radiation. Rocket data has played a key role in the latter experiments. The list seems endless, and the literature is enormous.[17]

The Future

Very few of the developments of the past fifteen years were foreseen in advance. None the less, certain trends seem likely to continue. With the use of computers, communications channels are likely to be used by adaptive systems which control type of modulation, frequency, and directivity to match the propagation conditions. Computer control will be a must in organizing the use of satellite relays. The properties of the interplanetary environment and the ionospheres of other planets will be studied, and communications techniques adapted to the scale of the solar system. Our understanding of sun-earth relationships will grow, and with it an elucidation of ionospheric storm and auroral mechanisms. Finally, we will probably not leave the atmosphere as nature made it, but will find new ways to modify its propagation properties to suit our needs; already we have seeded the ionosphere with chemicals, shocked it with bombs, heated it with radio waves, and filled it with satellites. Next? Only dreams will tell.

[13] R. A. Helliwell, "On the measurement of ionospheric virtual height at 100 kilocycles," *Phys. Rev.*, vol. 73, p. 77; January, 1948.

[14] M. Balser and C. Wagner, "Observations of earth-ionosphere cavity resonances," *Nature*, vol. 188, pp. 638–641; November 18, 1960.

[15] R. A. Helliwell and M. G. Morgan, "Atmospheric whistlers," Proc. IRE, vol. 47, pp. 200–208; February, 1959.

[16] J. A. Ratcliffe, "Physics of the Upper Atmosphere," Academic Press, New York, N. Y.; 1960.

[17] L. A. Manning, "Bibliography of the Ionosphere—An Annotated Survey through 1960," Stanford University Press, Stanford, Calif., 1962.

The Future of Antennas*

M. D. ADCOCK†, SENIOR MEMBER, IRE, R. E. HIATT‡, SENIOR MEMBER, IRE, AND K. M. SIEGEL||, SENIOR MEMBER, IRE

Summary—**Although the rate and direction of advancement of the antenna field depends to a large extent on new inventions and techniques, and on the discovery of new materials, it is possible to indicate some of the trends which may be expected. The major advances in the near future will probably come in the areas of data-processing antennas and dynamic antennas. In the data-processing antenna, increased information is obtained by processing the signals available at the antenna terminals; in dynamic antennas, antenna performance is enhanced by time-modulating the aperture transfer function. It is likely that the most significant applications of data-processing antennas will be in the fields of radio astronomy and communications.**

Electronic scanning, another well-known technique of beam pointing, should lend itself naturally to the data-processing antenna approach, particularly since, for other purposes, the arrays will be used with computers. Such antennas will supersede movable large reflector types, for communication with space vehicles. Very large electronic scanning antennas will probably be built on the ground or on a horizontal plane, for minimum wind resistance. It may be expected, further, that electronic scanning arrays will be built in various shapes for installation on missiles or satellites.

In order to achieve larger antennas for increased tracking and communications capability, greater transmitted power, larger apertures, and natural geographic features may be used as foundation elements of the antenna. Canyons and cliffs, for example, could be used as reflecting surfaces if suitably coated. Ultimately, however, such techniques will be superseded by newer and more efficient developments. Somewhat further in the future, antennas for space communications will become extremely important. Work will no doubt be done on self-erecting antennas for automatic placement on the moon or planets, on antennas for the transmission of power from one location to another, on integral solid-state antenna elements, and on microwave-to-optical-to-microwave interplanetary communications.

THE AUTHORS in choosing to write about as yet unrecorded events, are sorely tempted to hedge their scientifically-based statements by pointing out the distinct possibility of science being strongly influenced by its inseparable and more unpredictable bedfellows, economics and human psychology.

Perhaps, sufficient hedgings can be had by saying that it takes a reckless or clairvoyant individual to predict the direction of future antenna research and development. In the antenna field, as in many other scientific programs, the rate and direction of advance depends to a large extent on new inventions, the discovery of new materials or the leveling of new requirements. During World War II, none of us would have predicted the development of an electronic scanning antenna using ferrite phase shifters. Who would have made a prediction five years ago that our 1961 antenna effort would include so much attention to antennas for space vehicles?

In spite of our lack of clairvoyance, we will attempt some predictions on the growth and the direction of growth for antennas of the future. Without doubt, improvements in materials and techniques will result in improvements in many types of antennas. Some of our estimates on expected trends on several antenna types will be discussed later. We predict, however, that the major advance in antennas will come in the field of data-processing antennas.[1] Such antennas, in more elementary forms and under different names, have been in use for years. An early example is the simultaneous lobing antenna in which the use of four feeds in a reflector provided more information on the location of a target than could be had from the use of a single feed. In this case and in many cases, data processing of the voltages available at the antenna terminal or terminals results in increased information on the nature of the target under surveillance or on the communication signal being received.

Other names used for closely related antenna work are "four-dimensional antennas"[2] and "dynamic antennas."[3] Work in this area also includes application of correlation techniques to antenna systems and the application of communication theory to antenna design. In fact, a symposium[4] was held under the latter name early in 1957. In dynamic antennas, the aperture transfer function is time modulated so as to obtain either enhanced performance characteristics or multiple beams or simultaneous scanning.

A notable example of past success in data-processing antennas is the synthetic aperture antenna system developed at The University of Michigan.[5] Here, the narrow antenna beam characteristic of a large aperture is obtained from a synthetic aperture produced by a small antenna. The energy received as the small antenna moves along a linear path is stored with the proper

* Received by the IRE, May 31, 1961.
† American Systems Incorporated, Inglewood, Calif.
‡ Radiation Laboratory, University of Michigan, Ann Arbor, Mich.
|| Conductron Corporation, Ann Arbor, Mich.

[1] S. Silver, "Antennas in Space Communication," presented at Inst. of Sci. and Tech., University of Michigan, Ann Arbor; May 25–26, 1960.

[2] H. E. Shanks and R. W. Bickmore, "Four-dimensional electromagnetic radiators," *Canadian J. Phys.*, vol. 37, pp. 263–275; March, 1959.

[3] R. W. Bickmore and R. C. Hansen, "Antennas 1957–1959," *J. Res. NBS*, vol. 64D, pp. 731–741; November-December, 1960.

[4] *Proc. Symp. on Commun. Theory and Antenna Design*, Boston University, Mass., January, 1957. See also Air Force Cambridge Res. Ctr., Cambridge, Mass., Tech. Rept. No. 57-105, ASTIA Doc. No. AD 117067.

[5] L. J. Cutrona, *et al.*, "A high-resolution radar combat-surveillance system," IRE TRANS. ON MILITARY ELECTRONICS, vol. MIL-5, pp. 127–131; April, 1961.

phase factor. Periodically the stored energy is combined to yield the effect of a large aperture antenna.

In the four-dimensional antennas as discussed by Shanks and Bickmore,[2] time is used as the fourth dimension to obtain better antenna performance. One or more of the antenna parameters is time modulated to improve radiation characteristics of the antenna or to increase the information-handling capacity of the system. An improved radiation pattern can be achieved, for example, by time modulating the length of the antenna. This may be accomplished by switching the end elements of an array into or out of the antenna. The sidelobes of the shorter array may be made to coincide with the minimums of the longer array and will partially cancel each other when the two antenna patterns are combined or averaged over a time longer than the modulation period. A 10-db improvement has been achieved by a simple application of this technique. Shanks and Bickmore show that other modulation techniques can result in a multipattern antenna while another would allow for simultaneous scanning of a two-dimensional array.

Data-processing techniques can be applied to interferometer antennas quite effectively. Drane[6] has shown that it is possible to combine a linear array with a number of co-linearly aligned two-element interferometers and obtain a pattern having no ambiguities and a resolution equal to that obtainable from a continuous aperture as long as the entire system. This is accomplished by phase modulating the output signals from each two-element interferometer and adding them in a nonlinear detector to the linear-array output.

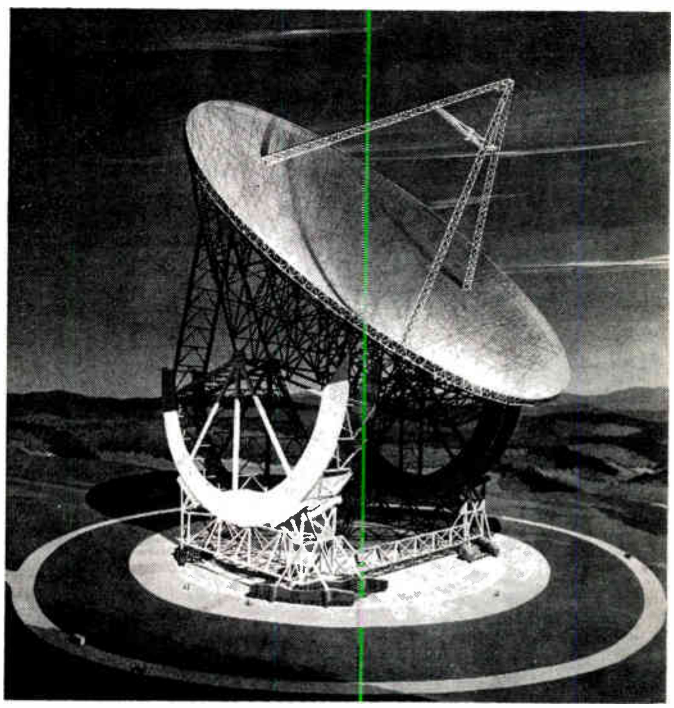
The basis, objectives, and applications of data-processing antennas have been widely discussed in the literature, only part of which is referenced in this paper. Our interest is in future developments in these areas. It is our belief that data-processing antennas will be used in many applications to add to the information that is obtained from a communication, radar or radio astronomy antenna. The area of greatest advances will be the areas of greatest payoff. One of these areas will be radio astronomy interferometer antennas; another will be in the closely related area of synthetic aperture antennas. It is believed also that these techniques will be used more and more in antennas for scatter communication. Electronic scanning antenna arrays should lend themselves quite naturally to data-processing techniques. Here it is possible to have information from each radiating element or from many combinations of elements. There is, therefore, infinitely more information available than may be ordinarily obtained from a single beam looking in a single direction. Data-processing innovations in such an antenna seem even more likely since it is probable that computers will be needed for other reasons in many future electronically scanned antennas.

Our opinion on the future of data-processing antenna techniques are in agreement with those of Shanks and Bickmore[2] who say,

> these concepts and techniques . . . represent one of the most interesting and promising new areas of investigation in the field of radiating systems. Without a doubt, the course of electromagnetic systems in the future will be influenced by this philosophy. However, effort in many different directions will be necessary to realize the full capabilities of these methods. Not only will extensive theoretical analysis be necessary on such problems as the accuracy, synthesis, procedures, etc., but in addition considerable component research and development is indicated.

Electronic scanning is the second broad area where we can expect great advances. Progress in electronic scanning during the past few years has been enormous but it is expected to continue in the future at a very rapid pace. This advance will be due, first, to the urgent need for large scanning antennas and, second, to the development and improvement of electronically controlled phase shifters. There is need for large electronically scanned antennas for communication with distant planets and space vehicles, for radio astronomy and for use in search and tracking of ICBM's.

Arrays which can be electronically scanned will be built to fill requirements which cannot be met by reflector-type antennas. With the 600-ft diameter parabolic antenna, Sugar Grove, W. Va. (Fig. 1), now being built, we are probably approaching the size limit of reflectors which can be scanned by rotating and tilting the reflector. Antennas of the future which must be capable of quickly beaming large amounts of power in



Courtesy of Naval Res. Lab.

Fig. 1—Artist's conception of NRL 600-ft parabolic antenna installation, Sugar Grove, W. Va.

[6] C. J. Drane, Jr., "Phase-Modulated Antennas," Air Force Cambridge Res. Ctr., Cambridge, Mass., Tech. Rept. No. 59-138, ASTIA Doc. No. AD 215374; April, 1959.

one direction or another will be electronically scanned arrays. The maximum practical size of antennas of this type which require no moving parts and which are capable of being built in modular form is much greater than for reflector-type antennas. In antenna arrays, the power maximum will not be set by the power-handling capability of the waveguide used, since it will be possible to have many power oscillators, one for each radiating element or for each group of elements. They can be properly phased by low-power control signals, which will also serve to scan the beam by some suitable arrangement of electronically controlled phase shifters. A digital computer, associated with the antenna, will be the means for providing the intricate control needed for the phase shifter; the computer can also be programmed to cause the beam to scan in any desired sequence. Electronically scanned antennas of this type will, of course, serve as receiving antennas by switching the power oscillator out of the circuit.

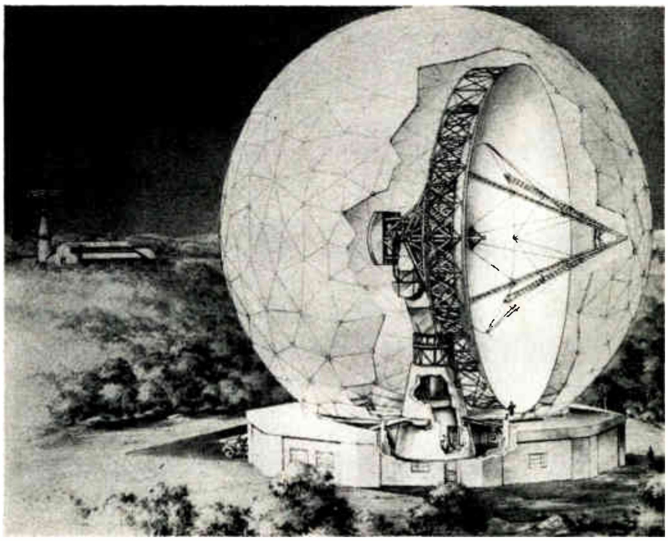
For many applications, the large scanning arrays can be built in the most convenient and economical form, that is, on the ground or on a horizontal ground plane. This provides both the advantages of minimum wind resistance and a low silhouette which in turn is advantageous for camouflage purposes.[7] Scanning arrays have been built in a form to cover part of a sphere, cylinder, or cone.[8] This trend will be carried further by designing scanning arrays suitable for use on surfaces of more complicated geometry such as are found on aircraft, missiles, or satellites. Here again it may be desirable to use a computer to schedule the necessary phase and current values for the various elements for the desired scanning sequence. The use of electronic scanning on satellites will be particularly desirable because of the low energy required for scanning and also because the scanning of such an antenna without physical motion will eliminate any tendency to upset the stability of the vehicle.

From these optimistic predictions about the future of electronic scanning one should not conclude that all electronic scanning problems have been solved. On the contrary, some very difficult problems remain,[9] just two of which will be mentioned here. Although the problem of mutual coupling between elements is better understood as a result of recent work, there is yet no attractive method of preventing high sidelobes and beam deterioration due to mutual coupling for wide-angle scans. More work is also required on phase shifters; although several types are now in use, all of them are far from ideal.

In emphasizing the future importance and growth of electronic scanning arrays, it is not intended that the importance of reflector-type antennas be de-emphasized. This antenna, which has been called the "workhorse of microwave communications" will always be in demand for many applications. At the present time it is probably the most popular antenna for radio astronomy use due in a large part to its broad-band frequency capability. The large reflector antenna is now and will continue to be widely used for keeping in contact with space vehicles.

At least two 1000-wavelength aperture antennas are in planning or being built. In discussing one of these, the Lincoln Laboratory Radio Propagation Research Facility Antenna (Fig. 2), Ward[10] says that, "most people who have studied the problem feel that one cannot build an antenna of equal precision that is much larger than this (1000-wavelength) without resorting to point-to-point surface control." He further says, "there will always be an insatiable demand by the customer for more and more tracking and communication capability with corresponding pressures for greater transmitted power, larger antenna apertures, and more sensitive receivers. . . . Antennas will continue to get bigger, demands will be made for more and more accurate radiating and reflecting surfaces on these antennas." Our position is somewhat similar to this—namely, that the practical size limit for scannable reflector-type antennas has been reached, but due to the demands for larger and larger antennas, we will continue to build larger ones even though they are not "practical."

As has been said, it is doubtful if moveable reflector-type antennas much larger than the 600-ft NRL-Sugar



Courtesy of Lincoln Lab., M.I.T.

Fig. 2—Artist's conception of radio propagation research facility.

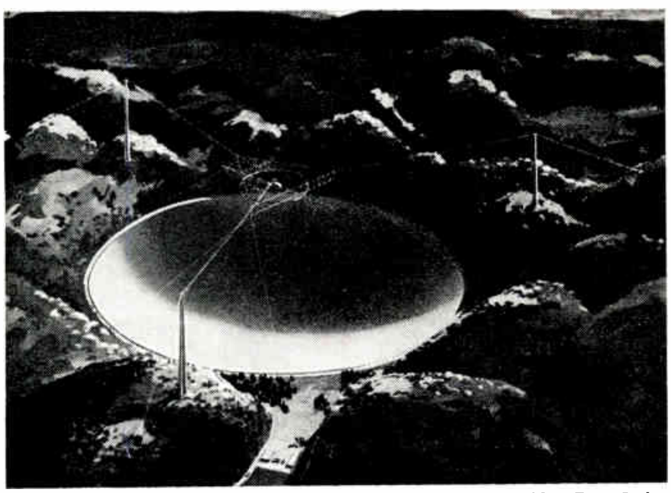
[7] C. J. Sletten, "Antennas—a strategic electronic component," *Microwave J.*, pp. 6–7; November-December, 1958.

[8] R. F. Goodrich, R. E. Kleinman, A. L. Maffett, C. E. Schensted, K. M. Siegel, M. G. Chernin, H. E. Shanks, and R. E. Plummer, "Radiation from slot arrays on cones," IRE Trans. on Antennas and Propagation, vol. AP-7, pp. 123–222; July, 1959.

[9] H. Snitkin, "Survey of electronically scanned antennas," *Microwave J.*, pt. 1, pp. 67–72, December, 1960; pt. 2, pp. 57–64; January, 1961.

[10] W. H. Ward, "An Account of the Factors that are Important in the Tracking of Space Vehicles," submitted to the Fifth Natl. Symp. on Space Electronics and Telemetry, Washington, D. C.; September, 1960.

Grove antenna will be built. The designers of the largest reflector-type antenna known, the 1000-ft antenna in Puerto Rico (Fig. 3), are taking advantage of Mother Nature to help them with their reflector. The 1000-ft spherical surface is formed in a hollow in the earth. Scanning is accomplished by moving the feed. A corrected line source feed is used to minimize spherical aberration. We predict that there will be more and more use of natural phenomena to assist the antenna designer. This use will not be limited to parabolic or spherical reflector-type antennas. One fascinating idea that should have a great future is the plan by Morgan[11] to use an island in the ocean as an antenna. This and other methods of using the contrasting dielectric properties of the earth and sea water should go far towards the solution of some of the VLF and ELF antenna problems. Other attempts have been made to utilize the ocean to solve an antenna problem. Zucker has studied the possibilities of its use as part of a surface-wave antenna. Here, the lack of an effective method of launching the surface wave has been a handicap. It is believed that this idea will be made to pay off in the future to provide solutions to some of the communication antenna problems. Mountains have been used to support wires for VLF antennas. In the future, mountains, canyons, or cliffs may be used also as reflecting surfaces—perhaps for a corner reflector antenna.[12] These and other ideas will be pursued in the future to solve some of the very difficult ELF and VLF antenna problems. Included among these problems is that of transmitting underground and through the ocean for certain communication requirements.



Courtesy of Air Force Cambridge Res. Labs.

Fig. 3—Artist's conception of Department of Defense ionospheric research facility. (Sponsored by Advanced Research Projects Agency.)

Since we have just entered the space age, we can with complete confidence predict rapid advances in antennas for space craft. For vehicles at altitudes well above the atmosphere where there are no problems with attenuation due either to the atmosphere or to the plasma sheath generated by the moving vehicle, antenna development will certainly include the use of VHF's. Millimeter and submillimeter frequencies will be used since these frequencies will make it possible to obtain high gain and narrow beams with antennas small and light enough to be compatible with space vehicle load limitations. These frequencies will be used for space-to-space communications and for radar for moon landings and for space-to-space vehicle rendezvous. The antennas used for these applications will in many cases be miniatures of present-day antennas. For space-to-earth communications, lower frequencies must be used. Low-gain antennas for such use are already being operated on current satellites at 20, 40, 108 Mc and higher frequencies. With the stabilization of satellites, higher-gain antennas will be used with frequencies well into the microwave band. Applications will include both communication and radar uses. Collapsible reflector-type antennas erected by gas pressure or mechanical means will be used.[13] It is expected that considerable use will be made of slot or dipole arrays distributed over the surface of the satellite.

For the satellite in a plasma environment, there exists a very difficult antenna problem. This is an area on which a large amount of future antenna effort will be expended. We can be sure of this because the problem is difficult, the need is urgent, and the payoff is high. To mention just one application, it is important for the Mercury space capsule to be able to maintain space-to-earth communications continuously and this is particularly necessary during re-entry.

The antenna problems that exist during the re-entry phase are well documented in the literature.[14] Of prime importance are solutions to the problems of power breakdown, reflections at both the inner and outer plasma boundaries, and the attenuation within the plasma. Solutions to these problems will include the use of higher frequencies and higher power, transpirational cooling, careful shaping of the vehicle, and care in choosing the antenna location. It is expected also that a way will be found to take advantage of the plasma properties, *e.g.*, by using the plasma as an impedance transformer to outer space. There will still exist, however, the need for a breakthrough to arrive at a successful solution to this problem. It may be found that it is possible to excite the ion column behind the satellite as a leaky waveguide antenna.

[11] M. G. Morgan, "An island as a natural very-low-frequency transmitting antenna," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-8, pp. 528–530; September, 1960.

[12] C. J. Sletten, "The world's biggest antenna-radio telescopes," *Proc. Conf. on Natl. Aeronautical Electronics*, Dayton, Ohio, pp. 9–13; May, 1960.

[13] P. D. Kennedy, R. F. Trainer, and J. W. Carr, "Progress in unfurlable antennas and related fields," *Proc. 10th Annual Symp. on USAF Antenna Res. and Dev.*; October 3–7, 1960.

[14] See for example W. Rotman and G. Meltz, "Electromagnetic Effects of Re-Entry," Pergamon Press, London, Eng.; 1961. Also Air Force Cambridge Res. Lab., Cambridge, Mass., Tech. Rept. No. 60–108.

There will also be work in the future on self-erecting antennas that can be dropped on the moon, or Mars, or Venus. No doubt some such antennas have been designed already. Future designs will be more sophisticated in mechanical design to provide needed gain and the ability to scan or perform other functions on command.

There will also be future studies on antennas for transmitting huge amounts of power from one location to another. These may be used, for example, to transmit power to a space ship. Such antennas will be characterized by having very high gain and the capability of handling unusually high power, probably by the use of multiple phased oscillators—one for each antenna element or group of elements.

Recently, it has been found possible to improve the performance of an array for certain applications by the use of nonuniform spacing for the elements.[15] This is a distinct advantage because it decreases the required number of elements and simplifies the feeding network and also results in lower sidelobes. This idea will be more fully investigated in the future.

Now that we have made speculations about the near future we might reach further out.

We foresee the use of arrays which will collect data. This data will automatically be processed in a digital computer and the array will be "rephased" and the information retransmitted. Thus we will invent systems which will first process the complete information received; second, determine the message of highest probability; and third, determine from the message and stored instructions its proper destination. We foresee arrays in which the signals from the various elements are not combined at radio frequencies, but where the usual coherent array phenomenon is duplicated by digital-data-processor techniques. Such an "array" will be as broadband as its individual elements; such an array will be ideally suited for nonplanar applications and, in addition, should provide versatility of operation superior in certain important respects to conventional arrays representing the present state-of-the-art.

We foresee the abolition of all metal radiating towers when they are used as dipole radiators. We foresee the use of ferrite-type materials for transmitting antennas to go with the present uses for receiving antennas. We believe the heat dissipation problem will be overcome with new materials and the "lost" heat will be partially recovered to be reused by the power source.

[15] M. G. Andreasen, "Arrays with Variable Interelement Spacings," presented at URSI Meeting, Boulder, Colo.; December, 1960.

We foresee the use of physical and chemical electron depletion processes at the conjugate point region in the ionosphere regulating the variation in electron density at the other conjugate point and thus producing lens effects which will allow us to focus energy into outer space. We will be able by these mechanisms to transmit energy into space at frequencies below the plasma frequency. We will also be able to focus energy into particular earth location from space at lower frequencies than now can penetrate the ionosphere.

Concerning reliability, we expect array elements individually backed up by solid-state blocks in such a way that several elements can be damaged or become defective without greatly degrading the system. Mutual coupling sensors will determine the malfunction and will automatically rephase the antenna to a new optimum.

Concerning sidelobes, we predict that we will find an irreducible minimum amount of energy that must be wasted by an antenna system. Antennas of the future will, of course, consider the surroundings as a part of the system.

As we pointed out earlier, we feel large reflectors will be on the way out and by fifty years from now will be "museum pieces" in the large. The search for natural reflectors will cease. We will decide for other reasons, possibly economic, possibly low noise of the background, where we want to locate our antenna. We will then obtain a precise three-dimensional map of the terrain and phase our array taking into account the geometry. Thus we will be able to obtain square miles of aperture at a cost much less than the large antennas presently being constructed. Twenty years from now we will be communicating with our explorers on Mars and Venus via advanced communication systems. It is not beyond reason to expect we will use microwave wavelengths to communicate to satellites which will retransmit the message at optical wavelengths to a satellite going around Venus which will retransmit at microwave wavelengths to the surface of Venus. We predict that in fifty years from now earth-to-earth communications will make use of an artificial ionosphere and replenishment rates with properties and gradients which will minimize the effect of solar activity and will be designed with communication reliability in mind. Thus the *golden years* of large reflector antennas will continue for perhaps ten to twenty years and then go the way of ancient mansions, well designed, but inefficient in a modern world. We hope this limb we have walked out on does not discourage sponsors of large antennas. There is much to be learned about the planets before we can maximize the probability of human survival on initial human landings on the moon and planets.

The Future of Propagation Research and Development*

HENRY G. BOOKER†, FELLOW, IRE

Summary—**It is predicted that the future of propagation research and development will involve: 1) world-wide radio communication via satellites, 2) greatly increased interest in the magnetosphere and the solar atmosphere, 3) greatly increased interest in the atmospheres and surfaces of planets, 4) a far wider interpretation of the notion of radio frequency, both at the lower and the upper ends of the spectrum, 5) extensive use of new methods for investigating the top side of the ionosphere by both reflection and scattering techniques, 6) the discovery of new phenomena by the use of large antennas, high-power transmitters and low-noise receivers, and, 7) greatly increased interest in the propagation of sonic and tidal waves in the atmosphere.**

PREDICTING the future in a field of research is unsafe to the point of being dangerous. However, it is an interesting intellectual exercise. That the science of radio propagation is undergoing important developments is illustrated by the reorganization that has recently taken place under the auspices of the U.S.A. Committee of the International Scientific Radio Union. A new U.S.A. Commission with the title "Magnetospheric Radio" has been established. The magnetosphere is the outermost part of the earth's plasma envelope where the distribution of charged particles is strongly controlled by the earth's magnetic field. The magnetosphere exists mainly above the level of maximum electron density in the F region. Ten years ago essentially nothing had been done to investigate the magnetosphere, whereas today it is the region where most satellites fly.

From a development standpoint the most striking advances in world-wide radio communication will undoubtedly be associated with the use of relay stations in satellites. Whether this will be true from the research standpoint, however, is more doubtful. Satellite communications will involve radio propagation problems of considerable interest extending from the lowest layers of the troposphere to the levels in the magnetosphere where relay satellites will fly. But to study these phenomena it will in all probability be better to make direct scientific investigations of the various levels of the atmosphere by methods not involving the satellite communication systems themselves.

One of the striking respects in which the field of radio propagation is expanding concerns the range of frequencies involved. Many of us can still remember the time when the radio frequency spectrum extended from something above 10 kc up to about 100 Mc or less. During World War II the frequency range was extended up to about 30,000 Mc. Now, however, we see the development of coherent sources of electromagnetic radiation at optical frequencies together with fibrous waveguides capable of handling these waves.

Nor is the extension in the range of radio frequencies solely in the upward direction. With the successful interpretation of whistlers[1] about ten years ago, we have been seriously concerned with the propagation of electromagnetic waves through the earth's atmosphere at audio frequencies. Moreover, at frequencies below the gyromagnetic frequency of atmospheric ions, these waves become the hydromagnetic waves studied originally by Alfvèn.[2] Hydromagnetic waves are now regarded as a likely feature of the geomagnetic storms that cause so much interference with radio communications. Pulsations in the earth's magnetic field are being actively investigated at frequencies well below 1 cps. It is clear that, in future, the expression "radio propagation" must cover the behavior of electromagnetic waves from optical frequencies down to a frequency of about 1 millicycle per second or less. Indeed we authors can derive some sly amusement from watching the editorial department of the PROCEEDINGS OF THE IRE wonder how to abbreviate "millicycles" to "mc" when for years they have been using this abbreviation for "megacycles."

The history of radio propagation has been punctuated by successive extensions in the range of communication beyond that originally anticipated. Except in the case of ionospheric scatter transmission,[3] these extensions in range have always been discovered experimentally, and a theoretical explanation has then followed. It is virtually certain that we may look forward to discovering a series of situations in which electromagnetic waves manage to do things not presently anticipated. Discovery of these unusual situations will be accelerated by the development of large antennas, of high-power transmitters and associated handling facilities, and of low-noise receivers.

Many of us can remember the time when all radio transmission took place below the level of maximum electron density in the ionosphere. The development of radio astronomy changed this situation somewhat, and the successful interpretation of whistlers brought about

* Received by the IRE, October 23, 1961.
† School of Elec. Engrg., Cornell University, Ithaca, N. Y.

[1] L. R. O. Storey, "An investigation of whistling atmospherics," *Phil. Trans. Roy. Soc.* (*London*), vol. 246, pp. 113–141; July, 1953.
[2] H. Alfvèn, "Cosmical Electrodynamics," Clarendon Press, Oxford, Eng.; 1950.
[3] D. K. Bailey, *et al.*, "A new kind of radio propagation at very high frequencies observable over long distances," *Phys. Rev.*, vol. 86, pp. 141–145; April, 1952.

a spectacular change. We now look forward to an era when the top side of the ionosphere will be measured and monitored by satellite-borne ionospheric recorders[4] with something approaching the detail in space and time already achieved on the under side of the ionosphere. Moreover, we can look forward to similar radar investigations of the atmospheres of the sun and the planets, and to investigations of the various plasma phenomenon of space. In addition, the same techniques will provide information about the roughness and movement of the surfaces of the planets in the manner already achieved for the moon.

History has shown that a newly acquired ability to detect weak signals is one of the ways in which research in radio wave-propagation progresses. Repeatedly we have found that what was considered the regular mode of propagation did not disappear into the noise with increase of distance but merely gave place to some other mode of propagation that had hitherto been disregarded. The discovery of tropospheric scatter propagation is an example of this.[5] The same is true of the use of backscattering from thermal fluctuations to measure the height distribution in the magnetosphere of electron density and other parameters. This method is currently under development by Gordon[6] and Bowles[7] and presages a new kind of ionospheric sounder that could be used on a world-wide scale in the future. Indeed, if we persevere, we may even succeed in comprehending the long-standing enigmas of "sporadic E" and "spread F!"

Guidance of electromagnetic waves by various natural phenomena will also be a profitable field of investigation. Guidance of microwaves by the semipermanent radio duct existing near the surface of the sea, especially over tropical oceans,[5] was discovered during World War II, but comparatively little has so far been done to make practical use of the phenomenon. Guidance of radio waves between the northern and southern hemispheres round field-aligned irregularities of ionization in the magnetosphere has recently been suggested both on theoretical[8] and experimental[9] grounds. This may prove to be an important method for scientific investigation of the magnetosphere. Even as an unusual method of radio communication such guidance cannot be completely ruled out at the present time; perfect guiding along the entire length of a tube of the magnetic field is not required if large antennas are used at both ends. What may be achieved with guidance of electromagnetic waves along ionization irregularities aligned by magnetic fields, both in the earth's atmosphere and in the sun's atmosphere, cannot easily be predicted at the present time.

A considerable field of investigation also lies ahead in connection with the propagation of sound waves and tidal waves in the earth's atmosphere. There are many refraction problems that require investigation, and new methods are now available for generating these waves with quite large amplitudes. Hines[10] has suggested that irregular motions in the lower ionosphere are simply the result of sonic noise propagated upwards from below. He is probably right, and the nonlinear interaction between these waves probably constitutes the "turbulence" of the lower ionosphere. Anyone interested in radio propagation in the earth's atmosphere will be well advised to study tidal waves, sound waves, hydromagnetic waves, electromagnetic waves and light waves over the entire frequency range from less than one millicycle per second upwards and at all levels from the earth's surface up to well above the conventional ionosphere.

It is clear that propagation in the earth's atmosphere and in space is a field of investigation that is in no danger of running dry; and even if it should, there is the whole field of plasma physics to which we can turn our attention.

[4] R. W. Kencht, *et al.*, "First pulsed radio soundings of the topside of the ionosphere," *J. Geophys. Rev.*, vol. 66, pp. 3078–3081; September, 1961.

[5] M. Katzin, *et al.*, "3- and 8-centimeter propagation n low ocean ducts," PROC. IRE, vol. 35, pp. 891–905; September, 1947.

[6] W. E. Gordon and L. M. LaLonde, "The design and capabilities of an ionospheric radar probe," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-9, pp. 17–22; January, 1961.

[7] K. L. Bowles, "Observation of vertical-incidence scatter from the ionosphere at 41 megacycles per second," *Phys. Rev. Lett.*, vol. 1, pp. 454–455; December, 1958.

[8] H. G. Booker, "Guidance of Radio and Hydromagnetic Waves in the Magnetosphere," presented at URSI Meeting, Boulder, Colo.; December, 1960.

[9] R. M. Gallet and W. F. Utlaut, "Evidence on the laminar nature of the exosphere obtained by means of guided high-frequency wave propagation," *Phys. Rev. Lett.*, vol. 6, pp. 591–594; June, 1961.

[10] C. O. Hines, "Internal atmospheric gravity waves at ionospheric heights," *Canad. J. Phys.*, vol. 38, pp. 1441–1481; November, 1960.

Section 3

AUDIO

Organized with the assistance of the IRE Professional Group on Audio

A Century of Microphones*

B. B. BAUER†, FELLOW, IRE

Summary—**Of the various manifestations of a sound wave, the action of pressure on a diaphragm still is the universal means for detecting the presence of sound. The diaphragm actuates a transducer converting its motions into equivalent electrical waves. Innumerable transducers have been tried, but five are pre-eminent: 1) carbon, 2) condenser, 3) piezoelectric, 4) moving conductor, 5) moving armature.**

Important microphone improvements during the late twenties and the thirties have come about as a result of application of equivalent circuit analysis to acoustical structures. The principle of pressure microphones, pressure gradient microphones, combination microphones and phase shift microphones are described. Each of these has found an important niche in modern microphone applications.

A small number of important applications require super-directional microphones. Here three approaches are used: 1) reflectors, refractors and diffractors, 2) line microphones, 3) higher-order combination microphones.

In the future, improvements in design of directional microphones will continue. Wireless microphones are bound to increase in popularity. New methods of transduction based upon solid-state technology appear to be imminent. Unconventional methods of sound pickup may find wide use in space communication.

* Received by the IRE, January 2, 1962.

† CBS Laboratories, Stamford, Conn.

Introduction

As a SENSOR which transforms sound into energy form suitable for amplification and transmission, a microphone is among the most common and useful technological servants of mankind. At this writing, a century of effort has been devoted to inventing and perfecting the microphone. This paper is intended to provide a record of the basic contributions made during that time as well as to survey the engineering principles employed in the present day microphones. A brief look into the future will also be attempted.

Plan of This Paper

From the scientific point of view a microphone may be designed to sense any of the manifestations of the sound wave and to convey it to a transducer which will transform it into electrical energy. A sound wave is accompanied by the presence of an alternating excess pressure called the sound pressure p; the particles of air are subject to a to-and-fro motion which may be described by their velocity u and since the medium follows the adiabatic law there exists an alternating change in temperature as well as corresponding changes in density, dielectric constant, magnetic susceptibility and index of refraction. This paper is confined to those microphones in which the sound pressure or sound-pressure gradient are transformed into a force F by use of a diaphragm which, together with an associated electromechanical transducer, is set into motion resulting in generation of electricity. This is the method employed in earliest microphones, and it is virtually the universal method for microphone operation today. Because of their importance to proper understanding of microphones, brief descriptions of typical diaphragms and their interaction with the medium have been included in this paper.

Some of the other functions of a sound wave that have found significant but limited application in microphones are 1) the combined action of the particle velocity and the alternating temperature upon a heated fine wire[1] and 2) the combined action of pressure and particle velocity upon a cloud of ions.[2,3] Other possibilities have been considered: the change in dielectric constant or magnetic susceptibility of the air could be used to modulate the frequency of an oscillator;[4] the varying refractive index may be caused to modulate a light beam,[5] for example. Some of these functions may hold a key to microphone developments of the future.

Every conceivable means of electro-mechanical transduction has been combined with the vibrating diaphragm in an effort to produce "new and better" microphones. In this paper five basic transducers are described, any one of which will be found in virtually all of the present-day microphones: 1) Loose contact (carbon), 2) Electrostatic (condenser), 3) Piezoelectric (Rochelle Salt and Ceramic), 4) Moving conductor (moving coil dynamic and ribbon), 5) Moving armature (magnetic or reluctance). Many other means of transduction have been studied, tested and patented, such as, variable fluid contact,[6] movable vacuum tube elements,[7] piezoresistivity[8] point-contact transistors,[9] etc. To this date, these have not been widely adopted, but again these and newer methods of transduction may become important in future microphones.

[1] G. Forbes, "A thermal telephone transmitter," *Proc. Roy. Soc. (London) A*, vol. 42, pp. 141–142; February 24, 1889.

[2] Early experiments are described in a paper by W. Duddel, "Rapid variations in the current through the direct-current arc," *The Electrician*, p. 271; December 14, 1900. Duddel credits the discovery to Simon whose experiments are recorded in *Ann. der Phys.*, vol. LXIV, No. 2, pp. 233–239; 1898. Also see L. de Forest, Brit. Patent 5258; 1906.

[3] S. Klein's ionophone described by J. C. Axtell, "Ionic loudspeakers," IRE Trans. on Audio, vol. AU-8, pp. 21–27; July, 1952.

[4] This possibility has come to the author's attention from time to time but it does not appear to have been explored.

[5] L. de Forest, U. S. Patent No. 1,726,299; 1924.

[6] A. G. Bell, March 10, 1876. See H. A. Frederick, "The development of the microphone," *J. Acoust. Soc. Am.*, vol. 3, pt. 2, p. 5; July, 1931.

[7] H. F. Olson, "Mechani-electronic transducers," *J. Acoust. Soc. Am.*, vol. 19, pp. 307–319; March, 1947.

[8] F. P. Burns, "Piezoresistive semiconductor microphone," *J. Acoust. Soc. Am.*, vol. 29, pp. 248–253; February, 1957.

[9] R. L. Hanson, "Transistor microphone," U. S. Patent No. 2,497,770; 1950.

Among the scientific tools of radio engineering, none has contributed as much to microphone development as the application of electrical circuit analysis to electroacoustical structures.[10] In employing the principles of this analysis, the operation of microphones is better understood and the groundwork is laid for future developments. It will be seen, for example, that some of the foregoing transducers are ***displacement responsive*** and others are ***velocity responsive*** (these terms arising from the generated voltage being dependent on the amplitude of displacement or the velocity of the diaphragm). Equivalent circuit analysis shows how to proportion microphone structures best to utilize these characteristics.

Diaphragms

Earliest among microphone diaphragms—perhaps because of its similarity to the eardrum—was a stretched flat membrane (actually a sausage skin) used by Reis[11] to actuate a loose metal-to-metal contact. A stretched flat membrane [Fig. 1(a)] made of metal or very thin metallized plastic is used in present-day electrostatic microphones. This diaphragm is typically clamped at its periphery by a ring 1.1 and stretched to any desired tension by a threaded ring 1.2.

The cross-sectional shape as a function of radius r taken on by a circular membrane of radius a made of nonrigid material uniformly stretched with tension T and loaded with a uniformly distributed pressure P is a paraboloid of revolution described by the equation

$$y = (Pa^2/4T)(1 - r^2/a^2) = y_{max}(1 - r^2/a^2), \qquad (1)$$

where y_{max} is the central or maximum displacement.[12] This equation is of interest since a stretched diaphragm used with condenser microphones commonly is subjected to uniform force of electrostatic attraction.

A flat diaphragm clamped between rings 1.3 and 1.4 or the like, is illustrated in Fig. 1(b). Used in early telephone receivers, this type of diaphragm is desirable where a great unbalanced pressure must be supported, but it also may be found in a modern electrostatic microphone.[13]

A more common arrangement for a flat plate diaphragm is shown in Fig. 1(c), where the diaphragm is held against a circular support edge 1.5 by a steady-state force, *e.g.*, magnetic attraction of transducer in a telephone receiver.

The previously described diaphragms are best adapted to drive a distributed acoustical load. When it

[10] For an excellent treatise see H. F. Olson, "Dynamical Analogies," D. Van Nostrand Co., Inc., New York, N. Y.; 1943. Also F. A. Firestone, "Twixt earth and sky with rod and tube; the mobility and classical impedance analogies," *J. Acoust. Soc. Am.*, vol. 28, pp. 1117–1153; November, 1956.

[11] J. P. Reis, "Ueber Telephone Durch Den Galvanischen Strom," Jahresbericht d. Physikalischen Vereins zu Frankfurt am Main, Germany, pp. 57–64; 1860–1861.

[12] I. B. Crandall, "Theory of Vibrating Systems and Sound," D. Van Nostrand Co., Inc., New York, N. Y., p. 20; 1927.

[13] J. K. Hilliard, "Miniature condenser microphone," *J. Soc. Mot. Pic. Telev. Engrs.*, vol. 54, pp. 303–314; March, 1950.

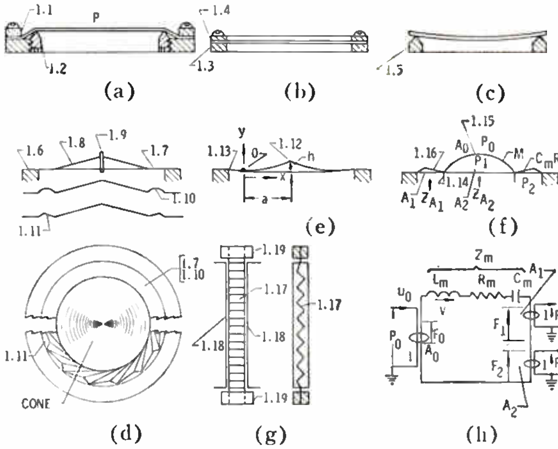


Fig. 1—Various types of diaphragms used in microphones.

becomes necessary to actuate a mechanical system from a point or a line, the diaphragm usually takes on a different shape so as to present an adequate *driving-point impedance* to the load.

Very commonly used for a point-drive is a cone diaphragm shown in three versions in Fig. 1(d). The edge of the diaphragm is effectively clamped or cemented against some support 1.6 leaving an annular portion 1.7 to flex in response to motions of the conical portion 1.8. The latter actuates a transducer through a drive-rod 1.9. The flat annulus gives way to a formed or corrugated annulus 1.10 when linearity of motion and freedom from spurious resonances at high frequency is required. A major advance in annulus design was achieved by Harrison[14] who invented a tangentially-corrugated annulus 1.11, shown at bottom of Fig. 1(d), and which is used at present in many moving-coil microphones and horn loudspeaker drivers.

A "curvilinear" diaphragm developed by the author for use with piezoelectric microphones a quarter of a century ago is now widely used with various point-drives. The goal is to provide a "nonbuckling" shape, that is, one that normally would be assumed by a pie-slice segment of a diaphragm supported at its apex and the circumferential edge and subjected to uniform pressure at one side. The desired shape may be defined approximately by the following equation:

$$y/h = (3/2)(x/a)^2 - (1/2)(x/a)^3, \qquad (2)$$

where the lowest point of the draw is at the origin *o* and *h* is the height at the apex 1.12. The contour may rise both toward the apex and toward the edge of support 1.13.

A "piston" diaphragm shown in Fig. 1(f) is practically universally used with moving-coil microphones and other transducers where force is transmitted at the circular line around the rim to a coil 1.14. The central portion of the "piston" 1.15 is of spherical shape. The annulus 1.16 commonly is tangentially corrugated after Harrison.

A ribbon diaphragm which also is a transducer was invented by Gerlach.[15] As used in a pressure-gradient microphone invented by Olson,[16] this transducer is made of corrugated aluminum ribbon 1.17 less than 0.0001-inch thick which either floats freely or is slightly stretched between two pole pieces 1.18. Electrical connections are made at the supports 1.19 to a high turns-ratio transformer.

From equivalent circuit point of view, the action of a diaphragm may be represented by Fig. 1(h).[17] The mechanical elements of the diaphragm, *i.e.*, mass M, compliance C_m and internal damping resistance R_m appear in the circuit as equivalent electrical elements to which forces derived from acoustical pressures are coupled by means of ideal $1:A$ transformers. The relationships between the pressure p and the force F developed upon an area A, and the volume velocity u and the linear velocity V resulting therefrom are correctly portrayed by the use of a transformer coupler, as may readily be verified from transformer equations. Normally the net areas on both sides of the diaphragm are equal, so that only two transformers (one for each side of the diaphragm) will be required. In the case of moving coil microphones the two sub-areas of the diaphragm in 1(f) separated by the coil from 1–14 are subjected to different pressures and are confronted by different acoustical impedances. In this case each independently acting area must be represented by its own coupling transformer. These transformers are merely aids to correct circuit analysis representation. Usually they can be deleted in the actual experimental circuit work.

Loose-Contact Transducers

Among the earliest devices intended for converting vibration into electrical impulses was Reis' loose metal-contact transducer[11] which is reported to have transmitted tones of different frequencies, but not intelligible speech. This latter event seems first to have been achieved by Bell, using a magnetic microphone, on June 3, 1875.[18] However, Bell's microphone proved not to be sufficiently sensitive for telephone work, and the experiments of Berliner,[19] Edison,[20] Hughes[21] and others soon thereafter introduced a long era of dominance for the loose-contact carbon transducer. To Edison goes the credit of being the first to design a transducer using granules of carbonized hard coal,[22] still used in present-day microphones.

[14] J. P. Maxfield and H. C. Harrison, "Methods of high quality recording and reproduction of music and speech based on telephone research," *Trans. AIEE*, vol. 45 (*Commun. and Electronics*, pp. 334–348; February, 1926.

[15] E. Gerlach, German Patent 421,038; 1925.

[16] H. F. Olson, U. S. Patent No. 1,885,001; 1932.

[17] B. B. Bauer, "Transformer analogs of diaphragms," *J. Acoust. Soc. Am.*, vol. 23, pp. 680–683; November, 1951.

[18] See, for example, F. H. Frederick, "The development of the microphone," *J. Acoust. Soc. Am.*, vol. 3, pt. 2, p. 3; July, 1931. Also, Alexander Graham Bell, U. S. Patent No. 174,465; 1876.

[19] E. Berliner, Caveat filed in U. S. Patent Off., April 14, 1877.

[20] T. A. Edison, U. S. Patent No. 474,230; filed April 27, 1877. Also, U. S. Patent Nos. 474–231–2.

[21] D. E. Hughes, "On the action of sonorous vibrations in varying the force of an electric current," *Proc. Roy. Soc .(London) A*, vol. 27, pp. 362–369; May 9, 1878.

[22] T. A. Edison, U. S. Patent No. 406,567; July 19, 1889.

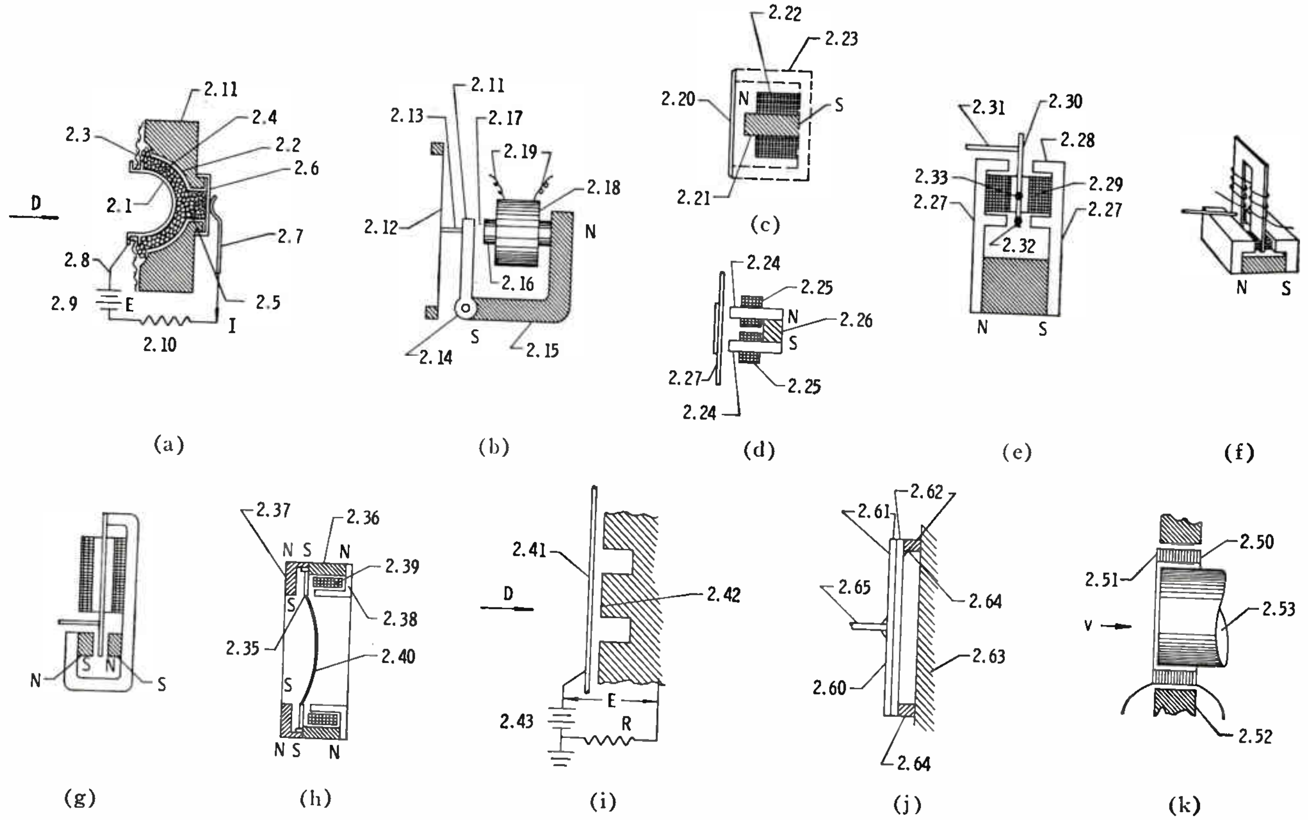


Fig. 2—Various types of transducers used in microphones.

The carbon granules are made of deep-black "anthraxylon" coal ground to pass a 60–80 mesh, treated chemically and roasted in several stages under a stream of hydrogen. This drives out volatile matter, washes out extraneous compounds, and carbonizes the coal. The last step of the process is magnetic and air-stream screening to eliminate iron-bearing and flat-shaped particles.[23]

Referring to Fig. 2(a), the modern carbon-granules transducer[24] is comprised of gold-plated metallic cups 2.1 and 2.2 attached to a diaphragm (not shown) and to a stationary back-plate 2.11, respectively. A fabric washer 2.3 encloses the carbon cavity which is filled with granules 2.4 through an aperture 2.5 capped with a contact 2.6. Leads 2.7 and 2.8 complete the circuit with a polarizing source of current 2.9 and a load impedance 2.10. Frequently, the load impedance is a primary winding of a step-up transformer. Variations of transducer resistance stemming from displacement D modulate the current I in the circuit. The incremental voltage developed across the load is proportional to displacement D.

The carbon-granules transducer not only has the distinction of being most widely used in microphones—every telephone in the world has one—but also of constituting its own amplifier of some 40- to 60-db gain. Its disadvantages are relatively high-noise level and distortion, instability caused by variation of the contact resistance of the granules with position and degree of packing, and a loss of sensitivity or "aging" under action of vibration. With the advent of economical and efficient solid-state amplifiers, the importance of the carbon transducer is bound eventually to diminish.

A brief mention should be made of a stretched diaphragm push-pull dual-button carbon transducer used in the early days of broadcasting because of its distortion-cancelling properties.[25] This microphone became outmoded during the early thirties as a result of advances in other types of microphones aided by electronic amplification.

Moving Armature Transducer

While claiming a record of first successful use for intelligible voice transmission, the "magnetic" transducer also can point with pride to continued service since its inception, principally in telephone receivers, and more recently again in microphones. Many different magnetic transducers have been designed. The type described in Bell's first patent application[18] of 1876 is shown in Fig. 2(b). An armature 2.11 is connected to a diaphragm 2.12 by a drive pin 2.13. The armature is hinged at 2.14 to a

[23] Production of carbon granules appears to a degree to be a "trade secret" but see, for example, J. R. Fisher, "Coal for transmitters," *Bell Labs. Rec.*, vol. 10, pp. 150–154; January, 1932. See also, W. E. Orvis, "Coal talks," *Bell Labs. Rec.*, vol. 10, pp. 200–204; February, 1932.

[24] W. C. Jones, "Instruments for the new telephone sets," *Trans. AIEE*, vol. 57 (*Commun. and Electronics*), pp. 559–564; October, 1938.

[25] W. C. Jones, "Condenser- and carbon microphones—their construction and use," *Bell Sys. Tech. J.*, vol. 10, pp. 46–62; January, 1931.

yoke 2.15. The yoke bears a pole-piece 2.16 forming an airgap 2.17 and carrying a coil 2.18 with terminals at 2.19. Bell's original idea was to interconnect two such transducers by means of a transmission line and a battery in the circuit which polarized the electro-magnets of both transducers. The generated signal voltage is proportional to armature velocity.

In 1877 Bell patented a notable improvement to the above structure in which he used a permanent magnet for purposes of polarization.[26]

The transducer of Bell is used to this day in telephone receivers in two modified forms shown here for reference. The one in Fig. 2(c) employs a combination diaphragm-armature 2.20 and a permanent magnet pole-piece 2.21 surrounded by a coil 2.22. A magnetic return cup 2.23 often is provided. A bi-polar form in Fig. 2(d) employs two pole-pieces 2.24, each provided with a coil 2.25 and a common permanent magnet 2.26. A Permindur pole-shoe 2.27 helps to carry the steady-state flux of the magnet. The above units have not been successful as microphones because the moving member requires sufficient heft to carry unbalanced dc flux and to support the steady-state forces which it produces.

A magnetically balanced-armature transducer, useful in microphones, was suggested by Siemens[27] and Watson,[28] but more definitely projected by Capps.[29] Shown in Fig. 2(e) an armature 2.30 within the coil carries the *differential flux only* stemming from motions imparted to it by the drive pin 2.31 connected to a diaphragm (not shown). The armature may be pivoted at a point 2.32 which results in a mechanically unbalanced structure, or at a point 2.33 which produces mechanical, as well as magnetic, balance.

In an attempt to dissociate as much as possible the steady state and the ac flux paths, the magnetic structure and the armature may be deformed, topologically speaking, until a straight-line pole-piece structure and a U-shaped armature form has been obtained, with great economy of dimensions.[30] This structure shown in Fig. 2(f) has found wide use in transistorized hearing aids in which miniaturization has become a most important virtue. A variation is shown in Fig. 2(g).

An improvement heretofore applied to a telephone receiver but with possible use in microphones is shown in Fig. 2(h).[31] In this arrangement, a ring armature 2.35 is maintained in an unsaturated condition by two ring magnets 2.36 and 2.37. The circuit comprising the alternating flux path includes a circular pole-shoe 2.38 and a circular coil 2.39. This transducer is well adapted to being driven by a piston diaphragm 2.40.

Magnetic transducers are characterized by the presence of a negative force-displacement function at the airgaps which has the dynamical form of ***negative stiffness***. The magnetization and saturation properties of the armature must be proportioned in such manner that the mechanical restoring stiffness of the armature and diaphragm are greater than the magnetic negative stiffness.[32]

Electrostatic Transducers

While Edison[33] and Dolbear[34] proposed the use of electrostatic transducers very early in this history of electroacoustics, it remained for Wente[35] to develop an electrostatic microphone that was truly a precision instrument. An electrostatic transducer is shown in schematic view in Fig. 2(i). A stretched flat conductive membrane 2.41 is arranged at a distance x from a back plate 2.42 defining an active area A. This produces a capacitor having a capacitance $C=kA/x$, where k is the dielectric constant of air. A polarizing potential difference E is provided from a source 2.43, connected to the electrodes through a very high resistance R. Thus a quasi-constant charge Q_0 is established on the capacitor where $Q_0=CE=kAE/x$. Solving for E,

$$E = (Q_0/kA)x. \tag{3}$$

Therefore the voltage across the condenser will vary linearly with the diaphragm displacement x. Because of the low-loss air dielectric, a capacitor transducer potentially is an extremely quiet, linear device.

In practice the spacing x is of the order of 0.001 inch, and the capacity of the microphone is around 25–50 $\mu\mu$f. Therefore, a very high impedance preamplifier at the transducer is required. The electrostatic transducer is most often used where highest quality is sought regardless of cost and inconvenience caused by the integral preamplifiers such as in recording and calibrations work. An electrostatic transducer also is characterized by presence of negative stiffness.

Moving Coil Transducer

While the early efforts and concepts in connection with moving-coil transducers are associated with the names of Cuttriss, Redding[36] and Siemens,[37] the credit for developing a wide-range practical moving-coil microphone goes to Wente and Thuras.[38] Referring to Fig.

[26] A. G. Bell, U. S. Patent No. 186,787; 1877.
[27] E. W. Siemans, German Patent No. 2355; 1878.
[28] T. A. Watson, U. S. Patent No. 266, 567; 1882.
[29] F. L. Capps, U. S. Patent No. 441,396; 1890.
[30] B. B. Bauer, U. S. Patent No. 2,454,425; 1948.
[31] E. E. Mott and R. C. Miner, "The ring armature telephone receiver," *Bell Sys. Tech. J.*, vol. 30, pp. 110–140; January, 1951.

[32] For example, see B. B. Bauer, "A miniature microphone for transistorized amplifiers," *J. Acoust. Soc. Am.*, vol. 25, pp. 867–869; September, 1953. In all transducers (other than loose-contact) any utilization of the generated electrical energy causes a reaction upon the transducer's mechanical impedance. In microphones, this effect is usually small, and is beyond the scope of this paper.
[33] Reported in G. B. Prescott, "The Speaking Telephone, Talking Phonograph and Other Novelties," D. Appleton and Co., New York, N. Y.; 1878.
[34] A. E. Dolbear, U. S. Patent Nos. 239,742 and 240,578; 1881.
[35] E. C. Wente, "A condenser transmitter as a uniformly sensitive instrument for the absolute measurement of sound intensity," *Phys. Rev.*, vol. 10, pp. 39–63; July, 1917.
[36] C. Cuttriss and J. Redding, U. S. Patent No. 242,816; 1881.
[37] E. W. Siemens, German Patent No. 2355; 1878.
[38] E. C. Wente and A. L. Thuras, "Moving coil telephone receivers and microphones," *J. Acoust. Soc. Am.*, vol. 3, pp. 44–55; July, 1931.

2(k), the moving-coil is circular in shape and is attached at the rim 2.51 to a diaphragm (not shown) being supported and centered thereby in an air gap between pole-pieces 2.52 and 2.53. If the length of the conductor in the air gap is l and the flux density is B, and the diaphragm velocity is v, the voltage generated in the coil is

$$E = Blv \tag{4}$$

and hence, a moving-coil transducer is a velocity-responsive device.

Wente and Thuras based the development of their microphone and receiver upon equivalent circuit analysis. It is interesting to note that the circuit developed by them for a moving-coil receiver (where the goal is constant diaphragm displacement as a function of input voltage) is identical with the circuit to be used for a displacement-responsive (*e.g.*, ceramic) microphone.

A moving coil transducer is sensitive, rugged, provides good frequency-response and low noise, and at present is the "workhorse" among the microphones used for broadcasting and public-address applications. The low-coil impedance is suitable for operation with long cables, followed with a step-up transformer at the pre-amplifier, although in many microphones, a built-in transformer provides the proper impedance transformation right at the microphone.

Piezoelectric Transducer

In 1820 Becquerel described and observed piezoelectric effects[39] although a systematic study leading to modern understanding of these effects is credited to the Curies.[40] Nevertheless, piezoelectric microphones had not become practical until the invention of the "bimorph" Rochelle Salt transducer by Sawyer.[41] The "bimorph" ushered a quarter of a century era of dominance for Rochelle Salt crystals in low cost microphones, which (because of relatively poor stability of Rochelle Salt in severe climates) subsequently was to be challenged by polycrystalline Barium Titanate ceramics of Gray[42] and more recently by Lead Zirconium Titanate ceramics of Jaffe.[43] The shapes taken on by these bodies is that of a sandwich designated 2.60 in Fig. 2(j) consisting of two slabs of piezoelectric material 2.61 and 2.62 which are joined into an integral unit with appropriate electrodes. The element is attached to a reference frame 2.63 by raised portions 2.64, and driven by means of a drive-unit 2.65 which is connected to a diaphragm. The tension and compression in the beam combine with the polarization mode of the individual slabs to produce a potential difference in the electrodes as a function of displacement. Rochelle Salt bimorphs are available for actuation by torsion or bending stresses, while the ceramic units are usually of the latter variety.

Microphone Structures

Hunt[44] refers to the 1870's as "vintage years for electroacoustics." The inventions of the telephone and phonograph together with innumerable transducers to implement them occurred during those years. In the same manner, the 1930's, having seen the development of many a modern microphone, may be thought of as vintage years for microphones.

In studying the historical development of microphones it becomes evident that control over their directional capabilities has become increasingly important with time. In the following sections, we describe 1) *pressure microphones* which respond to sound pressure at one exposed surface of the diaphragm and (because sound travels around corners) are more or less equally sensitive from all directions; 2) *gradient* or *pressure-difference microphones* in which the diaphragm is exposed for differential action by sound pressure equally at both surfaces to achieve a bi-directional operation; 3) *combination microphones* which unite pressure and gradient concepts to achieve unidirectional action and 4) *phase-shift microphones* which achieve unidirectional action with a single transducer and acoustical phase-shift networks. Diaphragms and transducers in endless combinations have been brought together to produce a multitude of such microphones, but only a few basic examples can be given here.

Pressure Microphones

The electrostatic microphone of Wente[35] is one of the simplest, and, with modern refinements, one of the most effective of microphones. Its basic form is shown in Fig. 3(a). The microphone is composed of a flat stretched conductive diaphragm 3.1 attached to a box 3.2 so as to expose one surface to the external sounds. A stationary electrode 3.3 inside the box, placed close to the diaphragm, forms the electrostatic transducer.

The equivalent network analogy of the electrostatic microphone is shown in Fig. 3(b). Sound pressure acts upon the diaphragm through an air-load radiation impedance portrayed by an inductance L_a in parallel with a resistance R_a.[45] Thin films of air between the diaphragm and the electrode are squeezed in and out as the diaphragm vibrates to-and-fro resulting in damping action, the collective effect of which is represented by R_b and L_b. The fluid motion finds its way into the volume

[39] A. C. Becquerel, *Bulletin des Sciences*, par la Societe Philomatique de Paris, France, vol. 7, pp. 149–155; March, 1820.

[40] J. and P. Curie, *Bulletin de la Societe Mineralogique de France*, vol. 3, pp. 90–93; April, 1880.

[41] C. B. Sawyer, "The use of Rochelle Salt crystals for electrical reproducers and microphones," Proc. IRE, vol. 19, pp. 2020–2029; November, 1931.

[42] Grey, U. S. Patent No. 2,486,560; 1949.

[43] Jaffe, U. S. Patent No. 2,708,244; 1955.

[44] F. V. Hunt, "Electroacoustics," Harvard University Press, Cambridge, Mass., p. 37; 1954.

[45] Means of approximating air-load impedance with fixed elements is an important tool in the bag of tricks of the electroacoustician. For example, see B. B. Bauer, "Notes on radiation impedance," *J. Acoust. Soc. Am.*, vol. 15, pp. 223–224; April, 1944; R. C. Jones, "A fifty horsepower siren," *J. Acoust. Soc. Am.*, vol. 18, pp. 371–387; October, 1946; F. B. Hunt, *op. cit.*, p. 158.

of the recesses 3.4 which taken together define an acoustical compliance C_b.

A simplified equivalent circuit is obtained by dividing the mechanical impedance of the diaphragm by the square of the area A, which allows the elimination of ideal $1:A$ transformers [Fig. 3(c)]. The input voltage E_p replaces the sound pressure p. Since diaphragm displacement D is equivalent to the charge on a condenser $Q=CE$, it is a requirement in the equivalent circuit that the voltage E_0 remain invariant with frequency for constant E_p. This result is achieved if the combined series compliance of the diaphragm (C_mA^2) and of the volume 3.4 (C_b) comprise the controlling circuit impedance. In the microphone of Wente this condition was obtained by stretching the diaphragm to a high resonance frequency. A similar effect may come about by reducing the dimensions of C_b until the spring of the air becomes the controlling factor.[46] R_b and L_b are selected to damp the diaphragm resonance, to provide a "flat" response at high frequency.

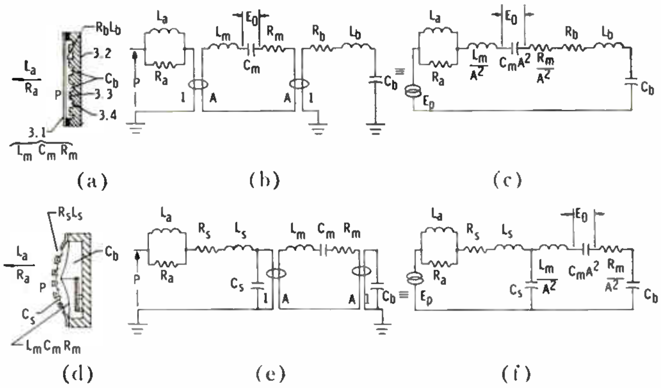


Fig. 3—Pressure microphone with displacement responsive transducer.

A piezoelectric microphone[41] in Fig. 3(d) is very similar in its equivalent circuit to the electrostatic microphone, except that a damping screen defining an acoustical resistance R_s and inertance L_s are added in the structure. The volume C_s between the screen and the diaphragm forms a part of the equivalent circuit mesh in Figs. 3(e) and 3(f). The mass, compliance, and resistance of the piezoelectric element and the diaphragm are lumped together and represented as L_m, C_m, R_m, respectively. A damping element or screen may be placed behind as well as in front of the diaphragm.

A different approach is taken in designing pressure microphones which use *velocity-responsive* transducers. Among the most elegant is the pressure-microphone portion of a combination microphone described by Olson in 1932.[47] In schematic cross section this microphone is shown in Fig. 4(a), its equivalent electrical network in Fig. 4(b), and a simplified version in 4(c). Because it is a design objective to make the *velocity* of the transducer invariant with frequency (and the electrical circuit counterpart of velocity is the current I), the circuit must be resistance controlled. This is achieved by the expedient of making all the mechanical and acoustical impedances small compared with the termination resistance R_b. The latter is obtained with a pipe or labyrinth filled with tufts of felt.

A more modern version of this microphone was described by Olson and Preston in 1950.[48] In this unit a pickup probe in the form of a small horn was added to the microphone to enhance its high-frequency response.

A major advance in pressure microphone design was

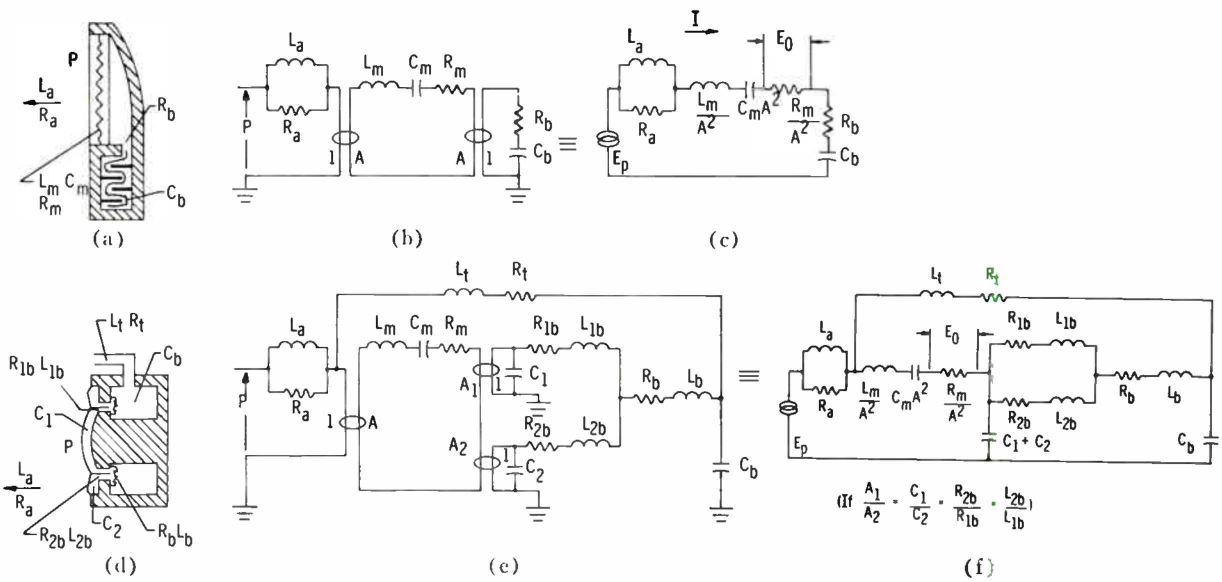


Fig. 4—Pressure microphone with velocity responsive trandsucer.

[46] The pressure-operated mode of Von Braunmühl and Weber microphone to be described is probably air-stiffness controlled, but also see T. J. Schultz, "Air-stiffness controlled condenser microphone," *J. Acoust. Soc. Am.*, vol. 28, pp. 337–342; May, 1956.

[47] H. F. Olson, "A unidirectional ribbon microphone," (abstract only), *J. Acoust. Soc. Am.*, vol. 3, p. 315; January, 1932.

[48] H. F. Olson and J. Preston, "Unobtrusive pressure microphone," *Audio Engrg.*, vol. 34, pp. 18–20; July, 1950.

achieved by Wente and Thuras[38] with the invention of a moving-coil microphone shown in Fig. 4(d). The equivalent circuit is given in Fig. 4(e), and a simplified circuit in 4(f). It should be noted that the acoustical compliances C_1 and C_2 confronting the two portions of the diaphragm are interconnected by the acoustic impedances R_{1b}, L_{1b}, R_{2b}, L_{2b} defined by the circular slits between the coil and the magnetic structure, which form a resonant circuit capable of producing spurious response. Wente and Thuras encountered this problem and solved it by addition of a separate internal circuit mesh. However, by choosing the acoustical constants in accordance with equations at the bottom of Fig. 4(f),[17] the spurious resonance is prevented, and the simplified circuit 4(f) then correctly portrays the operation of the microphone.

Gradient Microphones

A number of illustrations and patent drawings of early microphones show diaphragms open on both sides for access to the sound waves, and in the early part of the century Pridham and Jensen[49] and Meissner[50] invented noise-cancelling microphones in which access at both sides of the diaphragm was provided for entry of noise, with preferential access on one side for speech sounds. Notwithstanding, the invention of a pressure-gradient ribbon microphone ("ribbon velocity microphone") by Olson[16] was an outstanding contribution to the art. This microphone is shown in schematic elevation in Fig. 5(a) and in plan cross section in Fig. 5(b). Olson designed the pole-pieces 5.1 and 5.2 to form a small baffle with an effective front-to-back air path d equal to $\frac{1}{2}$ the shortest wavelength of sound to be received. A ribbon transducer 5.3 was installed for free motion therebetween. The resulting action is shown by the phasor diagram in Fig. 5(e). The front and back pressures, designated as p_1 and p_2 are displaced in phase by an angle $(\omega d/c)\cos\theta$.

The equivalent circuit of the ribbon microphone is shown in Fig. 5(c), and a simplified circuit in 5(d). It is noted that by making the mechanical compliance of the ribbon sufficiently large, and the damping resistance sufficiently small, the inductive (mass) elements will become controlling. Lumping these elements into a single constant L, the acoustic impedance of the transducer may be expressed as $j\omega L$. Therefore, the velocity v (and consequently the output voltage E_0) is expressed by:

$$\begin{aligned} v &= j(\omega d/c)pA\cos\theta/j\omega L \\ &= (d/cL)pA\cos\theta. \end{aligned} \tag{5}$$

The output of a ribbon microphone, therefore, is in phase with the sound pressure, invariant with frequency, and proportional to the cosine of the angle of sound incidence. The polar response is the "cosine" pattern $\rho=\rho_{max}\cos\theta$ shown in Fig. 5(f). The power response to random sounds for this pattern is $\frac{1}{3}$ or the response of an omnidirectional (circular) pattern exhibited by pressure microphones.

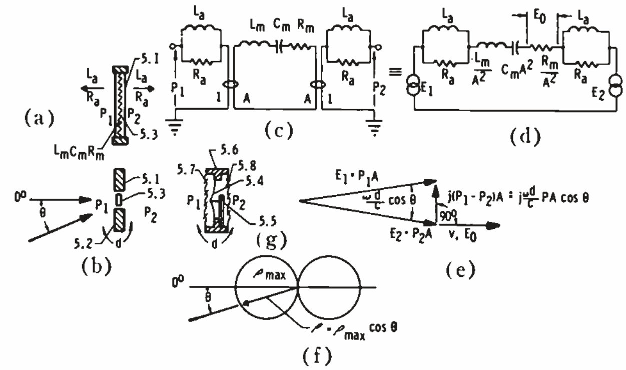


Fig. 5—Pressure gradient microphones.

A piezoelectric pressure-gradient microphone also can be constructed[51] as shown in schematic cross section in Fig. 5(g). A diaphragm 5.4 and a transducer 5.5 are housed in a round casing 5.6 with access to the atmosphere through damping screens 5.7 and 5.8. While this microphone had little commercial importance, it served as a stepping stone in the discovery of phase-shift microphones, later to be described.

Combination Microphones

The invention of the ribbon gradient microphone provided the necessary tool for the creation of a unidirectional microphone.[47,52] Such a microphone is shown in schematic cross section in Fig. 6(a). Two ribbons are provided with a common supporting frame 6.1. The ribbon 6.2 is freely accessible on both sides to form a pressure-gradient element with directional pattern expressed by the equation $\rho=\cos\theta$. The ribbon 6.3 is terminated by a damped pipe to form a nondirectional pressure microphone with directional pattern expressed by the equation $\rho=1$. Adding the two in equal half-and-half proportions produces a polar pattern $\rho=0.5+0.5\cos\theta$, which is a heart-shaped pattern or "cardioid." (The latter is a special case of the more general limaçon pattern $\rho=(1-k)+k\cos\theta$. The resulting directional characteristics are shown in Fig. 6(b).

A similar principle was employed by combining a piezoelectric pressure-gradient microphone with a piezoelectric pressure microphone to produce a cardioid pattern.[53] This latter unit incorporated a switch for selective choice of any of the three patterns. By combining a ribbon pressure-gradient with a moving coil pressure microphone, Marshall and Harry produced a very superior unidirectional microphone and endowed it

[49] Personal communication from the late P. Jensen.

[50] B. F. Meissner, U. S. Patent No. 1,507,081 (1924), filed March 12, 1919. In a recent personal communication, Meissner recounts attempts at intercommunication in open cockpit planes in 1916–17, leading to removal of back case from a Baldwin earphone (used as microphone) to provide equal noise access to both sides of the diaphragm.

[51] B. Baumzweiger (Bauer), U. S. Patent No. 2,198,424 (1940); filed November 4, 1937.

[52] T. Weinberger, H. F. Olson and F. Massa, "A unidirectional ribbon microphone," *J. Acoust. Soc. Am.*, vol. 5, pp. 139–147; October, 1933–1934.

[53] B. Baumzweiger (Bauer), U. S. Patent No. 2,184,247 (1939); filed December 20, 1937.

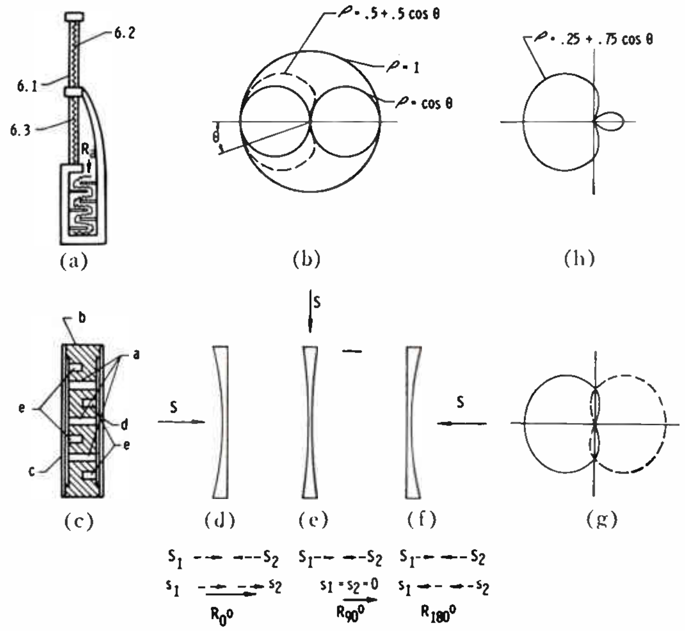


Fig. 6—Combination unidirectional microphones.

with six directional patterns.[54] It is to be noted that the pattern $\rho = 0.25 + 0.75 \cos \theta$, shown in Fig. 6(h) provides the lowest random energy pickup in the limaçon family: $\frac{1}{4}$ that of an omnidirectional pattern, while the pattern $\rho = 0.37 + 0.63 \cos \theta$ provides the greatest front-to-total random ratio of 93 per cent.[55]

The above microphones suffer from axial dissymmetry and production difficulty in the matching of two dissimilar units. In the newer designs they have been outmoded by simpler and more effective "phase-shift" microphones, later to be described.

A brilliant combination microphone based on the electrostatic principle was described by Von Braunmühl and Weber.[56] The principle of this microphone, according to the inventors, is as follows [Fig. 6(c)]: A brass circular body member b is provided with a series of holes *a through* the member and another series of holes *e part way through*. Two diaphragms c and d are fastened at the sides of the body, forming two electrostatic transducers. Assume first the condition of sound arriving at 90°. The sound pressure will merely push both membranes to-and-fro against the stiffness of the diaphragms and the air trapped within the body openings, by equal amounts denoted by the arrows S_1 and S_2 in Fig. 6(e). Now, let the sound arrive from the 0° direction; an additional pressure-gradient component will push both diaphragms and the air as a body (owing to the interconnection through the holes a) against the resistance of the film of air trapped between the diaphragm and the electrode faces. This latter effect is denoted by arrows s_1 and s_2. If the friction factor and the stiffness factors are suitably chosen, $S_1 = s_1$ and $S_2 = s_2$ and therefore only the front diaphragm c will move. By the same token, for sounds arriving from 180° direction only the rear diaphragm will be set into motion as shown in Fig. 6(f). The polar pattern exhibited by the front diaphragm, used by itself, will be a cardioid shown in solid lines in Fig. 6(g). If both diaphragms are connected in parallel, then the polar response of the combination will be omnidirectional or circular. While not so stated by the inventors, it is almost axiomatic that the rear diaphragm, by itself, will produce a reverse cardioid shown by the dotted line in Fig. 6(g); and if both diaphragms are oppositely polarized and their ac outputs summed, then a cosine pattern will emerge. The principle of Von Braunmühl and Weber is found to this day in electrostatic microphones used for recording and other high quality applications.

While Von Braunmühl and Weber envisioned the operation of their microphone as a combination of pressure and pressure-gradient functions, another way of looking at it, within certain limitations, is as a special case of a phase-shift microphone to be described next.

Phase-Shift Microphones

In attempting to balance the two damping screens of the structure in Fig. 5(g), the author noted that certain conditions of screen unbalance produced a small but decided unidirectional effect. In analyzing this phenomenon by means of equivalent circuit analysis it became apparent that acoustical phase-shift introduced by the networks was responsible. Soon thereafter the conditions were formulated for producing any directional pattern in the limaçon family with any transducer and an appropriate phase-shift network.

The invention is described in a parent patent[57] and four continuations-in-part.[58] The former outlines three different phase-shift networks which are the basis of practically all phase-shift microphones currently in use, and which are summarized in Fig. 7.

A piezoelectric phase-shift microphone is shown in Fig. 7(a). The microphone consists of a circular mounting plate 7.1 upon which is fastened a diaphragm 7.2 and a piezoelectric transducer 7.3. In the simplified equivalent circuit of Fig. 7(b) these are shown as defining an impedance Z_{am} which includes the air load. Sound waves for frontal (0° incidence) first impinge upon the diaphragm with a pressure p_1 traveling to the rear of the microphone through a distance d with a velocity c. The rear pressure p_2 lags behind p_1 by a phase angle $\phi = \omega d/c$. For any other angle of incidence θ, $\phi = (\omega d/c) \cos \theta$. Air flow into the volume 7.5 which defines an acoustical compliance C_2 is caused by pressure P_2 acting through the circumferential entry-port 7.4 which defines a resistance R_2 and inertance L_2. It is shown in the parent patent that *regardless* of the magnitude of Z_{am}, a cardioid pattern will be obtained if the

[54] W. R. Harry, "Six-way directional microphone," *Bell Labs. Rec.*, vol. 19, pp. 10–14; September, 1940.

[55] R. P. Glover, "A review of cardioid type unidirectional microphones," *J. Acoust. Soc. Am.*, vol. 11, pp. 296–32; January, 1940.

[56] Von Braunmühl and Weber, U. S. Patent No. 2,179,361 (1939); filed March 30, 1936.

[57] B. B. Bauer, U. S. Patent No. 2,237,298 (1941); filed September 29, 1938.

[58] B. B. Bauer, U. S. Patents No. 2,305,596 to 599 (1942); filed April 8, 1941.

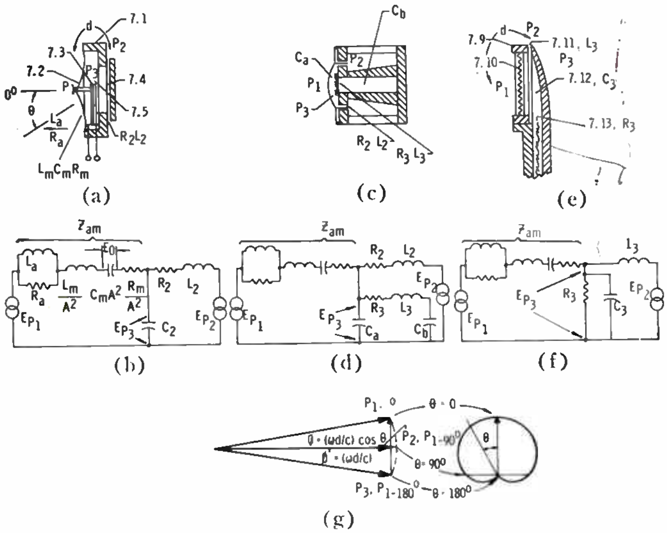


Fig. 7—Phase-shift unidirectional microphones.

elements of the network R_2, L_2, and C_2 are proportioned as follows:

$$R_2 = d/cC_2 \quad (6)$$

$$L_2 = C_2R_2^2/2. \quad (7)$$

In the above proportions, the elements R_2, L_2, and C_2 form a phase-shift network whereby the pressure P_3 within the microphone is equal to P_2 but is shifted in phase by an angle $\phi' = \omega d/c$. ϕ' remains unaffected by direction of arrival of sound. Referring to Fig. 7(g), and letting P_2 and P_3 remain stationary, as the source of sound rotates from the front to the back of the microphone, the phasor P_1 will describe a path from $P_{1-0°}$, to $P_{1-90°}$ and then to $P_{1-180°}$. The phasor connecting the ends of P_3 and P_1 plotted as a function of the angle θ will be a cardioid of revolution. By choosing properly the relative magnitudes of ϕ_0 and ϕ', any desired member of the limaçon family may be obtained. The above principle is employed in piezoelectric microphones intended for public address applications.

A moving-coil phase-shift microphone exhibiting cardioid operation is shown in Fig. 7(c), and its simplified equivalent circuit in 7(d). Here the impedances of the moving coil and the air load again are lumped together as Z_{am}. The phase-shift network is composed of the rear port resistance R_2 and inertance L_2, compliance of the volume under the diaphragm C_a, and within the magnet C_b, and the impedance of the interconnecting screen R_3 and L_3. Subsequently, Black[59] and Wiggins[60] modified this structure by providing multiple rear entry ports. This approach allows the use of stiffer diaphragm suspension than that used with the microphone in Fig. 7(d), and serves to improve sensitivity to mechanically transmitted noise.

[59] Black, U. S. Patent No. 2,401,328 (1946); filed January 16, 1943.

[60] A. M. Wiggins, "Unidirectional microphone utilizing a variable distance between the front and back of the diaphragm," *J. Acoust. Soc. Am.*, vol. 26, pp. 687–692; September, 1954.

Practically all unidirectional moving-coil microphones built today use one of the principles described in the preceding paragraph.

A third phase-shift network to be found in the parent patent is especially adapted for use with mass-controlled transducers, such as the ribbon transducer. Shown in Fig. 7(e) is a frame 7.9 and ribbon transducer 7.10, and a phase-shift network comprised of the following elements: entry port 7.11 which defines an acoustic inertance L_3, a volume behind the ribbon 7.12 which furnishes a compliance C_3, and a damped pipe which defines a resistance R_3. The front-to-back distance is again defined as d. This network can be solved analytically for the cardioid with the following result:[57]

$$L_3 = dR_3/c \quad (8)$$

$$C_3 = L_3/2R_3^2. \quad (9)$$

An improved version of this microphone was developed by Olson[61] in which the rear entry is adjustable for selection of polar pattern. Subsequently, Olson, Preston and Bleazey reported achieving a further improvement by taking a phase-shift ribbon microphone with a limaçon characteristic where $\rho = 0.3 + 0.7 \cos\theta$ and providing a damped cavity in the vicinity of the rear entry ports.[62] The above microphones have found wide use in public address and television broadcasting.

Super-Directional Microphones

Three approaches have been taken to provide microphones with directional characteristics sharper than those possible with limaçon patterns.

A. Reflectors, Refractors, Diffractors

From optical analogy, the idea of using a parabolic mirror for improved directivity must have occurred to various investigators. The microphone is placed at or near the focus of the reflector. The angular resolution for short wavelengths is given by Rayleigh's criterion,[63] as $\theta_r = 0.61\ \lambda/r$ radians, where λ is the wavelength, r the radius, and θ the resolution angle. The directional capability is very high at high frequency and nil at low frequency. Hansen[64] describes a parabolic reflector used with condenser microphones. Olson and Wolff[65] proposed a concentrator consisting of parabolic and conical sections arranged in the form of a horn. Aamodt and Harvey[66] devised a wide area electrostatic microphone which attains notable directivity simply because of its

[61] H. F. Olson, "Polydirectional microphone," Proc. IRE, vol. 32, pp. 77–82; February, 1944.

[62] H. F. Olson, J. Preston and J. C. Bleazey, "The uniaxial microphone," IRE Trans. on Audio, vol. AU-1, pp. 12–19; July–August, 1953.

[63] See, for example, G. S. Monk, "Light, Principles and Experiments," McGraw-Hill Book Co., Inc., New York, N. Y., p. 206; 1937.

[64] O. B. Hanson, "Microphone technology in radio broadcasting," *J. Acoust. Soc. Am.*, vol. 3, pp. 81–93; July, 1931.

[65] H. F. Olson and I. Wolff, "Sound concentrator for microphones," *J. Acoust. Soc. Am.*, vol. 1, pp. 410–417; April, 1930.

[66] T. Aamodt and F. K. Harvey, "A large area condenser type of transducer," *J. Acoust. Soc. Am.*, vol. 25, p. 825 (abstract only); July, 1953.

large size. Further improvements in performance can be obtained by combining an acoustical lens[67] with a conical horn.[68]

B. Line Microphones

In 1939, Mason and Marshall described a microphone attachment consisting of 50 small tubes whose lengths vary by equal increments from 3 cm to 150 cm. These are assembled into a circular bundle and coupled to the diaphragm of a pressure microphone.[69] This microphone is roughly equivalent in directional effects to a 3-foot parabolic reflector, but with considerably less bulk and frequency dependence. The same year Olson described an improved line microphone in which directional characteristics were substantially independent of frequency, obtained by combining several multipipe units, each designed for operation over a given frequency range.[70]

C. Higher-Order Combination Microphones

It has been seen that subtraction of pressures at two points in space produces a gradient mode of operation described by $\rho = \cos\theta$. Subtraction of two gradient modes at two points in space will produce a second-order gradient, $\rho = (\cos\theta)(\cos\theta) = \cos^2\theta$. By continuing this process, in theory infinite improvement in directivity could be obtained in theoretically infinitesimal space.

A microphone with higher mode of operation was developed in 1938 and described in the parent phase-shift microphone case, U. S. Patent No. 2,237,298.[57] By providing an appropriate electrical network with two gradient transducers a polar pattern defined by equation $\rho = (1+\cos\theta)(\cos\theta)$ was obtained. The reissue patent 2,305,599[58] describes how the same effect may be achieved by subtraction of outputs of two spaced-apart cardioid microphones. Olson and Preston have carried out this work further by combining two special phase-shift microphones[62] with electrical networks to obtain a polar pattern defined by $\rho = (0.3+0.7\cos\theta\cos\theta/3)(\cos\theta)$.[71]

A second-order gradient differential microphone employing a single diaphragm and a case of suitable configuration was developed by Wiggins, for speech transmission from noisy environment.[72]

Future Developments

The art of microphone design still taxes the ingenuity of the physicist and the radio scientist. Despite the century of progress many problems remain unsolved.

Improved directional characteristics will continue to receive considerable attention. A "zoom" microphone, in which directional pattern can be adjusted to conform, say, to the optical angle of a television camera may find important use in the broadcasting industry. Light and highly effective directional microphones would aid with picking out the desired sounds amidst crowd noise.

The problem of an effective, reliable and inexpensive multichannel wireless microphone is yet to be solved. Such a microphone would be a boon to broadcasting, entertainment and similar industries.

One feels intuitively that we should be due for a "breakthrough" in transducer technology. Microphones especially suitable for use with transistor amplifiers and with sufficient sensitivity and low noise, figure to be useful in broadcasting recording and sound level meter applications would be very welcome.

Unconventional methods of sound reception will be further explored: Throat microphones already have been widely used in military actions, but they provide poor articulation. Microphones placed within the mouth, attached to the teeth, inserted in the ear canal and otherwise coupled to the skeletal structure of the head have already received considerable study.

Ultimately, lest we forget, speech is merely an end-product of the thought processes, and there is no reason why eventually these should not be directly picked up without the intervening aerial vibrations. One should not be surprised to see an Astronaut, someday, with a radio "thought" transmitter permanently implanted in his cranium. But then, alas, all this microphone development would have been in vain.

Acknowledgment

A debt of gratitude is due to the writers and historians who have documented the work of previous investigators well enough to allow this paper to be written without need of exhaustive original research. H. A. Frederick[18] and F. V. Hunt[44] in their respective publications have provided a wealth of historical material. Olson's encyclopaedic "Acoustical Engineering"[73] describes a great variety of microphones from the technological point of view, and it was an invaluable reference.

Back volumes of the *Journal of the Acoustical Society of America* and the IRE Transactions on Audio have been most useful in reviewing modern developments. Friends and associates too numerous to mention have been helpful with the location of references. To those of them who are still young enough reasonably to expect to celebrate the 100th Anniversary of the IRE, this article will hopefully be an acceptable starting point in appraising the progress in microphones that will have taken place during the next 50 years!

[67] W. E. Kock and F. K. Harvey, "Refracting sound waves," *J. Acoust. Soc. Am.*, vol. 21, pp. 471–481; September, 1949.

[68] M. A. Clark, "An acoustic lens as a directional microphone," IRE Trans. on Audio, vol. AU-2, pp. 5–7; January–February, 1954.

[69] W. P. Mason and R. N. Marshall, "A tubular directional microphone," *J. Acoust. Soc. Am.*, vol. 10, pp. 206–215; January, 1939.

[70] H. F. Olson, "Line microphones," Proc. IRE, vol. 27, pp. 438–446; July, 1939.

[71] H. F. Olson and J. Preston, "Directional microphone," *RCA Rev.*, vol. 10, pp. 339–347; September, 1949.

[72] A. M. Wiggins, U. S. Patent No. 2,552,878 (1951), application September 24, 1947.

[73] H. F. Olson, "Acoustical Engineering," D. Van Nostrand Co., Inc., Princeton, N. J.; 1957.

Loudspeakers*

HARRY F. OLSON†, FELLOW, IRE

Summary—**A loudspeaker is an electroacoustic transducer intended to radiate acoustic power into the air, the acoustical waveform being essentially equivalent to that of the electrical input. Electrodynamic, electromagnetic and electrostatic driving systems have been used for loudspeakers operating in air. However, during the past two decades the electrodynamic has overwhelmingly predominated in all applications for loudspeakers operating in air. The almost universal use of the direct-radiator dynamic loudspeaker in radio receivers, phonographs, magnetic-tape reproducers, television receivers, announce and intercommunicating systems is due to the simplicity of construction, small space requirements and the relatively uniform response-frequency characteristics. For small-scale applications the low efficiency of the direct-radiator loudspeaker is not a handicap. However, for large-scale high-power sound applications, the high-efficiency horn loudspeaker is particularly suitable because the amplifier requirements are reduced by an order of magnitude. For the future, the promising developments in loudspeakers appear to be in the field of motion control.**

Introduction

A LOUDSPEAKER is an electroacoustic transducer intended to radiate acoustic power into the air, the acoustical waveform being essentially equivalent to that of the electrical input. The transformation of electrical waves into the corresponding acoustical waves may be accomplished in a multitude of ways. At the present time, while practically all loudspeakers may be classed as of the diaphragm type, the essential essential difference resides in the method of coupling between the diaphragm and the air and the means for driving the diaphragm. The two extreme ends of the audio-frequency range are the most difficult to reproduce with efficiency comparable to that of the mid audio-frequency range. In general, loss of coupling between the diaphragm and the air occurs in the low-frequency range. Among the common methods employed to increase the low audio-frequency radiation from diaphragms are, namely, the use of large diaphragms, groups of diaphragms and various shapes of baffles, cabinets and horns. Loss of efficiency in the high audio-frequency range is due to a relatively large mass reactance as compared to the sound radiation resistance of air. The effective mass reactance may be reduced by multiple loudspeakers and horn couplers.

There are many possible transducers which may be employed to drive the diaphragm as follows: dynamic, magnetic, electrostatic, piezoelectric and magnetostrictive. In the direct-coupled loudspeaker systems with no diaphragm there are various electronic expedients for producing compression or motion of the air. Some of these have been used commercially.

The purpose of this paper is to present a brief historical review of the loudspeaker art, a technical exposition of the present state of loudspeakers and future possibilities.

* Received by the IRE, June 21, 1961; revised manuscript received, October 19, 1961.

† RCA Laboratories, Princeton, N. J.

Historical Review

Thousands of investigators have contributed to the science of electroacoustic transducers over a period of more than one hundred years. Therefore, it is impossible in this limited presentation to provide a complete historical account of the loudspeaker art because the main objective is to present the significant technical aspects of loudspeakers. In view of the fact that the dynamic or moving-coil driving mechanism has been the most important in terms of modern commerical utilization, the historical presentation[1] in this paper will be confined to this system.

The motor mechanism consisting of a circular coil located in a radial magnetic field was first disclosed by Siemens.[2] Lodge,[3] Pridham and Jenson[4] and others contributed to the suspension system. However, there were very few developments in loudspeakers in the twenty-seven years following Lodge's disclosure. A breakthrough in the dynamic loudspeaker was made by Rice and Kellogg[5] in 1925. The success of the development was due to their recognition of three physical factors with relation to the action and design of a direct radiator loudspeaker. The first is that the sound-power output of a loudspeaker is the product of the mechanical resistance due to sound radiation and the square of the velocity of the diaphragm. The second is that sound radiation from a small vibrating diaphragm gives rise to a mechanical resistance which is proportional to the square of the frequency. The third is a vibrating system which is mass controlled. It follows then that, if the fundamental resonance occurs below the lowest frequency of interest, the complementary variations of the second and third factors which control the sound output as given by the first factor conspire to provide a uniform response up to the frequency region at which the assumptions begin to fail. This was the contribution of Rice and Kellogg, and it continues to be the basic precept that guides the design of all direct-radiator loudspeakers.

[1] For those interested in a complete historical account of loudspeakers, the reader is referred to F. V. Hunt, "Electroacoustics," John Wiley and Sons, Inc., New York, N. Y.; 1954.

[2] E. W. Siemens, German Patent, No. 2355, filed December 14, 1877.

[3] O. J. Lodge, British Patent, No. 9712, filed April 27, 1898.

[4] E. S. Pridham and P. L. Jenson, U. S. Patent No. 1,448,279, filed April 28, 1920.

[5] C. W. Rice and E. W. Kellogg, "Notes on the development of a new type of hornless loudspeaker," *Trans. AIEE*, vol. 44, pp. 461–475; April, 1925.

Horn loudspeakers are used for applications which require high efficiency and large power outputs. The first use of horns in connection with loudspeakers is somewhat obscure. However, Webster[6] was the first to introduce the concept of acoustical impedance and to apply the theory to horns to show that a horn may be used as a marching device between the relatively large acoustical impedance of the diaphragm and the low acoustical impedance of air in free space.

The preceding all-too-brief history of loudspeakers includes only a few of the thousands of investigators whose contributions during the past century have made modern sound reproduction possible.

Driving Systems

All loudspeakers employ an electromechanical or electroacoustic transducer, termed a driving system, for converting electrical variations into the corresponding acoustical variations. The most common driving systems in use today are the electrodynamic, the electromagnetic, the electrostatic, the magnetostrictive and the piezoelectric.

In the electrodynamic driving system depicted in Fig. 1A, the mechanical forces are developed by the interaction of the current in the conductor and the transverse magnetic field.

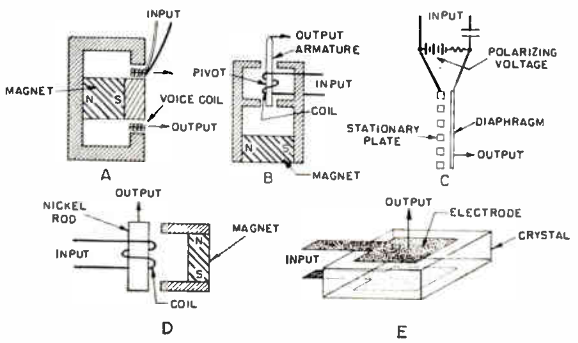


Fig. 1—Driving Systems. A) Electrodynamic. B) Electromagnetic. C) Electrostatic. D) Magnetostrictive. E) Piezoelectric.

In the electromagnetic driving system depicted in Fig. 1B the mechanical forces are developed from magnetic reactions. There are innumerable possibilities in the design of electromagnetic driving systems, but the most common arrangement for use in loudspeakers is the balanced armature system shown in Fig. 1B.

In the electrostatic driving system depicted in Fig. 1C the mechanical forces are developed from electrostatic reactions.

In the magnetostrictive driving system depicted in Fig. 1D the mechanical forces result from the deformation of a ferromagnetic material possessing direct magnetostriction properties.

[6] A. G. Webster, "Acoustical impedence and the theory of horns and of the phonograph," *Proc. Natl. Acad. Sci.*, vol. 5, pp. 275–282; 1919.

In the piezoelectric driving system depicted in Fig. 1E the mechanical forces result from the deformation of a crystal or ceramic having converse piezoelectric properties.

Electrodynamic, electromagnetic and electrostatic driving systems have been used for loudspeakers operating in air. However, during the past two decades the electrodynamic has overwhelmingly predominated in all applications for loudspeakers operating in air.

Direct-Radiator Dynamic Loudspeakers

The almost universal use of the direct-radiator dynamic loudspeaker in radio receivers, phonographs, magnetic-tape reproducers, television receivers, announce and intercommunicating systems is due to the simplicity of construction, small space requirements, and the relatively uniform response-frequency characteristic.

The vibrating system of a direct-radiator dynamic loudspeaker consists of a conical paper diaphragm coupled to a voice coil located in the air gap of a magnetic structure (Fig. 2). The voice coil consists of a cylindrical multilayer coil of aluminum or copper wire. In general, the resistance of the voice coil, depending upon the design, ranges from 1.5 to 100 ohms. The mass of the voice coil, depending upon the size of the loudspeakers, ranges from 0.1 to 4 grams. The permanent magnet is of the high-flux high-coercive type and may be an alloy of aluminum, cobalt, nickel and iron, or a ferrite of iron, cobalt, barium and nickel. The flux density in the air gap of commerical loudspeakers ranges from 3000 to 20,000 Gauss. The diameter of the cone of commerical direct-radiator loudspeakers ranges from 1 to 18 in. The mass of the cone from the smallest to the largest covers a range of 0.1 to 100 grams.

The inherent characteristics of the elements of the vibrating system and radiation load of a direct-radiator dynamic loudspeaker which conspire to provide a uniform response-frequency characteristic may be illustrated as follows.

A cross-sectional view, the voice coil circuit and the mechanical circuit of a direct-radiator dynamic loudspeaker are shown in Fig. 2. The total mechanical impedance z_{MT} in mechanical ohms, of the vibrating system at the voice coil is

$$z_{MT} = r_{MS} + r_{MA} + j\omega m_C + j\omega m_A - \frac{j}{\omega C_{MS}}. \tag{1}$$

where

r_{MS} = mechanical resistance of the suspension system (in mechanical ohms)

r_{MA} = mechanical resistance of the air load (in mechanical ohms)

m_C = mass of the cone and the voice coil (in grams)

m_A = mass of the air load (in grams)

C_{MS} = compliance of the suspension system (in cm/dyne).

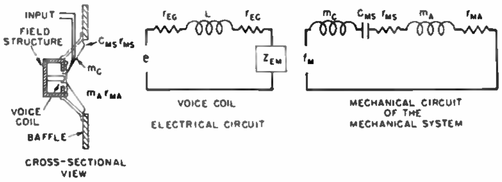


Fig. 2—Cross-sectional view, electrical and mechanical circuits of a direct-radiator dynamic loudspeaker mounted in a baffle. In the electrical circuit: e = the internal voltage of the generator. r_{EG} = the electrical resistance of the generator. r_{EC} and L = the electrical resistance and the inductance of the voice coil. z_{EM} = the motional electrical imedance. In the mechanical circuit: f_M = the driving force. m_C = the mass of the cone. C_{MS} = the compliance of the suspension system. r_{MS} = mechanical resistance of the suspension system. m_A = the mass of the air load. r_{MA} = the mechanical resistance of the air load.

Eq. 1 may be written as follows:

$$z_{MT} = r_{MS} + r_{MA} + jx_{MC} + jx_{MA} - jx_{MS}, \tag{2}$$

where

r_{MS} = mechanical resistance of the suspension system (in mechanical ohms)

r_{MA} = mechanical resistance of the air load (in mechanical ohms)

$x_{MC} = \omega m_C$ = mechanical reactance of the voice coil and cone (in mechanical ohms)

$x_{MA} = \omega m_A$ = mechanical reactance of the air load (in mechanical ohms)

$x_{MS} = 1/\omega C_{MS}$ = mechanical reactance of the suspension system (in mechanical ohms).

The mechanical resistance and mechancal rieactance of the air load may be obtained from Fig. 3.

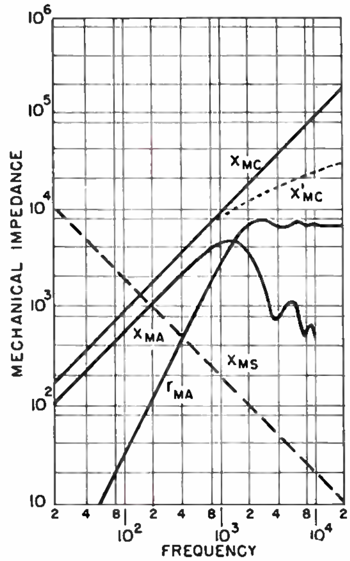


Fig. 3—Mechanical impedance-frequency characteristics of a direct-radiator dynamic loudspeaker with a cone 4 in in diameter. x_{MC} = the mechanical reactance of the cone and coil. x_{MA} = the mechanical reactance of the air load. r_{MA} = the mechanical resistance of the air load. x_{MS} = the mechanical reactance of the suspension system. x'_{MC} = the mechanical reactance of the cone and coil where wave motion occurs in the cone.

where

r_{ER} = component of the motional electrical resistance due to the radiation of sound (in abohms)

r_{EM} = total motional electrical resistance (in abohms)

r_{EC} = damped electrical resistance of the voice coil (in abohms).

The components r_{ER} and r_{EM} may be obtained from (2) and (3).

From (2)–(4) the efficiency (in per cent) of the loud speaker is

$$\mu = \frac{(Bl)^2 r_{MA}}{(Bl)^2(r_{MS} + r_{MA}) + r_{EC}\left[(r_{MC} + r_{MA})^2 + (x_{MA} + x_{MC} - x_{MS})^2\right]}. \tag{5}$$

The motional electrical impedance z_{EM} (in abohms) of the mechanical system is

$$z_{EM} = \frac{(Bl)^2}{z_{MT}}, \tag{3}$$

where

B = flux density in air gap, (in gauss)

l = length of the conductor in the voice coil (in centimeters)

z_{MT} = total mechanical impedance of the mechanical system (in mechanical ohms).

The efficiency of the loudspeaker is the ratio of the sound-power output to the electrical-power input. The efficiency, μ (in per cent) may be obtained from the voice-coil circuit of Fig. 2 and expressed as follows:

$$\mu = \frac{r_{ER}}{r_{EC} + r_{EM}} \times 100, \tag{4}$$

To simplify the discussion, assume that the mechanical reactance and mechanical resistance of the suspension system are zero. The mechanical-impedance characteristics of the mechanical system are shown in Fig. 3. Since r_{MA} is small compare to x_{MA} and x_{MC}, (5) becomes

$$\mu = \frac{(Bl)^2 r_{MA}}{r_{EC}(x_{MA} + x_{MC})^2} \times 100. \tag{6}$$

Referring to Fig. 3, it will be seen that x_{MA} and x_{MC} are proportional to the frequency, and r_{MA} is proportional to the square of the frequency. Under these conditions (6) shows that the output of a direct-radiator loud speaker is independent of the frequency from the fundamental resonant frequency up to the frequency at which the ultimate resistance obtains.

The efficiency characteristic μ_1 of Fig. 4 depicts the foregoing in a graphical manner. The efficiency characteristic μ_1 assumes a resonant frequency at 0 cps. If the mechanical reactance of the suspension system is included, the efficiency will be as depicted by the efficiency characteristic μ_2 of Fig. 4. In the high-frequency region the cone no longer vibrates as a single unit, and wave action takes place which lowers the mechanical impedance, as shown by the dotted characteristic x'_{MC} (Fig. 3). As a result, the efficiency in the high-frequency range is increased, as shown by the efficiency characteristic μ_3 (Fig. 4). The preceding illustration substantiates the physical principles of operation of the dynamic direct radiator loudspeaker as outlined in the historical section of this paper.

Another important characteristic of a loudspeaker is the directional characteristic, which depicts the response as a function of the angle with respect to some axis of the system. The directional pattern of a vibrating cone depends upon three principal factors: the cone diameter, the cone angle and the frequency. Other factors such as the paper pulp, the fabrication and processing of the cone, the voice-coil diameter and the cone suspension also influence the directional pattern, but to a lesser degree.

The directional characteristics of direct-radiator dynamic loudspeakers for cones of various angles as a function of the ratio of the diameter to the wavelength are shown in Fig. 5. The directional pattern becomes sharper with increasing frequency. However, the pattern is broader than that of a vibrating piston of the same diameter due to the relatively low velocity of sound propagation in the paper cone. A further consideration shows that the directional pattern becomes broader at the higher frequencies as the cone angle is increased, which is due to the delay between the sound emitted from the outside edge and the center of the cone which will increase as the angle of the cone is increased. As a result, the directional pattern will be broadest for the cone with the widest angle.

Frequency discrimination will be introduced for various angles with respect to the axis of the cone if the directional pattern varies with respect to frequency. A consideration of the data of Fig. 5 shows that a relatively uniform directivity pattern may be obtained by the use of different size cones graduated with respect to frequency. This is graphically illustrated in Fig. 6 in which a broad and uniform directivity pattern is obtained by employing a 15-in diameter cone to cover the frequency range up to 1200 cps and a $2\frac{1}{2}$-in diameter cone to cover the frequency range above 1200 cps.

As in all electronic systems, a loudspeaker exhibits nonlinear distortion. The principal sources of nonlinear distortion are as follows: nonlinear cone suspension system, inhomogeneity of the air gap flux in which the voice coil moves, frequency modulation of a high-frequency sound radiation due to large motion of the cone in the low-frequency range, and nonlinear distortion of the air due to the high sound level in close proximity to the cone.

The most troublesome, and in most cases the largest, source of nonlinear distortion in direct-radiator dynamic loudspeakers employing cone-type radiators is the suspension system. The nonlinear distortion due to the suspension system may be reduced to a negligible value by employing a low-resonance loudspeaker mounted in a completely enclosed cabinet, as shown in Fig. 7. If the compliance C_{MS} of the suspension is made large compared to the compliance of the air in the cabinet C_{MB}, the controlling negative mechanical reactance will be that of the compliance of the air in the cabinet. In this way the distortion introduced by the nonlinear characteristics of the suspension C_{MS} will be negligible.

In general, all of the other nonlinear distortions can be reduced to a negligible quantity. However, as in most electronic systems, the problem involves economics.

Speech and music are transient in character. Therefore, in order to reproduce speech and music faithfully the loudspeaker should exhibit excellent transient response. A tone burst simulates, to a first approximation, the transient nature of speech and music. Therefore, if an electrical tone burst is applied to the loudspeaker, the sound output will represent the transient response. The response of a loudspeaker for the range 400 to 3000 cps is shown in Fig. 8. There is a peak in the response at 800 cps and a dip in the response at 1100 cps. The sound output of this loudspeaker for an electrically applied tone burst is shown in Fig. 8. In the case of a peak in the response-frequency characteristic, there is a slow growth and a slow decay in the response of the loudspeaker to the applied tone burst. In the case of a dip in the response frequency characteristic, there is a rapid growth followed by a decrease in the output and then by an increase in the output. A measure of the transient distortion may be designated as the remaining response after the tone-burst signal has been cut off. In general, if the response frequency characteristic is smooth, the transient response will be good.

The diaphragms or cones of practically all direct-radiator loudspeakers are made of paper by a felting process employing a master screen having the shape of the diaphragm. The mixture of pulp and water is drawn through the screen leaving a thin deposit of compressed pulp. When this deposit is dried, it can be removed from the screen and the result is the finished diaphragm. The outside suspension system can also be felted as part of the cone.

The simple direct-radiator dynamic loudspeaker shown in Fig. 2 represents the most common and enjoys the greatest widespread use. However, there are many versions of the direct-radiator loudspeaker in which the principal objective is to provide improved performance over a wide frequency range. The most common of these systems is depicted in Fig. 9.

The direct-radiator loudspeaker mechanism is usually mounted in a flat baffle or some type of cabinet enclosure as shown in Fig. 10. A direct-radiator loudspeaker mounted in a flat baffle is shown in Fig. 10A. A rela-

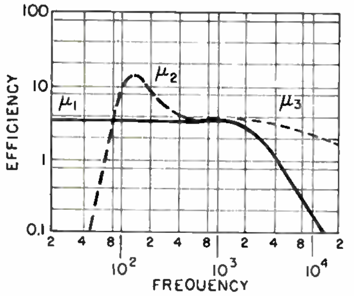


Fig. 4—Efficiency-frequency characteristics for the mechanical impedance characteristics of Fig. 3. μ_1 = the efficiency for $x_{MS}=0$. μ_2 = the efficiency for x_{MS} as depicted in Fig. 3. μ_3 = the efficiency for x_{MS} and X'_{MC} as depicted in Fig. 3.

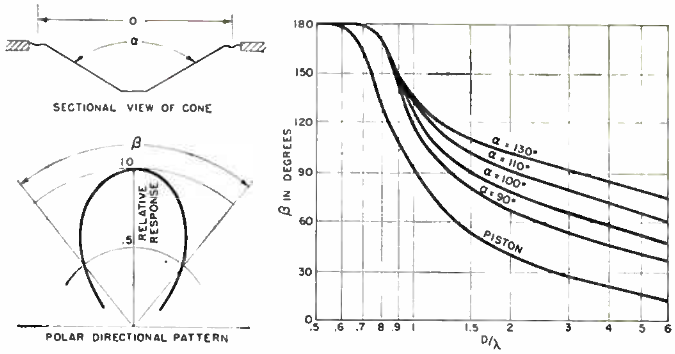


Fig. 5—Directional characteristics of direct-radiator dynamic loudspeakers for cones of various angles α as a function of D/λ, where D = effective diameter of the cone and λ = wavelength of the radiated sound. The directional characteristic of a rigid piston is shown for reference. The angle β is the total angle for which the radiation is down 6 db at the extremes.

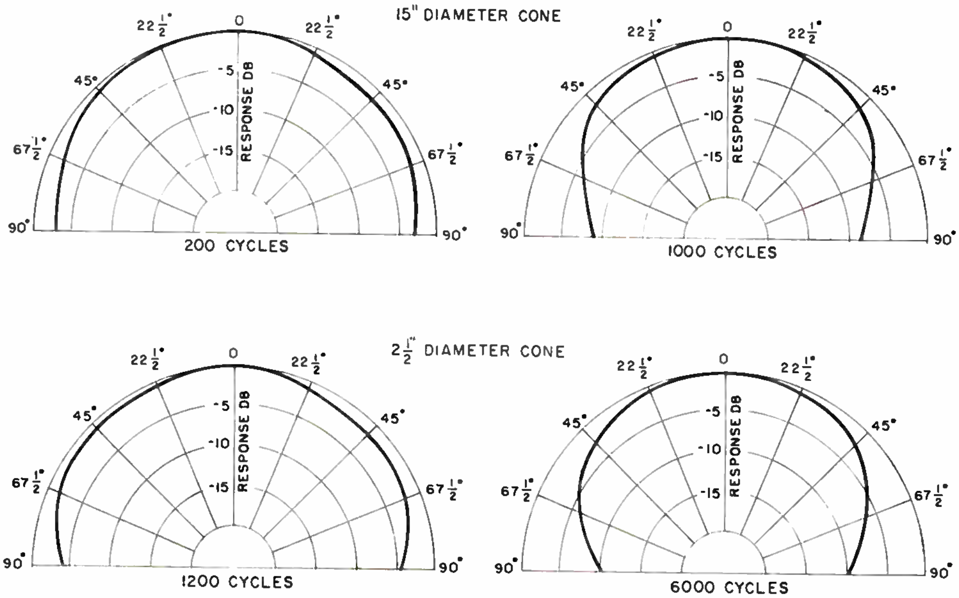


Fig. 6—Directional characteristics of direct-radiator dynamic loudspeakers with cone diameters of 15 in and $2\frac{1}{2}$ in.

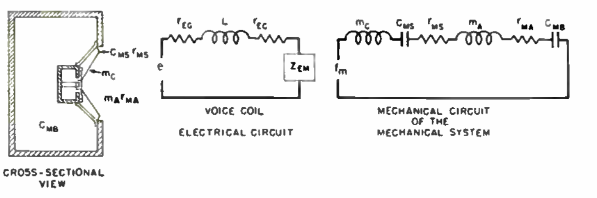


Fig. 7—Cross-sectional view, electrical and mechanical circuits of a direct-radiator dynamic loudspeaker mounted in a completely enclosed cabinet. In the electrical circuit: e = the internal voltage of the generator, r_{EG} = the electrical resistance of the generator. r_{EC} and L = the electrical resistance and the inductance of the voice coil. z_{EM} = the motional electrical impedance. In the mechanical circuit: f_M = the driving force. m_C = the mass of the cone. C_{MS} = the compliance of the suspension system. r_{MS} = mechanical resistance of the suspension system. m_A = the mass of the air load. r_{MA} = the mechanical resistance of the air load. C_{MB} = compliance of the air in the cabinet.

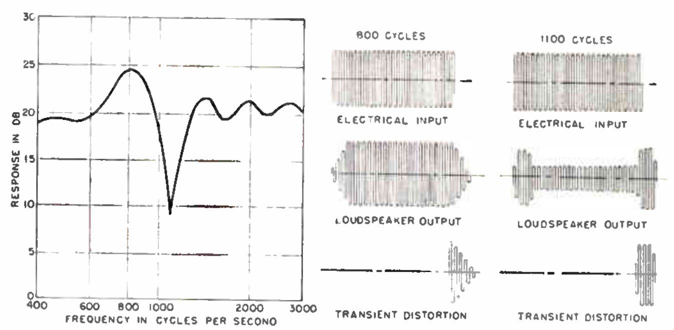


Fig. 8—Response-frequency characteristic of a direct-radiator dynamic loudspeaker, and the transient response to tone bursts of 800 and 1100 cps. The waves show the electrical tone burst input to the loudspeaker, the sound output from the loudspeaker and the sound output after the electrical input to the loudspeaker has been stopped.

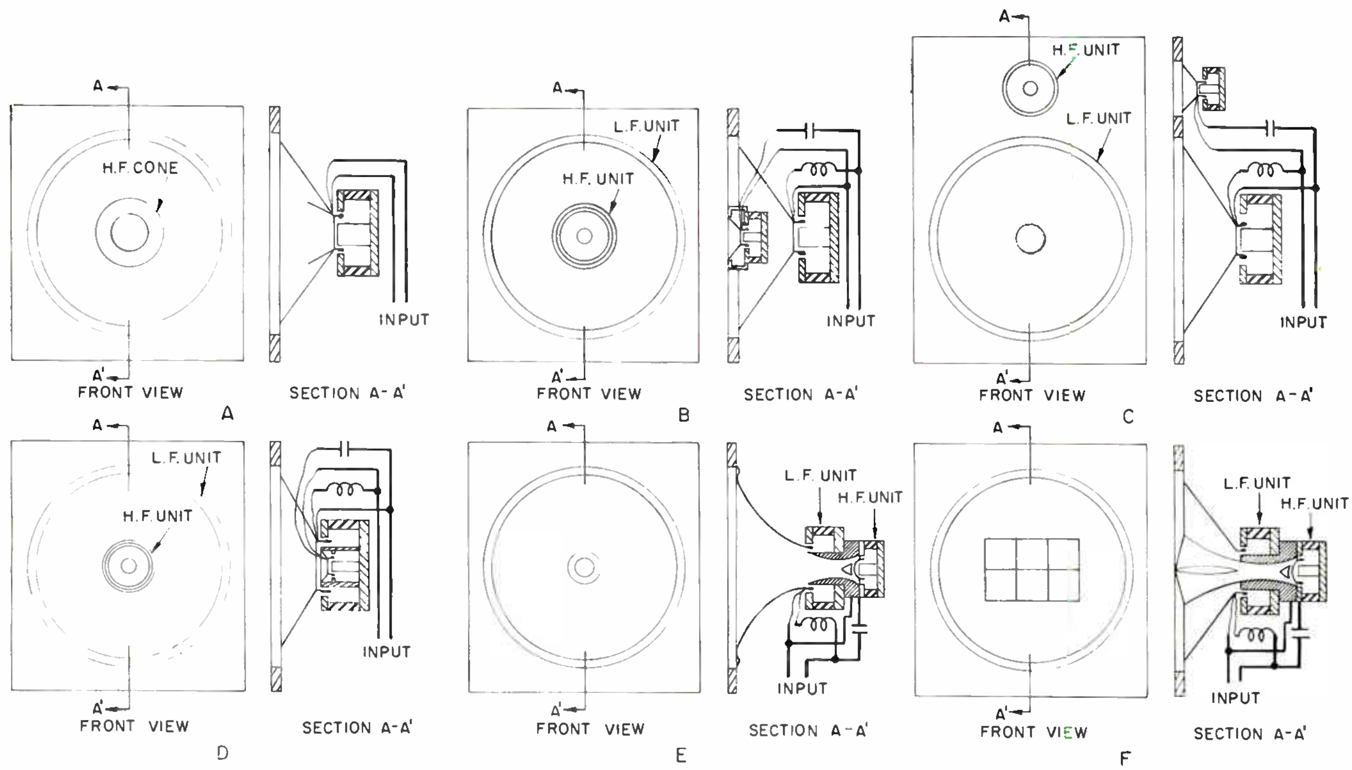


Fig. 9—Multiple direct-radiator dynamic loudspeaker systems. A) Double cone. B), C), and D) two separate direct-radiator units. E) and F) Direct-radiator and horn combination.

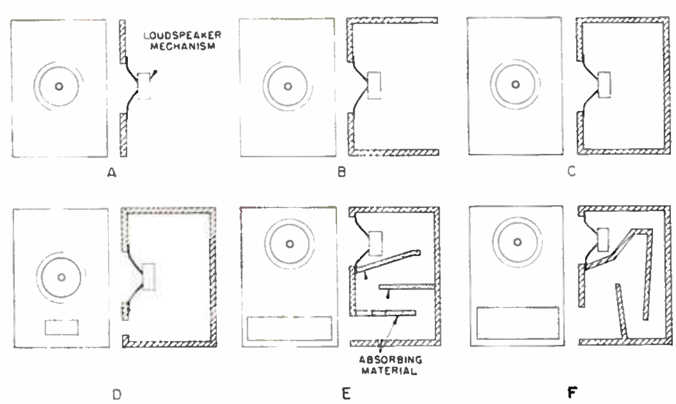


Fig. 10—Mountings and enclosures for direct-radiator dynamic loudspeakers. A) Flat baffle. B) Open-back cabinet. C) Closed cabinet. D) Closed cabinet with a port. E) Labyrinth. F) Horn.

tively large baffle must be used to obtain adequate low-frequency response, because the baffle cutoff is determined by the distance from the front to the rear of the mechanism. When this distance is less than a quarter wavelength of the sound wave, the response falls off rapidly below this frequency. In order to obtain good response to 70 cps a four-foot baffle is required. A direct-radiator loudspeaker mechanism mounted in an open-back cabinet is shown in Fig. 10B. This system is used in practically all radio receivers, phonographs and television receivers of both the table-model and console types. The fundamental resonance of the cabinet, acting as a pipe, connected to a loudspeaker mechanism produces considerable accentuation of the low-frequency response, and is termed *cabinet resonance*. A direct-radiator loudspeaker mechanism mounted in a completely enclosed cabinet is shown in Fig. 10C. This system will exhibit uniform response to about one-half octave below the fundamental resonance frequency of the combination of the mechanism and the cabinet. The accentuated response at the resonant frequency of the loudspeaker mechanism and cabinet can be reduced by the addition of sound absorbing material in the cabinet so that practically uniform response is obtained. A direct-radiator loudspeaker mechanism mounted in a cabinet completely enclosed save for a port is shown in Fig. 10D. The addition of the port accentuates the response in that portion of the frequency range in which the motion of the cone is in phase with motion of the air in the port. A direct-radiator loudspeaker mechanism coupled to one end of a long damped pipe in the form of labyrinth with the other end of the pipe opening in the front of the cabinet is shown in Fig. 10E. At the first half wavelength resonance, the velocity at the open end is in phase with that of the front of the cone. The radiation then is additive and the response is increased. An increase in response can be obtained over an octave. The increase of absorption with frequency can be used to control the radiation from the pipe in the frequency region in which the output of the loudspeaker and pipe are out of phase. A combination horn and direct-radiator loudspeaker are shown in Fig. 10F. The sound radiation in the low-frequency range takes place from the horn coupled to the back of the cone. The mid- and high-frequency sound radiation takes place from the front of the direct-radiator loudspeaker.

The efficiency of a direct-radiator dynamic loudspeaker is relatively low, being about 1 to 5 per cent. Therefore, the direct-radiator loudspeaker is not particularly suitable for installations requiring large amounts of sound power, as, for example, in large-scale high-power sound motion picture, sound reinforcing and public address systems, because the power amplifiers required to drive the loudspeaker become too large. Under these conditions, it is more economical to employ loudspeakers with high efficiency such as, for example, horn loudspeakers. With a horn loudspeaker it is possible to obtain an efficiency of 20 to 50 per cent. However, for reproduction in the average living room a relatively small amount of power is required to obtain an adequate sound level employing a direct-radiator loudspeaker. For the average listening sound-pressure level of 80 db the input power required is only 0.05 watt. A level of 100 db can be obtained with an input of 5 watts. Thus it will be seen that the direct-radiator dynamic loudspeaker is quite satisfactory for small scale applications.

During the past three and one half decades, more than 500,000,000 direct-radiator dynamic loudspeakers have been sold in the United States.

Horn Loudspeakers

A horn loudspeaker consists of an electrically or a mechanically driven diaphragm coupled to a horn. The principal virtue of a horn resides in the possibility of presenting practically any desired value of acoustical resistance to the driving system. The horn provides a matching means between the large acoustical impedance of a relatively dense diaphragm and the small acoustical impedance of the relatively light air.

The mechanisms in use today for coupling to horn loudspeakers are of the permanent-magnet dynamic type. These mechanisms may be classed as the large- and small-throat types. The large-throat mechanism shown in Fig. 11A resembles the direct-radiator mechanism. The diameter of the diaphragm ranges from 10 to 15 in. The voice coil is usually about 2 to 3 in in diameter. The flux density in the air gap is usually maintained at a high level of 15,000 Gauss to obtain a high order of efficiency. In general, the large-throat mechanism is used for reproduction in the low-frequency range. The small-throat type mechanism is shown in Fig. 11B. There are many ramifications of this design. The diameter of the diaphragm ranges from 1 to 4 in. The diameter at the throat of the horn ranges from $\frac{1}{4}$ to 1 in. The flux density in the air gap ranges from 15,000 to 22,000 Gauss. In some applications the mechanism shown in Fig. 12B may be used to reproduce the upper portion of the audio-frequency range, while the mechanism of Fig. 11A is used to reproduce the lower portion of the audio-frequency range. Under these conditions, an electrical network is used to allocate the power input to the loudspeakers. With a suitable horn, either of the mechanisms of Fig. 11 may be designed to cover the entire audio-frequency range.

Many types of horns are used in horn loudspeakers. A straight-axis large-throat horn is shown in Fig. 12B. A folded large-throat horn is shown in Fig. 12A. A straight axis small-throat horn is shown in Fig. 13B. A folded small-throat horn is shown in Fig. 13A.

The efficiency of horn loudspeakers may range from 20 to 50 per cent. Horn loudspeakers have been built which will deliver as much as 1 kw of the sound output.

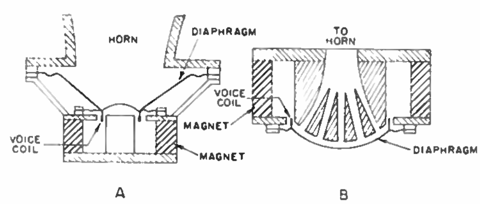


Fig. 11—Mechanisms for horn loudspeakers. A) Large throat. B) Small throat.

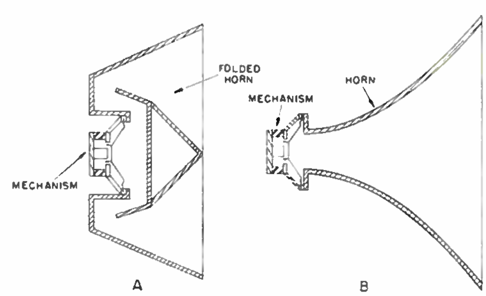


Fig. 12—Large-throat horns for horn loudspeakers. A) Folded horn. B) Straight-axis horn.

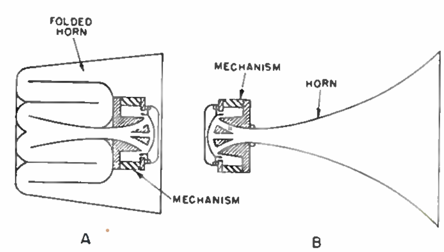


Fig. 13—Small-throat horns. A) Folded horn. B) Straight-axis horn.

Miscellaneous Loudspeakers

Direct-radiator and horn loudspeakers employing electrodynamic driving systems have dominated the loudspeaker field for more than three decades. Nevertheless a brief mention of other loudspeaker systems which have been commercialized is appropriate.

The magnetic driving system of Fig. 1B, coupled to a diaphragm, has been applied to both direct-radiator and horn systems. The magnetic loudspeaker was used in large quantities in the late twenties and early thirties. However, with the advent of high-coercive high-energy permanent magnets required in dynamic systems, the magnetic loudspeaker was displaced by superior dynamic loudspeakers.

The electrostatic or condenser loudspeaker has been in sporadic and limited commercial use for more than three decades. The electrostatic driving system was shown in Fig. 1C.

Commerical electrostatic loudspeakers for the frequency range above 7500 cps are built with a diaphragm of 0.001 in plastic with a thin metallic coating and a diaphragm area of a few square inches. The diaphragm rests directly upon the perforated metal back plate. The electrostatic loudspeaker is coupled directly to the vacuum tube in which the plate voltage supplies the polarizing potential. With an effective spacing of about 0.001 in, the sensitivity is comparable to conventional dynamic loudspeakers. This small spacing limits the maximum amplitude and confines the operation to a frequency range above 7500 cps.

Bilateral or push-pull electrostatic loudspeakers have also been commercialized. A diaphragm of one square foot is placed between two perforated plates with a spacing of 1/16 in. A polarizing voltage of 3000 volts is used. Since the directivity pattern of a diaphragm of these dimensions is quite sharp in the high-frequency region, two or more units are used and directed so as to obtain the required coverage. The use of the bilateral design makes it possible to extend the response to lower frequencies.

A throttled air-flow loudspeaker consists of a valve mechanism actuated by a dynamic or magnetic driving system. The valve mechanism modulates the steady air stream so that the undulations in the air stream correspond to the variations in the electrical input. The throttled air stream is usually coupled to a horn to improve the efficiency of the system. The throttled air-flow loudspeaker has not been commercialized to a significant extent.

A commerical thermoelectronic loudspeaker consists of an audio-modulated confined corona discharge coupled to a horn. The volume of air is expanded by the gas. Therefore, the system is essentially a constant amplitude device. For this reason the useful response range is confined to the high-frequency portion of the audio-frequency range.

Future Developments

The theoretical and practical superiority of the dynamic loudspeaker, established and maintained over a period of three decades, would appear to preclude the probability of any possible replacement by another system. Many improvements have been made in the performance of the dynamic loundspeaker in the past three decades. These have resulted in smoother response-frequency characteristics, more uniform directional patterns, lower nonlinear distortion, and superior response to transients. Work will continue in these directions. Among possible developments will be a feedback system for controlling the sound output. The most promising design for the low-frequency range is the pressure-control system shown in Fig. 14. The response can be controlled over the low portion of the frequency range which presents the greatest problem on obtaining low nonlinear distortion. The most promising design for the high-frequency range appears to be the motion control system shown in Fig. 15.

Other possible applications involve thermoelectronic and dynaelectronic systems, all of which involve the ionization of air. The thermoelectronic system has been described under miscellaneous loudspeakers. In the dynaelectronic system the air is ionized, and the ionized volume is set in motion by the application of an electric field corresponding to the signal.

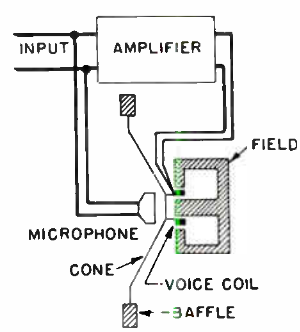


Fig. 14—Loudspeaker, amplifier and microphone feedback system.

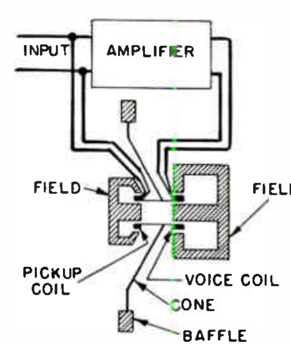


Fig. 15—Loudspeaker, amplifier and pickup-coil feedback system.

Disk Recording and Reproduction*

W. S. BACHMAN†, FELLOW, IRE, B. B. BAUER‡, FELLOW, IRE, AND
P. C. GOLDMARK‡, FELLOW, IRE

Summary—**In 1878, Thomas Alva Edison built the first working phonograph but it remained for Berliner to bring together the various ideas which charted its future course. Major electroacoustical advances responsible for vast improvement in disk recording and reproduction took place during and after World War I, and into the twenties, such as the concept of mechanical and acoustical impedance, "electrical recording," folded exponential horn, and electromechanical pickup.**

Further advances occurred in the thirties: utilization of Rochelle salt crystals in phonograph pickups, the use of negative feedback in recording, as well as significant contributions to the theory of groove-stylus relationships. The end of World War II marked the beginning of "high fidelity" as a household word including developments of improved magnetic pickups, advances in measurements and calibrations, utilization of barium titanate transducers and improved record materials, and culminating with the introduction of long-playing records.

The fifties saw an ever widening use of magnetic tape for mastering, development of hot-stylus and variable-pitch recording techniques, and continued trend toward lightweight pickup design. These improvements, combined with ideas on binaural and stereophonic recording that have been dormant since the twenties and thirties paved the way to the introduction of stereophonic disks in 1957.

In the hands of the artist and the composer the advances in disk recording are leading to new art forms as well as to more faithful reproduction of great classics.

Introduction

IN AN ARTICLE for the *North American Review* of June, 1878, Thomas Alva Edison forecast that the phonograph would be used in ten ways. It was a remarkable prediction. Disk recording systems have been used in all of the applications, and multimillion dollar industries have been built around at least two of them, the reproduction of music, and office dictation.

Recording of telephone conversations, the teaching of elocution and other subjects, and talking books for the blind, these were also foreseen by Edison, but he long persisted in regarding the phonograph as primarily an office dictating machine. We believe that he would be shocked to know that the phonograph is used today for plays and poetry, comedians and preachers, sportscar roars and train whistles, and advertising and sales messages. Edison certainly would be aghast at records to relax by, to reduce by, and for dispersing birds or capturing mosquitoes, yet all these have been produced.

Specialized applications (even in dictating machines) are secondary in cultural and economic importance to the use of disks for recording and reproducing music. Our consideration here, then, shall be mainly entertainment disk recording and reproduction, its history and growth from an idea which sprang almost simultaneously in the minds of several men, and developed, through decades of strife, into today's giant record industry, which amounts to more than half a billion dollars a year.

Historical Note

Alexander Graham Bell's invention of the telephone, in 1876, drew attention to the problems of the reproduction of speech. One of those attracted was Thomas Alva Edison, who was all the more interested in the study of sound because he was partially deaf. Sometime in the fall of 1877, Edison wrapped a sheet of tinfoil around a cylinder, set a needle in contact with it, turned a crank to rotate the cylinder, and into a mouthpiece attached to the needle he shouted the nursery rhyme that begins, "Mary had a little lamb." After making a few changes, he cranked the cylinder again, and from the horn of the instrument he heard a recognizable reproduction of his voice. Thus was born the first phonograph. The date of invention was later recollected by Edison as August 13, 1877, but there is some question as to how precise that date is. We know, however, that the patent application was filed December 24, 1877, and the patent was granted on February 19, 1878.[1]

After almost a decade of preoccupation with other interests, during which he produced the incandescent lamp, Edison returned to work on the phonograph.[2] His improved model of 1888 employed an electric motor powered by heavy-duty batteries. Its advantages were derived from a constant speed for recording and playback, together with the superior qualities of wax rather than tinfoil as a recording medium. In 1888, the pianist, Joseph Hofmann, age 12, visited the Edison Laboratories. He became the first recognized artist to make a phonograph record.

On May 4, 1887, Emile Berliner applied for a patent[3] on what he called a "Gramophone," to distinguish it from Edison's phonograph of ten years earlier and from Bell's and Tainter's Graphophone of one year earlier.[4] The first figure of the drawings in Berliner's patent shows a record wound on a cylindrical support but by

* Received by the IRE, November 17, 1961; revised manuscript received December 7, 1961.
† Columbia Records, New York, N. Y.
‡ CBS Laboratories, Stamford, Conn.

[1] U. S. Patent No. 200,521.
[2] R. Gellatt, "The Fabulous Phonograph," J. B. Lippincott Co., Philadelphia, Pa., and New York, N. Y., p. 33; 1955.
[3] U. S. Patent No. 372,786, granted November 8, 1887.
[4] U. S. Patent No. 341,214.

1888, when he introduced his first model, he had changed to a flat-disk record.[5] The groove in this record had a lateral side-to-side movement, as against the vertical "hill and dale" system which had been employed by others. This lateral recording process was reminiscent of Leon Scott's phonoantograph, which used a diaphragm and hog bristle to trace a record of sound vibrations on lamp-blacked paper some thirty years earlier.

Berliner also used lamp-black as the recording medium, and combined this method with an etching process which permitted transfer of the original engraving to copper or nickel.[6] Thus Berliner achieved a permanent master recording, and for the first time mass duplication of records was possible.[7] No longer did artists have to repeat each number endless times.

By 1895, Berliner had developed a system utilizing many ideas of his own and others: Scott's lateral groove, his own flat disk, and a coating of Bell's and Tainter's wax. The system stood up as the industry standard for half a century, thus Berliner deserves a mantle as the father of disk recording and reproduction.[8]

Edison was again active in phonograph work at the turn of the century. He had tackled the problem of the wax mold used for the original, too. As a nonconductor, this master could not be electroplated. By 1903, he had worked out an ingenious duplicating process under which the original record mold of wax was placed in a vacuum chamber, with a piece of gold leaf suspended on either side. High tension electricity was then discharged between the two gold-leaf electrodes, while the record mold revolved. Upon this gold film, which faithfully adhered to all the original record grooves, a heavier deposit of another metal was then electroplated. The resulting original mold was a solid, durable negative, from which almost any number of positive copies could be duplicated.

Another early contributor to phonograph progress was Eldridge Johnson, who in the 1890's ran a small machine shop in Camden, N. J. By 1901, Johnson had so improved the disk phonograph that it sounded as good as the wax cylinder machine and had much greater volume.[9]

The conflicts during the early years of the phonograph industry, involving charges, litigation, intrigue, and publicity, are too numerous and too complex to treat here. Finally in 1902, all the companies and individuals who had patentable interests got together. Their settlement resulted in three dominant companies, Columbia (the Bell-Tainter interests), Victor (Berliner-Johnson), and Edison.

Electroacoustic Advances

The early efforts of the pioneers, Edison, Berliner, Bell, and others, were based on the empirical approach. Between the beginning of the twentieth century and the end of World War I, there was little progress made in the phonograph art. The War witnessed many advances in acoustical theory; the acoustical horn analysis of Webster,[10] and the concept of mechanical and acoustical impedance by Kennelly[11] and others ushered in the era of scientific electroacoustics; the phonograph was to be its most important beneficiary. The concurrent development of the electronic amplifier provided a further impetus to the art of recording and reproduction.

The best known early effort in that direction was the design of a phonograph recording and reproduction system by Maxfield and Harrison.[12] Their recorder was a magnetic transducer employing a loaded rubber transmission line as a damping element. Records of much improved modulation and clarity were now possible. With the new method of "electrical recording," the sound to be recorded was picked up by means of a microphone, amplified, and applied to the coil of the recorder or "cutter."

The reproducer of Maxfield and Harrison (subsequently licensed to the Columbia Phonograph Company and the Victor Talking Machine Company) was embodied in the so-called "Orthophonic Victrola," introduced in 1925. Together with the "rubber-line recorder" and a folded exponential horn, the system reproduced a frequency range of 100–5000 cps constituting a significant advance over other phonographs then available. It was not long, however, before the Brunswick and Majestic Companies began to offer phonographs employing electromechanical pickups which began to replace the acoustical phonograph. An electromagnetic pickup of the type used in the early phonographs was described by Kellogg.[13] The early pickups bore upon the record with a stylus force of several ounces, and represented but a slight improvement over the Orthophonic Reproducer. Their virtue resided in the ready control of volume and tone possible through the use of electronic amplification. Improved electromagnetic transducers

[5] L. N. Reddie, "The gramaphone and the mechanical recording and reproduction of musical sounds," *J. Roy. Soc. Arts*, vol. LVI, pp. 633–649; May, 8 1908.

[6] F. W. Wile, "Emile Berliner: Maker of the Microphone," The Bobbs-Merrill Co., Indianapolis, Ind., p. 187; 1926.

[7] E. T. Canby, *et al.*, "The Saturday Review Home Book of Recorded Music and Sound Reproduction," Prentice-Hall, Inc., Englewood Cliffs, N. J., p. 7; 1952–1956.

[8] O. Read and W. L. Welch, "From Tin Foil to Stereo: Evolution of the Phonograph," Howard W. Sams and Co., Inc., Indianapolis, Ind., The Bobbs-Merrill Co., Inc., Indianapolis, Ind., and New York, N. Y., p. 119; 1959.

[9] R. Wallace, "First It Said Mary," *Life*, pp. 87–101; November 17, 1952.

[10] A. G. Webster, "Acoustical impedance and the theory of horns and of the phonograph," *Proc. Natl. Acad. Sci.*, vol. 5, pp. 275–282; 1919.

[11] A. E. Kennelly, "Electrical Vibration Instruments," Macmillan Co., New York, N. Y.; 1923.

[12] J. P. Maxfield and H. C. Harrison, "Methods of high quality recording and reproducing of music and speech based on telephone research," *Trans. AIEE*, vol. 45-I (*Commun. and Electronics*), pp. 334–348; February, 1926.

[13] E. W. Kellogg, "Electrical reproduction from phonograph records," *J. AIEE*, vol. 46, pp. 1041–1049; October, 1927.

for disk recording and reproduction were developed during the thirties by Hasbrouck,[14,15] and others.

Vertical vs Lateral Recording

The lateral recording system adopted by Berliner and espoused by Maxfield and Harrison in the Orthophonic system established the lateral mode of modulation as dominant in home phonographs. The advantage of the moving coil system made itself strongly felt in the design of recorders and reproducers for high-quality applications, such as motion pictures and radio broadcasting. The type of system employed in these devices is exemplified by the moving-coil pickup described by Frederick.[16] The construction is similar to that of a moving-coil microphone, except that a conical diaphragm is used at the apex of which is installed a jewel-stylus tip which may be adapted to record or reproduce a phonograph groove. Since the moving coil normally moves axially the choice of vertical modulation is obvious. The moving-coil pickup was capable of tracking record groove modulation at forces of the order of ½ oz.

The use of negative feedback including the electromechanical system again must be credited to Maxfield and Harrison.[17] This principle was applied to the vertical recorder of Vieth and Wiebush in the mid-thirties[18] introducing the era of truly high-quality recording and reproduction of sound. There is no doubt in the light of the present-day knowledge that the presumed superiority of vertical over lateral recording in the mid-thirties stemmed more from the moving-coil feedback cutter and the moving-coil pickup, than from any basic advantage of the vertical system. This was made evident by the brilliant work of Pierce and Hunt published in 1938.[19] It was well realized by earlier workers that a problem arose because recording is accomplished with a flat-faced tool while playback takes place with a spherical-tip stylus. Pierce and Hunt made an important contribution by recognizing that any form of groove modulation could be represented by a normal sinusoidal modulation of the individual groove walls combined in appropriate phase relationship. They named the motion resulting from the tracing of a sinusoidally modulated surface by a spherical-stylus tip a "poid" and analyzed its harmonic components. With lateral modulation, the odd harmonic components of the poid are retained in the horizontal mode, while the even harmonic components cause a vertical motion of the stylus, which is commonly referred to as "pinch effect." With vertical modulation, all the harmonics, odd and even, remain. Simultaneously, Hunt and Pierce developed a moving-conductor lateral pickup capable of tracking 78 rpm records with a force of 2 grams, and envisioned the possibility of playing back 19 minutes of music on a 12-in 33-rpm record with a 0.75-mil-radius stylus. Further advances in the theory of tracing distortion were published by Lewis and Hunt in 1941.[20]

Hunt, Pierce, and Lewis had assumed for simplicity that the record material is unyielding. In 1941 Kornei[21] used the Hertz equation for calculating the deformation between the cylindrical groove-wall modulation and spherical-stylus tip, thus arriving at the playback loss (or gain) due to the dynamic interaction between the two. This work was extended further by Miller,[22,23] using a similar approach but correcting some of the assumptions of the Kornei analysis. Miller's result established the existence of a translation loss function G which is dependent on wavelength, and a pickup response function H which is dependent on playback frequency, thus providing a good fundamental basis for analysis of the playback process.

Development of Phonograph Reproducers

The role that was to be played by lightweight phonograph pickups in diminishing record wear and improving the performance of phonograph systems was already recognized during the early thirties. The perpendicular force required to hold a pickup in proper tracking engagement with the groove is directly proportional to the mechanical impedance at the stylus tip. This impedance is determined by the tip compliance, its mass, and its damping resistance. Development of low mechanical impedance pickups was greatly assisted by the introduction of the Bimorph Rochelle salt transducer element by Sawyer.[24] By coupling the pickup stylus to the Bimorph element using various types of elastomers and levers, a number of "crystal" pickups capable of tracking at 1–2 oz. were placed on the market.[25–27] Most of these had sufficient output to drive the first audio stage of a radio

[14] H. J. Hasbrouck, "Improving the fidelity of disc records for direct playback," *J. Soc. Motion Picture Engrg.*, vol. 32, pp. 246–252; March, 1939.

[15] H. J. Hasbrouck, "Lateral disk recording for immediate playback with extended frequency and volume range," Proc. IRE, vol. 27, pp. 184–187; March, 1939.

[16] H. A. Frederick, "Vertical sound records," *J. Soc. Motion Picture Engrg.*, vol. 18, pp. 141–163, February, 1932.

[17] U. S. Patent No. 1,535,538, issued April 28, 1925.

[18] L. Veith and C. F. Wiebusch, "Recent developments in Hill and Dale recorders," *J. Soc. Motion Picture Engrg.*, vol. 30, no. 1; January, 1938.

[19] J. A. Pierce and F. V. Hunt, "On distortion in sound reproduction from phonograph records," *J. Acoust. Soc. Am.*, vol. 10, pp. 14–28; July, 1938.

[20] W. D. Lewis and F. V. Hunt, "A theory of tracing distortion in sound reproduction from phonograph records, *J. Acoust. Soc. Am.*, vol. 12, pp. 348–356; January, 1941.

[21] O. Kornei, "On the playback loss in the reproduction of phonograph records," *J. Soc. Motion Picture Engrg.*, vol. 37, pp. 569–590; December, 1941.

[22] "Stylus-groove relations in phonograph records," Office of Naval Res., Contract N5 ORI-76, Project Order X, Tech. Memo. No. 20; March 15, 1950.

[23] Summarized by F. V. Hunt in "Stylus-groove relations in the phonograph playback process," *Acoustics*, vol. 4, pp. 33–35; 1954.

[24] C. B. Sawyer, "The use of Rochelle salt crystals for electrical reproducers and microphones," Proc. IRE, vol. 19, pp. 2020–2029; November, 1931.

[25] R. P. Glover, "A record saving pickup," *Electronics*, vol. 10, pp. 31–32; February, 1937.

[26] A. L. Williams, "New piezoelectric devices of interest to the motion picture industry," *J. Soc. Motion Picture Engrg.*, vol. 32, pp. 552–557; May, 1939.

[27] A. D. Burt, "The reproduction of record noise by pickup design," *Electronics*, vol. 16, pp. 90–93; January, 1943.

set, thus ushering in the era of the popularly priced radio-phonographs.

After World War II, the trend toward lower force pickups was resumed. The relative ease with which substantial amplification was obtainable made it unnecessary to design the pickups primarily with high output in mind, and low mechanical impedance became an important goal of the designers, especially for the so-called High Fidelity applications. An important step in that direction was a variable reluctance pickup[28] capable of tracking at but a few grams. The armature took the form of a light elongated ferromagnetic member with the stylus attached at the free end and the magnetic pole pieces placed close enough to the stylus to generate, in effect, an output proportional to the velocity of stylus motion. A moving magnet pickup for 2-gram operation was described in 1957.[29] Subsequently, this moving magnet principle was applied to stereophonic pickups.

As a result of wartime research, piezoelectric properties of barium titanate ceramics were discovered and used in pickups,[30] as well as other transducers. These elements have properties similar to Rochelle salt, but without the temperature and humidity sensitivity of the latter. Subsequently, lead zirconium titanate ceramic was introduced by Gray,[31] providing a sensitivity and dielectric constant almost equivalent to Rochelle salt.

With the emphasis on low force pickups, the measurement of mechanical impedance at the stylus tip became of interest to pickup designers. Stylus-tip mass could be ascertained by weighing and calculation, but compliance and damping required measurement. A meter for ascertaining the compliance and damping factor by resonating the stylus compliance with a known force-driven mass was described in 1947.[32]

The tone arm which carries the pickup cartridge became the subject of intensive study from the geometrical and dynamical point of view. Even prior to World War II, it had been realized that the vibration of the playback stylus in a pivoted arm did not follow precisely the original vibrations of a cutter stylus moving in a straight line along the ways of a recording lathe. The difference could be minimized however, by curving the arm or offsetting the pickup head and mounting it appropriately with respect to the center of the record.[33] A mathematical analysis of this effect was contributed by Baerwald.[34] Subsequently, the geometry applicable to a variety of practical conditions was derived.[35]

With the development of "High Fidelity" instrumentation, the importance of the dynamics of the arm carrying the pickup head across the record was better appreciated. In 1951, an arm was developed in which damping in the form of a viscous substance was provided between two spherical surfaces eliminating resonance at low frequency.[36] "Dynamic" damping by provision of a viscous substance between the arm and the counterweight was introduced in 1957.[29]

The Development of the Lp[37] Record

In June 1948, the development of the 33⅓ rpm Long-Playing Record System was announced.[38] The relative specifications of the "Lp" records and the then current 78.26 rpm records are shown in Table I.

TABLE I

RPM	78.26 rpm	33⅓ rpm
Groove width	6 mil	2.6 mil
Groove bottom radius	1.5 mil	0.2 mil
Groove include angle	90°	90°
Stylus radius	2.5–3.0 mil	0.8–1.1 mil
Grooves per inch of radius	100	260 ave
Maximum groove/diameter		
12-in record	11½ in	11½ in
10-in record	9½ in	9½ in
Minimum groove diameter	3.75 in	4.75 in
Maximum playing time	5 min	22½ min
Tracking force	2 oz. max	7 gram max

Coincident with this announcement, a player was made available which met the required new specifications and which could be connected to existing phonographs. The development of the Lp system involved the combination of ideas and processes previously offered, as well as some new ones. New pickup lightweight and low cost drive mechanisms minimizing wow and rumble had to be developed to make the system practical for home use, including dual-stylus pickups to play either the 78 or the Lp records.

New methods of recording and mastering were evolved together with a new recording characteristic. Two important improvements followed: the "hot-stylus" recording technique[39] and the variable-pitch system of recording,[40] to be discussed later, which ultimately extended the recording time to 30 minutes per side of a 12-in Lp record.

[28] W. S. Bachman, U. S. Patent No. 2,511,633.

[29] B. B. Bauer and L. Gunter, Jr., "A high fidelity phonograph reproducer," 1957 IRE National Convention Record, pt. 7, pp. 76–81.

[30] H. W. Koren, "Applications of activated ceramics to transducers," *J. Acoust. Soc. Am.*, vol. 21, pp. 198–201; May, 1949.

[31] U. S. Patent No. 2,486,560.

[32] B. B. Bauer, "Measurement of mechanical compliance and damping of phonograph pickups," *J. Acoust. Soc. Am.*, vol. 19, pp. 319–321; March, 1947.

[33] See Glover[25]; also B. Olney, "The acoustical labyrinth," *Electronics*, vol. 10, pp. 24–27, 36; April, 1937.

[34] H. G. Baerwald, "Analytic treatment of tracking error and notes on pickup design," *J. Soc. Motion Picture Engrg.*, vol. XXXVII, pp. 591–622; December, 1941.

[35] B. B. Bauer, "Tracking angle in phonograph pickups," *Electronics*, vol. 18, pp. 110–115; March, 1945; "Pickup placement," *Electronics*, p. 87; June, 1949.

[36] W. Bachman, "The Application of Damping to Phonograph Reproducer Arms," presented at IRE National Convention, New York, N. Y.; March, 1951.

[37] Trademark of Columbia Broadcasting System, Inc.

[38] P. C. Goldmark, R. Snepvangers and W. S. Bachman, "The Columbia Long-Playing Microgroove Recording System," Proc. IRE, vol. 37, pp. 923–927; August, 1949.

[39] W. S. Bachman, "The Columbia Hot Stylus, Recording Technique," *Audio Engrg.*, vol. 34, pp. 11–13; June, 1950.

[40] W. S. Bachman, U. S. Patent No. 2,738,385.

Gradually, the Lp record became adopted for serious music and shows, since it allowed the playing of whole movements of symphonies without interruption.

Shortly after introduction of the $33\frac{1}{3}$ rpm record, the RCA Victor Company announced a 45 rpm, 7-in record playing 4 minutes per side, using groove and stylus specifications similar to the Lp.[41] The 45 rpm records became widely adopted for popular music.

Standardization of Response Characteristics

During the early days of phonograph development there were no means for performing acoustical measurements and most of the development work was done by listening and experimentation. Numerous phonograph "tone chambers" and styli made of various metals, cacti, etc., were sold under the claim of "producing a mellow tone." Not until the work of Maxfield and Harrison[12] were attempts made to establish the transmission characteristic of a phonograph system. Groove modulation was measured by means of a calibrated microscope. A more convenient method for measuring groove modulation velocity by observing the width of a light pattern reflected from a test record was described by Buchmann and Meyer.[42] Later, this method was modified, taking cognizance of the distance of the record from the observer[43] and employing diffraction[44] and interference patterns[45] of light to enhance the accuracy of measurements.

However, the question of what recording characteristic actually was used remained a trade secret for many years. It was not until the recording companies realized the mutual advantages of standardization that various standardizing bodies established Standard Recording Characteristics.[46] The result of these standards was not to diminish the performer's or director's artistic freedom, but rather to insure that the listener would be able to hear what was originally intended. In a parallel effort, means for evaluating and controlling distortion stemming from groove-stylus relationships were studied by Roys[47] and others.[48]

Stereophonic Recording

The possibility of modulating a groove in orthogonal mode so as to produce two tracks for independent modulation was envisioned by early workers, as evidenced by acoustical phonograph patents filed in the twenties.[49,50] Another attempt at two-channel recording, using electrical methods of recording and reproduction took place in the thirties.[51] In the instruments described, it was envisioned that one program channel would be recorded as a vertical modulation, and the other as horizontal modulation.

In a brilliant contribution to stereophonic recording in the thirties, Blumlein[52] not only envisioned the modern $45-45°$ recording method (where one channel is identified with a 45° motion of the inside groove wall, and the other channel with the $-45°$ motion of the outside groove wall), but also laid down the foundations of improved methods of stereophonic sound pickup and broadcasting. A. C. Keller and I. S. Rafuse[53] suggested a number of configurations for cutters and pickups for this system.

Development work which led to the final commercial introduction of stereophonic records as they are known today appears to have been going on simultaneously in various laboratories. Vertical-lateral equipment was under development and it had reached the commercial stage in 1957. The first to announce a commercial $45-45°$ recording system were Davis and Frayne, in 1957,[54] and it was adopted by the Standards Committee of the Record Industries Association of America (RIAA) as being technically superior to the vertical-lateral record.

In the $45-45°$ system, the two channels of stereophonic information are in effect two perpendicularly modulated channels, and therefore they contain all the harmonic components of the poid. To diminish this effect, a 0.7-mil-radius stylus is used for playing stereophonic records. Stereophonic records are recorded with the RIAA characteristics. Significantly, the stereophonic record standards were adopted prior to the commercial introduction of the system.

Artistic and Musical Considerations

In the original acoustical recording system, the nature of the recording equipment placed severe limitations on the choice of repertoire, and soloists were more readily accommodated than large ensembles. Electrical recording removed many of the limitations of acoustical recording and permitted successful recording of large orches-

41 B. R. Carson, A. D. Burt, and H. I. Reiskind, "A record changer and record of complementary design," *RCA Rev.*, vol. 10, pp. 173–190; June, 1949.

42 G. Buchmann and E. Meyer, "A new optical measuring method for phonograph records," *Elekt. Nachrichten. Tech.*, vol. 7, pp. 147–152; April, 1930.

43 B. B. Bauer, "Measurement of recording characteristics by means of light patterns," *J. Acoust. Soc. Am.*, vol. 18, pp. 387–395; October, 1946.

44 P. E. Axon and W. K. E. Geddes, "The calibration of disc records by light measurements," *Proc. IEE*, vol. 100, pt. III, pp. 217–227; July, 1953.

45 B. B. Bauer, "Calibration of test records by interference patterns," *J. Acoust. Soc. Am.*, vol. 27, pp. 586–594; May, 1955.

46 "Standard Recording and Reproducing Characteristic," Record Industry Assoc. of America, Inc., New York, N. Y.; June, 1954.

47 H. E. Roys, "Analysis by the two-frequency intermodulation method of tracing distortion encountered in phonograph reproduction," *RCA Rev.*, vol. 10, pp. 254–269; June, 1949.

48 B. B. Bauer, "Notes on distortion in phonograph reproduction caused by needle wear," *J. Acoust. Soc. Am.*, vol. 16, pp. 246–253; April, 1945.

49 C. M. Heck, U. S. Patent No. 1,342,442.

50 S. S. Waters, U. S. Patent No. 1,520,378.

51 W. B. Jones, U. S. Patent No. 1,855,150.

52 British Patent No. 394,325.

53 A. C. Keller, *et al.*, U. S. Patent No. 2,114,471.

54 C. C. Davis and J. G. Frayne, "Westrox Stereo disc system," demonstration presented at 1957 National Convention of Audio Engrg. Soc., New York, N. Y.; October 8–12, 1957.

tral groups as well as soloists. The early electrical recordings used a single microphone, and balance between the various performers was obtained by their relative positions and spacings.[55]

Recording in the early electrical era in America was influenced greatly by radio broadcast techniques which introduced the mixing of multiple microphones and the use of highly absorbent studios for both broadcast and recording. In contrast, many of the orchestral recordings made in Europe during this period were done in large reverberant concert halls, and the reproduced sound quality was generally considered superior to nonreverberant sound. Now the use of large reverberant halls and studios, as well as the mixing of artificial reverberation, has become almost a universal practice.

The artistic and musical aspects of recording appear to have evolved as a result of the technical facilities available which were developed for recording. In the acoustical recording days, there was little opportunity to consider whether the object of a record was to recreate an original performance on the basis of the listener being, in effect, transported to the hall or the performance being presented in the listening room. The advent of high quality microphones and loudspeakers made it possible to realize more nearly these philosophical approaches. Spectacular demonstrations of stereophonic sound were made by reproducing a program originating in one concert hall and carried by wire lines to corresponding loudspeakers in another hall where the listeners were located. This very closely approached the objective of placing the performers in the listening room.

Experiments with two microphones mounted in an artificial head and transmitting the individual signals to corresponding earphones, closely simulated the placing of the listener in the concert hall. The technique used in most musical recording lies somewhere between these two cases, in that the microphones are placed closer to the source than listeners would wish to sit during the actual performance. Many successful monophonic recordings have been made, however, with a single microphone quite distant from the sound source. In general, this technique lends itself to halls which are large and have agreeable reverberation characteristics.

The use of multiple microphones in recording raises a point of artistic judgment. The balance between choirs and soloists can be manipulated far beyond the possible range in a live performance. In some ways this may make it possible to approach more nearly the composer's wishes, although it can be argued that it may seriously violate his intent. The manipulation of the balance makes the recording an art form in itself. This is most evident in popular recording, and the artistic justification for it is seen in the widespread use of electronic aids in the live performance.

Mastering

In the early acoustical and electrical recording era the masters were made directly at the time of the performance. Frequently a duplicate master was made for playback, the unplayed one being processed to produce a master-matrix. The original recording was cut on wax, which would be seriously degraded by playback. Soon after the introduction of electrical recording, lacquer-coated metal disks instead of wax were used. The lacquer disks were more satisfactory for playback and with the development of improved reproducing equipment, it became practical to make original recordings on lacquers and re-record from them to make the master for processing. This was an important consideration since a replacement master could be provided in the not infrequent case where the original was destroyed in processing.

After World War II, magnetic tape machines became available in this country. Magnetic tape is a nearly ideal medium for studio recording because of the ease of starting, stopping, and playback without damage to the master, and its relatively long playing time. It was quickly adopted as the universal medium for original recording.

While most record companies made multitrack recordings of most important repertoire, their commercial exploitation was limited to pre-recorded tapes and it was not until the debut of stereophonic disks that intensive development of studio techniques for stereophonic recording were begun. Currently, it is common to do original tape recording on three-track machines, from which the stereophonic master as well as the monophonic master are derived. In the former, it is usual to feed the center-channel information in equal proportion to the left and right channels. The combination of editing and re-recording makes possible the addition of reverberation, the alteration of frequency characteristics, as well as changes in sound balance and acoustic perspective in the preparation of the final master tape.

Master Cutting and Processing

Early acoustical and even electrical recordings were cut mostly on shaved cake wax. Later, flowed wax on glass disks supplanted the cake wax. Following the flowed wax were lacquer-coated metallic disks. The latter were much more satisfactory for playback and gave added flexibility in handling.

The recording stylus made of sapphire is shaped much as a lathe thread cutting tool. It was found that in order to get a quiet cut on lacquers it was necessary to apply "dubbing" facets to the recording stylus.[56] Since these facets are an appreciable fraction of the recorded wavelength at the inner diameters, they interfered with satis-

[55] See for example, H. F. Olson and F. Massa, "Applied Acoustics," Blakiston's Son and Co., Philadelphia, Pa., 1934.

[56] C. J. LeBel, "Properties of the dulled lacquer cutting stylus," *J. Acoust. Soc. Am.*, vol. 13, pp. 265–273; January, 1942.

factory cutting of the highest frequencies. This problem became acute in the case of the Lp record. By applying heat to the cutting stylus,[39] it was found to be possible to use extremely small facets on the cutting edge of the stylus, providing a quiet groove with good, high frequency performance.

Processing

To produce records in quantity, it is necessary to make metal parts from the original cut disk, which is done by electro-deposition. This required the application of graphite powder or gold sputtering to render the surfaces conductive. The graphite process gave way to a far superior one in which pan silvering was employed. In this method, silver is precipitated on the master record much in the way it is done on mirrors. This process has been superseded by spray silvering, which is chemically similar to pan silvering, but is much more readily controlled. It is common practice to deposit nickel on the silvered master a few thousandths of an inch thick, and then build it up to 30–40 thousandths thickness with copper by electroplating. The master matrix is then stripped from the original. It can be used to press records, but it is usually passivated and put back in the bath to develop a metal positive which is called a mother. After a play test to determine its quality, it is passivated, put back in the electrolytic bath, and stamper plates are grown. While other metals are frequently used, it is common to make mothers and stampers of solid nickel. The quality of the surface produced in this electrodeposition process is quite remarkable. For an SNR of 60 db, the magnitude of surface irregularities must be less than one millionth of an inch. This SNR is closely approached in commercial pressings.

Specialized Applications

Dictation Machines

The use of phonograph recording and reproduction for dictation is one of the uses first envisioned by Edison for the phonograph, and indeed, until recent days the wax cylinders which were shaven after use were a common sight in business offices. Today most office dictation (other than that done on magnetic material) is done by embossing, rather than a cutting process. A blunt stylus presses against a moving vinylite disk or cellulose acetate belt, and is modulated by the voice of the user speaking into a microphone. A pickup converts the displacements into electrical currents for playback. The quality of the embossed groove so produced falls short from that of a cut groove, but is sufficient for the purpose intended.

Automobile Phonograph

In 1955, an automobile phonograph system was developed.[57] A special 7-in record, turning at $16\frac{2}{3}$ rpm, and containing up to 45 minutes of music, or one hour of speech on each side was used. 500 and 600 grooves/in were employed and the stylus of $\frac{1}{3}$-mil radius exerted a force of approximately 2 grams on the disk. The pickup cartridge was mounted in a balanced and heavily damped arm, and pivoted to permit handling records without scratching. The phonograph was suspended on elastic mounts to permit operation in the moving vehicle.

Talking Books

One of the ways in which Edison, in 1878, forecasted that the phonograph would benefit mankind was phonographic books, which would speak to blind people without effort on their part. Today, two major organizations are making disk records available to the blind. The Library of Congress, Division for the Blind, handles literary material, but generally excludes textbooks. Recording for the Blind, Inc., a national, nonprofit organization, records textbooks for blind students, primarily college and postgraduate. The two organizations, which cooperate but do not duplicate services, sponsored the development in 1958 of an $8\frac{1}{3}$ rpm Talking Book system.[58] The Talking Book record plays two hours on each side. The record has approximately 650 grooves/in and uses a stylus of $\frac{1}{4}$-mil radius and approximately 2-gram force. The frequency response of this experimental Talking Book system extends to approximately 5000 cycles, suitable for speech.

[57] P. C. Goldmark, "Highway Hi-Fi," *Audio*, vol. 63, pp. 15–17; December, 1955.

[58] P. C. Goldmark, "Latest advances in extra fine groove recording," *J. Audio Engrg. Soc.*, vol. 6, pp. 152–153; July, 1958.

Film Recording and Reproduction*

M. C. BATSEL†, FELLOW, IRE, AND G. L. DIMMICK‡

Summary—**In a historical summary of motion-picture sound recording and reproduction, some of the technical problems and solutions are discussed, starting with the conversion of silent motion-picture studios to sound studios. During the early years, both the sound-on-disk system (Warner Vitaphone) and the sound-on-film system (Fox Movietone) were used. The film recording method was in its early development stage and the sound quality was relatively poor, but it had many inherent advantages which ultimately led to its universal acceptance.**

The first commercially successful photographic sound recording system used a variable intensity modulator known as the Aeo-light. This was later replaced by the ERPI light valve which is still in use today. The optical system used with the light valve is shown, and certain photographic requirements for linear recording are discussed.

The variable area recording system is the outgrowth of the Duddell oscillograph. The oil damped galvanometer was replaced by a magnetic-type optical modulator, and a more effective optical system was developed. Several sound tracks produced by this unit are shown. Noise reduction methods required by both the variable area and variable density recording systems are described.

An unusual type of sibilant distortion caused considerable difficulty in the early variable area sound tracks. The nature of this distortion and the means of correcting it are given.

Although magnetic recording is one of the oldest methods known, it came into its own during World War II. Its potential importance to the motion-picture industry was quickly recognized. The editing, re-recording and reproducing systems play an important part in the over-all process. Many problems had to be covercome before satisfactory performance was assured.

Introduction

SOUND-PICTURE recording and reproduction techniques employ many branches of physics and technology. The early conversion of silent motion-picture studios to sound studios required the development of compatible processes and equipment, and the training of technicians, as well as the construction of new facilities. New stages were required to exclude outside noise and to eliminate undesired reverberation. Cameras were noisy, requiring the development of camera enclosures that were effective noise suppressors but permitted required mobility. Laboratory processing of films had to be established on a controlled basis. Special microphones and booms for moving them, to follow action, were necessary developments. Systems for operating cameras and recorders in synchronism were required.

Extensive facilities for re-recording from several sound films synchronously to make new composite negatives were required. Special editing equipment was necessary to make it possible to cut and splice sound and picture films, retaining synchronization.

Reproducing equipment had to be developed that would be suitable for attachment to motion-picture projectors. The problem of obtaining constant speed at the sound-reproduction point had to be solved. Loudspeaker systems had to be developed that could reproduce speech in auditoriums and theaters with the effect of intimate audience relationship with actors in close-up scenes. This was necessary because the acoustics of large auditoriums and especially existing motion-picture theaters produced disturbing reflections or echoes. Many were too reverberant and required modifications.

Because of the cost of establishing facilities and trained personnel and the uncertainty of studio management regarding the popular appeal of motion pictures with sound, the large producers were not enthusiastic about undertaking the expensive step. As a result of lagging interest in the motion pictures by the public, Warner Bros., in 1926, decided to test the popularity of sound pictures. To minimize the cost of the venture, this studio arranged to have Western Electric develop the necessary equipment to synchronize disk-recording machines with cameras that were housed in booths to suppress the camera noise. Arrangements were made with the Victor Talking Machine Company to do the recording in their facilities and with their personnel. The Victor Talking Machine Company was a Western Electric licensee, and their studios were equipped with Western Electric recording equipment.

Western Electric developed motor drives for theater projectors and disc turntables. These were mechanically connected to the same constant speed motor system. Essentially standard public address system amplifiers and loudspeakers were used. The first picture produced was "Don Juan." In October, 1927, "Jazz Singer" followed and was a success.

The public reaction was so enthusiastic that the large theater chains wanted equipment immediately to play the pictures. Western Electric agreed to lease equipment to them. As a result of the success of the first pictures, Warner Bros. installed disk-recording equipment in their studios. This system was called Vitaphone. It was destined to be supplanted by systems that recorded the sound as photographic images on the same film the picture was printed on. Having demonstrated the popularity of sound pictures and developed the equipment, the industry proceeded with great speed to convert studios for sound-picture production.

In 1928 many "sound stages" were built in Hollywood, new men, new tools and new technicians entered the

* Received by the IRE, July 14, 1961; revised manuscript received, November 2, 1961.

† Consultant, 1460 Via Tuscany, Winter Park, Fla. Formerly with Defense Electronic Products, RCA, Camden, N. J.

‡ Defense Electronic Products, RCA, Camden, N. J.

scene. In early 1929 there were 16 recording channels in Hollywood. By the end of the year there were 116.

During the early years of sound-motion picture, both the sound-on-disk system (Warner Vitaphone) and the sound-on-film-system (Fox Movietone) were used. Disk recording was much more highly developed, and was immediately capable of high quality reproduction in the theater. From the standpoint of production flexibility, ease of handling, and assured synchronization of sound and picture, the disk method left much to be desired. The film recording method was in its early development stage and the sound quality was relatively poor at the start, but it had many inherent advantages which ultimately led to its universal acceptance.

Variable Density Recording

The first commercially successful photographic sound recording system (Fox Movietone News) used a variable intensity method of modulating a beam of light to expose the film negative. The gas-filled lamp known as the Aeo-light had an oxide coated cathode, and its intensity could be modulated over a considerable range by varying the anode voltage, at audio frequencies, between 200 and 400 volts. The Aeo-light was mounted in a tube which entered the camera at the back. Directly against the film was a light restricting slit which passed a beam about a tenth inch long and 0.001 inch high, placed between the picture and the sprocket holes. The Aeo-light could produce a sufficiently high intensity to expose the sensitive negative films used for picture taking. The system worked quite well for news photography, where the sound and picture were taken simultaneously on the same camera. For studio photography, however, there was need for separate but synchronized picture and sound cameras to permit the flexibility of multiple picture cameras, and the separation of the sound and picture recording facilities. The separate sound camera could use higher definition, lower sensitivity film, and thus improve the quality of the sound.

To meet the need for a higher intensity, higher quality variable density sound recording system, Electrical Research Products, Inc. (ERPI), a subsidiary of Western Electric, introduced in 1927 a complete recording equipment utilizing a light valve (Fig. 1) as the sound modulator. An optical system (Fig. 2) permitted the illumination of the light valve ribbons by the image of a high-intensity incandescent lamp. The space between the ribbons gave the effect of a variable width slit focused upon the film by means of a high resolution lens. Modulation of the valve occurred when current passed up through one ribbon and down through the other, and vice versa.

After sound had thus been recorded upon a film negative and this had been developed, it could be recovered by scanning the track with a fine light beam and permitting the transmitted light to fall upon a photocell. This process did not necessarily result in high-quality

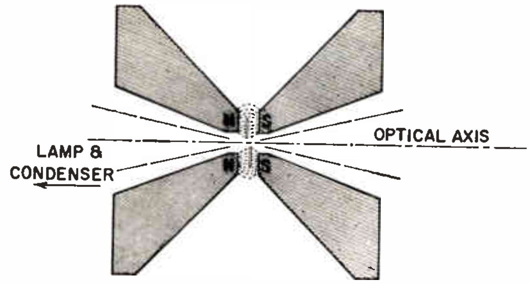


Fig. 1—Light valve ribbon and pole piece arrangement.

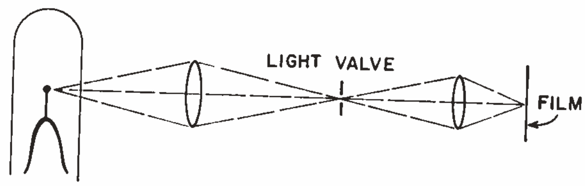


Fig. 2—Optical system for light valve recording.

sound recording because the transmission of the negative did not bear a linear relation to the exposure. Instead, T_n varies as $(1/E_n)^{\lambda n}$ where E_n is the negative exposure, T_n is the negative transmission and λn is an exponent referred to as the negative "gamma" or development factor. However, there is a tendency for the nonlinearity or distortion in the negative to be reduced in contact printing. It can be shown that the transmission of the print T_p varies as $E_n^{\lambda_n \lambda_p}$ where λp is the "gamma" of the prints. If the product $\lambda n \lambda p$ can be made to equal unity, the print transmission will bear a linear relation to negative exposure, and distortion will be eliminated. In normal practice the negative is developed to a gamma of about 0.4 while the print is developed to a gamma of about 2.5, thus satisfying the above conditions for low distortion.

Variable-Area Recording

In 1928 RCA Photophone, Inc., was formed to exploit the sound-on-film developments that had taken place at the General Electric Company and the Westinghouse Company. The variable-area system developed at GE was chosen as the basis for a product design. Variable-area recording systems may be considered to be an outgrowth of the Duddell oscillograph in which a tiny galvanometer swings a spot of light back and forth across a moving film and traces a picture of the wave shapes of the electric currents sent through the galvanometer. Fig. 3 shows the essential optical arrangement of the oscillograph. For sound recording the narrow wavy line was not satisfactory, and a modification of the optical system was made to produce a track like that shown in Fig. 4. When in reproduction a track of this type is scanned by a fine uniformly illuminated beam of light, the transmission to the photocell bears a linear relationship to the modulating current in the galvanometer.

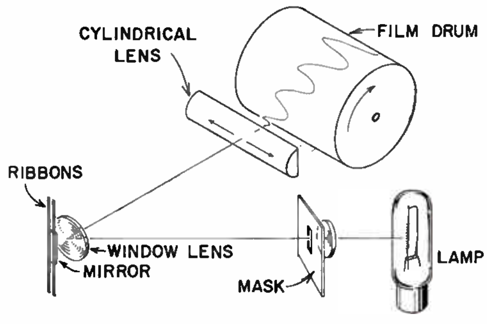


Fig. 3—Optical system of Duddell oscillograph.

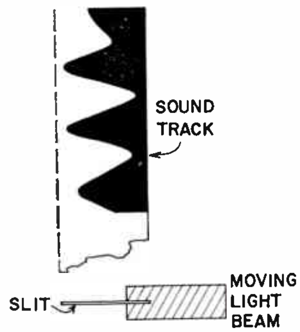


Fig. 4—Variable-area sound track.

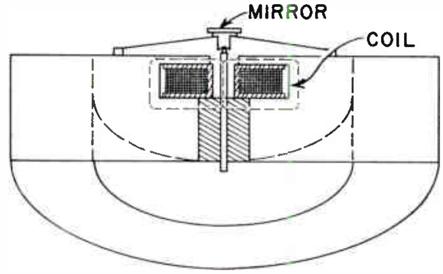


Fig. 5—Magnetic recording galvanometer.

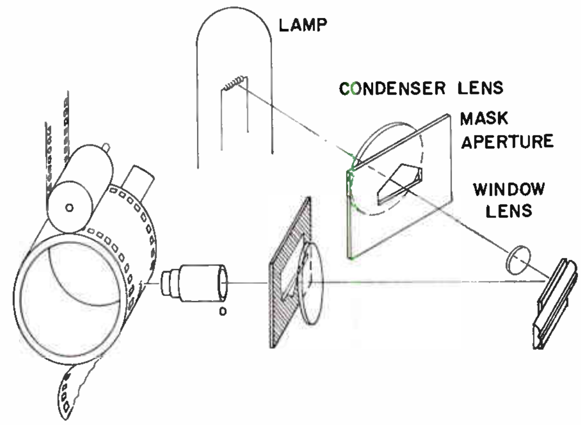


Fig. 6—Optical system for variable-area recording.

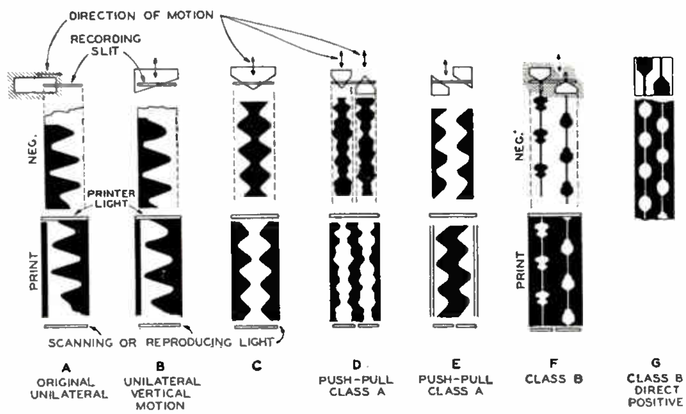


Fig. 7—Types of variable-area sound tracks.

The first generation of variable-area recorders used an oil-damped Duddell oscillograph galvanometer and an optical system to produce a track like Fig. 4. It soon became apparent, however, that the very small 0.017″ ×0.060″ oscillograph mirror was seriously limiting both the brightness and contrast of the optical image used to record the sound track. The oil damping was also very sensitive to temperature, and this caused difficulty when recording in hot or cold climates. Intensive development over the next several years resulted in a new recording galvanometer of the magnetic type (Fig. 5) with a modulating mirror having about twelve times the area of the mirror on previous galvanometers. To take full advantage of this new modulator, a new optical system (Fig. 6) was developed to provide a high-quality recording image with sufficient intensity to record on slow, high-resolution film. Further investigation showed that even higher resolution of the recorded sound track could be obtained by the use of an ultraviolet filter in the light path of the optical system. Since the film emulsion was sensitive to ultraviolet light, but heavily absorbed it, the end result of the UV filter was to concentrate the image on the surface of the emulsion and to prevent scattering of light and reflection from the film base.

An important advantage of the recording optical system shown in Fig. 6 was that the system was readily adaptable to making various types of track. The light spot moved transversely to the slit instead of parallel to it. By making apertures of different shapes, and focusing these upon the slit, many types of sound track were made on the same recording system. These are shown in Fig. 7.

Noise Reduction

The signal-to-noise ratio of a normal variable-density or variable-area sound track is only about 35 db, and the noise increases with film usage. It was recognized from the start that a much higher SNR is needed to accommodate the dynamic range of sound called for in the theater. A great deal of effort has gone into the solution of this problem. In the case of variable density recording, the principal cause of noise is the "graininess" of the developed film. This causes random variations in the transmitted light, which gives rise to a soft hiss in the reproduction. For a typical film and developing condition, the relation between noise level and film density is shown in Fig. 8. Density is defined as the logarithm of the reciprocal of the transmission. It will be noticed that on the straight portion of the curve, 0.3 change in density (2 to 1 change in transmission) results in a 6 db change in noise.

In the case of the variable area sound track, the dark portion transmits a negligible amount of light and contributes in a minor way to the noise. The clear area of the track is not completely transparent, but contains some grains of silver, imbedded particles of dirt and scratches. These are the sources of the noise, which is of a distinctly different character from the noise in a variable density track. The dirt and abrasions give rise to numerous "ticks" and "pops," the number of which vary directly with the width of the clear area. The measured rms value of the noise changes 3 db for each 2 to 1 change in clear area.

The light transmission through an unmodulated sound track is analogous to the RF carrier in an unmodulated radio transmission. This is normally 50 per cent of the transmission required for the sound peaks. It is obvious, however, that if the recorded amplitude is only 10 per cent of the peak value, the average value of the carrier need only be 5 per cent. The noise resulting from a 5 per cent carrier is much less than that resulting from a 50 per cent carrier. The noise reduction system, which has been in constant use over the past 32 years, depends upon an automatic circuit and mechanism for supplying just enough carrier to support the amplitude of modulation being recorded. When there is no modulation, the track transmission is reduced to a minimum value of approximately 5 per cent of its peak. The effectiveness of this system depends upon the fact that ground noise is most objectionable when the modulation is low. An improvement of 10 or 12 db in SNR is obtained by this method.

Two variable area sound tracks with ground noise reduction are shown in Fig. 9. One of these [Fig. 9(B)] was made with an optical system which employed a bilateral shutter to restrict the clear area of the track is required. The other track [Fig. 9(A)] was made with the optical system shown in Fig. 6. A special "bias" winding on the recording galvanometer permits the clear area to be reduced as the modulation decreases. Current for activating either the shutter or the bias winding is obtained by applying the audio-frequency voltages to a rectifier, and then filtering the rectified current, so that at no time does it change rapidly enough to contribute audible noise in the reproduced sound. The design of the filtering system involves a compromise between clipping the tops of the waves (in the case of a quick increase in sound amplitude) and moving the shutter or bias current fast enough so that the motion itself becomes audible. In variable density systems, noise reduction is accomplished by automatically darkening the sound track print when the modulation is low. The light valve shown in Fig. 2 accomplishes this by bringing the ribbons of the valves closer together by means of a biasing current obtained as described above.

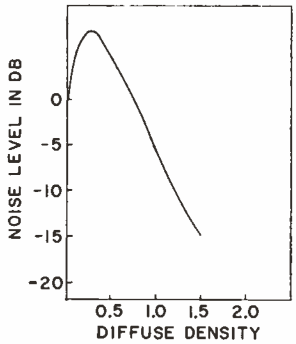


Fig. 8—Relation between density and noise.

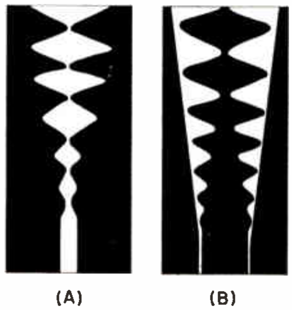


Fig. 9—Variable-area sound tracks with noise reduction.

Optimal Processing of the Variable Area Negative and Print

From the beginning it was recognized that the processing of a variable-density sound record must be carefully controlled in order to maintain a uniform sound level and to minimize harmonic distortion. Many of the early investigators believed, however, that the variable area method avoided this problem because the modulated track was a true graphical picture of the sound waves. In the low- and mid-frequency range this argument was valid, but at high frequencies, where image spreading is of a magnitude comparable with the recorded wavelength, careful control of the processing is needed to optimize the frequency response and to eliminate cross modulation. Fig. 10 shows how the shape of a recorded HF sine wave is affected by exposure of the negative. When this negative is printed, two opposite effects are present. The negative wave-shape distortion is reversed in the printing operation, and more distortion is introduced by image spreading in the print emulsion. Since these two effects are of opposite sign, they may be made to cancel at certain negative and print densities.

The harmonics produced by wave-shape distortion of the higher frequencies have practically no effect on the quality of the reproduced sound. The resultant change in the average transmission does however give rise to a very serious impairment of sound due to cross modulation. When a HF wave is modulated at a LF rate, which happens during the sibilant sounds in speech, cross modulation results in the introduction of a LF signal not present in the original speech. When variable area sound is not correctly processed, the "S" sounds are accompanied by an explosive LF blast which is intolerable. Soon after the nature and cause of this distortion were understood, a method was quickly devised for eliminating it through optimal processing of the negative and print.

An electronic test set was built to provide a 9000 c sine wave, amplitude modulated at 400 c rate. At the end of a normal sound recording "take," a few seconds of this special test signal is recorded. After the negative is processed, the test film is cut off and put through the printing operation several times at different print densities. The test print is then played through a 400 c filter and the 400 c amplitude is plotted against print density. The correct print density is that which gives a minimum of 400 c output. This is a very sensitive and a very practical test method which has been used at many movie studios and sound-processing laboratories for more than twenty years.

Magnetic Recording

Although magnetic recording was one of the oldest methods known, it was not until World War II that this form of recording came into its own. During this period there was developed in Germany a fine grain, low-noise, magnetic oxide, and a process for uniformly coating it on a thin flexible base ¼ inch in width. Use of this new tape in properly designed recorders and reproducers resulted in sound quality which was higher than had previously been obtained from either the film or the disk method. Immediately after the war some of the German recorders were demonstrated in this country, and the potential impact of magnetic recording on the motion-picture industry was quickly recognized. The theory of magnetic recording is covered in a separate article[1] and will not be covered here.

Early in 1948, oxide coated 35 mm film became available for use in motion-picture sound recording. It was then possible to convert photographic sound recorders to combination units capable of recording either magnetic or photographic sound. Many recorders were converted as quickly as possible and were tried in motion-picture sound studios for original "takes." Fig. 11 shows the arrangement of the "tight loop" film path, the optical recording barrel and the magnetic heads, in an early recorder built by RCA. The tests of magnetic recording in the studios were immediately successful, not only because of its high-quality and large dynamic range, but also because magnetic film provided a more flexible and a more economical means of recording. Since re-recording of all original "takes" to a composite photographic negative was already the accepted practice in the industry, little inconvenience was caused by the change.

[1] M. Camras, "Magnetic Recording and Reproduction," this issue, p. 751.

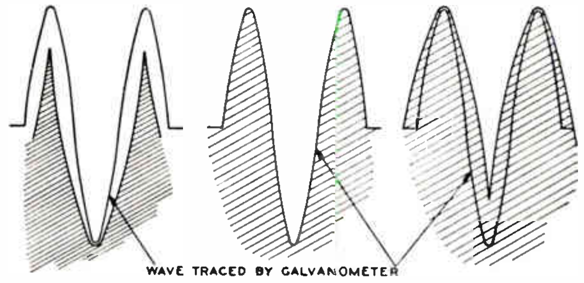


Fig. 10—Effect of exposure on the negative wave shape.

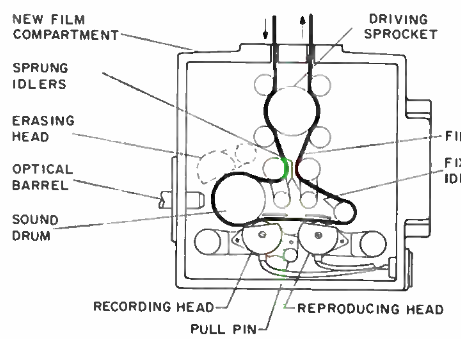


Fig. 11—RCA optical-magnetic recorder.

Editing and Re-Recording

Motion pictures are produced by photographing and recording short scenes or takes. Usually more than one take is made of each scene. These takes are viewed by the production staff daily and a final selection of preferred takes is made. There is then the task of assembling these takes to produce the final picture.

The sound is recorded on a separate film from the picture. At the start of each take synchronizing marks are made on both picture and sound films. These are usually made by means of "slap sticks." After the camera and recorder are up to synchronous speed, each being driven by a sync motor on a common power supply, the sticks are photographed while being struck together. The sound is recorded simultaneously. The take number and picture identification are also displayed in the picture when the "sticks" are photographed.

The editor and film cutters receive these rolls of sound track and picture after they are developed and printed. Some convenient means of viewing the picture and listening to the sound is necessary. For this purpose a small projector and sound reproducer was developed. This became known as a Movieola. It is standard throughout the industry. It is operated by the film cutter. It can be run backward and forward at variable speed as well as the standard 24 frames/second speed, and is convenient to thread.

The sound recording that is made simultaneously with the picture is usually dialogue only. Sound effects and music are recorded separately; in fact libraries of sound-effect recordings were established for frequently needed sounds. Facilities were necessary for re-recording the original sound (dialogue) and simultaneously adding sound effects and music as required for the finished production. Music is played on a scoring stage. This is usually a stage of adequate size to provide, when properly treated with absorbing and reflecting materials, acoustical characteristics to satisfy the needs for various types of music.

A re-recording installation consists of a number of sound reproducers or play-back machines, mixers, amplifiers, monitors and a recording machine. Electrical filters are available for modifying the response-frequency characteristic of the audio-frequency band. The use of a high-pass filter in transmission circuits for dialogue recorded on a stage, for example, may be used to eliminate rumble caused by traffic, etc., thus eliminating the need for a stage that is completely vibration-free. Reverberation chambers are employed to enhance certain sound effects. In these cases the amplified output of a reproducer is supplied to a loudspeaker in a reverberant chamber. A microphone is installed in the chamber and its amplified output is available at the mixer controls.

During each re-recording operation one reel of dialogue film is run, along with all sound effect and music films that have been supplied by the cutting room. Sound effects and music are spliced in blank film to make the same total footage as the dialogue reel, and to insure that each passage is reproduced at the proper time. The re-recording mixer controls are manipulated so as to produce the desired sound in the monitor. This usually requires running the films a number of times while suitable adjustments of gain and audio characteristics are determined. When these tests show the results to be satisfactory, the films are rerun and the combined output is supplied to a recording machine where the negative for the release print is made. The several reproducing machines and the recorder are driven by Selsyn motors which lock all rotors in a fixed angular position at stand still, and maintain this relationship while accelerating and running. This permits threading all films to start marks for proper synchronization.

Film stock used for sound recording is similar in characteristics to stock used for positive prints of pictures. It is fine grained in comparison to picture negative film. It is designed to reduce light scattering in the emulsion, thus improving the resolution. The equipment used in the re-recording installation is the result of a great amount of development work to produce the needed characteristics and to optimize the operating procedures. The design of gain controls and volume indicators are examples of items that are very important. Filter designs, methods of insertion, etc., are also important aspects of mixer and control facilities.

Reproducing Systems

The most difficult problem to solve in reproducing systems was that of uniform motion of the film at the sound take-off point. The film is a flexible material subject to warpage and shrinkage. It is driven by sprockets that engage perforations whose spacing may vary with shrinkage. When film is drawn past the take-off point by a sprocket, there is usually a modulation of the sound at a frequency corresponding to the rate of sprocket hole engagements. The shaft on which the sprocket is mounted must be driven through a mechanical filter to eliminate speed variations introduced by gear teeth involved in the driving system. The first systems employed sprockets at the film take-off point. Musical reproduction was not smooth, however, and development of mechanisms to improve film drives continued. The solution was to wrap the film around a smooth drum attached to a shaft having a flywheel on it. The flywheel-drum assembly was driven by the film which was held in contact with the drum by a pressure roller or by tension. When a pressure roller was used in connection with very low friction bearings, the stiffness of the film in the conventional film-path provided adequate elasticity to absorb frequency variations introduced by the driving sprockets. The flywheel assembly was damped by having a flywheel coupled to the assembly by means of an oil film. This damper prevented LF oscillation of the system.

In systems employing tension to obtain traction of film on the drum, spring-pressed idlers were employed. The springs provided elasticity to absorb variations in speed. The flywheel was solidly secured to the drum shaft. Damping of system oscillation was obtained by dash pots attached to the idler arms or other means of providing viscous friction.

Photoelectric cells of the caesium variety were employed to convert the varying light produced by the shuttering effect on the scanning light beam. Incandescent lamps were employed to produce the light. The development of caesium photocells, with high sensitivity in the red, permitted the lamps to be burned at low temperature, thus resulting in long life. Before the caesium photocell was available, potassium cells were used. Lamp life was a serious problem because the cells were only sensitive in the blue region.

Acoustical Problems in Recording

To simplify the problem of recording dialogue as the action is photographed, sound effects, etc., are generally excluded, so that the most effective microphone placement for natural quality and perspective can be obtained.

Sets usually are assemblies of flat surfaces of convenient dimensions for ease of assembly. Many of these are hard surfaced and reflect sound. Reflections can cause distortion or echoes. Through the use of microphone booms to provide flexibility and rapid movement of the microphone, and through the use of microphones having suitable directional properties, the operator who is responsible for the recorded sound can obtain an appropriate perspective. When it is difficult or impossible to do so, the sound may be recorded after the picture is made by having the actors watch the picture and speak the lines so their lip movements and sounds are synchronized.

The development of microphones having a cardiod-shaped directional pattern was a valuable contribution to the solution of the problems encountered in dialogue recording. The insensitivity to sounds to the rear of the microphone served to reduce the pick-up of "set" noise on the stage due to manipulation of equipment during the action. The uniform sensitivity of the microphone throughout a wide front angle made the placement of the speakers less critical.

The sound level of speech in theaters is reproduced considerably higher than the original sound because of the audience noise and other noise present in theaters. This change in sound level results in an increase in LF loudness to the listener. Also, reverberation in theaters reinforces the vowel sounds in comparison to the sibilants. Compressors that act with sufficient rapidity to suppress vowel sounds, and to increase gain during sibilants, have been effective in producing recordings that are easy to understand when heard in theaters. Before compressors were used, recordings that were excellent in small rooms were frequently very difficult to understand when heard in motion-picture theaters.

It is necessary to change the frequency characteristics of the recording system when the original sound is to be amplified as heard by the audience. Loud sound is satisfactory when reproduced through a linear system. Speech spoken softly, when amplified, is not natural. The low frequencies should be reduced as well as the higher-frequency sibilants. Some actors or actresses, when speaking softly, emit sibilant or hissing sounds that have a larger amplitude than the vowels. The required frequency characteristics are effected automatically through the use of a proper arrangement of filters and a fast-acting compressor amplifier.

Conclusion

The trend today is to employ magnetic recording for original takes. Magnetic materials are applied to perforated film, thus requiring a minimum modification of systems. Essentially the change is to replace the optical systems and modulator with the magnetic recording head, or in the reproducer, to replace the lamp, optical system and photocell with a magnetic reproducing head.

The magnetic recording is used to produce a photographic negative by re-recording, to make prints for general release to theaters. For release prints to theaters having stereophonic reproducing systems, the prints have magnetic sound recorded on strips of magnetic material that have been applied to the film, and which are capable of withstanding the picture developing process.

Current Problems in Magnetic Recording*

MARVIN CAMRAS†, FELLOW, IRE

Summary—**The past decade has seen great advances in techniques and in applications of magnetic recording until, at present, many of the empirical methods have been forced to their limit. Important future developments depend on a better understanding of fundamental processes which, as yet, have hardly been explored. This paper discusses some of the explanations and theories that are of current interest with regards to the playback process, the recording process, the record medium, ac bias, and noise. Limitations of the theories and avenues for further exploration are suggested.**

ALTHOUGH the principles of magnetic recording are easy to explain qualitatively, there is great difficulty in constructing a quantitative picture. Yet a quantitative or at least a semiquantitative understanding is necessary in order to guide one into fruitful areas for further improvement. There has been considerable work in this field, but much remains to be done Difficulties arise from:

1) The complexity of magnetic field distribution in and near tapes and heads.
2) The nonlinear nonreversible character of magnetic record materials.
3) The action of high-frequency bias combined with the recording signal.
4) Interaction of the above.

As a note of caution, it should be pointed out that many analyses are derived for a specific configuration of head and tape, and may not apply to other configurations. Simplifying assumptions are often made at the outset, which are forgotten by the time one has emerged from several pages of mathematical ritual; and the results thereby acquire an unjustified status.

* Received by the IRE, December 11, 1961.
† Armour Research Foundation of Illinois Institute of Technology, Chicago, Ill.

For simplicity, the examples in this paper deal with the most usual case of powder coated tapes and coil type confronting-pole heads as in Fig. 1. Transverse and vertical recording, flux sensitive heads, solid metal tapes, etc., may be handled in a similar manner.

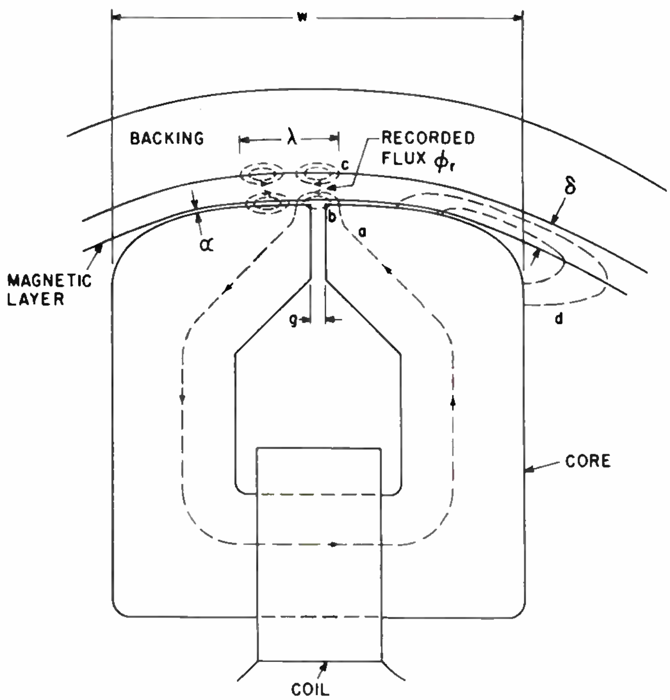


Fig. 1—Head-tape configuration.

The Playback Process

The fundamental concept of the playback process is very easy to understand. Our record is a multipolar permanent magnet. We move this past a transducer head which gives us an electrical output corresponding to the magnetic field of the record. The most common form of transducer has a coil with N turns of wire arranged so that it links *part* of the record flux ϕ_r. Let m be the fractional part of the recorded flux that threads the coil. The induced voltage is then

$$E = mN \frac{d\phi_r}{dt}.$$

The real difficulty comes in trying to predict m; and this is what papers in the field are all about. It turns out that m varies with frequency, wavelength, gap size, geometry of polepieces in the head, permeability of the head, permeability of the record, thickness of the record, direction of magnetization, space between head and record, and on many other factors.

Typical geometry for a popular configuration of tape and head is shown in Fig. 1. Even prior to an exact analysis, one notes that of the total retained flux, only the part a links the coil, while b and c take alternative paths. When the recorded wavelength λ is decreased in size so that it approaches the gap dimension g, there is a cancellation effect that causes the output to go through zero. On the other hand, if the recorded wavelength becomes longer than the gross dimension of the head w, an appreciable portion of the coil flux must complete its path through air as at d and much of it is lost altogether, so that the head becomes very inefficient. One expects also that the part of the tape layer in direct contact with the head is more effective than the outermost layer, and that spacing the record away from the gap is detrimental to high-frequency response.

The playback process has been treated mathematically by making a few simplifying assumptions. Usually these are:

1) That the remanent flux is low enough so that only linear changes take place in the pickup head.
2) That the remanent flux is unaffected by presence of the playback head; or that it is affected a constant amount which does not depend on the variables being studied.

In addition, it is sometimes assumed:

3) That the recorded layer has negligible thickness; or that layers of different spacings from the head act independently of other layers.
4) That the magnetic permeability of the head core structure is infinite, while that of the record member is unity.
5) That sine wave analysis is adequate for describing complex wave recordings.

A powerful tool in theoretical analysis is the reciprocity principle, which states that the sensitivity function of a head on playback is proportional to the recording field function of the same head with unity current through its coil. The relation has been used by Wallace,[1] Westmijze,[2] Gilbert,[3] Fan,[4] and others, and in its most general form may be expressed by the relation

$$E = kvN \frac{d}{dt} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \overline{M} \cdot \overline{H} \, dx \, dy \, dz$$

where

E is the playback voltage,
$\overline{M}$ is the vector magnetization at any point,
$\overline{H}$ is the vector field of the head at any point for unit current in its coil.
The dot product represents the product of the components of each vector in a given direction.
N = number of turns on the pickup coil,
v = velocity of the tape,
k = constant of proportionality.

[1] R. L. Wallace, Jr., "The reproduction of magnetically recorded signals," *Bell Sys. Tech. J.*, vol. 30, pp. 1145–1173; October, 1951.
[2] W. K. Westmijze, "Studies on Magnetic Recording," Phillips Res. Repts., vol. 8, pp. 161–183, 245–269; June, 1953.
[3] T. L. Gilbert, "Theoretical Aspects of Noise in Magnetic Recorders," Armour Res. Foundation, Chicago, Ill., Bulletin No. 94, pp. 43–73; 1956.
[4] G. J. Y. Fan, "A study of the playback process of a magnetic ring head," *IBM J.*, vol. 5, pp. 321–325; October, 1961.

The integration is taken throughout all space symbolically, but the only region of interest is within the magnetized tape layer. The case for a wide tape record is simplified by treating it as a two-dimensional problem; or even further simplified to a one-dimensional problem by assuming negligible thickness of the tape layer.

One may consider tape effects separately from head effects. The tape magnetization is quite complex, as will be appreciated from the next section. The head effects are a geometrical problem in field mapping.

One of the most interesting of head phenomena is the gap effect, since the finite gap size is often the most important limitation on recorded density. This effect was previously analyzed for scanning apertures in photographic sound on film,[5] and a similar function was also found for magnetic recording. In its most elementary form the relation is

$$m' = \frac{\sin \frac{\lambda}{g}}{\lambda/g}.$$

Westmijze[2] derived a more exact relationship for gap loss in the head of Fig. 1. His result is rather complex and is best expressed graphically, as in Fig. 2, where it is compared with the simpler formula. According to Westmijze's formula the output drops to zero when the gap length g is about 0.89 times the recorded wavelength, making the effective gap length about 1.12 times its physical length. It is also interesting that this result is independent of spacing between the tape and head. In practice, the factor of 1.12 is unimportant since the mechanical tolerances of small gaps may give variations greater than this.

As indicated by the maxima near 1.5 λ, 2.5 λ, etc., a head will give output at wavelengths shorter than the scanning gap. Use has been made of this principle in working out a "long-gap" method for measuring surface induction of tapes. From theoretical considerations[2] the response in this case is as shown in Fig. 3, where the output rises at a six db/octave rate for frequencies below the first null. At higher frequencies the line connecting the peaks rises at a two db/octave rate. Unfortunately it turns out that for best results the straightness and parallelism of the gap-defining edges of such a long-gap head must be as exact as for a short-gap head of similar short wave response; so that in most practical applications one might as well make the gap as small as possible and avoid the dips.

In playback of very short wavelengths another important factor is the loss due to even the slightest separation between the record and the head gap.

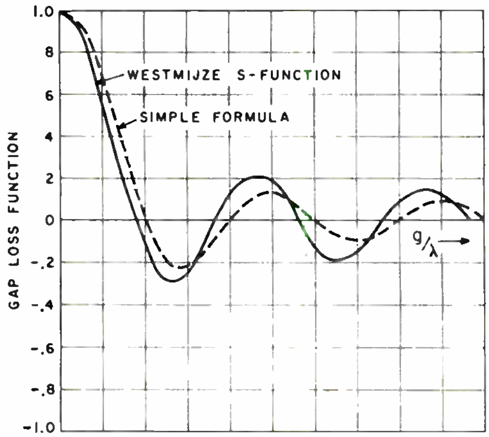


Fig. 2—Loss in output at short wavelengths due to finite gap size.[2]

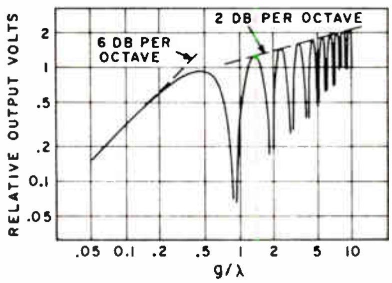


Fig. 3—Output of a long gap head.[2]

Wallace[1] derived a simple relation for this effect:

$$\text{spacing loss, db} = 54.6 \frac{\text{separation}}{\text{recorded wavelength}}.$$

This means that if a 15-kc response is desired at a record speed of 1 7/8 in per second (λ = 0.000126 in), a spacing of 0.00005 in gives a 22-db loss.

Experimental work in recent years has given results that are explainable by the presence of an 0.00006 in spacing.[6] There are several theories as to how this comes about. One explanation is that a surface layer of low permeability forms on heads made of permalloy and similar stress-sensitive core materials. Another is that an air film forms between the head and tape, especially at high speeds. However, such a separation has been observed even in a vacuum; so it is possible that the tape bounces enough to assume a net separation from the head.[7] Usually such an elaborate explanation is not required; mechanical imperfection in both the head and the tape are sufficient to prevent perfect contact.

Electrical engineers are familiar with an analogous situation where it is a rule-of-thumb that a butt-joint in a magnetic circuit acts as if it were at least 0.001 in long, no matter how carefully made. The usual explana-

[5] E. D. Cook, "The aperture effect," *J. Soc. Motion Picture Engrs.*, vol. 14, pp. 650–652; June, 1930.

[6] J. J. Brophy, "High-density magnetic recording," IRE Trans. on Audio, vol. AU-8, pp. 58–61; March–April, 1960.

[7] M. Camras, "Information storage density of magnetic recording and other systems," IRE Trans. on Audio, vol. AU-9, pp. 174–179; September–October, 1961.

tions are surface effects, stresses, etc. But the machining and grinding accuracies on transformer core sections are also of this order, and mask the more subtle effects. When the cores are really made accurately, as in head manufacture, gaps of less than 0.001 in are obtained.

If the magnetic record has appreciable thickness, then the very short wavelengths are not coupled efficiently to the head gap and there is, in effect, a "spacing loss" within the record. This has been termed the *thickness loss*[1]. When the above described losses are taken into account, one obtains for the head of Fig. 1:[2]

$$\phi = \phi_r \left[\frac{1 - e^{-2\pi\delta/\lambda}}{2\pi\delta/\lambda} \right] \left[e^{-2\pi\alpha/\lambda} \right] \left[S\left(\frac{\pi g}{\lambda}\right) \right] \cos 2\pi \frac{v}{\lambda} t,$$

where

ϕ is the flux which threads the coil for a sinusoidal signal,
ϕ_r is the maximum remanent flux (assumed constant for all wavelengths),
S is the gap loss function of Fig. 2,
α and δ are the spacing and tape thickness, as in Fig. 1,
v is the tape velocity,
t is time.

The first bracketed term accounts for the thickness loss. The second term is the spacing loss, and third term is the gap loss.

The effects on the response curve of different factors has been plotted in Fig. 4. An ideal record should have a rising response as in curve 1, diminished by appropriate factors from curves 2, 3, and 4; the result being a curve that resembles Fig. 3 to the left of the first null. But a perfect record is not obtained in practice because of the recording process, erasing effects, self demagnetization, core losses, etc. Hence an actual over-all response is always below that of Fig. 3, especially at high frequencies.

Certain applications, such as memories for computers, depend on the recording and playback of step functions, or of pulses. A very interesting analysis by Hoagland[8] showed that when a step function is played back the output waveform gives a picture of the field function of the head. Eldridge[9] used this relation to investigate the field of a head, especially the effect of spacing, tape thickness, and field strength on the sharpness and location of pulses. He concluded that the horizontal component of magnetization was by far the most important, the vertical components having little effect. It appears that pulses are a useful tool for investigating

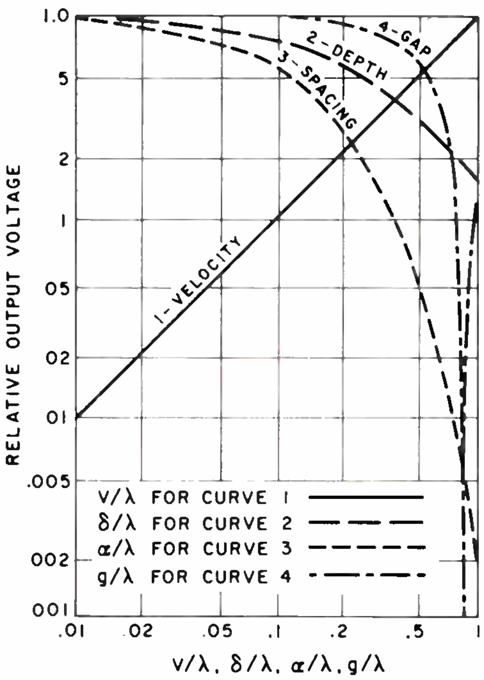


Fig. 4—Effect of tape velocity, thickness, spacing, and gap size on output.

the recording process even where one's interests are in sine wave recording.

On the other hand Kostyshyn[10] used the equations for sinusoidal recordings, but applied Fourier analysis to handle pulse recording. His method was adapted to machine computation, and he calculated the pulse width, recording demagnetization, and output EMF for changes in gap size, tape thickness, bit density, and head-to-tape spacing.

To summarize: the tools for analyzing the playback process are well developed, and the problem is mainly to apply them to realistic situations. This cannot be done if, for example, the magnetization within the record is not accurately known.

The Recording Process

To record on a tape, one applies to it a magnetic field strong enough to leave it permanently magnetized. For maximum information density the recorded regions should be as small as possible, and it has been found that the arrangement of Fig. 1 is excellent for this purpose. The degree of magnetization should be proportional to the input signal if distortion is to be avoided, and this is accomplished by a biasing arrangement. Direct current biasing is well understood,[2] but has disadvantages of noise and distortion, so that ac bias is almost universal in high-quality recorders. For the special cases of step or pulse recording, linearity is not required, and bias is omitted.

[8] A. S. Hoagland, "Magnetic data recording theory: head design," *AIEE Trans.*, vol. 75, pp. 506–512; November, 1956.
[9] D. F. Eldridge, "Magnetic recording and reproduction of pulses," IRE Trans. on Audio, vol. AU-8, pp. 42–57; March–April, 1960.
[10] B. Kostyshyn, "A harmonic analysis of saturation recording in a magnetic medium," 1961 IRE International Convention Record, pt. 2, pp. 112–127.

Unlike the playback process where linear reversibility makes it the joy of the mathematician, recording depends on hysteresis of ferromagnetic materials whose very nature is nonlinear and nonreversible. In addition, there are head and tape geometry effects and the nature of high-frequency bias. These must be considered simultaneously. It is not sufficient to analyze them separately and to superimpose the results.

One approach is a step by step analysis for a definite tape, head, bias, and signal. This is laborious, especially since it must be repeated a large number of times if one is to ascertain the effects of each variable. A situation of this kind is ideal for computer programming. Computer studies of this kind have been made for specialized instances of biasless pulse recording,[10] but not for the more difficult and more interesting general case where sine waves or complex waves are recorded with the use of high-frequency bias.

The results of such analyses should check closely with test data. Here again much work remains to be done, since agreement is not good. We may only guess at the reasons: perhaps not all the important elements were considered; possibly some of the theoretical relations do not hold in practice; perhaps the values of some parameters were incorrect.

Of all the factors that enter into magnetic recording the action of ac or high-frequency bias is the most intriguing, because intuitively one would not expect it to work, and because its operation is complex and obscure. Almost every investigator in the field has been dissatisfied with previous explanations, and the more venturesome have contributed their own versions which have in turn been rejected by their successors.

What takes place is as follows: If a signal field is applied directly to an unmagnetized record the residual magnetization vs applied field characteristic (B_r vs H) has the form shown in Fig. 5, which is an unsatisfactory output-input curve for linear recording. But if we add a weak signal field to a strong high-frequency bias field, and apply the resultant as in Fig. 6(c) to the record, then we obtain the input-output curve of Fig. 7, which has excellent linearity. The bias frequency, generally about 30 to 200 kc, does not appear in the output.

In explaining ac bias two approaches have been most common. The earliest suggestions were based on a physical picture of what might be happening to the magnetic elements in the record. Carlson and Carpenter[11] said that the bias had "the effect of agitating the recording element so as to greatly increase the sensitiveness of said element to feeble signal impressions." They were interested primarily in sensitivity to wireless-radio telegraph signals, and apparently did not realize that there might be conditions where the retained flux would be linearly related to the input signal.

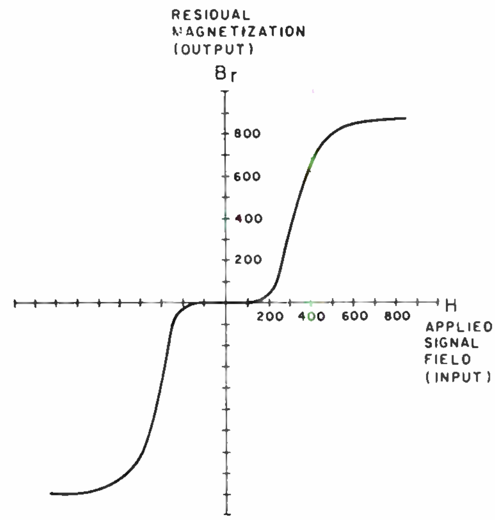


Fig. 5—Output-input characteristics of a magnetic record without bias.

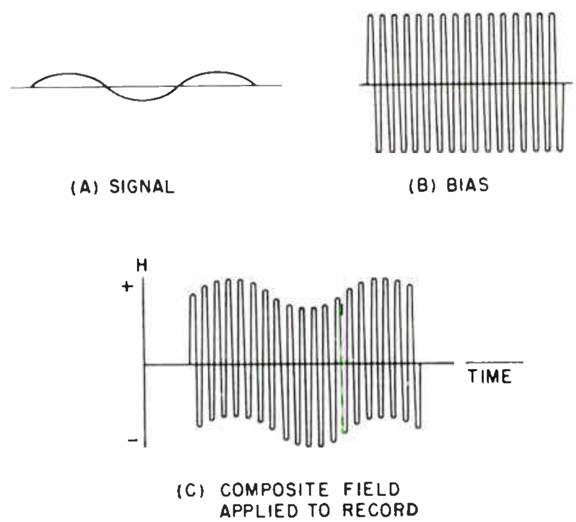


Fig. 6—Fields used for ac-biased recordings.

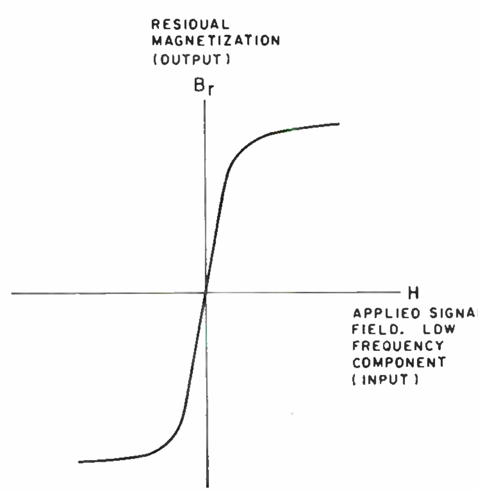


Fig. 7—Output-input characteristics of a magnetic record when ac bias is added to the signal.

[11] W. L. Carlson and G. W. Carpenter, "Radio telegraph system," U. S. Patent No. 1,640,881; August, 1927.

Later analyses were made in terms of the hysteresis loops of the magnetic record material,[12] [13] or its B_r vs H (residual induction vs applied field) characteristics.[14-17] For these it was not necessary to consider the mechanism of what took place at the domain level. The problem was reduced to: Given a family of magnetic characteristic curves and a complex sequence of applied magnetic fields. Find the magnetism retained by the record after it has been subjected to these fields. An analysis along these lines shows that the addition of a strong high-frequency field to the signal gives a linear output vs input characteristic.

Some explanations use the B_r–H curve such as Fig. 5 as their starting point. In so doing they assume (often tacitly) that the record is moving fast enough to present fresh tape at the recording point even for the bias frequency; in other words, the bias is recorded together with the signal. The reason for this is that the B_r–H curve is valid only if the polarity of the applied field never reverses; if the field does reverse, a B–H loop forms and the B_r–H curve no longer applies.

Holmes and Clark[14] used the B_r–H characteristic to plot the waveform of output, compared with a sinusoidal signal input. Their example showed distortion which they indicated might be reduced "by a more judicious choice of bias and audio signals." Camras[15] made a more complete graphical analysis on the basis of actual magnetic B_r-H data of a recording material, deriving the output-input curves, and showing how they varied in linearity and sensitivity as the bias and signal levels were changed. Montani[16] observed that if the rising portions of the B_r-H branches could be approximated by second degree functions, then the distortions would cancel out when the positive and negative cycle contributions were averaged. Zenner[17] set up a polynomial equation that fitted a B_r-H curve, and then operated on it with a function representing the sum of bias plus signal frequency terms. The result, after eliminating the high-frequency components, was a faithful reproduction of the input signal. Recent analyses by Bedford[18] and by Sebestyen and Takacs[19] are based on recording of the bias frequency, and on the pulse width modulation of the recorded bias which results when the signal is added. Such pulse width modulation is inherent also in the model of Bauer and Mee.[20]

While these explanations are valid as far as they go, they do not tell why ac bias still operates, and operates very well, under conditions where the bias is not recorded. In most recorders the resolving power of the head-tape system is inadequate to register the high-frequency bias, which goes through several reversals during the time an element of tape is in the recording gap field. One analysis which does not require that the bias be recorded is given in Camras[12] where minor hysteresis loops at bias frequency were followed as the bias plus signal decayed when passing out of the head-gap. A plot of the final magnetization versus input gave a straight line as in Fig. 7. A different analysis, based on the sloping sides of major hysteresis loops was made by Toomin and Wildfewer.[13]

There are those who feel that manipulation of characteristic curve can never get to the heart of the matter, and that a satisfactory understanding of ac bias is possible only in terms of electron spins and molecular interactions. So we return again to a physical picture, but an incredibly more complicated one than in early days.

In a magnetic record we have an aggregate of magnetic domains which are reversing in direction of magnetization for each reversal of the alternating bias field as long as the field strength is well above the coercive force of the material. As the record leaves the head, the field weakens until, at a certain point, it can no longer reverse the domains. If the domains were magnetically identical we should expect them all to remain saturated in the direction of the last cycle preceding this critical field strength. This is contrary to experience, since a record which has been subjected to a decaying ac field retains practically zero net magnetization. With the above record, if we should add a signal which varied slowly with respect to the bias frequency, the record would be left in a state of either positive or negative saturation depending on the polarity of the added signal. The output vs input curve would then be as in Fig. 8 which is also contrary to experience.

The output versus input curve actually obtained has a finite slope as in Fig. 7. A logical explanation might be that the domains are not all alike magnetically, but have a statistical range of coercive forces. Closer examination reveals the surprising fact that even with a range of coercive forces we still get the infinite slope of Fig. 8 rather than the finite linear range of Fig. 7. This may be seen from a practical example. Let us consider a mixture of magnetic domains having all possible coercive forces from zero to infinity. We now subject them to an ac bias field of 300-oersteds peak value combined with a signal field which is momentarily positive by 30

[12] M. Camras, "Methods and means of magnetic recording," U. S. Patent No. 2,351,004; June, 1944.

[13] H. Toomin and D. Wildfewer, "The mechanism of supersonic frequencies as applied to magnetic recording," PROC. IRE, vol. 32, pp. 664–668; November, 1944.

[14] L. C. Holmes and D. L. Clark, "Supersonic bias for magnetic recording," *Electronics*, vol. 18, pp. 126–136; July, 1945.

[15] M. Camras, "Graphical analysis of linear magnetic recording using high-frequency excitation," PROC. IRE, vol. 37, pp. 569–573; May, 1949.

[16] A. Montani, "The Mechanism of Supersonic Bias in Magnetic Recording," AIEE, New York, N. Y., Tech. Paper No. 49-54; January 31–February 4, 1949.

[17] R. E. Zenner, "Magnetic recording with ac bias," PROC. IRE, vol. 39, pp. 141–146; February, 1951.

[18] L. H. Bedford, "Magnetic tape recording," *Electronic Radio Engr.*, vol. 36, pp. 320–322; September, 1959.

[19] L. G. Sebestyen and J. Takacs, "Magnetic recording theory of tape magnetization," *Electronic Technology*, pp. 274–278; August, 1961.

[20] B. B. Bauer and C. D. Mee, "A new model for magnetic recording," IRE TRANS. ON AUDIO, vol. AU-9, pp. 139–145; September–October, 1961.

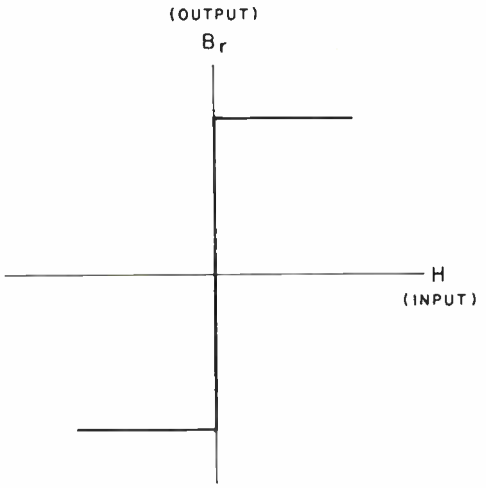


Fig. 8—Output-input characteristics showing infinite slope at origin expected under certain conditions.

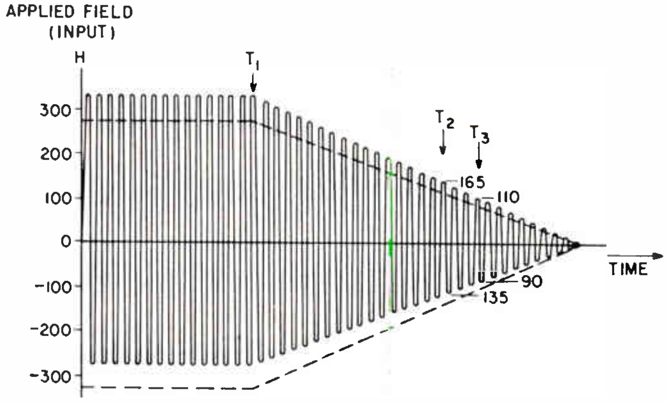


Fig. 9—Example of combined bias and signal field encountered by a record element leaving the head.

oersteds, and reduce both fields to zero at a linear (or any other) rate as shown in Fig. 9 over an indefinitely large number of ac cycles. Those particles with a coercive force above 330 need not be considered, since they are not altered by the maximum field and are small in population. Particles with a coercive force of 330 will be magnetized by the positive cycles which reach 330, but will not be reversed by the negative cycles of 270. Hence they will remain positively magnetized. A particle with coercive force of 135 undergoes a series of reversals up to the time T_2; after this time the negative field is too weak to cause further reversals, but the positive peaks of 165 are effective, so the particle with a coercive force of 135 remains positively magnetized. Similarly a particle with a coercive force of 90 will remain positive after time T_3. It is apparent that under these conditions the input versus output curve will still be as in Fig. 8. A more careful analysis including such factors as higher coercive particles, thermal agitation, and rapid field decrement, will indeed make the slope less than infinite, but hardly enough to account for the experimental results.

Early studies of the action of combined ac and dc fields on ferromagnetic materials were made by Steinhaus and Gumlich.[21] Their conditions were different from the ones encountered in magnetic recording, but they obtained characteristics resembling the ones of Fig. 7, which they called ideal magnetization curves. Their explanation followed a line of reasoning similar to the one given above, and they arrived at the conclusion that the curve should have an infinite slope as in Fig. 8.

Westmijze[2] pointed out the differences between ac bias recording and ideal magnetization curves. He gave a very illuminating explanation for the finite linear slope of the Fig. 7 characteristic, by supposing that the magnetic elements could assume various degrees of permanent magnetization. Corresponding to a given strength of retained magnetization, each element has an internal demagnetizing field opposite to the field that caused the retained magnetization. Under the influence of combined bias and signal fields, the element acquires just enough permanent magnetization so that the internal demagnetizing field cancels the external signal field. This concept gives a satisfying qualitative picture as to why the retained magnetization should be directly proportional to the signal; for as the signal changes, the recorded magnetization must adjust itself accordingly to maintain an equilibrium of zero field inside the element.

Independent single domain particles which are saturated in one direction or the other, do not satisfy the condition that permanent magnetization is acquired by degrees, and some other mechanism must be relied on. Woodward and Della Torre[22] filled this gap by showing that in recording tapes the particles are not independent, but have considerable magnetic interaction, and that such interaction explains the shape of the Fig. 7 characteristic. They made use of an ingenious technique which allowed them to measure and to plot both the interaction and the net magnetic differences among particles in actual tapes, as a statistical distribution. To do this they applied the method of Preisach,[23] where each magnetic particle is represented on a diagram by a point whose abcissa is the field required to switch it negatively (H_-), and whose ordinate is the field required to switch it positively (H_+). Fig. 10 shows the coordinate system and indicates where particles of different hysteresis loops would be plotted. Particles not under unbalanced influences are represented by points lying along a 45° line, with the ones requiring stronger switching fields being farther from the origin as indicated by progressively wider hysteresis loops *B*, *A*, and *C*.

[21] W. Steinhaus and E. Gumlich, "An approximate law of susceptibility," *Verhandl. deut. physik. Ges.*, vol. 17, pp. 369–384; August 1915.

[22] J. G. Woodward and E. Della Torre, "Particle interaction in magnetic recording tapes," *J. Appl. Phys.*, vol. 31, pp. 56–62; January, 1960.

[23] F. Preisach, "Magnetic after-effect," *Z. Physik*, vol. 94, pp. 277–302; April, 1935.

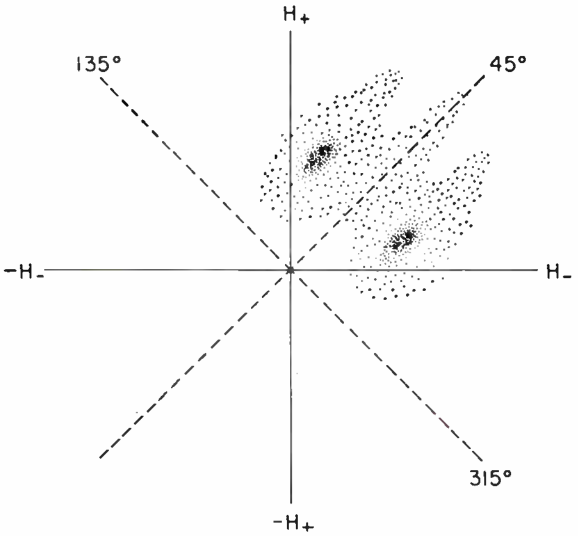


Fig. 10—Preisach coordinate system showing hysteresis loops of particles plotted in various areas.

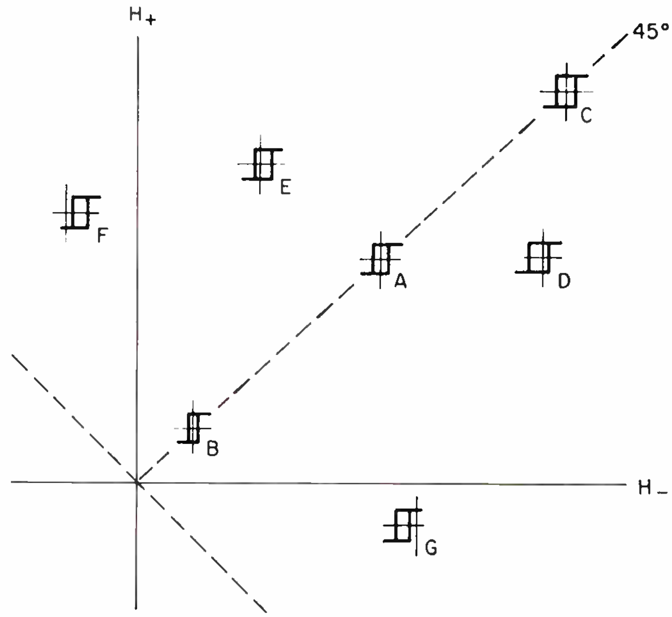


Fig. 11—Distribution function of particles in a magnetic tape (after Woodward and Della Torre).

In general, every particle will have neighbors whose magnetization modifies the over-all externally applied field. The influence of such neighbors may give rise to unbalanced conditions represented in regions D and E in Fig. 10, where the positive and the negative switching fields are unequal. In extreme cases, the unbalance may be high enough to displace the hysteresis loop from its axis, so that remanence in the region F is always positive, and in the region G is always negative. Since the magnetization of the neighbors is also changing during the magnetizing process, we cannot assign a permanent point on the diagram to any individual particle, but statistically a certain percentage will be in a selected area.

Once the distribution function given by such a diagram is known for a magnetic tape, its characteristic curve as of Fig. 7 can be plotted. The distribution function for samples of commercial gamma oxide tape was derived by Woodward and Della Torre from experimental data obtained by measuring the residual magnetization on a tape after it had been subjected to various combinations of ac and dc fields. They obtained results, shown roughly in Fig. 11, with marked peaks in regions well removed from the 45° line, signifying a high degree of unbalanced interaction among particles.

Some points lie in the second and fourth quadrants, just across the H_+ and H_- lines. Mathematically, there is also the area between 135° and 315° in the second, third, and fourth quadrants where the magnetization switches in a direction opposite to the applied field. Whether this is a forbidden region is not certain. Conceivably the applied field might upset a group of domains in such a way that their strong local influence would switch a domain oppositely to the applied field. At least one plot of Woodward and Della Torre showed a population in the third quadrant, which they obtained by extrapolation from the first quadrant.

Daniel and Levine[24,25] also carried out extensive work, both theoretical and experimental, with anhysteretic magnetization and with Preisach diagrams. They measured the "bulk" anhysteretic properties of recording tapes and came up with a set of constants that could be used to predict the recording sensitivity, the ac bias field required, and the maximum output. These are:

η_m—the maximum anhysteretic susceptibility, defined by maximum slope of the m vs H curve

H_0—the minimum ac bias field amplitude that produces an η of half η_m

S—the saturation value of m.

The Preisach distribution derived from their bulk measurements on recording tape, and shown roughly in Fig. 12, has a pear-shaped distribution with a single region of maximum density on the 45° line near a field of 240 oersteds.

Eldridge[26] has pointed out that the demagnetizing field relied on by Westmijze would only subtract from the applied field, but would not overcome it, and hence any applied field should still produce saturation were it not for the interaction fields. He reasoned that the anhysteretic magnetization curve of a material (which resembles that of Fig. 7) was really a plot of the cumulative distribution of interaction fields, since the abcissa represented an applied field which overcame interaction

[24] E. D. Daniel and I. Levine, "Experimental and theoretical investigations of the magnetic properties of iron oxide recording tape," *J. Acoust. Soc. Am.*, vol. 32, pp. 1–15; January, 1960.

[25] ——, "Determination of the recording performance of a tape from its magnetic properties," *J. Acoust. Soc. Am.*, vol. 32, pp. 258–267; February, 1960.

[26] D. F. Eldridge, "The mechanism of ac biased magnetic recording," IRE Trans. on Audio, vol. AU-9, pp. 155–158; September–October, 1961.

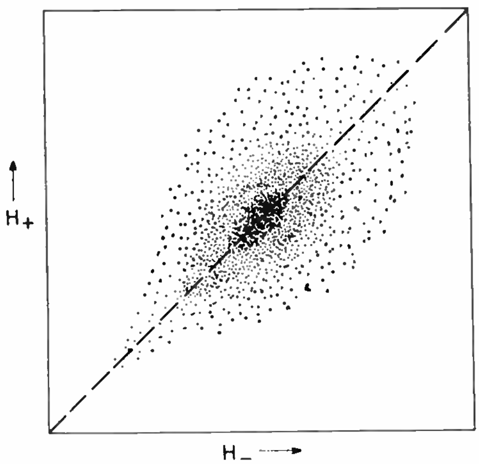


Fig. 12—Distribution function of particles in a magnetic tape (after Daniel and Levine).

fields of a certain number of domains and the ordinate represented the number of domains which had been overcome. He derived a distribution function from theoretical considerations, and showed that its form agreed closely with experimental measurements on tape. He found that the root mean square value of interaction field σ_{H_i} was proportional only to the magnetization intensity I_s of the bulk material:

$$\sigma_{H_i} = 2.9\, I_s \text{ total, or}$$

$$\sigma_{H_i}(x) = 1.7\, I_s \text{ in the } x \text{ direction.}$$

Summarizing, we may say that the premise that the domains are not alike magnetically does not, by itself, explain the B_r–H curve; but when the interaction between such domains is considered, a very good approximation to actual conditions is obtained. In the example of Fig. 9, at high ac bias fields all the domains are reversing their magnetic polarity together, but as the field decreases some of them become "set" and remain magnetized in one direction, while others of lower coercive force are still reversing. The first domains to become set will favor the polarity of the signal field (which may be considered constant since it changes slowly compared to the ac bias field). These first magnetized domains will now produce local fields which oppose the signal field; so that neighboring domains are subjected to the combined influence of the signal field and the local field of adjacent domains. Where the local fields overcome the signal field, a domain will become magnetized in a direction opposite the signal field, and a certain percentage of these opposite domains will exist. If the signal field is made stronger it will upset a greater and greater number of local situations in its own favor. If the signal field becomes weaker the self-balancing action of local fields tends to keep the net magnetization low, but statistically always in the direction of the signal field.

The result is that after being subjected to a field function as in Fig. 9, a record member is not saturated in one direction, but is magnetized in proportion to the signal field. To verify that it is a linear proportion in the region of the origin Eldridge[26] plotted his theoretically derived cumulative distribution function, and found it to be very close to experimental values for tape.

Another subject of prime interest is the magnetic field geometry in the tape during recording, and particularly the sequence of fields encountered by an element when it passes across the recording gap. As in the playback situtation, this must be analyzed for each type of head. The type shown in Fig. 1 is usually chosen because of its importance. Its external field has already been analyzed for the playback case. Unlike the playback case, the record on an element is not proportional to the integrated field but is influenced mainly by the highest field and by the last field encountered before it leaves the influence of the gap. This has given rise to the rule that head gaps need not be as small for recording as for playback, since in recording "only the trailing edge counts."

Armed with the head field configuration, the magnetic properties of the tape, and an assortment of simplifying assumptions, one may proceed to analyze the over-all recording process. A straightforward approach was used by Begun[27] who plotted the field variation versus time for different distances into the magnetic recording layer, and used the B_r–H curve to predict the resultant magnetization. This is an extension of the graphical analyses, which though accurate, does not give a physical picture that satisfies everyone. Stein[28] analyzed the recording of sine waves without bias, and came up with the interesting observation that the recording point begins at the trailing edge of the gap and progresses to the leading edge during the course of each half-cycle of recording field. Thus the magnetization moves contrary to the direction of tape travel, then jumps forward to begin another half-cycle. A similar effect was shown by Bauer and Mee[20] who devised a model to illustrate the recording process. At the recording gap they consider a "bubble" of magnetization. Inside this bubble magnetization can take place; outside the bubble the field is too low for recording. As the recording field increases and decreases, the radius of the bubble becomes larger or smaller. As seen from Fig. 13, when the bubble is expanding, it reaches out to the left, beyond the trailing polepiece, and upwards, thus magnetizing portions of the tape which would not have been magnetized with a constant radius of bubble. (It also expands to the right, beyond the leading edge, but here the tape would have become magnetized anyway as it passed through the bubble.) When the bubble is shrinking the leading edge becomes the important one, defining a boundary where tape is not magnetized which would have been magnetized with a constant radius of bubble. It might be

[27] S. J. Begun, "Magnetic field distribution of a ring recording head," *Audio Engrg.*, vol. 32, pp. 11–13, 39; December, 1948.
[28] I. Stein, "Analysis of the recording of sine waves," IRE Trans. on Audio, vol. AU-9, pp. 146–155; September–October, 1961.

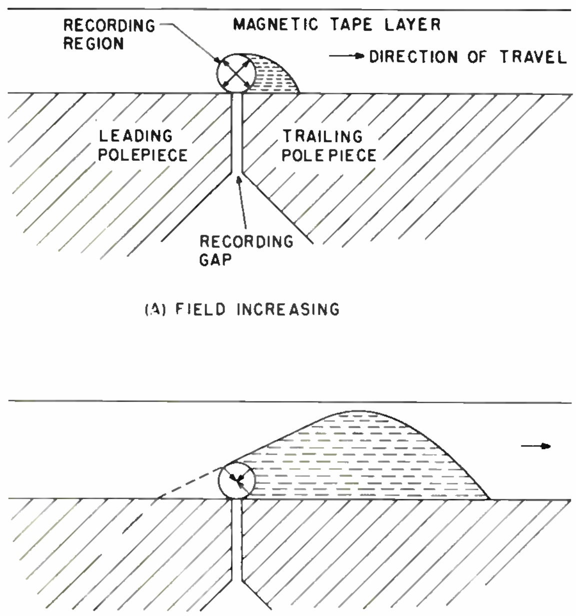


Fig. 13—Sine wave recording without bias (Bauer and Mee model).

expected that at short wavelengths, corresponding to rapid movement of the bubble wall compared to tape speed, some interesting effects would occur. Stein[28] analyzed the results of movement of the recording point under somewhat different assumptions.

When the Bauer and Mee model is used to show the effects of ac bias they obtain a pattern such as Fig. 14, where the signal expands the width of either the positive or the negative bias pulses. The model gives a qualitative explanation of output vs bias, but predicts a rapid drop in output if the bias amplitude or the bias frequency are increased, which does not conform to experiment. The authors admit that in these instances the model has gone beyond its limits of usefulness. Despite its limitations, the model shows that ac-biased recording on a moving tape can take place regardless of the mechanisms of finite anhysteretic magnetization, although we know such magnetization to be a fact from direct measurements on bulk material.

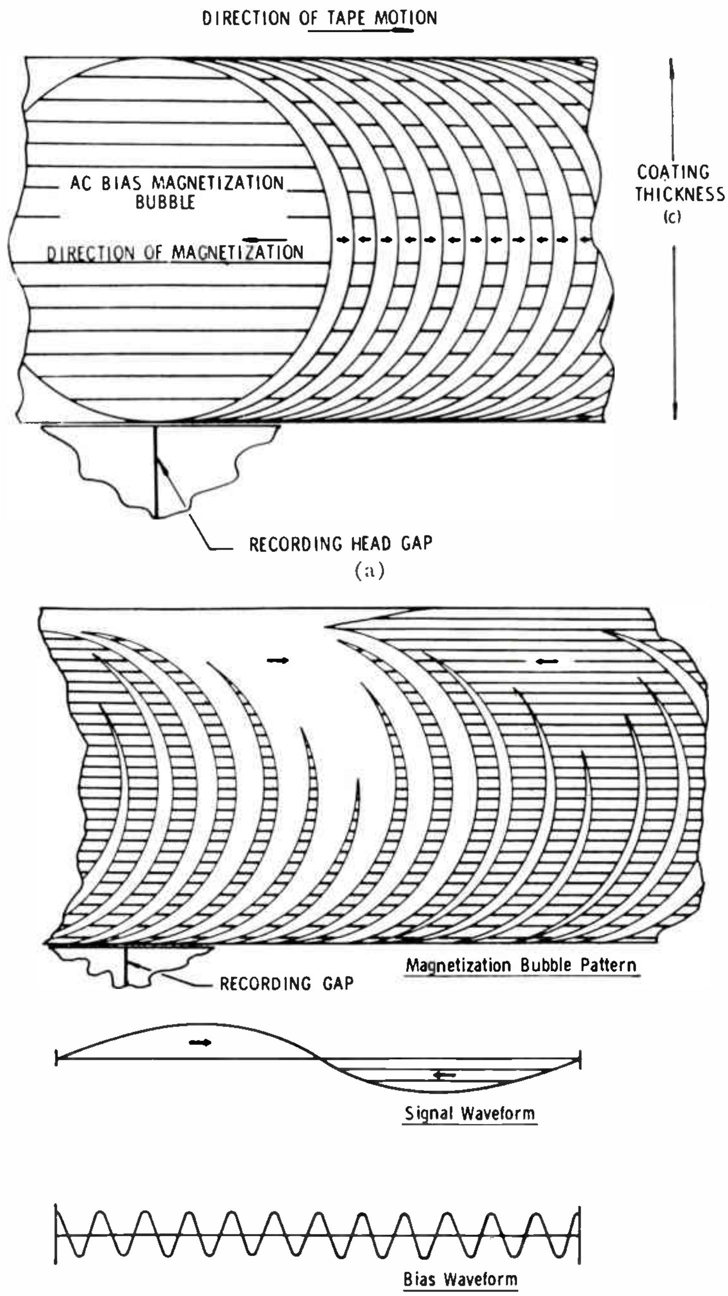


Fig. 14—AC bias recording. (a) Zero signal. (b) Long wavelength signal. (From Bauer and Mee.[20])

Noise in Magnetic Recording

When an unrecorded tape is played through a reproducing system, the noise level at the output will increase, compared to the noise level under identical conditions where the tape is not being scanned. Certain effects are noteworthy:

1) The least amount of noise is generated by a virgin tape or by a "bulk-erased" tape which has been demagnetized by subjected the entire reel of tape to a slowly decreasing field of power-line frequency.
2) Tapes erased with the erase head of a recorder may be several db noisier than bulk-erased tapes.
3) Tapes which have been scanned with a biased recording head (no signal present) are usually about 2 to 10 db noisier than either of the above.
4) The presence of a signal (including dc) will usually increase the noise in proportion to the signal level, typically about 10 db, but in severe cases as much as 20 or 30 db. This is called modulation-noise, and disappears when the signal is removed.
5) If the track width is changed the noise changes as the square-root of the width, while the signal changes directly as the width. Cutting a track in half reduces the signal-to-noise ratio 3 db.

6) Increasing the tape speed increases the over-all noise level and the high-frequency portion of the noise spectrum.

Even under best conditions, noise level from the tape itself is the limiting factor of a system, the amplifiers being considerably quieter.

A logical approach to analysis of noise is to find the response elicited by a single pulse, and then to sum the effects of a random assortment of such pulses forming a noise distribution. An early study of noise in magnetic recording was made on this basis by Wooldridge,[29] who derived the noise spectrum (noise power versus frequency) from theoretical considerations. He also derived the frequency response curve in the same way, and predicted that it would have the same form as the noise spectrum, provided that: a) the noise was caused by many random magnetic irregularities passing by the pickup gap; b) the pickup polepieces had negligible effect on tape magnetization; c) the random irregularities in the tape had small dimensions compared to the size of pickup gap. Measurements of frequency response and the noise spectra showed these to be alike, so that Wooldridge concluded that assumptions a), b), and c) were valid.

He then calculated the noise voltage that should result from a tape made up of saturated domains having completely random directions of magnetization. Under such conditions the noise would be about 1000 times as high as was actually observed. This indicated that the domains were not randomly magnetized, but in fact were arranged so that they cancelled out 99.9 per cent of each other's fields when the tape was demagnetized. Saturating the tape increased the noise 10 to 20 db without changing its spectrum. This was accounted for by postulating a greater degree of local disorder among the domains.

A theoretical analysis of the effects of variations in head-tape separation and in tape thickness during playback was made by Gilbert.[3] He also analyzed the case of sinusoidal tape irregularities, which could be useful for complex irregularities by application of Fourier series.

Wetzel[30] showed noise spectra taken in the presence of a recorded signal in which bands of noise were present near the signal frequency components. These might be explained by modulation noise effects. Periodic speed flunctuations in the tape drive under such conditions can also superimpose frequency modulation effects which modify the spectral distribution.

Glendon[31] obtained a great deal of data on noise spectra of modern tapes in a modern recorder, giving the results as a function of type of erase, backing and coating thickness, mylar vs acetate material, orientation, and tape speed. He reached no unexpected conclusions, but his results are useful in checking different noise theories.

In a recent paper Hammon[32] discussed noise problems occasioned by the trend to narrower and more closely spaced tracks. He was especially interested in modulation noise and showed spectra obtained under various conditions, including longitudinal vibration of the tape at 5 kc in the vicinity of the head.

Snow[33] pointed out that in practical situations the over-all noise spectrum of a recorder used for audio must be considered in relation to masking noises which are always present in the ultimate listening environment. Under proper conditions a recorder with a signal-to-noise rating of 50 db might effectively have a 70-db ratio.

Conclusions

Of all problems in the analysis of magnetic recording, the theory of the playback process has been treated most satisfactorily and most adequately. Others such as the recording process and noise are currently being developed, but a definitive treatment has not yet been evolved.

[29] D. E. Woolridge, "Signal and noise levels in magnetic tape recording," *Elec. Engrg.*, vol. 65, pp. 343–352; June, 1946.

[30] W. W. Wetzel, "Review of the present status of magnetic recording theory," *Audio Engrg.*, vol. 31, pp. 14–17, November, 1947; vol. 31, pp. 12–16, December, 1947; vol. 32, pp. 26–30, January, 1948.

[31] R. E. Glendon, "An Analysis of Tape Noise in a 100 kc Bandwidth," presented at Audio Engrg. Society Meeting, New York, N. Y.; October 8–12, 1957.

[32] F. Hammon, "Noise in magnetic sound recording," *Radio Mentor*, vol. 27, pp. 864–867; October, 1951. (In German.)

[33] W. B. Snow, "Characteristics of tape noise," *Audio*, vol. 45, pp. 26, 82–83; February, 1961.

Electroacoustic Measuring Equipment and Techniques*

LEO L. BERANEK†, FELLOW, IRE

Summary—**Since 1876 electroacoustic measuring equipment and techniques for its use have advanced from a primitive state to an advanced engineering science. This paper discusses sound-level measurements and primary techniques for the calibration of microphones. It covers systems and techniques for the measurement of the steady-state and transient response of amplifiers, microphones and loudspeakers. Means for shock-wave calibration of microphones and for measurement of nonlinear distortion and complex waveforms are also described.**

ELECTROACOUSTICS as a field of engineering was born in 1876 with the invention of the telephone. Microphones and earphones developed rapidly after 1880, following the invention during the previous decade of the carbon-button, moving-iron and moving-coil types of transducer. The moving-coil loudspeaker was invented in 1878. The most prevalent instrument for measuring sound was the human ear, which measured sound intensity, frequency response and nonlinear distortion better than the more objective means of that day. It was not until after the turn of the century that electroacoustic measurements as we know them today came of age.

I. Sound-Level Measurements

In 1908 the first report on the relative measurement of sound pressure by electronic equipment was published. G. W. Pierce, in an experiment described in the literature,[1] used a Bell magneto-telephone receiver (also a carbon button) as a microphone, together with a tuning condenser, a step-up transformer, a molybdenite (MoS_2) rectifier, and a galvanometer. He said that with this apparatus he had examined the sound "intensity" in different parts of various auditoriums and had observed the whistle of a distant train on the galvanometer scale.

A precision method for measuring the mean-square magnitude of the particle velocity in a plane-wave sound field was invented by Lord Rayleigh in 1882 and is still being used for free-field primary calibrations of laboratory pressure-gradient microphones.[2] This method makes use of the fact that a steady torque is exerted on a thin, small disk suspended in a sound wave by a fine quartz fiber or phosphor bronze wire. A Rayleigh disk is not practical for field measurements, because of its sensitivity to drafts and winds. Nor was it very useful in laboratories, until oscillators and loudspeakers were devised to create relatively intense sound waves.

High-quality equipment for the measurement of sound became practical with the advent of the vacuum-tube amplifier (about 1915). Within two years engineers at the Western Electric Laboratories (later the Bell Telephone Laboratories) had perfected the electrostatic (condenser) microphone[3] and an absolute technique for calibrating it, the thermophone method.[4] The condenser microphone has very low internal noise compared with a carbon-button microphone, and a more uniform response over a wider frequency range than a magneto-telephone receiver. This equipment was valuable in the laboratory, but at first it was delicate to transport and unreliable in the presence of high humidity.

By 1925 sturdy vacuum-tube amplifiers were available, and the condenser microphone had been conditioned against the effects of high humidity. Portable sound level meters were developed.[5,6] In 1930[7] the range of these instruments extended from 60 to 9000 cps. However, they were expensive and heavy.

Until about 1932 the only light-weight instruments for the field measurement of sound were the audiometer and the tuning fork, with which the subjective loudness of a sound could be compared with the loudness produced at the ear by a tone of known and adjustable sound pressure level.

The modern low-cost portable sound-level meter came into use prior to 1935. It was made possible by the advent of the piezoelectric (crystal) microphone of the diaphragm type, and by improved batteries and smaller-sized vacuum tubes. This type of sound-level meter has remained standard for 25 years. Improved types using transistorized circuits, miniature batteries and condenser microphones are recently available. Sound-pressure levels can now be measured with an accuracy of one to two decibels between 20 and 15,000 cps with a light-weight sound-level meter having an average battery life of 100 hours.

* Received by the IRE, August 3, 1961; revised manuscript received, October 13, 1961.

† Bolt Beranek and Newman, Inc., Cambridge, Mass.

[1] G. W. Pierce, "A simple method of measuring the intensity of sound," *Proc. Am. Acad. Arts and Sci.*, vol. 43, pp. 377–395; 1908.

[2] Lord Rayleigh, "On an instrument capable of measuring the intensity of aerial vibrations," *Phil. Mag.*, vol. 14, pp. 186–187; 1882. See also, L. L. Beranek, "Acoustic Measurements," John Wiley and Sons, Inc., New York, N. Y., pp. 148–158; 1949.

[3] E. C. Wente, "A condenser transformer as a uniformly sensitive instrument for the absolute measurement of sound intensity," *Phys. Rev.*, vol. 10, pp. 39–63; July, 1917.

[4] H. D. Arnold and I. B. Crandall, *Phys. Rev.*, vol. 10, pp. 22–38; July, 1917.

[5] F. A. Firestone, "Technique of sound measurements," *J. Sox. Automotive Engrs.*, vol. 19, pp. 461–466; November, 1926.

[6] R. F. Norris, *Projection Engrg.*, vol. 1, p. 43; 1929.

[7] E. E. Free, "Practical methods of noise measurement," *J. Acoust. Soc. Am.*, vol. 2, pp. 18–29; July, 1930.

Electronic means for analyzing sound waves were available in the early 1930's and have been improved continuously since that time. These include continuously variable, narrow-band, heterodyne frequency analyzers with crystal filters, step-wise variable octave, one-half-octave and one-third-octave band filter sets and constant-percentage, continuously variable, adjustable-bandwidth filters achieved by means of feedback circuits.

The most significant advance in the accurate measurement of sound pressure was the reciprocity method for calibration of linear, reversible microphones (transducers).[8-11] As seen in Fig. 1, the simplest embodiment of this principle[12] is the production of a spherical wave by the transducer (acting as a source) and the detection of this wave after reflection from a nonabsorbent plane wall by the same transducer (acting as a microphone). It can be shown mathematically[12] that the ratio of the open-circuit voltage e_{oc} produced by the transducer (when acting as a microphone) to the free-field sound pressure p_{ff} (that would exist in the reflected sound wave if the microphone were not there) is given by

$$\frac{e_{oc}}{p_{ff}} = \sqrt{\frac{e_{oc}}{i_T}\,\frac{4\lambda d}{\rho_0 c}}\, e^{-j(\pi/4-kd)}, \qquad (1)$$

where e_{oc} is the open-circuit voltage in the receive condition; p_{ff} is the free-field sound pressure acting to produce the open-circuit voltage; i_T is the current in the transducer acting to produce the wave in the transmit condition; λ is the wavelength of sound; $\rho_0 c$ is the characteristic impedance of air; d is the distance between the acoustic center of the transducer and the wall; and $k=2\pi/\lambda$.

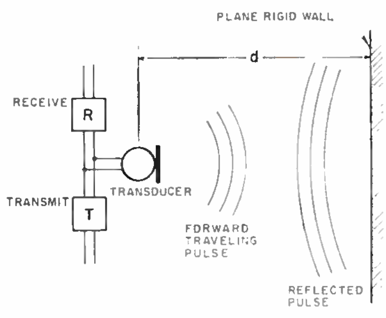


Fig. 1—Experimental arrangement for calibrating a transducer by the self-reciprocity method.

[8] W. Schottky, "Das gesetz des tiefempfangs in der Akustik und electroakustik," *Z. Physik*, vol. 35, pp. 689–736; 1926. (In German.)

[9] W. R. McLean, "Absolute measurement of sound without a primary standard, *J. Acoust. Soc. Am.*, vol. 12, pp. 140–146; July, 1940.

[10] R. K. Cook, "Absolute pressure calibrations of microphones," *J. Res. NBS*, vol. 25, pp. 489–505; November, 1940. See also, *J. Acoust. Soc. Am.*, vol. 12, pp. 415–420; November, 1941.

[11] A. L. DiMattia and F. M. Wiener, "On the absolute pressure calibration of condenser microphones by the reciprocity method," *J. Acoust. Soc. Am.*, vol. 18, pp. 341–344; October, 1946.

[12] L. L. Beranek, "Acoustics," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 377–387; 1954. See also Beranek,[14] pp. 113–148.

The accuracy of the calibration depends only on the length of the path d and the measurement of the ratio of a current to a voltage.

II. Steady-State Response of Electroacoustic Equipment

Amplifiers and Filters

The standard method of measuring the response of an amplifier as a function of frequency is to connect a signal generator to the input and a meter to the output of the equipment under test. The sinusoidal signal is set to each frequency of interest at a constant level and the indicated output noted. These data can also be taken automatically by sweeping the signal source over the desired frequency spectrum and feeding the output to a graphic-level recorder which is linked to the sweep. For systems that are substantially flat within the frequency range of interest, this method is satisfactory. When measuring systems with a large attenuation at some frequencies in the range of interest, an instrument is used which comprises a signal generator connected to the input, and a wave analyzer, and graphic-level recorder connected to the output.

In order for an amplifier to produce minimum distortion of complex signals not only must it have a frequency response that is essentially flat over a wide frequency range, but it must also have a linear phase shift as a function of frequency. Equipment for the direct measurement of phase in the audio range is available from many manufacturers.[13]

Microphones

The following measurements are frequently made on microphones:[14,15]

1) Free-field frequency response, *i.e.*, open-circuit voltage vs frequency for a constant free-field sound pressure.

2) Directivity vs frequency.

3) Electrical impedance vs frequency.

4) Self noise vs frequency.

5) Linearity vs level.

6) Stability with time and environmental conditions.

The *frequency response* of a microphone is almost always determined by measuring its output after placing it in a sound field which previously had been adjusted to a known value with the aid of a standard microphone. The sound field is generally either a free, progressive,

[13] Typical instruments for the direct measurement of phase are the AD-YU Phase Meter of the Electronics Laboratory, Inc., of Passaic, N. J., and Acton Laboratories, Acton, Mass.

[14] L. L. Beranek, "Acoustic Measurements," John Wiley and Sons, Inc., New York, N. Y., pp. 636–660; 1949.

[15] W. B. Snow, "Calibration and rating of microphones," IRE Trans. on Audio, vol. AU-8, pp. 5–13, January, 1960.

TABLE I
CHARACTERISTICS OF SOME GENERAL-PURPOSE MICROPHONES

Manufacturer	Altec			Shure			Bruel and Kjaer	Bruel and Kjaer	Western Electric
Model	633A	21BR150	21BR200	9898	9899	98108	4132	4134	640AA
Open-circuit sensitivity, db re 1 volt for sound-pressure excitation of 1 μbar	−90	−60	−90	−56	−59	−59	−46	−57	−50
Impedance or capacitance, μμf	30 ohms	6	4	2,000	700	460	60	20	50
Maximum pressure level for linear response, db re 0.0002 μbar	110 150	160	190	155	155	155	155	164	155
Maximum temperature, °F	150	500	500	110	110	210	400	400	400
Stability vs temperature	Fair	Poor	Good	Fair	Fair	Good	Good	Good	Good
Operation in high humidity	Good	Fair	Fair	Fair	Fair	Good	Fair	Fair	Poor
Frequency range, cps, for ±1 db re 400-cps deviation, random-incidence sound field	200–600	20–3,000	20–10,000	20–2,000	20–6,000	20–8,000	10–8,000	20–25,000	10–10,000
Frequency range, cps, for ±3 db re 400-cps deviation, random-incidence sound field	50–5,000	10–12,000	10–16,000	20–3,000	10–8,000	10–8,500	1–10,000	10–40,000	1–12,000

plane wave or is diffuse.[16] Free, progressive plane waves are produced by a loudspeaker in an anechoic (echo-free) chamber or outdoors. A diffuse sound field is obtained in properly designed reverberation chambers. Measurements of the open-circuit output voltage are made as a function of frequency with the sound-pressure level (measured before introduction of the microphone) held constant. The phase response is seldom measured.

Directivity patterns are measured in a free, progressive plane-wave sound field with the microphone oriented in space at all angles θ and ϕ with respect to the direction of travel of the wave. For those microphones that have symmetry of response about an axis, the directivity patterns are determined for angles θ alone.

The *electrical impedance* of a microphone is measured with an appropriate impedance bridge. Because the impedance includes the electrical resistance and reactance plus the reflected mechanical and radiation resistance and reactance, the measurements must be made with the microphone in a defined acoustical environment. The magnitude of the electrical impedance as a function of frequency can be determined by supplying a constant current to the microphone and measuring the voltage drop across the terminals.

Self noise is the limiting noise arising from the resistive portion of the electrical impedance. It is generally more easy to calculate it from a formula than to measure it:

$$e_{\text{total}} = \sqrt{4kT \int_{f_1}^{f_2} R\, df}, \qquad (2)$$

where

e_{total} = rms open-circuit self-noise voltage in the total band between f_1 and f_2 cps

k = Boltzmann gas constant = 1.37×10^{-23} joule/°K

T = absolute temperature in degrees Kelvin (°K)

R = resistive component of the electrical impedance of the microphone in ohms, in general, a function of frequency.

Microphone *linearity* is generally measured only for instances where sound-pressure levels well above 100 db *re* 0.0002 dyne/cm² are to be encountered. Apparatus for producing high sound levels, relatively free of distortion at the diaphragm of a microphone, have been described in the literature.[17]

The characteristics of some general-purpose microphones commonly used in audio measurements today are shown in Table I.[18]

[16] H. G. Diestel, "Reciprocity calibration of microphones in a diffuse sound field," *J. Acoust. Soc. Am.*, vol. 33, pp. 514–518; April, 1961.

[17] Beranek, *op. cit.*, pp. 433–437.

[18] L. L. Beranek, Ed., "Noise Reduction," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 103–107; 1960.

Loudspeakers

Loudspeakers are more difficult to measure than any other component of an audio system. The difficulty arises from the radiation and resonance characteristics of the diaphragm from which the sound emanates. Typically the diaphragm (if a direct-radiator loudspeaker) has a diameter of the order of one foot. If the diaphragm were a flat surface in an infinite baffle with all parts vibrating in phase, it would begin to show directivity at frequencies near 400 cps. At frequencies above 1500 cps such a diaphragm would become very directional and would exhibit secondary radiation lobes. Actual loudspeaker diaphragms do not vibrate in one phase, but exhibit mechanical resonance conditions at frequencies above several hundred cycles.[19] These resonances are evidenced by lines of zero motion and by areas of out-of-phase motion on the diaphragm. The observable acoustical results are irregularity with frequency in the power output and in the directivity patterns. The type of baffle (enclosure) used also affects the frequency response and the directivity patterns because of resonances inside and sound diffraction outside the enclosure.[20] Horn loudspeakers exhibit similar difficulties, but at higher frequencies. Finally, the room in which a loundspeaker is located affects its response and directivity, so that there is no one set of measurements that characterizes a loudspeaker in all environments.

The steady-state measurements commonly performed on loudspeakers are as follows:

1) Sound pressure level on axis vs frequency, measured in an anechoic chamber or outdoors in free space.
2) Directivity vs frequency.
3) Impedance vs frequency.
4) Power-available efficiency vs frequency.

To determine the on-axis frequency response of a loudspeaker, measurements are generally made in an anechoic chamber.[21–24] Measurements in a space free of room characteristics have the advantage of being well defined and repeatable. For large loudspeaker systems, the open space in the room should approximate 25 ft in all directions, and the sound absorption lining should not be less than 4 to 5 ft deep on all walls. The test microphone is generally located between 5 and 15 ft from the loudspeaker. Typical response curves for an 8-in loudspeaker are shown in Fig. 2.[24] The frequency scale on the graph paper shown is logarithmic below 2000 cps and linear above. (This is not standard.)

The directivity patterns are determined at each frequency either by rotation of the loudspeaker on a turntable or by movement of the microphone on a spherical surface surrounding the loudspeaker. Typical directivity patterns for an 8- to 12-in loudspeaker and a 30-in horn are shown in Fig. 3.

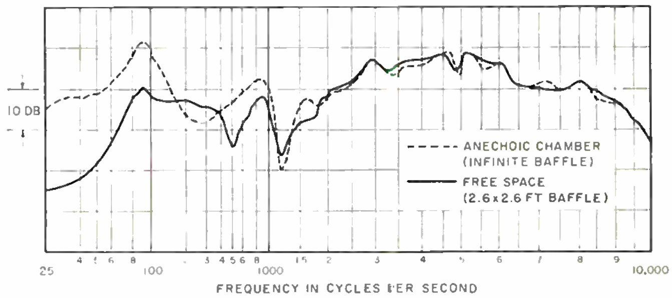


Fig. 2—Frequency-response curves of an electrodynamic loud-loudspeaker; microphone 1 m from the reference plane, constant voice-coil voltage.

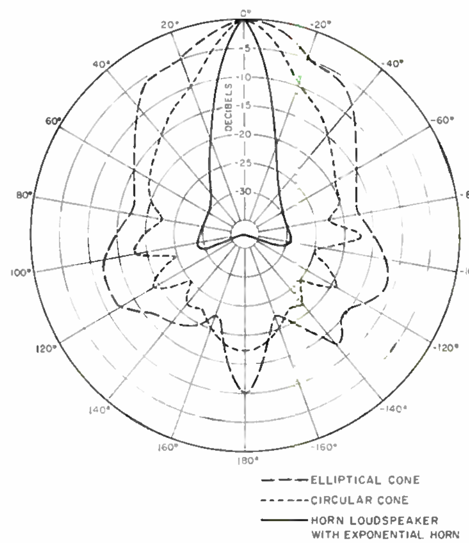


Fig. 3—Directivity patterns of three types of loudspeakers at 5000 cps.

The electrical impendence (including the reflected mechanical and radiation impedance) may be measured using a suitable impedance bridge. If only the magnitude is desired, the voltage drop across the terminals may be measured as a function of frequency for a constant signal current applied to the terminals.

Power-available efficiency in per cent is defined as 100 times the ratio of the total acoustic power radiated by the loudspeaker to the power developed by the driving amplifier in its nominal load resistance. This quantity is determined by, first, measuring the electrical power available from the amplifier for a particular gain setting. For example, if the 16-ohm terminals of an amplifier are to be connected to a nominal 16-ohm loudspeaker, the

[19] Beranek, *op. cit.*,[12] pp. 101–103 and 199–201.
[20] Beranek, *op. cit.*,[11] pp. 106–110; and Beranek, *op. cit.*,[12] pp. 216–222 and pp. 254–258.
[21] L. L. Beranek, and H. P. Sleeper, "The design and construction of anechoic sound chambers," *J. Acoust. Soc. Am.*, vol. 18, pp. 140–150; July, 1946.
[22] E. Meyer, *et al.*, "Ein Neuer Grosser Reflections freier Raum für Schallwellen und Kurze Electromagnetische Wellen," *Acustica*, vol. 3, pp. 409–420; 1953.
[23] G. Kamperman, paper in preparation, to be presented to Acoustical Society of America.
[24] P. Chavasse and R. Lehmann, "Procedures for loudspeaker measurements," IRE TRANS. ON AUDIO, vol. AU-6, pp. 56–57; May–June, 1958.

load resistance R connected to the amplifier should be 16 ohms. The power available is defined as

$$P_e = \frac{e^2}{R}, \quad (3)$$

where e is the voltage developed across the resistance R. The total acoustic power radiated is determined by removing the resistance R from the terminals and substituting the loudspeaker for it without changing the gain of the amplifier. The sound intensity I in watts/cm^2 is determined at a number of points in space[25] by measuring the mean-square sound pressure p^2 in (dynes/cm^2)^2 and using the formula

$$I = \frac{p^2}{\rho c} 10^{-7} \text{ watts/cm}^2, \quad (4)$$

where ρ is the density of air and c is the speed of sound in cgs units ($\rho c \doteq 41$ dyne-sec/cm^3 at 72° F and 750 mm Hg). Then the total acoustic power W_a radiated is

$$W_a = (1/\rho c)10^{-7} \int p^2 dS \text{ watts},$$

where dS is an element of area on a spherical surface enclosing the loudspeaker, and the integration is carried out between 0 and 2π in the general case or between 0 and π if hemispherical radiation only is being determined.

III. Transient Response of Electroacoustic Equipment

Amplifiers

The need for measurement of the transient response of amplifiers has decreased considerably because of the improvement in components and circuitry. Transient behavior is easily calculated from steady-state frequency response and phase measurements. With modern feedback circuits linearity over a wide dynamic range is possible.

Microphones

The transient response of most modern microphones is superior to that of loudspeakers. Microphones used to measure sonic boom and explosive sound waves are calibrated by a shock-tube technique.[26]

Loudspeakers

Transient measurements on loudspeakers are of great importance because loudspeakers generally exhibit resonance peaks in their response curves and widely varying phase characteristics. One manner in which transient distortion is measured was first advanced by MacLachlan[27] and refined by Shorter[28] and Corrington.[29] By this method tone bursts of 4 to 20 cycles are produced at a given frequency (see Fig. 4). The electrical input signal is turned on and off as the sine wave goes through zero. The duration of time between any two tone bursts equals the length of one burst. The loudspeaker must be operated from a source with the same internal impedance at the amplifier with which it will be used, because the internal resistance of the amplifier usually affects the duration of the transient, at least at low frequencies. The electrical input to the amplifier and the acoustical output of the loudspeaker are recorded by an oscillograph. The transient distortion is determined by recording the output wave in the interval of time between two input bursts. The energy of this distorted wave is determined and compared with the energy in the output wave in the interval of time when an input tone burst was on. The relative transient distortion is the difference in decibels between the output energies in the two equal-length time intervals. The test is repeated for each frequency of interest.

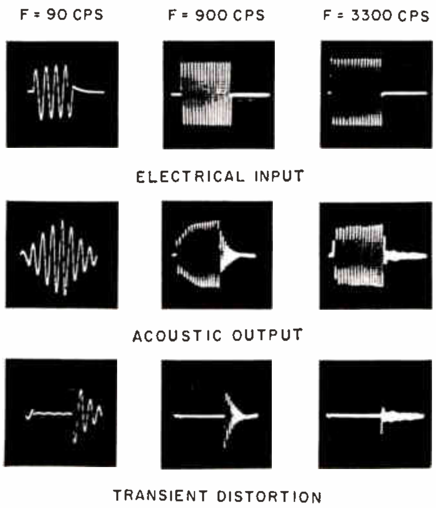


Fig. 4—Transient distortion for an 8-in direct-radiator loudspeaker.

Corrington reports very good correlation between such measurements of transient distortion on loudspeakers and subjective listening tests. He says that, when the transient distortion is more than 18 to 20 db down by the above test (using a 4-cycle pulse below 300 cps and a 16-cycle pulse above) the loudspeaker is comparable to the best obtainable. In addition, whenever

[25] Beranek, *op. cit.*, pp. 109–111.

[26] R. Bowersox, "Calibration of high-frequency response pressure transducer," *ISA J.*, vol. 5, pp. 98–103; November, 1958.

[27] N. W. MacLachlan, "Loudspeakers," Oxford University Press, Oxford, Eng.; 1934.

[28] D. E. L. Shorter, *BBC Quart.*, no. 3, October, 1946.

[29] M. S. Corrington, "Correlation of transient measurements on loudspeakers with listening tests," *J. Audio Engr. Soc.*, vol. 3, pp. 35–39; January, 1955. See also, M. S. Corrington, "Transient testing of loudspeakers," *Audio Engr.*, vol. 34, pp. 9–13, August, 1950; F. Bruner, "Untersuchungen an lautsprechern," *Oesterr. Z. Telegraphen-Telephon-Fornsehtech.*, vol. 8, pp. 1–7, 1954; and R. J. Larson and A. J. Adducci, "Transient distortion in loudspeakers," IRE Trans. on Audio, vol. AU-9, pp. 79–85, May–June, 1951.

the curve of transient distortion vs frequency exhibits peaks, listening tests reveal an unpleasant sound, even though the frequency-response curve might be quite smooth. This test is not sensitive enough for use on high-quality amplifiers.

By Shorter's technique the decay time of the transient following the cutoff of the input signal is measured. Obviously, there is a positive correlation between his measurements and those of Corrington, and his subjective listening tests also bear this out.

IV. Measurement of Nonlinear Distortion in Electroacoustic Equipment

Methods for measuring nonlinear distortion in electroacoustic equipment may be divided into two major groups, *harmonic* and *intermodulation*. For the harmonic method [see (a) of Fig. 5][30] a single sinusoidal signal A is supplied to the equipment under test. Signals C, D, etc., in the output are harmonically related to the input signal and are measured and used to specify the harmonic distortion. For example, the second-order (C/A), the third-order (D/A) or the total rms distortion $[(C^2+D^2)/(A^2+C^2+D^2)]^{1/2}$ may be determined. In the intermodulation methods[31–33] *two* sinusoidal signals are applied to the input [see (b) and (c) of Fig. 5]. Then all signals in the output with frequencies different from those at the input are measured. If the system is nonlinear, harmonics of the sine-wave frequencies will be produced and, in addition, components which are not harmonically related to either of the driving signals. Since speech and music already contain a number of harmonic components, additional harmonics introduced by the audio system are probably not as objectionable as new components not harmonically related to any of the original components. By the SMPTE method, signals C_1, D_1, C_2 and D_2 of Fig. 5(b) are measured. By the CCIF method, signals C and D_1 are measured. The harmonic and nonharmonic components appear not to bear a fixed relation to each other for some types of audio systems.[31] Hence, it is generally agreed that one or the other of the two intermodulation methods is preferable to the harmonic method.

Aagaard concludes from analysis and listening tests that the SMPTE intermodulation method is the logical choice. The CCIF difference-frequency method, using only the (f_2-f_1) term, may be grossly misleading for a system with a symmetrical transfer characteristic. On the other hand, if the second-order component $2f_2-f_1$,

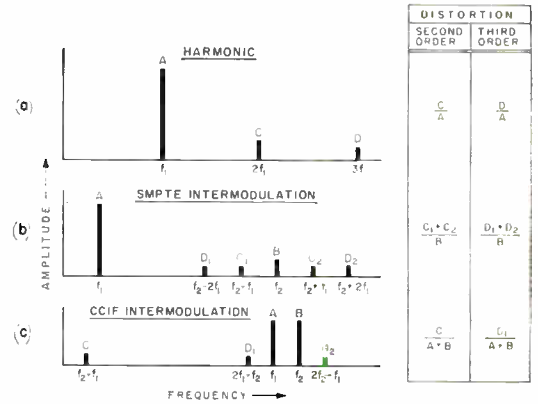


Fig. 5—Methods of measuring nonlinear distortion. (a) Harmonic. (b) Intermodulation method of SMPTE. (c) Intermodulation method of the Comité Consultatif International des Communications Téléphoniques à Grande Distance.

i.e., D_2 of Fig. 5(c), is measured, the CCIF method is the only really satisfactory way to determine nonlinear distortion near the upper-frequency limit of a system. In evaluating the results allowance must be made for the fact that the SMPTE method gives a numerical value 5 to 6 times that obtained with the CCIF procedure.

On highly nonlinear systems, or on systems that are amplitude limited (peak-clipped), such as hearing aids, a rough measure of distortion is the nonlinearity of the plot of the output signal level vs the input signal level. The point of nonlinearity in this curve is most quickly observed if a peak meter is employed.

In all cases it is wise to supplement the distortion measurement with oscillographic observation of the output waveform. This procedure not only gives an indication of the shape of the transfer characteristic, but it also reveals other defects such as instability.

V. Measurement of Random and Impulse Signals

Three types of rectifier characteristics commonly used when measuring random noise are 1) mean-square, 2) linear, and 3) pseudo-peak, all three normally calibrated to read alike when measuring a single sine wave.

Mean-Square (Square-Law) Indicating Instruments

A square-law device delivers an average current to a load that is 0.707 of the peak value of a sine wave. It indicates correctly the square of the standard deviation of the probability distribution curve of a random noise.

Average Indicating Instrument

A linear full-wave device delivers an average current to a load that is 0.636 of the peak value of a sine wave. If its output is adjusted upward by a factor of 1.11, it will, when measuring a sine wave, deliver the same output to a load as a square-law rectifier. When so adjusted and when measuring a random noise, it will deliver to a load a current that is about a factor of 1.13 (1.05 db) lower than that delivered by an rms meter.

[30] J. S. Aagaard, "An improved method for the measurement of nonlinear audio distortion," IRE Trans. on Audio, vol. AU-6, pp. 121–130; November–December, 1958.

[31] J. G. Frayne and R. R. Scoville, "Analysis and measurement of distortion in variable density recording," *J. SMPTE*, vol. 32, pp. 648–673; June, 1939.

[32] J. K. Hilliard, "Distortion tests by the intermodulation method," Proc. IRE, vol. 29, pp. 614–620; December, 1941.

[33] H. H. Scott, "Intermodulation measurements," *J. Audio Engr. Soc.*, vol. 1, pp. 56–61; January, 1953; and A. P. G. Peterson, "The measurement of non-linear distortion," General Radio Co., Concord, Mass., Tech. Publ. B-3, March, 1949.

Peak Indicating Instrument

A basic type of peak indicating instrument is shown in Fig. 6. Let us assume that E is a signal comprising a number of inharmonically related sine waves or a random noise. Then the ratio of the voltage E_b at the output of the peak indicating instrument to the true rms value of the signal E is shown by Fig. 7.[34] We see that for a high ratio of rectifier to series resistance (R_b/R_r) a pseudo-peak meter when measuring a random noise reads 5 to 8 db higher than an rms meter, provided both are calibrated to read alike on a sine wave ($n=1$). Only for $R_r/R_b \doteq 0.1$ would an rms and a pseudo-peak meter read approximately alike for a random-noise input.

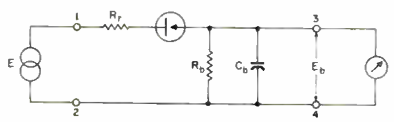


Fig. 6—Schematic diagram of a peak-indicating voltmeter.

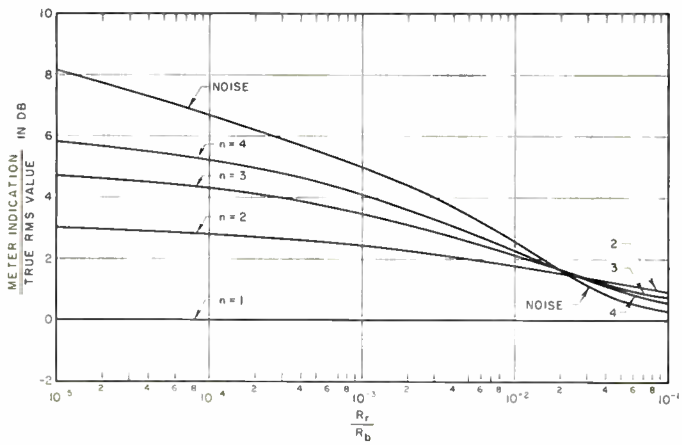


Fig. 7—Indication of a peak-reading meter relative to the true rms value of n equal, inharmonically related sine waves, and of white noise. The source impedance was assumed zero in constructing these curves, but can be taken into account by including it as part of R_r.

Impulse (Impact) Noise Meter

The time history of an impulse noise can be obtained from high-speed oscillographs or by photographing the trace on an appropriately calibrated oscilloscope. The energy-density spectrum can be obtained by appropriate filtering and integration. An impact (impulse) meter, such as the General Radio type 1556A, is relatively compact and can be used in the field or at the output of a magnetic tape recorder.[35] It measures 1) the peak level, and 2) a time-averaged level of the sound pressure. From 1) and 2) an estimate is obtained of the duration (decay time constant) of an impulse.

VI. Conclusion

A present-day view of the broad subject of electroacoustic equipment and techniques for measurement in the audible range of frequencies has been presented. Measurements of sound fields, loudspeakers, microphones, amplifiers and filters have been discussed. The quantities measured were the steady-state, transient and nonlinear correlates of the behavior of systems as judged by listeners. The methods are also applicable to situations where listeners are not involved. No attempt has been made to review subjective means for appraising audio systems.

In the next decade we may expect the digital computing machine to change radically our concepts of measurement. To fully utilize the capabilities of such machines, we need better electroacoustic transducers—particularly loudspeakers—with uniform frequency response and broad directivity patterns over a wide range of frequencies.

[34] G. R. Partridge, "Principles of Electronic Instruments," Prentice-Hall, Inc., Englewood Cliffs, N. J., p. 26; 1958. This graph was taken from data on pp. 457–464 and 475–484 of Beranek, *op. cit.*[14]

[35] *General Radio Experimenter*, vol. 30, no. 9, General Radio Co., Cambridge, Mass.; February, 1956.

Speech Communication Systems*

WINSTON E. KOCK†, FELLOW, IRE

Summary—**Telephone, radio broadcasting, public-address and bandwidth conserving systems are discussed with particular attention being given to the latter two. The relation of certain properties of hearing (*e.g.*, the Haas effect) to public address system design is reviewed along with several bandwidth conserving techniques including speech interpolation systems and vocoders.**

Introduction

SPEECH communication systems can be divided conveniently into four categories: telephone, radio broadcasting, public-address and bandwidth conserving systems. Telephone systems seek high intelligibility and reasonable naturalness while maintaining economy in bandwidth. Radio broadcasting, first limited in bandwidth by channel spacing, later achieved high fidelity capability with the advent of FM and TV broadcasting. Public-address speech systems possess no serious bandwidth limitations but encounter problems in room acoustics. Speech-bandwidth compression systems utilize certain properties of speech to achieve an economy in bandwidth. These four categories will be discussed in turn.

Telephone Communication Systems [1]

The first telephone systems were limited in bandwidth by the transducer capabilities. Highly resonant telephone receivers achieved high sensitivity but imposed a limitation on the intelligibility of the speech transmitted. The first vacuum tube amplifier was installed between New York and Baltimore in October, 1913, and this type of amplifier later led to the development of carrier-current telephony, whereby several telephone conversations could be transmitted simultaneously in different parts of the frequency spectrum. A typical carrier system (*e.g.*, the type K) provides 12 channels within a bandwidth extending from 12 kc to 60 kc. The actual speech band within each of the 12 transmission circuits extends from approximately 300 to 3400 cps. With the advent of the coaxial cable and its extremely wide-band capability, a new carrier system (the type L) was developed, capable of varrying 600 voice channels on one coaxial circuit. Further development of the coaxial cable and the introduction of the microwave relay systems such as the TD-2 (a typical tower is shown in Fig. 1), led to even greater sophistication in carrier concepts, typically the L-3 carrying 1800 voice channels in one circuit.

* Received by the IRE, July 24, 1961; revised manuscript received, October 25, 1961.

† Research Laboratories Division, Bendix Corporation, Southfield, Mich.

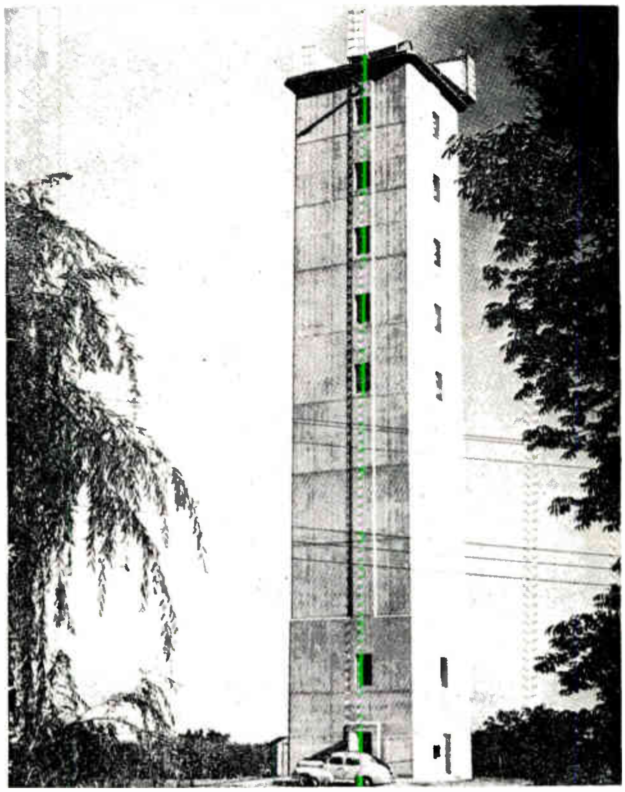


Fig. 1—Microwave radio relay circuits, exemplified by this Bell System TD-2 repeater tower, carry thousands of speech transmission channels on one radio circuit. The horn-lens antennas atop the tower, with a gain of 40 db and a beamwidth of 2 degrees [2], employ delay lenses [3] which were later found to be effective for directing the acoustic speech waves issuing from loudspeakers [4].

Early radio broadcasting with its 5-kc frequency band acquainted radio listeners with rather good quality electronic speech and an improvement in telephone speech quality accordingly seemed desirable. Better carbon microphones and improved telephone receivers such as the ring-armature receiver [5] were developed to improve the naturalness and intelligibility of telephone speech.

During this period of improvement in speech fidelity many experiments were conducted in various acoustical laboratories to determine the relations between audio bandwidth and intelligibility. Fig. 2 shows how the upper and lower limits of bandwidths affect intelligibility or articulation [6]. Monosyllabic word lists, referred to as phonetically balanced or PB lists, are used in listener tests to determine the quality of various transmission circuits. Such tests are particularly useful in evaluating speech bandwidth compression systems to be discussed later.

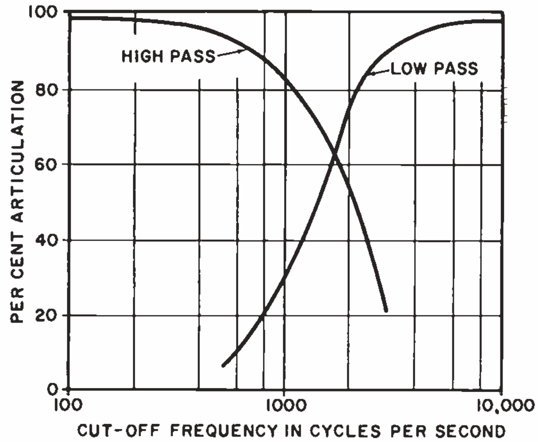


Fig. 2—Band-limited speech exhibits fairly high intelligibility (per cent syllable articulation) even with the low-frequency cutoff at 400 cps and the high-frequency cutoff at 4000 cps. [6]

Radio Broadcasting Systems

The spacing of radio channels set 5000 cycles as the limit of the audio spectrum to be encompassed in radio broadcasting. This constituted a moderately sizeable extension of the bandwidth used in telephone circuits, and, because of the state of the art in audio reproducing systems, it appeared adequate at the time. Recent developments in microphones, recording systems, and loudspeakers, however, permit electrical sound reproducing systems to encompass the full range of audibility of the ear, and this has made quite evident the limitations imposed by the 5000 cycle radio broadcasting channels. When frequency modulation broadcasting came into being, an audio channel of 50 to 15,000 cycles was selected; the audio channel for television broadcasting also covers this range. Obviously, such a frequency bandwidth can only be used effectively if the receiver loudspeakers and associated electronics are adequate. Since this is often not the case, network TV programs frequently have their 15,000 cycle audio channel transmitted over the narrower AM voice circuits. With the arrival of two-channel stereophonic recording and reproducing, it was only a matter of time before FM broadcasting standards were decided upon (April, 1961) permitting two-channel stereo to be broadcast (compatibly!) over FM stations.

Public-Address Systems

In the transmission of speech from a microphone on a lecture stand to loudspeakers in an auditorium, the problems of achieving good fidelity rest not in the bandwidth requirements, but rather in the acoustics associated with the hall or auditorium. Several recent developments which are pertinent to this problem are worthy of discussion here.

An effect called the precedence or Haas effect is now in wide use. It had long been recognized that both the relative nearness and the relative loudness of a sound source were pertinent in causing it to appear dominant over another source. Only recently, however, were the factors involved in this phenomenon thoroughly investigated. In 1951, H. Haas determined that if two sound sources are radiating the same signal, and the signal from one is delayed 5 to 35 msec relative to the other signal, the ear perceives the sound as coming almost entirely from the undelayed source [7] (Fig. 3). So great is this effect that the delayed source in many instances must be increased in intensity by many decibels in order that the listener be cognizant of its existence.

One of the first applications of this effect was its use in the public-address system installed in St. Paul's Cathedral in London [8], [9]. In previous systems the use of multiple loudspeakers resulted in large portions of the audience being aware that the amplified sound came from the loudspeakers instead of from the lecturer himself. In the St. Paul's installation the signals transmitted by the microphone in the pulpit are delayed (before being amplified and fed to the many distributed loudspeakers) so that the direct signal from the pulpit arrives at all audience positions in the cathedral before any loudspeaker signal can arrive there (Fig. 4). A listener thus imagines that the perceived sound comes from the pulpit rather than from a nearby loudspeaker. In addition to maintaining realism, the St. Paul's installation improved word intelligibility at the rear of the Cathedral from 52 per cent to 85 per cent [8].

Another effect which has been found useful in its application in public-address systems is the use of directional loudspeakers or directional-loudspeaker arrays. Many loudspeakers radiate in all directions and in an auditorium they create an excessive amount of reverberation through sound striking the ceiling and other undesirable areas. By modifying the radiation pattern of public-address loudspeakers, much can be done to increase the direct signal to the listener. In some instances the directivity of loudspeakers has been improved through the use of acoustic lenses [4], some of which are modifications of lenses originally developed for focusing radio waves [10] (Fig. 5). Another approach has been to employ vertical-line arrays of loudspeakers

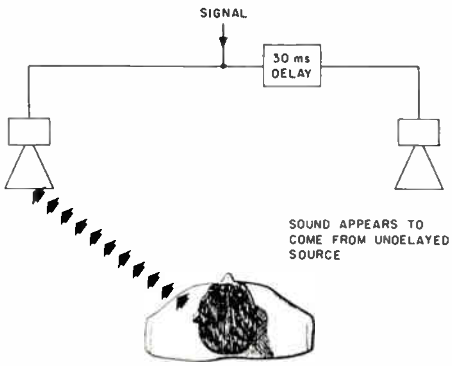


Fig. 3—Because of the precedence or Haas effect, the same sound issuing from two sources appears to come from the one with least delay.

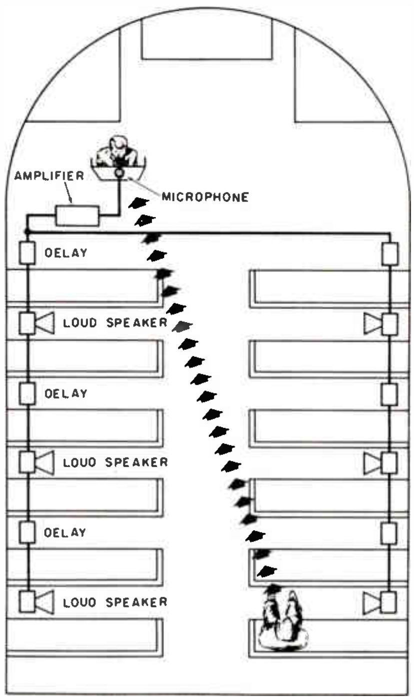


Fig. 4—An auditorium public-address system provides greater realism and intelligibility of delays as required by the Haas effect are introduced in the loudspeaker lines.

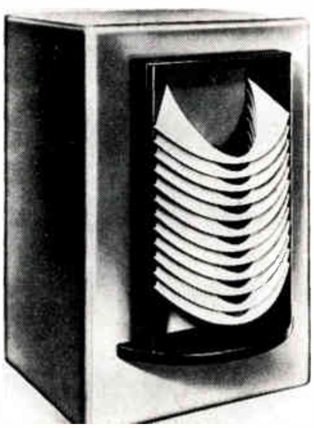


Fig. 5—Public-address system loudspeakers can provide maximum coverage with low reverberation if their directional patterns are properly modified as, for example, by the acoustic lens portrayed here.

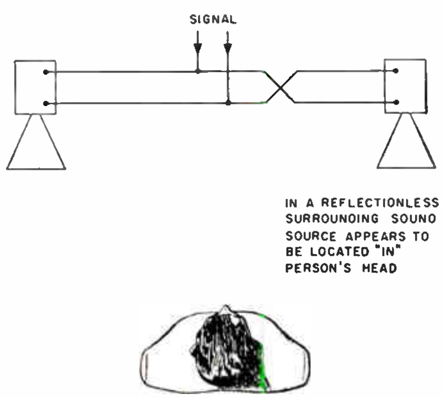


Fig. 6—In surroundings with low reflection, improper poling of two public-address system loudspeakers can create queer listening conditions for persons located along the center line.

permitting high directivity to be achieved in the vertical plane (thus concentrating the sound in the plane in which the audience is located) and low directivity in the horizontal plane (thereby covering the entire audience). Such loudspeaker arrays are employed in the St. Paul's Cathedral installation mentioned above [9].

Another effect which has been investigated recently is of importance in public-address systems employing two spaced loudspeakers. A listener equidistant from the loudspeakers observes that the quality of the speech he hears is quite dependent upon the relative poling of the two loudspeakers; as indicated in Fig. 6, the words actually seem to originate within the listener's head when the speakers are oppositely poled. The effect is most noticeable in a free space room or outdoors. When single frequency tones are employed in this situation, "phantom" sources can be localized [11], [12] but for broad-band signals the listener is quite confused [13]. R. L. Hanson and the author attributed the quite striking effect to the presence of a pronounced null for all frequencies along the exact center line plus the many high-frequency maxima and minima located near the center line. At certain frequencies this creates high intensity signals in one ear relative to the other, an effect otherwise achieved only by the sound being extremely close to one ear.[1]

Speech-Bandwidth Compression Systems

Savings in speech-channel capacity benefit expensive transmission systems (such as the transatlantic voice cable) and those communication systems where added voice-channel capacity cannot otherwise be obtained (such as in aircraft communications). Thus in the case of the transatlantic cable, a doubling of the number of speech channels would be equivalent to installing a second cable; the economic importance is obvious. In aviation communications, added channels could mean added traffic capability which might not otherwise be possible.

Bandwidth conserving systems can be divided into two categories: those dependent upon properties of the speech envelope and those dependent upon the frequency structure of speech. The first exploits the nature of speech conversations, in particular, the fact that in a normal two-way conversation each talker uses his circuit for only a fraction of the time. Accordingly, if the free channel time of an individual talker is interpolated among additional talkers an appreciable increase in the number of channels can be achieved [16], [17]. The second technique takes advantage of certain properties of the speech signal so as to permit adequate information on the quality and characteristics of the talker's voice to be transmitted over a narrower frequency band. When one considers the voice mechanism, one is not surprised to find that voiced speech is a relatively simple form of signal containing only one fundamental and one series of harmonically related overtones. This fact has led many to explore various sorts of speech-bandwidth reducing schemes.

Inasmuch as the time assignment interpolation system is now being employed on transatlantic voice cables [18], it will be reviewed first. Measurements on actual telephone working circuits indicate that the average activity is no more than 35 per cent to 40 per cent of the time that the circuit is busy at the switchboard. Since long distance circuits use separate pairs of wires and separate carrier channels for the two directions of transmission, it follows that each one-way channel is on the average free for 60 to 65 per cent of the time. When a large number of channels are operated as a group, variations in individual activity time average out and the law of probability serves to minimize mutual interference. In the operation of a system called TASI (time assignment speech interpolation) each of the lines from the toll switchboard is equipped with a special detector capable of recognizing, within 5 msec or less, the presence of speech sounds. When this detector operates, a switching network automatically connects that particular line to an idle channel and simultaneously sends a coded signal instructing the distant receiving terminal to connect that channel to the proper listener. This time sharing plan, connecting a talker to an available transmission channel for the duration of a talk spurt, has been able effectively to double the number of channels available in the 36 channel undersea telephone cable [19]. The first TASI system was put in service between White Plains, New York and London, England in May, 1960.

We turn now to systems making use of the frequency structure of speech. The first steps in this direction were taken almost 40 years ago when K. W. Wagner first created electronically the vowel sounds of speech [20]–[22]. It was about that time that experimenters were discovering that vowel sounds are dependent upon the shape of the resonant cavities in the vocal tract. Wagner employed a periodic series of pulses to excite resonant electrical circuits whose frequency and damping corresponded to the resonances present in the vowel sound. Fig. 7 portrays how vocal resonances (called formants) are excited by the glottal puffs originating at the vocal cords. A similar effect is present in many musical instruments [23]. During the 1930's the pioneering work of Wagner was extended by Homer Dudley who developed techniques for creating speech manually (the voder) [24] and for recreating speech from electrical signals obtained from an electronic speech analyzer (the vocoder) [25]. Dudley's vocoder system, shown in Fig. 8, might be likened to two high-speed electronic computers, one at the sending end and one at the receiving end. At the sending end, an analyzer ascertains what frequencies are present in the incoming speech signal, decides whether the sound is voiced or unvoiced, and measures the pitch of the signal if voiced. Since transitions from one sound to another occur quite slowly, the information

[1] Schodder [14] employed the Lauridsen effect [15] for another explanation of what he called the "Kocksche Wiedergabeweise (!)."

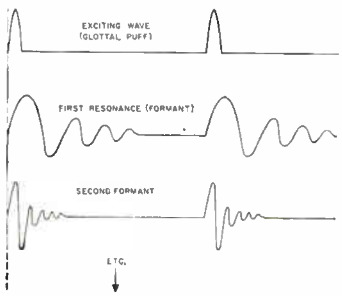


Fig. 7—Most speech-bandwidth compression systems make use of the fact that voiced sounds are combinations of the vocal cord wave (glottal puff) and several "formants," generated by the glottal puff acting on the resonant air chambers in the vocal tract.

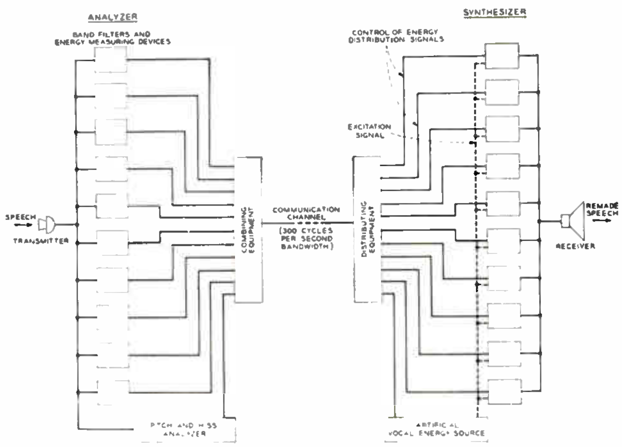


Fig. 8—The transmitting analyzer of the Dudley vocoders ends low-frequency signals over a narrow-band line to instruct the synthesizer on how to reconstruct the speech correctly.

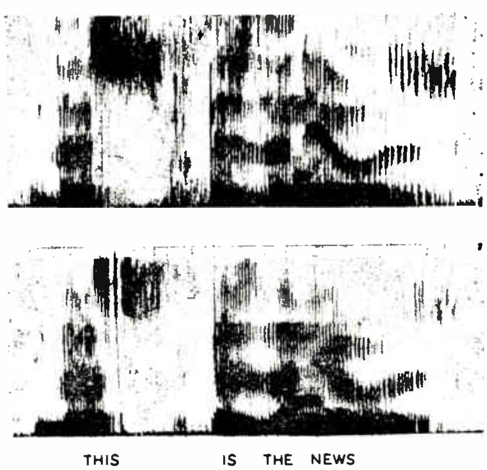


Fig. 9—Visible speech spectrograms of radio commentator Edward R. Murrow's words, "This is the news," show little difference between the original radio broadcast (top) and the synthetic vocoder version (bottom).

acquired by the analyzer can be transmitted over channels which are only 15 or 20 cycles wide. At the receiving end the synthesizer rebuilds the original speech from the information it receives, matching an artificial energy source (an electrical oscillator) in frequency content and pitch to the original speech. Factors of five or ten in bandwidth saving are possible. Early vocoders assigned equal bandwidth to the band filters in the analyzer and synthesizers. More recently R. L. Miller experimented with unequal bandwidth filters using narrower bands at the low-frequency end and wider bands at the high-frequency end of the speech spectrum. Using 16 analyzing filters he achieved vocoder speech of very high quality [26]. Fig. 9 illustrates the remarkable accuracy of speech reproduction obtained by this particular vocoder.

A somewhat allied type of vocoder developed from experiments at Bell Telephone Laboratories on the automatic recognition of speech sounds. In these experiments a recognition system designed to recognize any one of 10 speech sounds was used in an attempt to permit recognition of 10 complete words (the 10 digits). For this purpose, 10 electrical circuits were employed, each having frequency response characteristics which matched one of 10 selected phonemes (Fig. 10). Fig. 11 shows one speech sound (phoneme) being recognized in a word. This device (called Audrey) "recognized" individual sounds rather well and this suggested its use in a type of vocoder [27] as sketched in Fig. 12. Here only the information that a particular sound has been uttered is sent through the communication channel. At the receiving end the proper one of a series of spectrum modifiers is then activated to recreate the original sound. This type of vocoder may be promising where large amounts of bandwidth reduction (greater than 20 times) are desirable at some sacrifice of naturalness and perhaps intelligibility. The Audrey and "Audrey vocoder" experiments were conducted by K. H. Davis, R. Biddulph, S. Balashek [28] and the author [27], [29]. Earlier Dudley [30] had considered a vocoder related not only to the phonetic recognition vocoder but also, because of its use of variable resonant circuits, to the formant or resonance vocoder to be discussed shortly.

The acquiring of accurate fundamental frequency of the transmitted signal is one problem confronting all vocoder systems utilizing a synthetic energy source at the receiving end. In telephone circuits the fundamental component is usually missing and this aggravates the problem. One solution is to transmit the low-frequency end of the speech signal directly, coding only the higher frequency portion for bandwidth saving purposes [31]. Recently M. R. Schroeder and E. E. David sent a speech band of 10 kc in this way over a 3.5-kc channel [32].

The formant or resonance vocoder operates by transmitting information on the position in the frequency scale of the first three speech formants [33]–[36]. Fig. 13 portrays the motion of these formants in a section of

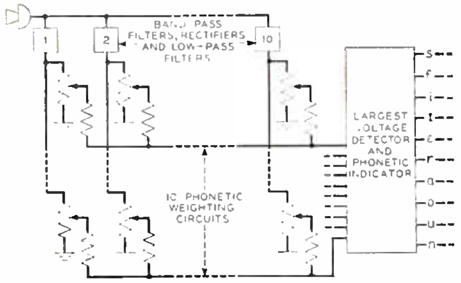


Fig. 10—The output of a filter whose frequency-amplitude characteristic matches that of a particular speech sound will be greatest when that sound enters the microphone. In this way, sounds can be "recognized" electronically.

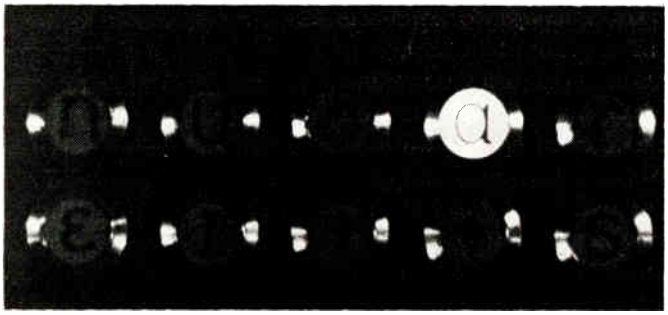


Fig. 11—The device of Fig. 10 has just "recognized" the sound *a* as in *father*, and caused a light to illuminate this symbol.

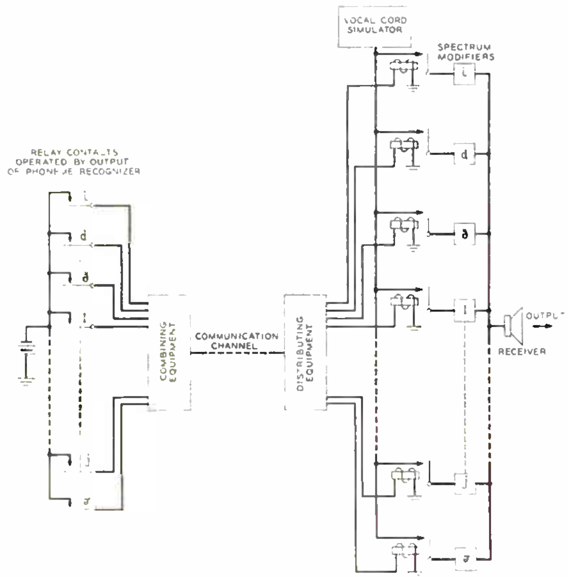


Fig. 12—An electronic analyzer (left) which recognizes successive speech sounds as categories in its memory can signal a speech synthesizer (right) for reconstructing the original speech. Since successive phonetic sounds in speech occur at rates less than 10 per second, the communication channel can be extremely narrow.

speech; the dotted lines suggest how the resonant filter circuits of a "formant tracker" at the transmitting end could control the position of similar filter circuits at the receiving end for reconstituting the speech sounds [37]. J. L. Flanagan applied to such a vocoder the technique mentioned earlier of transmitting a portion of the original speech in an uncoded way [38] (Fig. 14).

Recently a vocoder employing correlation techniques has been investigated. The "sound spectrogram" techniques of R. K. Potter (power spectra versus time) [39] were supplemented by W. R. Bennett's "correlatogram" techniques (auto-correlation functions vs time) [40]. This recognition of the possibility of representing the characteristics of speech sounds either by power spectrum or correlation function [41]–[43] led several to consider the correlation process in conjunction with vocoders. Inspired by a report [49] of R. M. Fano's "time domain vocoder" (which used a noise source as a signal for reconstituting the speech), B. P. Bogert and the author proposed in 1949 the correlation vocoder of Fig. 15 [44]. Here the speech signal is first distorted or "spectrum flattened" [45] so as to create a peak signal resembling the wave form at the top in Fig. 7. This is transmitted down a delay line and, at successive points in time, cross correlated with the original speech signal. Since the multiplication process of the cross-correlation procedure yields a signal only during the peak or pulse interval the cross correlation can be considered a sort of "gating" process. At the receiving end a similar periodic pulse (whose repetition frequency matches the fundamental frequency of the original speech signal) is multiplied, with proper delay, with the signals arriving over the various transmission channels. These signals are slowly varying signals, so narrow-band transmission can again be employed with bandwidth savings comparable to these achieved by the Dudley vocoder. The sampling intervals must be the Nyquist interval for the highest frequency of the signal involved, but because speech sounds die out during the course of a pitch period (Fig. 7), the sampling need only take place during the first portion of each period. Several forms of correlation vocoders including one employing auto-correlation have recently been designed and built by Schroeder [46], [47].

It is interesting to note that the same improvement obtained by Miller [26] through the use of nonuniform filter widths in a spectrum vocoder should also be available to correlation vocoders through the process of unequal spacing of the delay times. Formant bandwidths increase only slightly at the higher frequencies [48] and since decay time is inversely proportional to frequency *squared*, the decay time of the higher formants should be appreciably less than that of the lower-frequency ones. Maximum sampling rate thus need only be employed at the very start of the cycle and wider spacings can be employed between succeeding samples. This procedure (accomplished by increasing the values of the higher numbered τ's on the two delay lines of Fig. 15) was described in the original correlation vocoder patent [44].

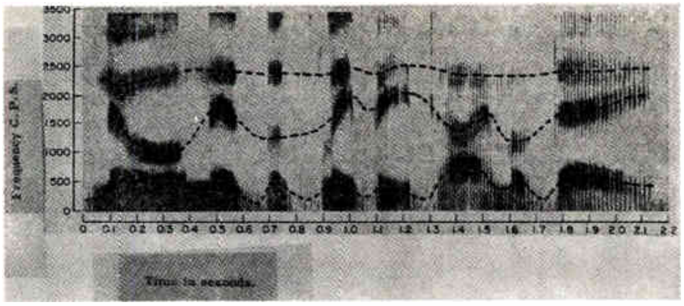


Fig. 13—A sound spectrogram of a section of speech portrays the varying motion of the resonances or formants. Formant "trackers," following the course of the important first three resonances can provide a speech synthesizer with the information needed to recreate the speech.

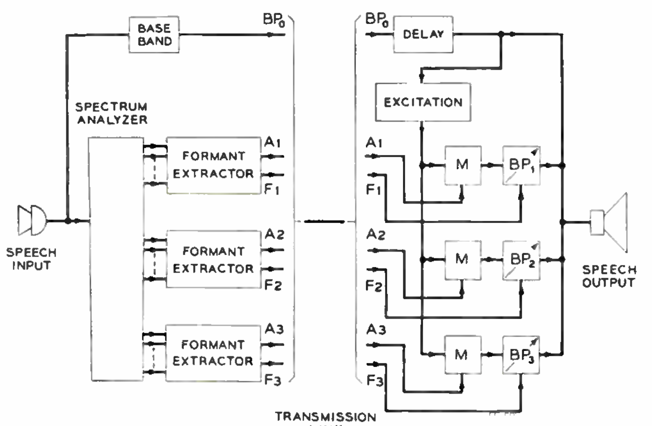


Fig. 14—The transmitting end (left) of a formant or resonance vocoder furnishes the receiving end (right) with information on the amplitude A and frequency F of the first three speech formants. At the receiver, modulators M and band-pass filters BP spectrum-modify the excitation signal to recreate the input speech. In this version some of the low-frequency portion of the original speech is transmitted in order to achieve better realism at some cost in bandwidth.

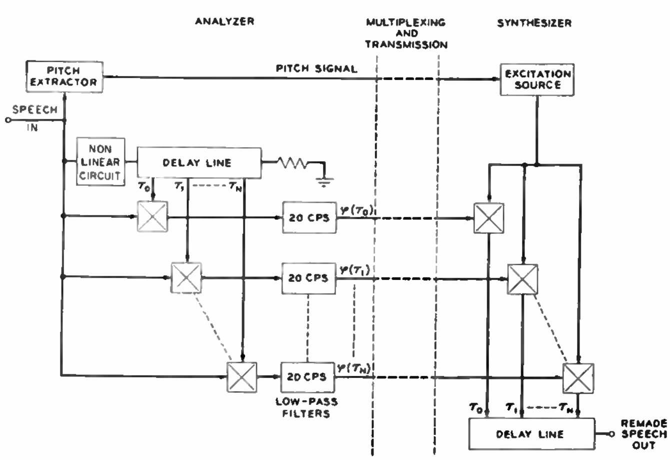


Fig. 15—Just as spectrum analysis has its counterpart in correlation or "time-domain" analysis, so spectrum vocoders have their counterparts in time-domain or correlation vocoders. This early version suggested by B. P. Bogert and the author, transmits signals in the time domain for reconstruction of the speech at the receiver.

Acknowledgment

The author wishes to express his deep appreciation to F. K. Harvey and M. R. Schroeder of the Bell Telephone Laboratories for assistance in securing reference material for this paper.

References

[1] R. B. Stecker, "A history of transmission systems," *Western Electric Engr.*, vol. 3, pp. 3–10; October, 1959.
[2] A. H. Lince, "Antennas for the TD-2," *Bell Labs. Record*, vol. 20, pp. 49–55; February, 1952.
[3] W. E. Kock, "Metalic delay lenses," *Bell Sys. Tech. J.*, vol. 27, pp. 58–82; January, 1948.
[4] —— and F. K. Harvey, "Refracting sound waves," *J. Acoust. Soc. Am.*, vol. 21, pp. 471–481; September, 1949.
[5] E. E. Mott and R. C. Miner, "The ring armature telephone receiver," *Bell Sys. Tech. J.*, vol. 30, pp. 110–140; January, 1951.
[6] J. C. Steinberg, "Effects of distortion upon the recognition of speech sounds," *J. Acoust. Soc. Am.*, vol. 1, p. 121; 1929.
[7] H. Haas, "Über den Einfluss Einfachechos auf die Hörsamkeit von Sprache," *Acoustica*, vol. 1, pp. 49–58; January, 1951.
[8] P. H. Parkin, "The application of the Haas effect to speech reinforcement systems," *Acustica*, vol. 4, pp. 87–89; January, 1954.
[9] —— and P. H. Taylor, "Speech reinforcement in St. Paul's Cathedral, "*Wireless World*, vol. 58, pp. 54–57, 109–111; February–March, 1952.
[10] W. E. Kock, "Path length microwave lenses," Proc. IRE, vol. 37, pp. 852–855; August, 1949.
[11] L. A. Jeffress, "Method for the study of the localization of airborne sound," (abstract), *J. Acoust. Soc. Am.*, vol. 27, p. 208; January, 1955.
[12] W. E. Feddersen, "The localization of sound from two sources in phase opposition," (abstract), *J. Acoust. Soc. Am.*, vol. 27, p. 208; January, 1955.
[13] R. L. Hanson, and W. E. Koci, "Interesting effect produced by two loudspeakers under free space conditions," *J. Acoust. Soc. Am.*, vol. 29, p. 145; January, 1957.
See also the description of this effect by W. A. Van Bergeijk, J. R. Pierce and E. E. David, Jr., "Waves and the Ear," Anchor Books, Doubleday and Co., Inc., Garden City, N. Y., pp. 100–101; 1960.
[14] G. R. Schodder, "Vortäuschungen eines akistischen Raumeindrucks," *Akust. Beih.*, vol. 6, pp. 482–488; February, 1956.
[15] H. Lauridsen, "Nogle forsog med forskellige former for sum akustick gengivelse," *Ingenioren*, vol. 47, p. 906; 1954.
[16] A. E. Melhose, "Multiples speech interpolation system," U. S. Patent No. 2,541,932; February 13, 1951.
[17] A. R. Kolding and G. N. Packard, "Speech interpolation system," U. S. Patent No. 2,957,946; October 25, 1960.
[18] K. Bullington and J. M. Fraser, "Engineering aspects of TASI," *Bell Sys. Tech. J.*, vol. 38, pp. 353–364; March, 1959.
[19] E. F. O'Neil, "TASI," *Bell Labs. Record*, vol. 39, pp. 83–87; March, 1959.
[20] K. W. Wagner, "Der Frequenz Bereich von Sprache und Musik," *Electrtech., Z.*, vol. 45, pp. 451–456; May, 1924.
[21] ——, "Uber die Elektrische Nachbildig von Vohalen," *Sitzber. preuss, Akad. Wiss., Physik,-math. Kl.*, Bd. 1, p. 1; January, 1933.
[22] ——, "Ein neues elecktrisches Sprechgerät zur Nachbildung der Menschlichen Vokale," *Abhandlung d. Preuss. Akad. d. Wissenschaft, Physik.-math.* Kl., Ab. 2, pp. 1–44; 1936.
[23] W. E. Kock, "Aehnlichkeit zwischen Vokalformanten und Formanten von Musikinstrumenten," *Elektrotech. Z.*, vol. 74, p. 166; March, 1953.
[24] H. Dudley, R. R. Riesz and S. S. A. Watkins, "A synthetic speaker," *J. Franklin Inst.*, vol. 227, pp. 739–764; June, 1939.
[25] H. Dudley, "The automatic synthesis of speech," *Proc. Natl. Acad. Sciences*, vol. 25, pp. 377–383; July, 1939.
[26] R. L. Miller, "Improvements in the vocoder," (abstract), *J. Acoust. Soc. Am.*, vol. 25, p. 832; August, 1953.
[27] W. E. Kock, "Speech bandwidth compression," *Bell Labs. Record*, vol. 34, pp. 81–85; March, 1956.
[28] K. H. Davis, R. Biddulph and S. Balashek, "Automatic recognition of spoken digits," *J. Acoust. Soc. Am.*, vol. 24, pp. 637–642; November, 1952.
[29] W. E. Kock, "The problem of selective voice control," *J. Acoust. Soc. Am.*, vol. 24, pp. 625–628; November, 1952.
[30] H. Dudley, "Production of artificial speech," U. S. Patent No. 2,243,527; May 27, 1941.
[31] F. Enkel and W. Meyer-Eppler, German Patent No. 957,235; April 21, 1953.
[32] M. R. Schroeder and E. E. David, Jr., "A vocoder for transmitting 10 kc/s speech over a 3.5 kc/s channel," *Acustica*, vol. 10, pp. 35–42; 960.

[33] J. L. Flanagan, "A speech analyzer for a formant coding compression system," Acoustics Lab., Mass. Inst. Tech., Cambridge, Sci. Rept. No. 4; May, 1955.
[34] G. C. M. Fant, "Speech communicators research," *Iva*, vol. 24, pp. 331–337; 1953.
[35] W. Lawrence, "Synthesis of Speech for Signals which have a Low Information Rate," in "Communication Theory," W. Jackson, Ed., Butterworths' Scientific Publications, London, Eng., ch. 6, pp. 460–471; 1953.
[36] J. C. Steinberg, "Speech analyzing and synthesizing communication system," U. S. Patent No. 2,635,146; December 15, 1959.
[37] J. L. Flanagan, "Bandwidth and channel capacity necessary to transmit the formant information of speech," *J. Acoust. Soc. Am.*, vol. 28, pp. 592–597; July, 1956.
[38] ——, "A resonance vocoder and baseband system, a hybrid system for speech transmission," 1959 IRE WESCON CONVENTION RECORD, pt. 7, pp. 5–11.
[39] P. K. Potter, G. A. Kopp, and H. C. Green, "Visible Speech," D. Van Nostrand Co., Inc., New York, N. Y.; 1947.
[40] W. R. Bennett, "The correlatograph: a machine for continuous display of short term correlation," *Bell Sys. Tech. J.*, vol. 32, pp. 1173–1185; September, 1953.
[41] R. M. Fano, "Short time autocorrelation functions and power spectra," *J. Acoust. Soc. Am.*, vol. 22, pp. 546–550; September, 1950.
[42] W. H. Huggins, "A note on autocorrelation analysis of speech sounds," *J. Acoust. Soc. Am.*, vol. 26, pp. 790–792; September, 1954.
[43] R. Biddulph, "Short term autocorrelation analysis and correlatograms of spoken digits," *J. Acoust. Soc. Am.*, vol. 26, pp. 539–541; July, 1954.
[44] B. P. Bogert and W. E. Kock, "Narrow-band transmission of speech," U. S. Patent No. 2,890,285; June 9, 1959.
[45] O. O. Gruenz, Jr., and L. O. Schott, "Extraction and portrayal of pitch of speech sounds," *J. Acoust. Soc. Am.*, vol. 21, pp. 487–495; September, 1949.
[46] M. R. Schroeder, "Correlation techniques for speech bandwidth compression," *J. Audio Engrg. Soc.*, presented at the 12th Annual Aud. Eng. Soc. Meeting, October 11–14, 1960 (to be published)
[47] ——, "Vocoders for military use," *Proc. AFCRC Speech Symp.*; September 28–30, 1959.
[48] B. P. Bogert, "On the bandwidth of vowel formants," *J. Acoust. Soc. Am.*, vol. 25, pp. 791–792; July, 1953.
[49] R. C. Mathes, verbal communication.

The History of Stereophonic Sound Reproduction*

JOHN K. HILLIARD†, FELLOW, IRE

Summary—**This paper reviews the history of stereophonic sound reproduction. It includes a comprehensive bibliography; the philosophical and psychological research on stereophony; methods used to make disc recordings and tape recordings on two channels; description of large auditorium motion picture systems for stereophonic effects; FM broadcasting of stereophonic programs on a multiplex basis and brief descriptions of pseudo-stereophonic reproduction systems for the home.**

INTRODUCTION

THE DEVELOPMENT of modern stereophonic sound reproduction in the home is generally represented as being recent. As shown by Offenhauser, (Reference Group A) however, it dates back to 1881, when two separate channels of telephone transmitters and earphones were used at the Paris Opera. G. W. Stewart spent years in research, and postulated in 1930, "The difference in phase is the important factor in determining the direction of sounds from 100 up to 1000 cps. Above this frequency, the intensity becomes a factor."

The reproduction of stereophonic sound thus had its foundations in the reproduction of "binaural" sound by telephone receivers worn on the head. The stereophonic reproduction of sound is preferably done on a high-fidelity basis. The groundwork for "hi-fi" reproduction may be found in the pioneering accomplishments of many scientists at the Bell Telephone Laboratories (Reference Group A), at the RCA Laboratories (Reference Group B) and elsewhere.

Stereophonic reproduction *per se* was pioneered almost simultaneously by Blumlein (Reference Group C) in Great Britain and at the Bell Telephone Laboratories (Reference Group D). Blumlein's contributions are presumed to be described in his patents. He showed a complete system applicable to sound-on-disc motion pictures, including microphone arrays utilizing bidirectional as well as omnidirectional microphones, transmission circuits, and disc recording systems utilizing simultaneous lateral and vertical recording. Economic difficulties are believed to have prevented completion and commercial exploitation of these systems.

The recognized early systems approach to large audience stereophonic reproduction was a public demonstration of Bell Telephone Laboratories equipment under the guidance of Dr. Harvey Fletcher on April 27, 1933. The Philadelphia Orchestra was in the Academy of Music in Philadelphia and it was reproduced in Constitution Hall, Washington, D. C. This was a three-channel system. The work was undertaken to evaluate a system which was capable of reproducing sound to a large audience and with a quality approaching "real performance." The basic equipment used is still at this

* Received by the IRE, July 10, 1961; revised manuscript received, October 18, 1961.
† Ling Altec Research, Div. of Ling-Temco Electronics, Inc., Anaheim, Calif.

date a tribute to the knowledge and technique of that group.

Further improvements in the system were subsequently worked out, as illustrated, for instance, by a patent to Snow, (Reference Group D) showing a grouping of high-frequency loudspeakers in the middle and a separation of low-frequency speakers toward the sides; and a patent to Bogert, (Reference Group D) showing various refinements in two-channel film systems, including a phantom third channel having a loudspeaker positioned in the middle and having the audio signals transmitted to it through a time delay network.

Philosophical and Psychological Research on Stereophony

Various pseudo-stereophonic concepts have arisen. For instance, the frequency range has been divided into two parts, divided around 600 cps, and reproduced by laterally-spaced loudspeakers. This provides a novel effect, but cannot be depended upon for separating voices from music in a consistent manner. Another system is to divide the frequency range into two groups of alternate, narrow bands. This "comb filter" system is described by Schroeder, and relates to an earlier time-delay system by Lauridsen.

The early phase vs intensity concept stated by Stewart above has been highly refined. For instance, Hanson (Reference Group E) has demonstrated various relationships between the apparent position of a sound source reproduced by two laterally-spaced loudspeakers, including the effects of delaying the signal to one loudspeaker and of altering the intensity. Moushegian and Jeffress (Reference Group E) have presented data relating interaural time and intensity differences with laterality. The results indicate that time can be traded for intensity in the lateralization of low-frequency tones.

In general, it has been found that the direction from which a given sound appears to come is a function first, of the time of incidence and second, of the intensity. From this concept, Becker (Reference Group E) has developed a two-channel system in which all of the spectral information occurs in each channel, but in which the sound from selected sources, as on the left, is timed to occur first in Channel A, and after a slight delay of approximately 10 msec, occurs in Channel B; and the sound from other selected sources, as on the right, is timed to occur first in Channel B, and after a similar delay occurs in Channel A. Sources which are to be reproduced as occurring in the middle are timed equally in each channel. A stereophonic system is thus produced which is comparable with either of two monophonic systems, the latter being individually available for use through different reproduction channels such as AM and FM broadcasting.

The necessity for utilizing two separate low-frequency loudspeakers, rather than reproducing "lows" by a common loudspeaker located in the middle or at one side has been investigated by Beaubien and Moore (Reference Group F). It was found if the low-frequency reproduction were localized in one loudspeaker below, for instance, 300 cps, a considerable loss in stereophonic effect could be noticed; and that if it were localized below 100 cps, a significant loss could be detected. Their bibliography is exhaustive, containing 88 papers.

Optimum conditions for reproduction of stereophonic sound in the home have been thoroughly reviewed by Olson (References Group E). It was shown that listeners prefer full frequency range to limited frequency range under conditions of low distortion, similar to the case of monophonic reproduction. Other investigations, by Olson and Belar, include the application of "information theory" to a study of the "amount of information" which can be transmitted by a stereophonic record as compared to that for a monophonic record, with the conclusion that the amount of information is increased approximately 90 per cent (Reference Group E).

Disc Recording of Two Channels

The cutting of a disc record with both lateral and vertical undulations was described in the Blumlein patent, and other forms of such cutters and reproducers are shown, for instance, in patents to Harrison; Henning; and Keller and Rafuse (Reference Group G).

The original recording is generally made by the use of a considerable number of microphones, each with a separate tape sound track, without regard to the channel in which the information may ultimately be recorded. These tracks are subsequently edited and combined into two channels, which are finally recorded by a cutter such as that mentioned above.

The modern cutting of disc records has been engineered by Davis and Frayne, and their associates (Reference Group G). Two independent, highly-degenerated cutters are disposed at 45° to the surface of the record and controlled by the respective channels. The concept is to cut one side of the groove with the information of one channel and the other side with that of the other channel. The discs are usually released for $33\frac{1}{3}$-rpm reproduction.

For several years all of the master or original recording, both for tape and records, has been made with tape running at 30 inches per second to reduce the distortion and maintain a high signal-to-noise ratio. Under controlled conditions the noise level from the disc record base is lower than the noise from tape but after a few plays the impulse noise is greater than with tape.

Tape Recording Developments

Pre-recorded stereophonic tapes have been available in both two- and four-track types. At the present time, however, it appears that the four-track method will be universally used in the home field because of the economy in tape cost. When the linear speed of tape is maintained sufficiently high, the high-frequency response of tape can be maintained somewhat above that of the disc records.

SINGLE-UNIT LOUDSPEAKERS FOR HOME STEREOPHONIC

It would appear obviously desirable to reproduce stereophonic sound by a single-unit loudspeaker assembly. A compromise between two separate loudspeakers and a single assembly was patented by de Boer, in which two loudspeakers are mounted in front panels which are adapted to be slid laterally outwards from the main body of a cabinet (Reference Group H). In a later patent, de Boer shows two cone loudspeakers mounted in the side walls of a cabinet and having especially designed reflective doors, hinged at the rear corners of the cabinet, and being adapted to reflect at least the short wavelength sound from the respective cone loudspeakers. This can be made to operate with reasonable satisfaction, especially since the longer wavelength sound is reflected from the walls of a room and gives a diffused effect. The effective separation between the sound sources is, however, substantially limited to the distance between the midpoints of the doors.

A structure which gives wider apparent separation of the channels is disclosed in a patent to M. Camras. In this patent two cone loudspeakers are mounted in a cabinet and directed diagonally laterally outwards and towards a rear wall. Sound is reflected from the rear wall towards the side walls and from there towards the listening area. The effective separation of the loudspeakers is enhanced by the reduction in direct radiation toward the listening area due to their being directed somewhat rearward.

Another arrangement giving wide apparent separation has been described by Levy, Sioles and their associates. Two loudspeakers are mounted in the side wall of a cabinet, and doors are mounted on the front corners. When opened, the doors impede the propagation of sound from the cabinet directly to the listening area and thus increase the apparent separation of the sound sources. Sound is diffused toward the listening area by reflection from the side walls.

Still another arrangement providing wide apparent separation has been described by Bauer and Sioles. In this structure a labyrinth is located behind each of two laterally mounted loudspeakers by which the sound from each one is directed laterally outwards from the respective loudspeakers, and reaches the listening area by reflection from the side walls, as in the above systems. All of these systems are somewhat dependent upon the presence of walls at a reasonable distance from the cabinet, and are not adapted, for instance, to outdoor use.

The work of the Bell Telephone Laboratories ultimately was adopted for large motion picture theater production, beginning with Walt Disney's "Fantasia" in 1940, Cinerama in 1952, Cinemascope in 1953, and Todd-A-O systems in 1955.

These systems vary from the basic three-channel Bell system up to five basic and two derived channels for Cinerama. The derived channels are used for special "surround" or effects in other parts of the auditorium apart from the stage or channels directly associated with the picture on the screen.

On April 20, 1961, the Federal Communications Commission issued rules and regulations Docket #13506 to permit FM broadcasting stations to transmit stereophonic programs on a multiplex basis. The system is a composite of stereophonic standards proposed by the Zenith Radio Corporation and GE Company, respectively. Of the eight stereophonic systems submitted, the adopted system had standards capable of rendering as high a quality of service as the art can provide consistent with economic and other technical factors without significant degradation of the present monophonic FM service.

Many papers on pseudo-stereophonic reproduction indicated the stereo contribution can be restricted to less than the full frequency range normally required for high-quality transmission. However, the evaluation of the stereophonic FM broadcasting system clearly indicated contribution to the stereophonic beyond the 100 cycle to 8-kc band. The significance of the extreme frequencies to the stereophonic reproduction may increase in importance with the passage of time for two reasons:

1) Improvements in technologies may come about in the recording or reproducing processes to permit more of these high and low frequencies to contribute to the stereophonic effect.

2) Listeners may become more sophisticated relative to stereophony and require all available information.

Several proposals have been made for stereophonic broadcasting on an AM carrier but as of this date the FCC has not standardized this transmission.

ACKNOWLEDGMENT

The author wishes to acknowledge the generous work contributed by R. W. Carlisle in the preparation of this material.

BIBLIOGRAPHY

Reference Group A

W. H. Offenhauser, Jr., "Binaural and stereophonic sound—in retrospect," *J. Audio Engrg. Soc.*, vol. 6, pp. 67–69; April, 1958.

J. P. Maxfield and H. C. Harrison, "Methods of high quality recording and reproduction of music and speech based on telephone research, Bell Sys. Tech. J., vol. V, pp. 493–501; July, 1926.

W. B. Snow, "Audible Frequency Ranges of Music, Speech and Noise," *Bell Telephone Sys. Monograph*, B-591; September, 1931.

Reference Group B

H. F. Olson, "Acoustical Engineering," D. Van Nostrand and Co., Inc., Princeton, N. J.; 1957. Note especially the data on pressure-gradient and unidirectional microphones (ch. 8).

Reference Group C

A. D. Blumlein, Brit. Patent No. 394,325, December 14, 1931; U. S. Patent No. 2,093,540.

Reference Group D

F. B. Jewett, *et al.*, "Perfect transmission and reproduction of symphonic music in auditory perspective," *Bell Telephone Quart.*, vol. 12, pp. 150–170; July, 1933.

"Symposium on Auditory Perspective":
H. Fletcher, "Basic Requirements."
J. C. Steinberg and W. B. Snow, "Physical Factors,"
E. C. Wente and A. L. Thuras, "Loudspeakers and Microphones."
E. O. Scriven, "Amplifiers."
H. A. Affel, *et al.*, "Transmission Lines."

E. H. Bedell and I. Kerney, "Auditory perspective-system adoption," *Elec. Engrg.*, vol. 53, pp. 216–219; January, 1934; also *Bell Sys. Tech. J.*, vol. 13, April, 1934. Much of this material is also presented in the *Bell Labs. Rec.*, vol. 12, March, 1934.

F. L. Hunt, "Acoustics of sound recording and reproduction for motion pictures," *Rev. Sci. Instr.*, vol. 7, pp. 323–328; September, 1936.

"Symposium on Stereophonic Sound-Film Theory":
H. Fletcher, "General Theory."
E. C. Wente, *et al.*, "Mechanical and Optical Equipment," and five other papers.

W. B. Snow, U. S. Patent No. 2,137,032; September 30, 1936.

B. P. Bogert, U. S. Patent No. 2,819,348; September 7, 1954.

Reference Group E

M. R. Schroeder, "An artificial stereophonie effect obtained from a single audio signal," *J. Audio Engrg. Soc.*, vol. 6, pp. 80–89; April, 1958.

R. L. Hanson, "Sound Localization," presented to the 57th Meeting of the Acoust. Soc. Am., at Ottawa, Can.; May 14, 1959.

G. Moushegian and L. A. Jeffress, "Localization of Low Frequency Tones," presented to the 57th Meeting of the Acoust. Soc. Am., at Ottawa, Can.; May 14, 1959.

R. L. Hanson, "Clues to Sound Localization," presented to the 58th Meeting of the Acoust. Soc. Am., at Cleveland, Ohio; October 24, 1959. Following the above demonstration, the speaker demonstrated that both time interval and pressure differential must be kept in a certain balance in order to preserve directionality. Various other papers presented at these two meetings were also of significance to this subject.

New Release from Bell Telephone Labs., October 6, 1959: "Full Stereo Sound Broadcasting Now Compatible with Single Channel Reception"; system developed by F. K. Becker utilizes "precedence effect." This system depends upon the utilization of a pair of ten millisecond time delay networks, in such a manner that sound picked up by one microphone are transmitted not only on the channel corresponding to that microphone, but also on a second channel after being delayed by transmission through one of the networks.

H. F. Olson, "Stereophonic sound reproduction in the home," *J. Audio Engrg., Soc.*, vol. 6, pp. 80–90; April, 1958.

H. F. Olson and H. Belar, "Acoustics of sound reproduction in the home," *J. Audio Engrg. Soc.*, vol. 8, pp. 7–11; January, 1960.

H. F. Olson, "Stereophonic Sound Considerations," presented at the 59th Meeting of the Acoust. Soc. Am., Providence, R. I.; June 10, 1960.

Reference Group F

W. H. Beaubien and H. B. Moore, "Perception of the stereophonic effect as a function of frequency," *J. Audio Engrg. Soc.*, vol. 8, pp. 76–86; April, 1960. A revision of this paper was published in IRE Trans. on Audio, vol. AU-8, pp. 144–153; September–October, 1960.

E. A. Aisberg, "Improved stereophony," *Wireless World*, vol. 46, pp. 327–330; September, 1950.

H. Alpert, "It's trinaural," *High Fidelity*, vol. 3, p. 42; September/October, 1953.

"The first—stereophonic—Noel," *High Fidelity*, vol. 6, pp. 60–61; December, 1956.

"Developments in stereophony," *J. Soc. Motion Picture and Television Engrs.*, vol. 61, Pt. 2, pp. 353–466; September, 1953.

"Stereophonic sound," *Electronics*, vol. 21, pp. 83–89; August, 1948.

"Fantasound," *Wireless World*, vol. 47, pp. 276–278; November, 1941.

"Fantasound," *Electronics*, vol. 14, pp. 18–21; March, 1941.

A. Balle, A. Lo Surdo, and G. Zanotelli, "Binaural localization sensitivity as a function of frequency," *Ricera Sci.*, vol. 17, p. 873; 1947.

H. Banister, "The effect of binaural phase differences on the localization of tones at various frequencies," *Brit. J. Psychol.*, vol. 15, pp. 280–307; 1925.

A. M. Belkin, "Characteristics of neural processes constituting the principle of the binaural effect," *Prolb. fiziol. Akust.*, vol. 2, pp. 65–71; 1950.

O. C. Bixler, "Practical binaural recording systems," IRE Trans. on Audio, vol. AU-1, pp. 14–22; January/February, 1953.

O. C. Bixler, "Commercial binaural recorder," *J. Soc. Motion and Television Engrs.*, vol. 59, pp. 109–117; August, 1952.

B. P. Bogert, "Stereophonic sound reproduction enhancement utilizing Haas effect," *J. Soc. Motion Picture and Television Engrs.*, vol. 64, pp. 308–309; June, 1955.

D. E. Broadbent, "Successive responses to simultaneous stimuli," *Quart. J. Exp. Psychol.*, vil. 8. p. 145; 1956.

D. E. Broadbent, "Growing points in multi-channel communication," *J. Acoust. Soc. Am.*, vol. 28, p. 533; 1956.

D. E. Broadbent, "A note on binaural fusion," *Quart. J. Exp. Psychol.*, vol. 7, pp. 46–47; 1955.

M. Camras, "Stereophonic magnetic recorder," Proc. IRE, vol. 37, pp. 442–447; April, 1949.

R. Caussee and P. Chavasse, "Differences between binaural and monaural threshold as a function of the frequency," *Bull. biol.*, vol. 136, pp. 301–302; 1942.

R. Caussee and P. Chavasse, "An experimental comparison of the binaural and monaural thresholds for audition," *Compt. Rend. Soc. Biol. Paris*, vol. 135, pp. 1271–1275; 1941.

R. Chocholle and J. Segal, "Binaural beats," *Compt. Rend. Soc. Biol., Paris*, vol. 141, p. 237; 1947.

H. A. M. Clark, G. F. Dutton, and P. B. Vanderlyn, "'Stereonsonic' recording and reproducing system," *Proc. IEE*, vol. 104. Pt. B, pp. 417–430, September, 1957; Discussion, pp. 430–432.

A. B. Cohen, "Reflections on having two ears," *High Fidelity*, vol. 4, p. 28; August, 1954.

N. H. Crowhurst, "Advantages, scope, and limitations of perspecta stereophonic system," *J. Soc. Motion Picture and Television Engrs.*, vol. 64, pp. 184–189; April, 1955.

N. H. Crowhurst, "Stereophonic Sound," John F. Rider Publisher, Inc., New York, N. Y.; 1957.

E. E. David, Jr., "Naturalness and distortion in speech—processing devices," *J. Acoust. Soc. Am.*, vol. 28, pp. 586–589; July, 1956.

K. de Boer, "Remarkable phenomenon with stereophonic sound reproduction," *Philips Tech. Rev.*, vol. 9, pp. 8–13; 1947.

K. de Boer, "Formation of stereophonic images," *Philips Tech. Rev.*, vol. 8, pp. 51–56; February, 1946.

K. de Boer, "Experiments with stereophonic records," *Philips Tech. Rev.*, vol. 5, pp. 182–186; June, 1940.

K. de Boer, "Stereophonic sound production," *Philips Tech. Rev.*, vol. 5, pp. 107–144; April, 1940.

W. B. Denny, "Dual-channel control amplifier for stereophonic music systems," *Audio*, vol. 38, pp. 17–19; May, 1954.

J. P. de Visser van Bloemen, "Stereophonic music in cinema," *Philips Tech. Rev.*, vol. 11, pp. 129–130; October, 1949.

H. Fletcher, "Stereophonic sound film system—general theory," *J. Acoust. Soc. Am.*, vol. 13, pp. 89–99; October, 1941.

C. Fowler, "Hi-Fi for two ears," *High Fidelity*, vol. 2, pp. 46; January/February, 1953.

J. G. Frayne, "Compatible photographic stereophonic sound system," *J. Soc. Motion Picture and Television Engrs.*, vol. 64, pp. 303–307; June, 1955.

H. Friess, "Stereophonic sound film; pickup and reproduction," *Funk. u Ton*, vol. 8, pp. 622–630; December, 1954.

W. E. Garity and J. N. A. Hawkins, "Fantasound," *J. Soc. Motion Picture Engrs.*, vol. 37, pp. 126–146; August, 1941.

W. R. Garner, "Accuracy of binaural loudness matching with repeated short tones," *J. Exp. Psychol.*, vol. 37, pp. 337–350; August, 1947.

C. Gerhardt, "The recording and reproduction of space," *High Fidelity*, vol. 7, p. 41; March, 1957.

L. S. Goodfriend, "High fidelity," *Audio Engrg.*, vol. 34, pp. 64–66; November, 1950; December, 1950.

R. Held, "Shifts in binaural localization after prolonged exposures to atypical combinations of stimuli," *Am. J. Psychol.*, vol. 68, pp. 526–548; 1955.

I. J. Hirsch, "Binaural summation—a century of investigation," *Psychol. Bull.*, vol. 45, pp. 193–206; 1948.

I. J. Hirsh, "Interaural Summation and Inhibition. A Comparison of Binaural and Monaural Thresholds," Ph.D. dissertation, Harvard University, Cambridge, Mass.; 1948.

I. J. Hirsch, "Binaural summation and interaural inhibition as a function of the level of making noise," *Am. J. Psychol.*, vol. 61, pp. 205–213; 1948.

J. W. Hughes, "Binaural localization with two notes differing in phase by 180°," *Brit. J. Psychol.*, vol. 30, pp. 52–56; 1939.

R. G. Klumpp and H. R. Eady, "A Technique for Improving the Bearing Accuracy of Binaural Listening Systems," U. S. Navy Electronics Lab., Res. Rept. No. 419, 5 pp.; 1953.

W. E. Kock, "Binaural localization and masking," *J. Acoust. Soc. Am.*, vol. 22, pp. 801–804; Month, 1950.

R. J. Kowalski, "RCA's 'Fantasound' system as used for Disney's 'Fantasia,'" *International Projectionist*, vol. 15, pp. 20–21; November, 1940.

H. Lambert, "Remarks on binaural hearing and the localization of sound," *Ann. Prothese auric.*, No. 5, pp. 6–12; 1935. (In French.)

C. E. Lane, "Binaural beats," *Phys. Rev.*, pp. 401–412; May, 1925.
D. M. Leakey and E. C. Cherry, "Influence of noise upon equivalence of intensity differences and small time delays in two-loudspeaker systems," *J. Acoust. Soc. Am.*, vol. 17, pp. 356–357; April, 1946.
H. B. Lee, "Stereophonic sound recording system," *J. Acoust. Soc. Am.*, vol. 17, pp. 356–357; 1946.
T. Lode, "Stereophonic reproduction," *Audio Engrg.*, vol. 34, pp. 46–47; January, 1950.
J. Moir, "Stereophonic reproduction," *Audio Engrg.*, vol. 36, pp. 26–28; October, 1952.
J. Moir and J. A. Leslie, "Stereophonic reproduction of speech and music," *Brit. IRE*, vol. 12, pp. 360–366; June, 1952.
J. Moir, "Stereophonic sound," *Wireless World*, vol. 57, pp. 84–87; March, 1951.
J. Nigro, Jr., "Stereodynamic multichannel amplifier for single or binaural input," *J. Audio Engrg. Soc.*, vol. 1, pp. 287–291; October, 1953.
H. F. Olson and F. Massa, "Realistic reproduction of sound with particular reference to sound motion pictures," *J. Soc. Motion Picture Engrs.*, vol. 23, pp. 63–80; 1934.
N. C. Pickering and E. Baender, "Two ears in three dimensions," *J. Audio Engrg. Soc.*, vol. 1, pp. 255–261; July, 1953.
G. Plass, "Stereophony from the outside in," *High Fidelity*, vol. 4, p. 78; April, 1954.
E. H. Plumb, "Future of 'Fantasound,'" *J. Soc. Motion Picture Engrs.*, vol. 39, pp. 16–21; July, 1942.
K. P. Pokryvalova, "Data for the physiology of mono- and binaural hearing," *Prolb. fiziol. Akust.*, vol. 2, pp. 57–64; 1950. (In Russian.)
I. Pollack, "Monaural and binaural threshold sensitivity for tones and for white noise," *J. Acoust. Soc. Am.*, vol. 20, pp. 52–57; January, 1948.
G. Revesz, "Colour mixture and sound mixture," *Acta Psychol.*, vol. 6, pp. 3–26; 1949.
M. R. Rosenzweig and J. W. Evertet, "Binaural interaction at the inferior colliculi," *J. Comp. Physiol. Psychol.*, vol. 48, p. 426; October, 1955.
F. C. Saic, "Fidelity of electroacoustic reproduction," *Elektrotechnik u. Maschinenbau*, vol. 64, pp. 85–91; May–June, 1947. (In German.)
W. A. Shaw, E. B. Newman, and I. J. Hirsch, "The difference between monaural and binaural thresholds," *J. Exptl. Psychol.*, vol. 37, pp. 220–242; 1947.
C. A. Silver, "A theory of binaural inhibition," *Dissertation Abstracts*, No. 15, p. 1265; 1955.
R. Singer, "Stereophonic broadcasting," *Electricien (Paris)*, vol. 79, pp. 23–28; February, 1951. (In French.)
M. B. Sleeper, "Sound movement and dimensions," *FM-TV*, vol. 11, pp. 48–49; November, 1951.
W. B. Snow, "Basic principles of stereophonic sound," IRE TRANS. ON AUDIO, vol. AU-3, pp. 42–53; March–April, 1955.
W. B. Snow, "Effect of arrival time on stereophonic localization," *J. Acoust. Soc. Am.*, vol. 26, pp. 1071–1074; November, 1954.
W. B. Snow, "Basic principles of stereophonic sound," *J. Soc. Motion Picture and Television Engrs.*, vol. 61, pp. 576–587; November, 1953.
W. B. Snow and A. R. Soffel, "Electrical equipment for stereophonic sound-film system," *J. Soc. Motion Picture Engrs.*, vol. 37, pp. 380–395; October, 1941.
R. H. Snyder, "History and development of stereophonic sound recording," *J. Audio Engrg. Soc.*, vol. 1, pp. 176–179; April, 1953.
M. J. Stateman, "Methods for achieving audio perspective," *Sylvania Technologist*, vol. 8, p. 48; April, 1955.
J. C. Steinberg, "Stereophonic sound film system—pre- and post-equalization of compandor systems," *J. Acoust. Soc. Am.*, vol. 13, pp. 107–114; October, 1941.
G. W. Stewart, "The function of intensity and phase in the binaural location of pure tones," *Phys. Rev.*, vol. 15, pp. 432–445; 1920.
E. W. Templin, "Recent developments in multichannel stereophonic recording systems," *J. Soc. Motion Picture and Television Engrs.*, vol. 66, pp. 53–58; February, 1957.
R. J. Tinkham, "Stereophonic recording equipment," *Elec. Engr.*, vol. 72, p. 1053; December, 1953.
R. J. Tinkham, "Binaural or stereophonic?" *Audio Engrg.*, vol. 37, pp. 22–24; January, 1953.
O. C. Trimble, "Intensity-difference and phase-difference as conditions of stimulation in binaural sound-localization," *Am. J. Psychol.*, vol. 47, pp. 264–274; 1935.
R. Vermeulen, "Stereo reverberation," IRE TRANS. on AUDIO, vol. AU-4, pp. 98–105; July-August, 1956.
R. Vermeulen, "Comparison between reproduced and 'live' music," *Philips Tech. Rev.*, vol. 17, pp. 171–177; December, 1955.
R. Vermeulen, "Stereophonic reproduction," *Audio*, vol. 38, p. 21; April, 1954.
H. Waechter, "The Dependence of Binaural Hearing on Intensity and Time Difference," KZ-Druckerei, Kiel, Germany, 47 pp.; 1937.
N. D. Webster and F. C. McPeak, "Experiments in listening," *Electronics*, vol. 20, pp. 90–95; April, 1947.
E. C. Wente, R. Biddulph, L. A. Elmer, and A. B. Anderson, "Mechanical and optical equipment for stereophonic sound films system," *J. Acoust. Soc. Am.*, vol. 13, pp. 100–106; October, 1941.
D. J. Wilson, "A simple binaural phase apparatus," *J. General Psychol.*, vol. 22, pp. 215; 1940.

Reference Group G

H. C. Harrison, U. S. Patent No. 2,027,168; September 29, 1934.
H. A. Henning, U. S. Patent No. 2,025,388; September 22, 1934.
A. C. Keller and I. S. Rafuse, U. S. Patent No. 2,114,471; June 20, 1936.
C. C. Davis and J. G. Frayne, "The Westrex stereo disc system," PROC. IRE, vol. 46, pp. 1686–1693; October, 1958.
J. G. Frayne and R. R. Davis, "Recent developments in disc recording," *J. Audio Engrg. Soc.*, vol. 7, pp. 147–152; October, 1959.

Reference Group H

K. de Boer, U. S. Patent No. 2,547,447, filed in The Netherlands, March 18, 1942.
——, U. S. Patent No. 2,610,694; filed in The Netherlands, May 20, 1944.
M. Camras, U. S. Patent No. 2,710,662; December 23, 1948.
S. Levy, *et al.*, "A compact, single cabinet, stereophonic loudspeaker system," *J. Audio Engrg. Soc.*, vol. 7, pp. 217–222; October, 1959.
S. Levy, *et al.*, "Listener reaction to stereophonic reproduction by reflected sound," *J. Acoust. Soc. Am.*, vol. 31, pp. 1256–1259; 1959.
B. B. Bauer and G. W. Sioles, "Stereophonic projection console," IRE TRANS. ON AUDIO, vol. AU-8, pp. 13–16; Janaury-February, 1960.
J. K. Hilliard, "Acoustical measurements on a home stereo installation," IRE TRANS. ON AUDIO, vol. AU-9, pp. 41–43; March-April, 1961.

Section 4

AUTOMATIC CONTROL

Organized with the assistance of the IRE Professional Group on Automatic Control

Review of Control Developments *by John G. Truxal*

Automatic Control and Electronics *by Harold Chestnut*

Review of Control Developments*

JOHN G. TRUXAL†, FELLOW, IRE

Summary—**Control engineering, erected on the foundations of feedback theory and linear system analysis, brings together the fundamental concepts of communication theory, network theory, nonlinear mechanics, optimization theory, and statistical design theory. These various elements are combined in the modern research on adaptive control systems—research which provides the basis from which control theory contributes to system engineering and to the understanding of the behavioral, social, and biological sciences.**

ONE OF THE MORE significant technological developments during the first half century of the Institute of Radio Engineers has been the birth and explosive growth of the field of feedback control or automatic control. Providing the technical foundation for the development of both industrial automation and military guidance and control systems, the field of feedback control has produced such diverse accomplishments as the automatic control of the operation of a steel-rolling mill with the sheet steel produced at a rate of 30 mph, the automatic control of traffic lights in many of the major cities of the nation, the phenomenal terminal guidance characteristics associated with the modern infrared-homing antiaircraft missiles, and the guidance and control equipment permitting realization of both the desired orbit and the required attitude stabilization for the Tiros satellites.

From a technical standpoint, feedback control has built on the fundamental areas of network theory, communication theory, information theory, nonlinear mechanics, and finite mathematics. In the selection of basic analytical techniques from these various topics, feedback control has welded these diverse topics into a unified theory of lumped-parameter system analysis and design erected around the focal concepts of feedback theory and the properties associated with the existence of feedback. Thus, as control theory has developed over the past three decades and the control engineer has attacked more and more complex physical systems and more difficult phenomenological situations, he has drawn to an ever-increasing extent on the basic theoretical techniques developed by the applied mathematician and the electrical engineer.

Control engineering has also provided the strongest impetus to the accelerating cooperation and correlation among the different engineering disciplines. Control of a space vehicle can be achieved only through a joint effort of the aeronautical, chemical, mechanical, civil, and electrical engineers involved in the design of the vehicle structure, the design of the propulsion system, the packaging of the components and the realization of the electromechanical systems for construction of the required control operations. As a result of this strongly interdisciplinary character of feedback control, fundamental contributions to pneumatic and hydraulic control have been made by electronic engineers, to electronic system design by engineers of other specializations. The breadth of control engineering, represented by the strong interests of not only the IRE, but also such organizations as the AIEE, AIChE, ASME, ARS, ISA, and SAE, has been a primary influence in the gradual erosion of the barriers, academic and professional, which have so rigidly compartmented the various engineering fields.

The Birth of Feedback Control Technology

Although the basic principles of feedback control have been utilized in engineering equipment for many centuries, the modern intensive interest in control has

* Received by the IRE, October 11, 1961.
† Electrical Engineering Department, Polytechnic Institute of Brooklyn, N. Y.

evolved directly from the work on electronic feedback amplifiers by Black and Nyquist at the Bell Telephone Laboratories during the 1920's. In the development of repeater amplifiers for long-distance telephony, Black investigated[1] the relations between configuration and the sensitivity of the over-all transmission to changes in specific internal parameters. For example, the basic feedback configuration of Fig. 1 yields the following relations for T, the over-all transmission, and for S_x^T, the sensitivity of T to changes in a particular parameter x which is contained within the block G

$$T \equiv \frac{Y}{X} = \frac{G}{1 + GH} \qquad (1)$$

$$S_x^T \equiv \frac{dT/T}{dx/x} = \frac{1}{1 + GH} S_x^G. \qquad (2)$$

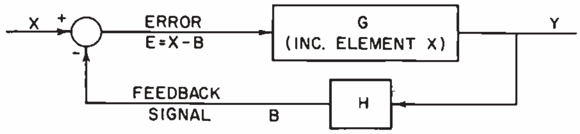


Fig. 1—Basic feedback configuration.

In the case of this simple, single-loop configuration, the sensitivity can be reduced to any extent desired if the designer is willing to accept a similar reduction in the over-all transmission; both functions are reduced by division by $(1+GH)$ as a result of the introduction of the feedback.

As soon as the essential property of sensitivity reduction is recognized as associated with the intentional insertion of feedback, the basic elements of feedback theory can be erected by consideration of other possible configurations in which similar results are achieved. For example, Fig. 2 displays two alternate configurations of slightly more complex structures in which sensitivity control is extended. System a) is designed to insert feedback around the feedback element H in order to reduce the sensitivity of T to changes in H. Configuration b) represents one realization of the advantages of a combination of the usual negative feedback (the G_cGH loop) and simultaneous positive feedback (the G_cG_mH loop). For this system

$$T = \frac{G_cG}{1 + G_cH(G - G_m)} \qquad (3)$$

$$S_G^T = \frac{1 - G_cG_mH}{1 + G_cH(G - G_m)} \cdot \qquad (4)$$

[1] Of particular interest as a retrospective view is the article by H. W. Bode, "Feedback—the history of an idea," *Proc. Symp. on Active Networks and Feedback Systems*, Polytechnic Inst. of Brooklyn, N. Y., 1961, pp. 1–18.

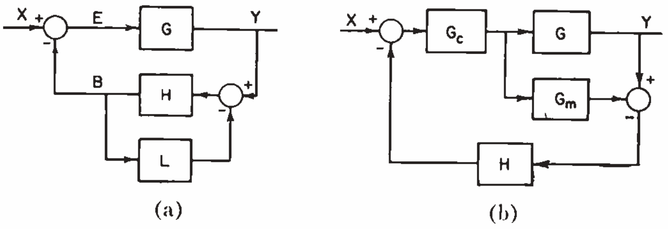


Fig. 2—Configurations for sensitivity control.

Hence, if G_m and H are selected so that $G_cG_mH=1$, the configuration realizes a zero sensitivity (or complete sensitivity control). Furthermore, if G_m is at least an approximate model of G, the over-all transmission is simply G_cG: the configuration is essentially an open-loop or nonfeedback system.

The effective removal of the feedback while still retaining sensitivity reduction is important because of the existence of the stability problem in the basic configuration of Fig. 1. As larger and larger values of the open-loop gain (GH) are utilized in order to realize smaller sensitivity functions, the system eventually exhibits uncontrolled oscillations (*i.e.*, instability). Following the presentation[2] of the Nyquist criterion (which establishes the relation between the transfer functions G and H and stability, and indicates experimental measurements on G and H to determine stability before the loop is closed), the highly developed techniques of equalizer design and network theory were applied to the development of methods for the design of compensation networks to stabilize feedback amplifiers[3] and then feedback control systems.[4]

In the same pattern characterizing so much of electronic engineering, the field of feedback control came of age during the rapid technical developments of World War II, particularly with the strong emphasis on the accelerated development of equipment for the control of radar antennas, antiaircraft turrets, and jet engines. During the early 1940's, the Nyquist-Bode-Hall techniques for design were fully developed, while at the same time a variety of new techniques were introduced which shaped the direction of control developments during the intervening 15 years and which provided the groundwork upon which modern control research was erected.

The awakening recognition of the importance of feedback control around 1940 is epitomized by the establishment of the Institute of Automatics and Telemechanics, in Moscow in 1939, the birth of the Servomechanisms Laboratory under G. S. Brown and the Instrumentation Laboratory under C. S. Draper at Massachusetts Institute of Technology, and the growth (during the

[2] H. Nyquist, "Regeneration theory," *Bell Sys. Tech. J.*, vol. 11, pp. 126–147; January 1932.
[3] H. W. Bode, "Network Analysis and Feedback Amplifier Design," D. Van Nostrand Co., Inc., New York, N. Y.; 1945.
[4] A. C. Hall, "Analysis and Synthesis of Linear Servomechanisms," Technology Press, Cambridge, Mass.; 1943.

War) of control activity in such locations as the Naval Research Laboratory, the Bell Telephone Laboratories, and the M.I.T. Radiation Laboratory. The impact on engineering education was not particularly noteworthy until just after the war, when control suddenly emerged as a primary field of both undergraduate and graduate interest.[5]

The Bases of Modern Control Theory

Largely as a result of this emphasis on control engineering during the 1940's, control theory involved significant portions of a wide variety of basic theoretical disciplines. As a consequence, modern control theory combines viewpoints and analytical and scientific techniques from many different areas, with perhaps the most important being:

1) Network theory
2) Feedback theory
3) Sampled-data theory and computer science
4) Nonlinear mechanics
5) Statistical design theory
6) Communication theory
7) System engineering theory.

In all of these areas, the intensive control research efforts have yielded interrelated developments which, viewed together, constitute the current status of control theory.

The work on *network theory*, previously stimulated and guided by the activities of the IRE Professional Group on Circuit Theory, has provided the theoretical foundation for the design of compensation and stabilization networks for feedback systems. The extensive development of network synthesis techniques by E. A. Guillemin[6] and his colleagues and the subsequent work on the realization of active networks for specified transfer functions have provided the basis for the realization of open-loop dynamic characteristics (*e.g.*, the open-loop transfer function $G(s)H(s)$ in Fig. 1) which correspond to an over-all feedback system which simultaneously satisfies the stability and sensitivity requirements, or such subsidiary and equivalent requirements as the ability of the system to reject corrupting signals entering at points other than the primary input.

Closely coupled with the research on the design techniques for the transfer elements constituting the over-all system, has been the work on *feedback theory*, the relation between the feedback configuration and the properties of the system. This work has grown from the original contributions of Black and Bode to an extensive body of knowledge centered around the concept of signal flow graphs as originally introduced by Mason.[7] The modern, complex weapon or industrial control system often involves a multitude of feedback loops, each involving transfer functions of high order. Feedback theory is concerned with the properties which can be associated with such a configuration, first by virtue of the configuration alone, then as a consequence of the nature of the specific transfer functions within that configuration. For example, a fundamental problem[8] in feedback theory is the determination of the extent to which a given parameter can change without causing instability of the system. The solution is found in an appropriate manipulation of the sensitivity of the transmission with respect to that parameter: an algebraic manipulation which is determined by the form of the feedback configuration around the element. Once this analysis problem is solved, the same techniques can be utilized to design a system in which the sensitivity of the stability to a particular parameter is controlled both by the choice of an appropriate configuration and by the selection of suitable transfer functions within that configuration.

Sampled-data theory and digital techniques[9] provide a close correlation between control theory and both computer technology and communication theory. As, over the last two decades, the specifications on control problems have become more stringent concurrent with the development of small-size, high-speed digital computing equipment, more and more of the control systems for both military and industrial applications incorporate a digital computer as one element of the control system, as indicated by the block diagram of Fig. 3. Here the analog error signal is modified by the active network with the transfer function $G_c(s)$ in order to obtain the analog signal $m(t)$. This $m(t)$ is sampled, quantized, and coded in the analog-digital converter, then modified by

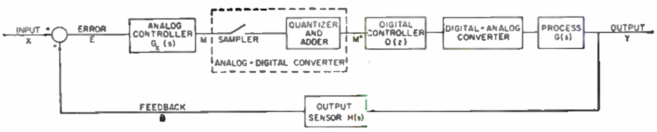


Fig. 3—Simplest computer-controlled feedback system.

[5] The full academic impact was not felt until after the publication of the first basic textbooks in the field: G. S. Brown and D. P. Campbell, "Principles of Servomechanisms," John Wiley and Sons, Inc., New York, N. Y., 1948; L. A. MacColl, "Fundamental Theory of Servomechanisms," D. Van Nostrand Co., Inc., New York, N. Y., 1945; H. Lauer, R. Lesnick, and L. E. Matson, "Servomechanism Fundamentals," McGraw-Hill Book Co., Inc., New York, N. Y.; 1947.

[6] E. A. Guillemin, "Synthesis of Passive Networks," John Wiley and Sons, Inc., New York, N. Y.; 1957.

[7] S. J. Mason, "Feedback theory—some properties of signal flow graphs," Proc. IRE, vol. 41, pp. 1144–1156, September, 1953; S. J. Mason and H. J. Zimmerman, "Electronic Circuits, Signals, and Systems," John Wiley and Sons, Inc., New York, N. Y.; 1960.

[8] W. A. Lynch and J. G. Truxal, "Sensitivity Considerations in Feedback Theory," Microwave Research Inst. Tech. Tept., Polytechnic Inst. of Brooklyn, N. Y.; June 1, 1961.

[9] H. M. James, N. B. Nichols, and R. S. Phillips, "Theory of Servomechanisms," MIT Radiation Laboratory Series, vol. 25, McGraw-Hill Book Co., Inc., New York, N. Y., 1947; J. R. Ragazzini and G. F. Franklin, "Sampled-data Control Systems," McGraw-Hill Book Co., Inc., New York, N. Y., 1958; C. T. Leondes, "Computer Control Systems Technology," McGraw-Hill Book Co., Inc., New York, N. Y., 1961.

the digital program $D(z)$ which yields an output number series related to the input number series $m^*(t)$. Following the digital-analog converter, the signal is utilized to actuate the output member or process being controlled. By insertion of the digital computer, the control system designer achieves the possibility of much greater flexibility in the transfer characteristics of the controller element, the potentiality of readily altering the controller program in order to automatically take into account changes in the environment or in the nature of the signals, and the possibility of programming entirely separate controller instructions for different operating conditions (*e.g.*, for start-up, shut-down, normal operation, and operation under temporary heavy overload). In addition, in many applications, the computer controller can be time shared for the simultaneous supervision of a variety of distinct control operations.

The field of *nonlinear mechanics* has provided the background for the development of nonlinear control theory,[10] a field in which the contributions of Russian control engineers have been particularly notable. By far the most important contribution in this area was the development of the describing-function approach for quasi-linearization,[11] in which a two-port nonlinear element is represented by an equivalent transfer function (or describing function) dependent upon the amplitude of the sinusoidal signal which is applied to the nonlinear device. In the analysis of the system described by the nonlinear differential equation

$$y'' + (3 + 2y^2)y' + 4y = x, \quad (5)$$

the equation is represented by the block diagram of Fig. 4. If the signal y at the input of the cubic element is assumed sinusoidal, the output contains the fundamental plus all odd-order harmonics. To an approximation sufficiently accurate for the majority of cases of control system design and evaluation, only the fundamental component of y^3 need be retained, and the system can be analyzed (for a single amplitude of the sinusoidal signal y) as a linear system only (the gain of the nonlinear block is the complex ratio of the amplitudes of the output and input sinusoids).

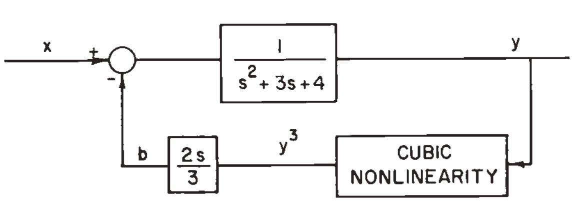


Fig. 4—Elementary nonlinear feedback system.

Perhaps the most significant work in control theory to evolve from the research efforts during the Second World War was the basic study on the design of optimum linear systems for the processing of *random signals.* In this work which was originally carried out by Wiener in studies of the antiaircraft fire-control problem,[12] several new concepts were introduced which have become of fundamental importance to the control engineer:

1) The utilization of correlation functions and numerical correlation techniques for the measurement of power density spectra of the signals which must actually be processed by the control system, rather than design on the basis of test signals such as the sinusoid or the step function.

2) The measurement of the transfer dynamics of a linear system by crosscorrelation of the input and output signals when the input is white noise (a measurement technique which provides the possibility of the measurement of industrial processes during operation and without disrupting the system for long periods of time).

3) The determination of optimum predictors and optimum filters as bases for the evaluation of the quality of performance of conventionally designed control systems.

Unquestionably, the primary result of this research was the entirely novel viewpoint which was introduced into control system design:-a viewpoint which encompassed recognition of the problems associated with the mathematical characterization of the actual signals upon which the system is to operate, and a viewpoint which represented a significant departure from the constraints imposed by the strong electronic engineering tradition of ac circuit analysis. Closely related to this work on the optimum design of linear systems is the more recent and as yet unapplied development by Wiener on the characterization of nonlinear processes through excitation by random signals.[13]

The aspects of *communication theory* which have been of particular interest to the control engineer are, in addition to statistical design theory, those which provide the bases for the evaluation of the ability of a system to transmit specified input waveforms with high fidelity, and those which are concerned with techniques for the identification problem, both signal identification (measuring the significant parameters such as the spectrum distribution of a given signal) and process identification (evaluation of an appropriate mathematical model for a physical system or subsystem). As an example of the former area of interest, the feedback control engineer is

[10] W. J. Cunningham, "Introduction to Nonlinear Analysis," McGraw-Hill Book Co., Inc., New York, N. Y.; 1958.
[11] R. J. Kochenburger, "A frequency response method for analyzing and synthesizing contactor servomechanisms," *Trans. AIEE,* vol. 69 (*Commun. and Electronics,* pp. 270–284, 1950; L. C. Goldfarb, "On some non-linear phenomena in regulatory systems," *Avtomatika i Telemekhanika,* vol. 8, pp. 349–383, 1947.

[12] N. Wiener, "The Extrapolation, Interpolation, and Smoothing of Stationary Time Series," John Wiley and Sons, Inc., New York, N. Y.; 1948.
[13] N. Wiener, "Nonlinear Problems in Random Theory," Technology Press, Cambridge, Mass., 1958.

involved with the design of the electronic feedback system for the control of the intermediate-frequency in a high-quality radar receiver: an operation which involves both frequency stabilization of the local oscillator and automatic compensation for Doppler frequency variations. The resulting feedback loop, involving frequency measurement equipment, limiter, and band-limited subsystems, poses a major feedback control problem which includes nonlinear theory, statistical design theory, electronic circuit theory, and the conventional control system evaluation and design theory.

The identification problem, one of the most important foci of modern control research, encompasses signal characterization in order to guide system design (*e.g.*, if the input signal can be represented as a finite sum of generalized exponentials, the sensitivity can be made zero at the corresponding natural frequencies without making the loop gain infinite over the sinusoidal frequency band of interest. Thus a zero-sensitivity characteristic can be achieved without the stability problem usually associated with systems involving both positive and negative feedback). Alternatively, signal characterization is required when the control system is to yield optimum performance in spite of specified power or saturation constraints or nonlinearities, in order that the controller characteristics may be changed automatically when the signals vary.

The second phase of identification is that of model evaluation: the determination of pertinent characteristics of the process which is to be controlled or other elements of the system. Successful design of a control system such as indicated in Fig. 3 requires that the engineer know at least approximately the dynamic transfer characteristics of the process G. For this purpose, extensive test procedures have been developed and extended by control engineers, including various techniques derived from impulse testing, crosscorrelation techniques requiring the addition to the process input of white noise uncorrelated with the primary input signal, and sinusoidal testing methods, based upon the measurement of the transfer function or describing function. The determination of the transfer function or the impulse response from normal operating records of the input and output signals is one of the key problems blocking the installation of automatic control equipment in industrial applications or the improvement of control equipment for military purposes. Past work by communication and control engineers in this area has centered on techniques for evaluating the amount of data required for a specified accuracy in the power density spectra,[14] investigation of the time required for measurements,[15] the utilization of matched filter techniques in such measurements,[16] and fundamental methods for the characterization of systems.[17]

Finally, control theory has built upon and utilized the applied mathematical work in the field of finite mathematics, in particular *statistical decision theory* and *dynamic programming* methods. The concept of dynamic programming provides an exceptionally powerful approach to a broad class of optimization methods,[18] when such problems can be visualized as multistage decision processes. Within the range of such problems fall the control engineering investigations of methods for the solution of terminal control problems (in which the only measure of the success of the control is the error at the end of the flight or the end of the control task), particularly when the control system is subject to constraints on quantities such as the total fuel consumed during the interval of control action. In more general terms, Kalman has made fundamental contributions[19] in the relation between the degree of controllability which can be achieved and the extent to which the state of the process can be accurately measured, as well as in the development of mathematical solutions to a broad class of optimization problems.

Concluding Comments

Thus, the control field is built upon the basic theoretical disciplines described briefly in the preceding paragraphs, with the different concepts and viewpoints focussed on the properties, design techniques, and evaluation methods of feedback control systems. Coupled with this theoretical foundation is the extensive development during the past two decades of continually improved control components, ranging from gryoscopic equipment for inertial navigation systems to hydraulic and pneumatic controllers for power amplification; from electronic circuits designed for improved reliability by the logical utilization of redundancy and self-checking techniques developed so extensively by the digital computer engineers to mechanical-electric transducers permitting the measurement of ever smaller displacements, pressures, velocities, accelerations, and temperatures in order to permit increasingly close control over the performance of physical equipment. Indeed, to a very considerable extent, major advances in control engineering are far more often associated with breakthroughs in component development than with the evolution of improved theories or design techniques. As

[14] R. B. Blackman and J. W. Tukey, "The measurement of power spectra from the point of view of communications engineering," *Bell Sys. Tech. J.*, pp. 185–282, 485–569; January–March, 1958.

[15] G. R. Cooper and J. C. Lindenlaub, "Limits on the Identification Time for Linear Systems," Purdue Research Foundation 2358-Tech. Rept. No. 2 for Wright Air Development Division, Lafayette, Ind., 1961.

[16] G. L. Turin, "On the estimation in the presence of noise of the impulse response of a random linear filter," IRE Trans. on Information Theory, vol. IT-3, pp. 5–10; March, 1957; cf. also the "Matched Filter Issue" of the IRE Trans. on Information Theory, vol. IT-6; June, 1960.

[17] L. A. Zadeh, "On the identification problem," IRE Trans. on Circuit Theory, vol. CT-3, pp. 277–281; December, 1956.

[18] R. E. Bellman, "Adaptive Control Processes: A Guided Tour," Princeton University Press, Princeton, N. J., 1961; V. G. Boltyanskü, R. V. Gamkrelidze, and L. S. Pontryagin, "Theory of optimal processes," *Izv. Akad. Nauk. SSSR*, 1959. (In Russian.)

[19] R. E. Kalman, "The theory of controllability," *Proc. First International Congress of IFAC*, Moscow, 1960, Butterworths, London, England; 1961.

the component technology progresses toward greater accuracy, better resolution, improved reliability, and smaller weight and size, the theory which underlies the field draws on associated areas (such as network theory, communication theory, and applied mathematics) to expand continually the viewpoint of the control engineer and to encompass an ever wider range of fundamental problems.

The current status of control engineering is perhaps best exemplified by the extensive work throughout the world during the past three years in the development of adaptive control systems, systems in which the controller is automatically adjusted to compensate for changes in the signals, the environment, or the process which is being controlled. Fig. 5 shows one possible block diagram of a general adaptive control system,[20] in which measurement of the process and identification of the input signals are combined in a computer with information about the environment and knowledge of the performance criterion (the mathematical basis upon which performance is to be measured). The computer, which in the general case includes learning in order to improve system performance with usage, determines the optimum form for the controller, which is then realized by the adjustment elements. Thus, a system of this form maintains optimum performance in the presence of variations in the process, the environment, and the signals.

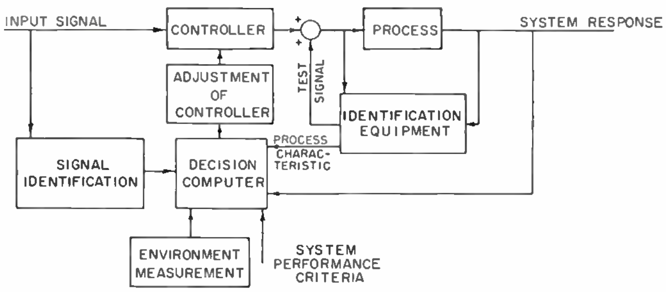


Fig. 5—Basic configuration for adaptive control.

Current research efforts are to a large measure centered on the various aspects of the design of a system such as that indicated in Fig. 5: the problems associated with identification; the problems related to the stability analysis of the nonlinear, adaptive loop involving the process, the controller, and the computer; the questions concerning the evaluation of various possible types of mathematical performance criteria; and the development of simple, feasible realizations of the concepts which are represented here.

Simultaneously, the control field is witnessing a major effort in the expansion of the fundamental concepts of feedback control to the study of related problems in the fields of physical science, social science, biological science and behavioral science. For example, feedback control theory and the related mathematics of systems engineering form a key element of the rapidly changing field of industrial engineering, with emphasis on the development of models for economic situations and the application of basic decision theory to such systems.[21] As a final example, the basic principles of feedback control are being utilized in the study of animal and human operations, as attempts are made to construct logical, feedback models representing the terminal characteristics of various biological organs.

Control engineering, in 1962 a middle-aged element of electronic engineering in the presence of the elderly discipline of communication engineering and the youthful field of solid-state devices, thus provides an essential element of technological developments which can be anticipated during the next half century of the IRE in both the industrial and the military fields.

[20] E. Mishkin and L. Braun, Jr., "Adaptive Control Systems," McGraw-Hill Book Co., Inc., New York, N. Y.; 1961.

[21] C. C. Holt, F. Modigliani, J. F. Muth, and H. A. Simon, "Planning Production, Inventories, and Work Force," Prentice-Hall, Inc., Englewood Cliffs, N. J.; 1960; J. W. Forrester, "Industrial Dynamics," M.I.T. Press and John Wiley and Sons, Inc., New York, N. Y., 1961.

Automatic Control and Electronics*

HAROLD CHESTNUT†, SENIOR MEMBER, IRE

Summary—**The past fifty years have seen the automatic control and electronic fields come close together and form an effective means for increasing man's productivity and his ability to control energy and materials. By extending automatic control concepts to new processes, by developing more flexible controls capable of changing their characteristics to optimize performance of the process being controlled, and by increasing the capability of the sensing means in difficult environments, man will be able to make even more effective his ability to control automatically in the years ahead.**

Electronics is increasingly able to provide physical means for providing the realization of automatic control principles and concepts. Increasing effort to achieve reliable electronic automatic control means must be continued in the years ahead to make possible the realization of the promised gains indicated by the automatic control theory. In addition, more use should be made of standardized design ranges of electrical and mechanical features so that all automatic control equipment can be made in less time and at a lower relative cost. The future appears bright for expanded use of automatic control and electronics as we look ahead for the next fifty years of the IRE.

IT IS SIGNIFICANT that on the occasion of the 50th Anniversary of the IRE, we think of Automatic Control and Electronics as being closely related. Today, automatic control has much of its technical structure firmly based on the same foundation of theory that was developed originally for electronic communications. A large proportion of automatic controls whether they be for the control of energy processes, materials processes, or information processes are electronic in nature. Further, the design, fabrication, and test of automatic controls are heavily dependent on the use of electronic computers both analog and digital.

In a somewhat parallel relationship, many of the applications of electronics are now considerably improved or enhanced by automatic control of such factors as voltage, frequency, or gain. Precision and high-speed fabrication methods for electronic devices and equipment are greatly facilitated by automatically controlled means of manufacture. In addition, many of the logic operations and other judgment criteria now being used more extensively for automatic control have served to make possible or increase the effectiveness of communication means such as automatic long-distance telephones and scatter transmission. Truly, the relationship between automatic control and electronics has developed into one that is most fruitful and mutually advantageous. Through use of some similar components these two fields have been able to capitalize to a certain extent on one another's practical developments, measurement equipment, and fabrication techniques.

* Received by the IRE, July 26, 1961.

† General Engineering Laboratory, General Electric Company, Schenectady, N. Y.

Developments During Past Fifty Years

At the same time of the inception of the IRE in 1912, however, neither electronics nor automatic control were terms that existed in the present sense of the words. It was essentially in the period of the 1920's and 1930's that these terms began to come into use, but the fields of engineering and science they represented tended to operate quite independently of each other.

Electronics in its early days was oriented largely in terms of its contribution to the higher frequencies of the communication and radio industries and operated in a physical environment that was in some respects more controlled and less rigorous than was that experienced by automatic control equipment used in heavy industry. However, as one reads the early IRE Proceedings in the 1913–1915 era, one has the feeling that environmental facts such as the vagaries of the transmission path for RF energy represented highly nonlinear and time-varying phenomena that are similar to those which challenge today's adaptive control.

Automatic control had its background in the speed regulation of steam engines and other power devices. As such, it was initially concerned with the control of high-energy processes of relatively slow operating speeds operating in strenuous environments which might experience high vibration and shock. Because of the high value of the equipment or material over which it exerted its influence, automatic control from its early days was forced to face high reliability requirements. With its lower signal levels and energy implications, electronics reliability requirements were initially less demanding but have grown to be very demanding because of the high military importance of electronics. Gradually, over the years electronics and automatic control have blended their interests and now share many common problems.

Tying Together of Automatic Control and Electronics

The 1930's saw the development by Nyquist,[1] Black,[2] and Bode[3,8] of the frequency response concept of feedback amplifiers for use for communication purposes. By

[1] H. Nyquist, "Regeneration theory," *Bell Sys. Tech. J.*, vol. 11, pp. 126–147; January, 1932.

[2] H. S. Black, "Stabilized feedback amplifiers," *Bell Sys. Tech. J.*, vol. 13, pp. 1–18; January, 1934.

[3] H. W. Bode, "Amplifiers," U. S. Patent No. 2,123,178; 1938.

the early 1940's the automatic control engineers spearheaded by Brown,[4,5] Hall[4,6] and others, began to embrace this method of analysis and design for feedback control system design. Wiener,[7] Shannon and Blackman,[8] and others during World War II tackled prediction and control problems in terms of communication and signal concepts and the intrinsic similarity between communication and control problems began to be established.

It was during the 1920's that the initial emphasis took place on making the electronic tubes then in use rugged enough for the industrial and military environments that would be encountered in many automatic control applications. The possibilities of using electronic tubes as part of the automatic controls for industry were attempted in a limited fashion prior to 1940. However, it was the military requirements of the early 1940's with the increased need for speed and accuracy that provided the greatest impetus for the wedding together of automatic control and electronic ideas.

Post War Developments

The emphasis on servomechanisms[9] and linear control theory during and immediately after World War II was followed by an effort to analyze nonlinear and sampled-data controls. The work of Kochenburger[10] on the describing function has been most helpful in understanding and designing nonlinear systems. In the field of sampled-data controls, Ragazzini,[11,12] Zadeh,[11] Franklin,[12] and later Jury[13] and Tou[14] have done much to clarify the basic analytical appreciation of this phenomenon so important to time-shared controls that are required by electronic digital computers. Truxal's[15] work in bringing together these various analytical methods into a clear perspective has been most helpful.

The widespread use of practical electronic computers, first analog in the 1940's and later digital in the 1950's, has greatly extended the control-systems engineer's analytical ability to understand and design automatic control systems. At the same time, some control systems were being built in which electronic computations were performed as part of the actual control process. The internal logic and control needs, such as floating decimal point and time sharing within some digital computers, have pointed the way to control principles that have been used for automatic controls quite apart from their computer contexts. In similar fashion to the electronic automatic gain control used since the early days of radio, the concept of variable gain or other adaptive controls has in the 1950's found its way into increasing control usage.[16]

Electronics in the form of radar or sonar had been developed during World War II to provide input, detection, and tracking signals for automatic search, ranging and tracking systems. As time has gone on, more extensive use is being found for electronic means of sensing in instrumentation systems both for display purposes and for inputs to automatic control systems. The high-speed, noncontacting aspect of electronic sensing as well as its capability of measuring certain quantities, as for example by X-ray or nuclear magnetic resonance, provide control with capabilities that it might otherwise not be able to attain. The recent development of analytical instruments making extensive use of electronics is providing still further tools with which to increase the capability of automatic control in the chemical and petroleum fields.

Present Status

Gradually, from this evolutionary blending of control, communication, computation, and instrumentation through the common concept of information as influenced greatly by the common medium of electronics, there is taking shape a new grouping of ideas and equipment that is variously called automatic control, instrumentation, or information processing. The basic building blocks of this new automatic control include the functions of sensing, converting, programming, communicating, regulating, computing, storing, actuating, and display as described in Table I.[17]

In the past there has been a considerable effort on individual systems designed to perform the automatic control functions indicated in Table I, and there will continue to be emphasis to improve such equipments.

[4] G. S. Brown and A. C. Hall, "Dynamic behavior and design of servomechanisms," *Trans. Am. Soc. Mech. Engrs.*, vol. 68, pp. 503–524; July, 1946.

[5] G. S. Brown and D. P. Campbell, "Principles of Servomechanisms," John Wiley and Sons, Inc., New York, N. Y.; 1948.

[6] A. C. Hall, "Application of circuit theory to the design of servomechanisms," *J. Franklin Inst.*, vol. 242, pp. 279–307; October, 1946.

[7] N. Wiener, "The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications," John Wiley and Sons, Inc., New York, N. Y.; 1949.

[8] R. B. Blackman, H. W. Bode, and C. E. Shannon, "Data-Smoothing and Prediction in Fire Control Systems," Research and Development Board, August, 1948, or see H. W. Bode and C. E. Shannon, "A simplified derivation of linear least square smoothing and prediction theory," PROC. IRE, vol. 38, pp. 417–425, April, 1950.

[9] H. M. James, N. B. Nichols, and R. S. Phillips, "Theory of Servomechanisms," McGraw-Hill Book Co., Inc., New York, N. Y.; 1947.

[10] R. J. Kochenburger, "Frequency response method of analyzing and synthesizing contactor servomechanisms," *Trans. AIEE*, vol. 69, pt. 1, pp. 270–294; 1950.

[11] J. R. Ragazzini and L. A. Zadeh, "The analysis of sampled data systems," *Trans. AIEE*, vol. 71, pt. 2, pp. 225–234; November, 1952.

[12] J. R. Ragazzini and G. F. Franklin, "Sampled-Data Control Systems," McGraw-Hill Book Co., Inc., New York, N. Y.; 1958.

[13] E. I. Jury, "Sampled-Data Control Systems," John Wiley and Sons, Inc., New York, N. Y.; 1958.

[14] J. T. Tou, "Digital and Sampled-Data Control Systems," McGraw-Hill Book Co., Inc., New York, N. Y.; 1959.

[15] J. G. Truxal, "Automatic Feedback Control System Synthesis," McGraw-Hill Book Co., Inc., New York, N. Y.; 1955.

[16] J. A. Aseltine, A. R. Mancini, and C. W. Sarture, "A survey of adaptive control systems," IRE TRANS. ON AUTOMATIC CONTROL, vol. AC-6, pp. 102–108; December, 1958.

[17] H. Chestnut and W. Mikelson, "The impact of information conversion on control," IRE TRANS. ON AUTOMATIC CONTROL, vol. AC-4, pp. 21–26; December, 1959.

TABLE I

TABLE OF DEFINITIONS OF BASIC AUTOMATIC CONTROL OR INFORMATION CONVERSION FUNCTIONS

Function	Definition or Explanation
Sensing	Generates primary data which describe phenomena or things.
Converting	Changes data from one form to another to facilitate its transmission, storage or manipulation.
Storing	Memorizes for short or long periods of time data, instructions, or programs.
Communicating	Transmits and receives data from one place to another.
Computing	Performs basic and more involved mathematical processes of comparing, adding, subtracting, multiplying, dividing, integrating, etc.
Programming	Schedules and directs an operation in accord with an over-all plan.
Regulating	Operates on final control elements of a process to maintain its controlled variable in accord with a reference quantity.
Actuating	Initiates, interrupts, or varies the transmission of power for purposes of controlling "energy conversion" or "materials" conversion processes.
Presenting	Displays data in a form useful for human intelligence.

At present, however, there is increasing emphasis on the understanding, design, and installation of increasingly complex systems that are a combination of a number of such equipments. The stability, performance, cost, reliability and maintainability of such systems are facets of concern for present-day automatic control activities. Such efforts represent one of the important directions in which automatic control work is proceeding.

National and International Cooperation

A recent interesting development in the organization of people and technical societies interested in automatic control has been the closer association of such individuals and groups on a national and an international basis. Because automatic control embraces such varied equipment as pneumatic, hydraulic, mechanical, electrical, as well as electronic devices and includes instruments, computers, regulators, and many other control means, the American Auotomatic Control Council (AACC) was formed in 1957 to represent the five major American technical societies interested in automatic control in the International Federation of Automatic Control (IFAC) which was also founded in 1957. In addition to the IRE, represented by the Professional Group on Automatic Control, AACC includes the American Institute of Electrical Engineers, the American Society of Mechanical Engineers, the American Institute of Chemical Engineers, and the Instrument Society of America. Through AACC, American control engineers are able to broaden their areas of interests to include a great number of fields of application of automatic control as well as to consider alternate physical means for accomplishing any given desired result. The benefits to be gained by having access to information on the developments in automatic control from all over the world are significant. Such increased technical contacts can do much to accelerate the future development of automatic control in the years ahead.

NEW AREAS OF CONTROL DEVELOPMENT

As we look to other ways in which automatic control developments are progressing, there are a number of encouraging activites taking place. Included in these are:

1) More extensive use of automatic control concepts in new fields where they have not been used to any appreciable degree.
2) Use of more flexible controls through the application of adaptive control concepts and including such means as self-learning or automatic optimizing.
3) Development of new and better sensors capable of measuring more quickly and accurately quantities that will be helpful to new controls.

New Fields

Some of the new fields where automatic control principles have not been used extensively, but appear to hold considerable promise, include the business and economic fields, the social, biological and political fields. Forrester[18] and others working in the field of industrial management have been able to develop dynamic models of individual businesses or industries and have shown the effects of feedback and various strategies on the stability and speed of response of such economic systems. The recent formation of many new countries and the development and growth of many existing ones are being considered from their possibilities of control even though the degree of automatic control is perhaps quite minor at present.

The possibility of using information, sensing, and control concepts to study social, biological, and even political situations is being tried to a limited extent. Data obtained here should be of value to automatic-control people in developing new general control concepts such as self-learning as well as for their value in developing control ideas for specific processes of these types. Uses of automatic control in the form of heart-pacers, synthetic limb controls, and medical operating aids are indicative of some of the directions which automatic control may develop. The use of over-all hospital information systems is another application of automatic control into the biological-social area.

Fig. 1 shows how information conversion is being

[18] J. W. Forrester, "The Impact of Feedback Control Concepts on the Management Sciences," The FIER Distinguished Lecture 1960, FIER, N. Y.; 1960.

used in conjunction with energy and/or materials processes to an increasing degree. Recent activities in the chemical and petroleum industry have been directed at the determination of the dynamic characteristics of the processes themselves and the most effective methods of controlling them. Significant cost reductions appear to be possible by more automatic *control of systems of such variables rather than of fixed control of the individual variables themselves.*

Experiences in the steel and electric utility industries are beginning to supply actual data from advanced automatic control systems involving control of materials, energy, and information as indicated in Fig. 2. In addition to the conventional automatic control shown by the solid lines, programmed inputs are supplied by the programmer, information about the actual performance of the process are recorded by the data logger, and modified inputs to the reference are determined by the computer operating on- or off-line. The need for extensive automatic control in such advanced energy processes as nuclear fission, magneto-fluid dynamics and power generation in space also represent new challenges to the control field.

An important aid in the application of automatic control in new fields will be the development of appropriate models for representing in a space-time sense the phenomenon taking place in the "process." These models must be quantitative to be most effective but the application of automatic control principles may make it possible for somewhat inexact or approximate models to be employed.

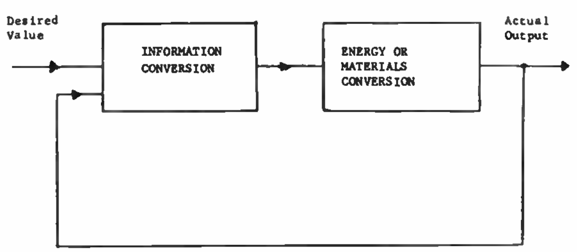


Fig. 1—Control system showing interrelationship of information, energy and materials conversion.

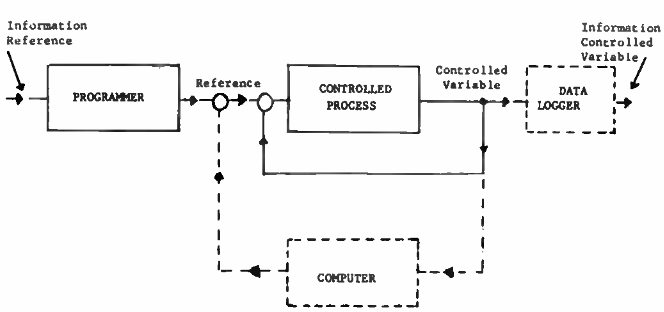


Fig. 2—Combined feedback control and information conversion system.

More Flexible Controls

The efforts at controlling such variable parameter processes as supersonic flight including that of space vehicles or combined materials and energy processes have highlighted the need for more flexible or adaptive controls. Fig. 3 serves to illustrate how changes in the character of the process being controlled can be used to provide an input to an adaptive control which in turn serves to modify the principal control of the process. Other forms of adaptive control use sensing and identification of the reference input or of the actuating error itself to provide a basis on which the adaptive control operates. Increasing sophistication, including extensive computation, is taking place in both the criteria being used as a basis for identifying the nature of the process as well as the nature of the changes that are made to take place in the principal control itself. Additional effort is required to provide a clear analytical understanding of the stability and performance phenomena of such highly nonlinear, time-varying, dynamic control systems.

Fig. 4 serves to illustrate the nature of the multivariable form of the adaptive control problem. By virtue of changing environments as well as the inherently nonlinear nature of the process itself, the control may be faced with the need for changing its logic and strategy in a fashion that can be only generally stated by the automatic control designer. Such controls need to possess a self-learning capability in which their characteristics are allowed fairly broad latitude if the process is so poorly known initially as to provide relatively little information to the designer. Based on quite general design criteria for acceptable performance, the control will use its memory as well as logic rules and endeavor to learn from its experience. By modifying its

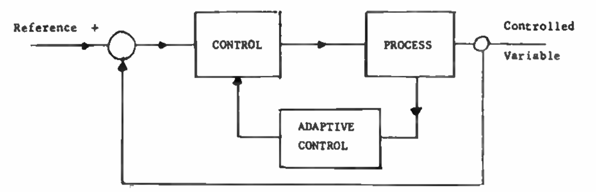


Fig. 3—Adaptive controls sought for processes which are not linear, constant, or fully known.

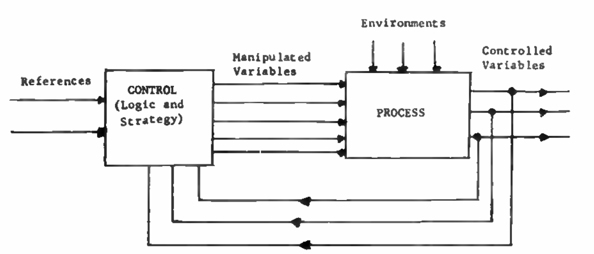


Fig. 4—More processes are multivariable with poorly controlled environments.

characteristics it will try to improve its performance in terms of some initially-agreed-upon objectives.

A variation of this problem is found in the information field itself where pattern or character recognition problems may possess a similar lack of definition for some information processes in their initial stages. Another form in which flexible adaptive control may prove increasingly effective is in the area of automatic design optimization. With engineering efforts representing a larger portion of the cost of producing new products and systems, more extensive computer methods for performing engineering evaluations are being employed. The expansion of control concepts of optimization more broadly to design[19] represents an attractive field for more flexible automatic-control concept application.

New Sensors

Paramount to the solution of any automatic control problem is the ability to sense the conditions that exist in the process. Although modern control theory is firmly based on the assumption that an adequate measure of the variables being controlled can be obtained, there are a number of materials and energy processes where sensors capable of making such measurements are not available. As Lord Kelvin has stated so well, it is necessary to measure things quantitatively in order to be able to understand and control them. Physical and chemical composition measurements in high and corrosive atmospheres are illustrative of specific problems in the sensor area. New analytical instruments capable of accurate and rapid measurements and suitable for on-line control applications will do much to further the usefulness of automatic control. Although creative thinking in conceiving new measurement concepts will be valuable, the development of new electronic components to operate in unfavorable environments such as these will also be very beneficial.

Associated Needs for Continued Control Progress

Accompanying the developments in the better understanding and broader application of automatic control, there exist associated needs for improved components and equipment to provide the means for accomplishing automatic control more reliably, quicker, and at lower relative cost. It is here that continued electronic development can be most effective.

Reliability[20]

The vast amount of work done in the past fifteen years on the reliability of electron tubes and their associated electronic components has resulted in marked electronic circuit improvement over the years and doubtless improvements will continue to be made. More extensive use of ceramic or relatively inert materials accompanied by strict observance of sound mechanical design principles have pointed the way to more reliable electronic and automatic control equipment. Conservative electrical designs in the sense of using components at operating conditions well below the design limits have likewise shown that improved reliability can be obtained.

The development of transistors and general solid-state electronics has provided smaller, lighter, and lower-power-consuming devices capable of high reliability that are finding increasing use in automatic control equipment. Certainly, the weight, size, and reliability needs in space vehicle control have emphasized the desirability of these miniature electronic devices. The possibility of micro-electronics for control in space appear to offer further advantages if the predicted reliability can be realized and adequate means for automatic controlled manufacture can be achieved at reasonable cost. Present thinking would seem to indicate however, that in the industrial area the need for extremely small size is not as important as high reliability and low cost.

Accompanying the efforts to improve the reliability of individual electronic elements or groups of elements, further emphasis is required in the judicious use of redundancy, fail-safe features, and/or other system means for providing over-all reliable operation. The large amount of electronic automatic control equipment currently being utilized in many industries and the large economic investment in processes being controlled automatically emphasize the present and future needs for automatic control of appreciably higher reliability.

Standardized Designs for Reduced Cost and Time

In many cases automatic control equipment has been designed from a custom-design point of view. Since the energy or materials portion of the system, *i.e.*, the controlled process, tended to be unique with each installation, the automatic control tended to be designed and tailored to the installation needs. This has resulted in the time required to produce automatic control being longer than if more standard control components were used. It is appropriate that increasing efforts be made to work toward a more standardized approach of automatic control system design in which appropriate voltage, frequency, impedance and power levels are established broadly across the control industry so that compatible components may be interconnected directly to a greater extent.

Compatible design should be directed toward achieving a greater measure of over-all electrical and mechanical uniformity for subassemblies and equipments so that shorter equipment delivery times can be realized and the cost for engineering and manufacturing per

[19] E. L. Harder, "Computers and automation," *Impact of Sci. on Soc.*, vol. 1, pp. 3–15; 1960.

[20] "Reliability of Military Electronic Equipment Report," Advisory Group on Reliability of Electronic Equipment (AGREE), U. S. Govt. Printing Office, Washington, D. C.; 1957.

individual equipment can be reduced. Automatic control and electronics are still changing at a sufficiently high rate that it does not appear appropriate that rigid standardization be introduced at this time. However, the rapid spread in the use of digital and logic control methods and the extensive use of transistor circuitry point up the desirability of having equipment with common characteristics for such electrical features as voltage level, impedances, and frequency in a number of ranges. By this means, fewer engineering designs and associated manufacturing fixtures will be required and more time and effort can be devoted to making sure the equipment is reliable and of suitable quality and performance. Industry efforts presently being undertaken in this direction should be strengthened and advanced more rapidly. With more compatible components the desirable objectives of reliable automatic control can be obtained more quickly and more cheaply.

Conclusions

The past 50 years have seen the automatic control and electronics fields come close together and form an effective means for increasing man's productivity and his ability to control energy and materials. By extending automatic control concepts to new processes, by developing more flexible controls capable of changing their characteristics to optimize the performance of the process being controlled, and by increasing the capability of the sensing means in difficult environments, man will be able to make even more effective his ability to control automatically in the years ahead.

Electronics is increasingly able to provide the physical means for the realization of automatic control principles and concepts. The increasing effort to achieve reliable electronic automatic control means must be continued in the years ahead to make possible the realization of the promised gains indicated by the automatic control theory. In addition, more use should be made of standardized design ranges of electrical and mechanical features so that automatic control equipment can be made in less time and at a lower relative cost. The future appears bright for expanded use of automatic control and electronics as we look ahead to the next 50 years of the IRE.

Section 5

BROADCAST AND TELEVISION RECEIVERS

Organized with the assistance of the IRE Professional Group on Broadcast and Television Receivers

The Development of the Art of Radio Receiving from the Early 1920's to the Present*

WILLIAM O. SWINYARD†, FELLOW, IRE

***Summary*—A brief résumé of important scientific discoveries leading to radio broadcasting is given. There follows a discussion of the most important types of home broadcast radio-receiving circuits which came into use after the advent of entertainment broadcasting. These circuits include crystal receivers, regenerative receivers, neutrodynes and superheterodynes. Attention is also given to the development of automatic volume control, automobile radios and FM receivers.**

An effort has been made to present the developments chronologically with attention primarily being given to the work of leading pioneers.

In the final part of the paper some of the more important innovations in the radio-receiving art are mentioned and, in some cases, briefly discussed.

Introduction

THE radio-receiving art owes a great debt to early pioneers—Ampere, Henry, Faraday, and, more specifically, Maxwell and Hertz. Credit is due Maxwell for mathematical studies which led to his electromagnetic theory of light and his prediction of the existence of electromagnetic waves of relatively low frequency, and Hertz for being the first to produce them—in 1887, some 25 years later.

The first crystal detector was designed in 1906 by Pickard. It consisted of a silicon crystal and a cat whisker. At about the same time Dunwoody developed a detector using a carborundum crystal clamped between two brass holders. In 1883 Edison, in seeking the cause of blackening of lamps, discovered current flow inside a vacuum tube. This led to the invention of the vacuum diode, the "valve," by Fleming in 1904 and to the invention of the triode vacuum tube, the "audion," by de Forest in 1907.

However, the triode remained a rather unsatisfactory device until 1912 when H. D. Arnold recognized that its imperfections stemmed largely from its defective vacuum and the use of a tantalum, instead of an oxide-coated filament. He improved the vacuum and Wehnelt invented the oxide-coated filament. With the removal of these shortcomings the triode became the cornerstone of modern radio communications.

At the conclusion of World War I many former radio amateurs, "hams," returned to civilian life with a keen interest in radio. This interest was greatly sharpened by the awareness of these men of the radio developments of the day. They, along with an ever-increasing number of other hams, became an enthusiastic group of potential radio listeners.

In 1920 F. Conrad of Westinghouse converted his amateur wireless station, which he built in 1916, to radiotelephony and returned to the air. He announced regular two-hour broadcasts on Wednesday and Saturday nights and the response in the Pittsburgh area was very enthusiastic. A local department store bought a supply of crystal sets which were quickly sold to the "amateur" public. This response paved the way for a decision to put the station on a regular broadcasting basis, and on November 2, 1920, Westinghouse station KDKA in East Pittsburgh, Pa., was officially opened and the

* Received by the IRE, October 20, 1961.
† Hazeltine Research, Inc., Chicago, Ill.

era of radio broadcasting began. The station's initial broadcast of the Harding-Cox election returns was sensational.[1]

Since that time radio broadcasting has shown a spectacular growth. Many of the companies which had been manufacturing military equipment started to produce "amateur" sets. Newspapers, department stores and educational institutions opened their own broadcasting stations. By 1922 there were 30 licensed broadcasting stations in the United States, and by 1924 there were over 500.

Crystal Receivers

Before and during World War I crystal sets were in wide usage by radio hams. It was only natural, therefore, that this type of receiver should be the first employed after the beginning of radio broadcasting. Tuning in crystal receivers was effected by varying either the tuning capacitance or the inductance.

The most commonly used crystals were silicon, carborundum (silicon carbide) and galena (lead sulphide). The receiver was operated by tuning the circuit to the station and then assuming the use of a silicon detector, probing the crystal surface with the cat whisker to find a sensitive spot, retuning, reprobing, etc. The cat whisker was a sharp, firm metal wire and in the case of the widely-used silicon detector was sometimes made of antimony. With a good straight-wire outdoor antenna, the range of a crystal set for satisfactory response from a broadcasting station was limited to a radius of 5 to 50 miles. All crystal sets suffered from a lack of sensitivity and crystal instability. However, they were available and were relatively inexpensive and easy to produce.

Regenerative Receivers

Crystal sets were soon replaced by regenerative receivers.[2,3] The regenerative circuit (Fig. 1) was invented in 1912 by de Forest, Armstrong and Langmuir in the United States, and by Meissner in Germany. After 20 years of litigation, the United States Supreme Court finally decided in de Forest's favor. Meanwhile, in 1917 Armstrong received the first IRE Medal of Honor for his work on regeneration and the production of oscillations. After the Supreme Court decision, Armstrong returned the Medal of Honor to IRE, but the Board of Directors unanimously declined to accept it and reaffirmed the original award.

In the regenerative detector circuit RF energy is fed back from the anode circuit to the grid circuit to give positive feedback at the carrier frequency, thereby increasing the sensitivity of the circuit.

Regenerative receivers marked a big step forward in providing greatly increased sensitivity. Inherently they provided large amplification of small signals and small amplification of large signals. By 1922 they had reached the high point in their development and had almost entirely superseded crystal sets.

Three sets captured a large share of the market: 1) Westinghouse (RC) Radiola, 2) Clapp-Easthan, and 3) Grebe. Each of these receivers included a detector and two stages of audio amplification. In some cases the audio amplifier was sold separately. At that time only two frequencies were allocated to broadcasting: 300 and 360 meters. However, everyone wanted to receive ship signals on 600 meters (500 kc), and hams on the "short-wave" band, 200–250 meters (1200 to 1500 kc) and this demand determined the frequency coverage provided by these early sets.

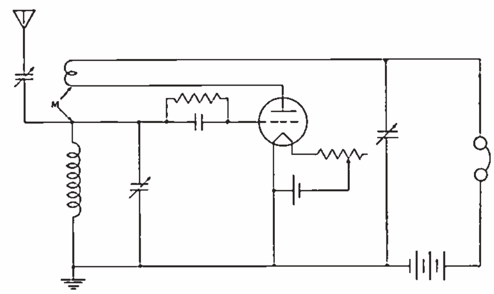


Fig. 1—Armstrong regenerative receiver.

Regenerative receivers, however, were subject to serious difficulties since they could easily be made to oscillate if the regeneration control were turned up too far. When the receiver was in an oscillating condition it produced heterodyne whistles in nearby receivers when it was tuned to, or through, stations to which they were tuned. This device soon became a public nuisance and threatened the development of the radio art. Regenerative receivers were also hard to adjust, and it was almost an impossibility to operate them so that they would always receive the same station at the same point on the dial. Logging stations was, therefore, impracticable.

Neutrodyne Receivers

The neutrodyne circuit was invented in 1918 by Hazeltine.[4–6] Basically this was a tuned radio-frequency (TRF) amplifier which employed a specific type of neutralization. A current obtained from the plate circuit was fed back into the grid circuit in the proper magnitude and phase to balance out, or neutralize, the effect of grid-to-plate capacitance inside the tube, and thus it achieved stability and prevented oscillation. The grid-to-plate capacitance of the tubes used in the early

[1] W. R. MacLaurin, "Invention and Innovation in the Radio Industry," The MacMillan Company, New York, N. Y., p. 112; 1949.

[2] E. H. Armstrong, "Some recent developments of regenerative circuits," Proc. IRE, vol. 10, pp. 244–260; August, 1922.

[3] E. H. Armstrong, "The regenerative circuit," *Proc. Radio Club of America*, April, 1915.

[4] F. E. Terman, "Radio Engineers Handbook," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 468–469; 1943.

[5] MacLaurin, *op. cit.*, pp. 127–129.

[6] Louis A. Hazeltine, "Tuned radio frequency amplification with neutralization of capacity coupling," *Proc. Radio Club of America*, vol. 2; March, 1923.

sets was high, and in the case of the type 201A was between 8 and 12 pf. Hazeltine originally proposed this circuit to stop squealing in audio amplifiers, stating that it was no doubt also applicable to radio-frequency circuits. In 1922 Hazeltine designed a radio receiver on paper incorporating his neutralizing circuit. This receiver was a reflex neutrodyne employing three tubes and three tuned circuits, each tuned by a separate tuning dial. Magnetic coupling between the unshielded tuned RF transformers was eliminated by mounting them in line with their axes inclined 54.7° with respect to the line of centers. The circuit diagram is shown in Fig. 2. A modified version in which a separate audio amplifier was used was placed on the market by Freed-Eisemann in 1923. This was a five-tube set and was the first set having high RF amplification and high RF selectivity. By 1924 neutrodyne receivers were selling by the hundreds of thousands.

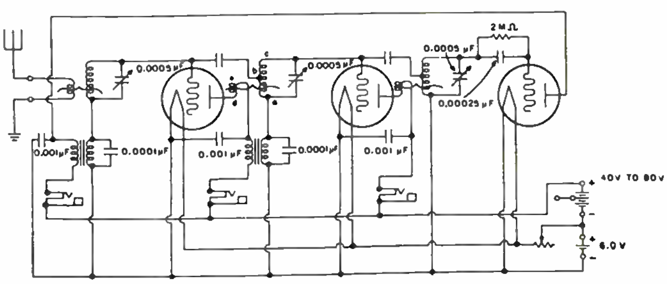


Fig. 2—Hazeltine reflex neutrodyne receiver.

Hazeltine developed the neutrodyne circuit to a complete solution from a mathematical point of view. For this work, along with his conception of vacuum-tube parameters, specifically mutual conductance,[7] he was awarded the Armstrong Medal by the Radio Club of America in 1937.

Neutrodyne receivers had stable characteristics; tuning dials could be logged when stations were tuned in and they were easy to operate.

In 1926 a neutrodyne receiver was designed which employed shielding between the coils, the shielded neutrodyne.[8] This receiver marked the first attempt at unicontrol. It had two dials (in addition to the volume control); one tuned the antenna circuit and the other tuned three stages of RF amplification. It was offered by Stromberg-Carlson and others.

For a number of years any set carrying the name neutrodyne was certain to sell. However, with the development of the screen-grid tube in 1928 with its low grid-to-plate capacitance, the necessity for neutralization disappeared and the demand for this receiver declined rapidly.

Many sets, neutrodynes and other types, having tuning capacitors ganged together so that they could be tuned by means of a single dial, were manufactured and sold in the late twenties. However, all fell short of true unicontrol operation since the volume control, or RF gain control, or both, had to be continuously readjusted as the set was tuned over the band in order to avoid inadequate volume or overload (blasting).

Automatic Volume Control

In 1925 the stage was set for the invention of a practical automatic volume-control circuit, and on January 2, 1926, Wheeler invented his diode AVC and linear diode-detector circuit.[9,10] This circuit was first incorporated in the Philco Model 95 receiver which he designed at the Hazeltine laboratory and which was announced about September, 1929.[11]

This receiver employed 9 tubes, two type-224 tuned RF amplifiers, one type-224 untuned RF amplifier, a type-227 connected as a diode detector, a type-227 resistance-coupled first AF stage, a type-227 transformer-coupled AF driver, and a push-pull audio output stage using type-245 tubes. The rectifier was a type 80.

Two coupled tuned circuits were employed preceding the first tube to prevent cross-talk from nearby local stations. The antenna circuit employed a high-inductance primary and was designed to provide tuning which was independent of antenna size, and uniform gain and selectivity over the band.

A type-227 tube with its grid and plate tied together served as a two-element (diode) detector which supplied both an AF signal to the audio amplifier and a dc bias voltage to the RF amplifiers. Since a two-element detector is linear over practically its entire rectification curve when a high dc impedance is used in its return circuit, as was the case here, it provided the greatly desired linear detection characteristic. In this receiver the useful range of linearity of the detector was from 1 to 10 volts, approximately.

Full AVC bias voltage was applied to the first two RF tubes, and, to prevent distortion in the third RF stage, half AVC bias voltage was applied to that stage. The automatic volume-control action was sufficiently gradual to permit accurate tuning by ear, and it was unnecessary to touch the volume control once it was adjusted. With all four tuning capacitors ganged to a common tuning dial, true single-knob operation was provided for the first time in any receiver. The sensitivity of the receiver was about 5 μv/m across the band,

[7] L. A. Hazeltine, "Oscillating audion circuits," Proc. IRE, vol. 6, pp. 63–98; April, 1918.

[8] J. F. Dreyer, Jr., and R. H. Manson, "The shielded neutrodyne receiver," Proc. IRE, vol. 14, pp. 217–247; April, 1926. "Discussion," L. A. Hazeltine, pp. 395–412; June, 1926.

[9] H. A. Wheeler, "Automatic volume control for radio receiving sets," Proc. IRE, vol. 16, pp. 30–39; January, 1928.

[10] H. A. Wheeler, "Design formulas for diode detectors," Proc. IRE, vol. 26, pp. 745–780; June, 1938.

[11] W. E. Holland and W. A. MacDonald, "The Philco '95' screen-grid plus," *Radio Broadcast*, vol. 16, pp. 111–112; December, 1929.

and the signal input ratio over which the automatic volume control was effective was 300 to 1. With a local-distance switch this was increased to 10,000 to 1.

By 1932 practically every receiver on the market used a diode AVC circuit employing a single-diode detector to supply both the audio signal and the AVC voltage. This situation remains unchanged today, insofar as electron-tube receivers are concerned.

The Superheterodyne

A number of engineers worked on the problem of the production of superaudible frequencies by means of heterodyning. Pioneers in this work were Levy in France, Round in England, and Armstrong in America.[12] The direct outcome of Armstrong's work was the development of the superheterodyne receiver.[13]

In the superheterodyne receiver the signal frequency is mixed with the locally-generated oscillator frequency in a nonlinear device to produce two new frequencies corresponding, respectively, to the sum of, and the difference between, the two frequencies. The difference frequency is fed to the IF amplifier where it is amplified and then rectified in the detector where the audio-frequency component is taken off and fed to the audio amplifier. In early receivers of this type the tube in which the signal and oscillator frequencies were combined was called the *first detector*, and the tube which separated the audio signal from the modulated IF signal was called the *second detector*. Later these terms were changed to *converter*, or *mixer*, and *detector*, respectively. More precisely, the term *converter* applies to the combined oscillator and mixer functions, which may be performed in the same tube.

The superheterodyne obtains only a relatively small part of its gain and selectivity in the antenna and RF amplifier circuits; by far the larger part is obtained at a lower frequency in the IF amplifier where gain is more manageable. Since tuning of the IF amplifier is fixed, a more desirable band-pass selectivity characteristic can readily be achieved. These factors combine to make it relatively easy to secure uniform gain, selectivity and fidelity over the band.

The first publicly sold superheterodyne receivers were made by the General Electric Company and placed on the market in March, 1924, by RCA.[14] These receivers were known as the Radiola Superheterodyne Second Harmonic, a table model, and the Radiola Super VIII, a console model. The table model employed a fixed built-in loop antenna, but did not include a loudspeaker. Batteries were located in side compartments of the cabinet. The console model included a built-in rotatable loop antenna and a horn loudspeaker. Both models employed the same chassis which utilized a circuit developed by H. Houck. This circuit employed six UV-199 tubes and included an untuned RF stage, second-harmonic converter stage, two IF stages, one of which was reflexed with the RF stage, an audio detector and two AF stages. Second-harmonic operation of the oscillator was employed to prevent tuning interaction between signal frequency and oscillator circuits. The circuit diagram is shown in Fig. 3.

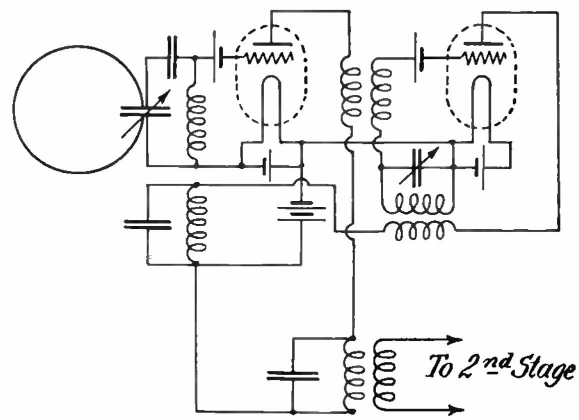


Fig. 3—Basic superheterodyne second harmonic circuit.

The chassis featured a component assembly called the *catacomb* which included the IF, RF and AF components and was potted in wax. The RF and IF transformers employed 3-mil silicon steel laminations, and two of the IF transformers had both primary and secondary tuned to 42 kc and were critically coupled to provide a band-pass selectivity characteristic. Regeneration was reduced by capacitive neutralization and by magnetic neutralization through the common silicon steel core.

Later superheterodynes were designed to operate on the fundamental of the oscillator, and the IF was raised to about 175 kc. By 1930 they began to take over, and by 1932 they were in universal use, as they are today. In 1938 the IF was set at 455 kc by the RMA Standards Committee.

Automobile Radios

Transitone began to produce automobile radios in about 1927 on a custom-made or low-production basis. In 1931 this company was bought by Philco, and the product was designated Philco-Transitone. However, Motorola claims to have turned out the first mass-produced commercial auto radio in 1929. This set employed a tuned RF circuit consisting of three gang-tuned stages, an antenna and two RF stages employing type-224 tubes, a type-224 detector, a type-201A audio driver, and push-pull type-112A's in the output stage. The 6-volt car battery supplied the filament current and "B" voltage was supplied by three 45-volt batteries. The receiver included a dynamic loudspeaker with a 6-volt field coil. Signal pickup was effected by means of a

[12] D. McNicol, "Radio's Conquest of Space," Murray Hill Books, Inc., New York, N. Y., p. 272; 1946.

[13] E. H. Armstrong, "The superheterodyne—its origin, development, and some recent improvements," Proc. IRE, vol. 12, pp. 549–552; October, 1924.

[14] Prior to 1929, RCA was not a manufacturer. It was a patent licensing and sales organization owned by General Electric, Westinghouse and AT&T. See McLaurin, *op. cit.*, p. 147.

screen-type antenna mounted in the roof of the car. The three-section gang capacitor and the volume control were remotely operated by means of flexible shafts running from the metal box housing the receiver to a control box mounted on the steering column.

The superheterodyne circuit was first used in an auto set in 1931. This set also employed batteries to supply filament, plate and screen potentials. The intermediate frequency was 175 kc. It was replaced the same year by a set which utilized a wired-in vibrator power pack.

In 1932 Motorola marketed what is claimed to be the first auto set which employed a power supply with a plug-in vibrator to furnish plate and screen voltages. This receiver utilized a superheterodyne circuit and employed 8 tubes. The intermediate frequency was 175 kc. It incorporated circuits to filter the battery and plate-supply voltages to reduce ignition noise and vibrator hash.

After World War II, ganged capacitors were scarce, and permeability tuning came into wide usage in home receivers and in auto sets. Because of the improvement in SNR and freedom from microphonism provided by this type of tuning-in auto sets, the industry stayed with "perm" tuning in receivers of this type. The improvement in SNR came about because the "perm-tuned" antenna circuit could be more efficiently coupled to a low-capacitance automobile antenna, thus providing higher antenna gain.

The transistor, first publicly announced in 1948, is said to have been first applied in entertainment-type automobile receivers in a set made by Motorola and marketed in about 1956. In the audio-output stage this set incorporated a power transistor mounted on a heat sink outside the metal cabinet. All other functions requiring electron devices utilized tubes designed to operate with 12-volt plate, screen and filament potentials which were supplied directly by the car battery. Sets of this type are called *hybrid* sets because of the use of both tubes and transistors.

The main reason for the existence of the hybrid set is based on economic considerations, transistors being too high priced to make completely transistorized auto sets competitive for the commercial market. Additionally, a transistor was used in the power output stage because it delivered greater power output than could be secured with a 12-volt audio-output tube.

Some manufacturers made hybrid auto sets which employed conventional tubes whose plate and screen voltages were supplied by a transistorized power oscillator.[15]

Hybrid auto sets have been very popular from 1956 to the present. However, they are gradually giving way to all-transistor receivers as the price of transistors comes down. It is estimated that by 1963 almost all automobile sets will be fully transistorized.

The first all-transistor auto set is said to have been placed on the market by Philco in about 1956. It employed a conventional superheterodyne circuit and 12 transistors. The intermediate frequency was 262.5 kc, the power output 3.5 watts and the nominal battery current drain 0.4 ampere at 13.8 volts.

In connection with hybrid and fully-transistorized auto sets, it should be pointed out that the usual problems of vibrator hum and "hash" no longer exist.

Most early auto sets used antennas which were mounted in the car roof or under the car or running board. After "touring" cars were replaced by sedans having metal tops, present day whip-type antennas came into general use.

Frequency-Modulation Receivers

The credit for promoting FM as a broadcast service goes to E. H. Armstrong. For years he had been seeking a way to reduce static, and finally he turned his attention to FM. Toward the end of 1933 he had perfected a system of wide-band frequency modulation which seemed to overcome natural and many forms of man-made static.[16] In this system the carrier was frequency modulated ± 75 kc by audio components up to 15 kc.

In 1938 General Electric placed the first FM receiver on the market—their Model GM 125, a superheterodyne employing 11 tubes plus a rectifier.[17] It was tunable over a range of 41–44 Mc; the IF was about 3 Mc, and the bandwidth was sufficient to handle a frequency deviation of ± 100 kc. It employed a type-6SJ7, as a grid-current limiter to remove amplitude modulation, followed by a type-6H6 double diode in a Foster-Seeley discriminator circuit to provide FM detection. It is interesting to note that this discriminator circuit was originally proposed in 1936 for use as an automatic-frequency-control detector and came into rather wide usage in this application.[18]

In 1945 the ratio detector invented by Seeley came into use as an FM detector and was widely adopted, particularly in medium- and low-priced FM receivers, because it did not need to be preceded by a limiter stage.[19]

On January 1, 1941, the FM band of 42–50 Mc came into general use in accordance with the FCC decision of May, 1940, and in August, 1945, the FCC allocated the present band—88–108 Mc. In 1946 the Radio Manufacturers Association (now Electronic Industries Association, EIA) proposed 10.7 Mc as the standard FM intermediate frequency and proposed an antenna-to-set transmission-line impedance of 300 ohms.

[15] C. C. Hsu, "Development of the 12 volt plate-voltage hybrid automobile radio receivers—AM, signal seeker and FM," IRE Trans. on Broadcast and Television Receivers, vol. BTR-4, pp. 1–31; March, 1958.

[16] S. W. Seeley, "Frequency modulation," *RCA Rev.*, vol. 5, pp. 468–480; April, 1941.

[17] G. W. Fyler and J. A. Worcester, "Frequency modulation in radio broadcasting—a new Armstrong frequency-modulated wave receiver," *Proc. Radio Club of America*, vol. 16; July, 1939.

[18] D. E. Foster and S. W. Seeley, "Automatic tuning, simplified circuits, and design practices," Proc. IRE, vol. 25, pp. 289–313; March, 1937.

[19] S. W. Seeley and J. Avins, "The ratio detector," *RCA Rev.*, vol. 8, pp. 201–236; June, 1947.

Innovations

There have been many innovations introduced since radio-broadcast-receiver manufacturing began. Some were worthwhile and contributed to the development of the art, but many were only short-lived gimmicks.

All early radio receivers employed headphones. However, the need for greater sound output soon resulted in attaching horns to telephone-type receivers. These were followed in 1922–1924 by cone-type speakers such as those made by Western Electric and Farrand, which employed cones attached to moving armatures. In 1924 Magnavox designed and marketed a balanced armature magnetic-type horn loudspeaker.

In 1925 C. W. Rice and E. W. Kellogg of General Electric developed the electrodynamic loudspeaker, marketed by RCA, and in 1931 Jensen put on the market a dynamic loudspeaker with a permanent magnet made of chrome steel, the forerunner of the present Alnico V and ceramic magnets.

Some of the early radio receivers employed loop antennas. Later, in an effort to receive signals from greater distances, outdoor antennas of many types came into general use. As advances were made in tubes, circuit design and loop-antenna design, it became feasible to use loop antennas without too much loss of sensitivity. In 1939 loops came into general use, and some receivers appeared on the market with permeability-tuned loops. In 1950 the use of ferrite-rod antennas instead of loops was proposed, and gradually this type of antenna, the *loop-stick*, replaced conventional loops.

Early receivers all employed battery operation since tubes with ac-operated filaments were not available. Then, in 1924 the Philadelphia Storage Battery Company marketed "B" battery eliminators which came into wide usage. Shortly thereafter "A" battery eliminators were developed; these also gained wide public acceptance. In 1926 Zenith built a receiver with both "A" and "B" battery eliminators built in as an integral part of the receiver. The wide usage of the ac receiver followed the manufacture in 1926 of tubes with filaments which could be operated on alternating current. Since there were many areas in the country where only dc power was available, the present day ac-dc circuit was developed by R. Weurfel of International Radio to meet this need. This was first marketed by International Radio in 1931, and by 1935 about 20 per cent of the total sets made were ac-dc.

In 1939 portable receivers using battery tubes with 1.4-volt filaments became popular and ac-dc-battery sets appeared on the market. One of the outstanding developments of 1940 was the miniature receiver which weighed less than five pounds and which was powered by readily-replaceable batteries. Further attempts at miniaturization resulted in a pocket radio employing subminiature tubes and especially developed components. One of the first of such sets was produced by Belmont late in 1946.

In 1952 the use of printed circuits employing a dip-soldering process was extended to small radio receivers. These steps along with the use in the laboratory of transistors in radio receivers for the first time high-lighted developments in that year. In 1954 Regency marketed the first transistorized radio.

Over the years since the early twenties there have been many other innovations—all-wave receivers, including the "weather" band, "spread-band" receivers, AM-FM receivers and "Fremodynes."[20] Some of these receivers employed IF amplifiers with adjustable selectivity, automatic bass boost, noise limiters, tuning eyes and inverse feedback. There have been big tuning dials, little tuning dials, round dials, long dials, short dials, no dials, push buttons and automatic tuning.

These and many other innovations, together with the circuit developments previously described, have resulted in the sale of over 300,000,000 broadcast receivers at a factory price of about $8,000,000,000, in the United States to the end of 1960.[21]

[20] "Hazeltine Fremodyne FM circuit," *Tele-Tech.*, vol. 6, pp. 41, 85–86; December, 1947.

[21] Paul Stone, Ed., "Television Factbook," Triangle Publications, Inc., Radnor, Pa., Spring/Summer ed., p. 28; 1961.

A Half Century of Television Reception*

F. J. BINGLEY†, FELLOW, IRE

Summary—**The paper traces the development of television receiving, starting with a description of the earliest demonstrations of mechanical television. While the major emphasis is on reception, enough discussion of television system standards is included to support the development of the receiver story. The progression of events which led from the early mechanical systems, through the various phases of electronic television using cathode-ray picture tubes, up to the modern compatible color television receivers, is covered in some detail. The development of circuitry for IF, VF, and deflection is covered. A complete bibliography, containing some of the great historic as well as the most recent contributions, is included.**

Introduction

THE DEVELOPMENT of the art of television receiving is inextricably bound up with that of television transmission; there is a "lock-and-key" relationship between the transmitter and receiver. Thus, while the emphasis of this paper will properly be placed on receiving techniques, it will be necessary in many cases to supplement the discussion with sufficient explanation of the transmitter and over-all system philosophy.

The earliest development of television was directed towards generation of television signals, and their display as television pictures without the intervention of radio transmission and reception circuits. The earliest demonstrations of television used equipment comprising mechanical devices for scanning the image. It is fitting to begin with a discussion of them.

I. Early Mechanical TV History

So far as the author is aware, the first demonstrations of actual television images were made in the year 1925 by C. F. Jenkins in Washington [1] and by J. L. Baird in London. They worked independently of each other, but it appears not to be possible to state to whom should be given the credit of priority. The pictures shown by Baird were not transmitted and received by radio, but rather over a connecting wire. Similarly he avoided the synchronizing problem by using mechanical scanning with a common drive connecting the scanning devices at both "transmitter" and "receiver." Jenkins used radio transmission, and synchronous motors for synchronization.

By the fall of 1927, Baird had progressed to the point of making experimental broadcasts from a radio station near London. His scanning equipment included an 8 foot diameter lens disk carrying 30 lenses associated with a photocell. The monitor display was an 8 foot, Nipkow scanning disk with 30 apertures, scanning vertically. A neon lamp tube served as the light source. Synchronization at the studio end of the system was achieved by the solid shaft connecting the lens disk with the Nipkow disk. The broadcasts were received at White Plains, New York. History does not record how well the pictures were received, or how the synchronization was accomplished.

A great forward step was taken in 1927 by Bell Telephone Laboratories [2], who set up a system for transmitting television from Washington, D. C., to New York, a distance of 250 miles. The transmission could be carried alternatively by radio or by wire circuit. There were 50 scanning lines per frame and 16 frames per second. The receiver displays used were of two types; one produced a small image about $2 \times 2\frac{1}{2}$ inches, the other a large image $2 \times 2\frac{1}{2}$ ft. The smaller picture was produced by a scanning disk revolving in front of a special neon bulb having a flat cathode; the disk was driven by a synchronous motor fed and synchronized from a separately transmitted signal. The large display consisted of a set of 50 parallel tubes containing neon and a set of 2500 electrodes, arranged in a sequence of 50 along each neon tube. These electrodes were connected by 2500 wires to a 2500 segment commutator run by a synchronous motor similar to that used for the scanning disk. The video signal was used to modulate a 500 kc RF carrier, and this was distributed by the commutator which thus provided the scanning function. When the signal was received over wire facilities, the circuit was carefully corrected for amplitude and phase up to a maximum frequency of 20 kc. When it was received by radio, a special superheterodyne receiver was used for the TV signals (the transmitted carrier was 1575 kc); separate receivers were also used for speech (1450 kc) and synchronizing signals (185 kc).

Another early pioneer was E. F. W. Alexanderson who, working at General Electric, produced a projection display using a drum of mirrors associated with a modulated light source [3]. The image of the light source was formed at the screen after reflection by the mirrors of the drum; as the drum rotated the image was caused to move vertically on the screen. Because, in addition to their angular spacing around the drum, the mirrors were tilted progressively with respect to its axis of rotation, the images reflected by successive mirrors were displaced sideways on the screen as each successive mirror came into play. The result was that the receiving screen was scanned. In order to produce more scanning lines from a drum having a given number of mirrors (24 in the early experiments), he used seven light sources so that in effect each mirror scanned seven lines in parallel.

* Received by the IRE, September 21, 1961; revised manuscript received, November 16, 1961.
† Government and Industrial Division, Philco Corporation, Philadelphia, Pa.

Thus the total effective number of lines was 168. In later experiments the multispot scanning was abandoned and a simple sequential scanning of 48 lines at 15 frames per second was used; the picture was displayed on a screen about 3 ft square.

II. Early Electronic Television Receivers

The arrival of electronic television may be dated at approximately the year 1930. Much prior to this year Dieckmann in Germany (1906), Rosing in Russia (1907) and Campbell Swinton in England (1911) [3]–[5] discussed and described electronic television receivers employing cathode-ray tubes as display devices. Dieckmann actually had a crude unit in operation in 1906; however his system transmitted only silhouettes, so there is some reluctance to give him full credit for having produced the first electronic receiver. No further work on electronic television appears to have been done until the early 1920's, when the work of Philo Farnsworth and Dr. V. K. Zworykin became known. The importance of their work was not recognized until 1930 however, perhaps because some of the mechanical approaches to television were beginning to show real promise. It was only when the overriding limitations of mechanical TV were finally, if reluctantly, admitted that electronic television received the research and development effort it merited.

In about the year 1929 Farnsworth demonstrated an all electronic TV system in which the display device was a cathode-ray tube, magnetically deflected and focused [6]–[8]. In 1931 Farnsworth joined Philco Corporation, where he continued his work for about two years during which time early experiments in broadcast TV reception on frequencies in the order of 60 Mc were made. Transmissions from W3XE (the Philco experimental station) were received in Mt. Airy, Philadelphia, a distance of about 6 miles, in 1932.

The work of Dr. V. K. Zworykin was originally carried out at Westinghouse Research Laboratories. He is reported to have demonstrated an all-electronic system in 1924. The earliest published account of his experiments appeared in *Radio Engineering* in December, 1929 [9]. The display device at the receiver was a 7-inch cathode-ray tube with ES-EM deflection, and ES focusing. For beam modulation a control grid was used. When Dr. Zworykin joined RCA in 1930, this work was continued by him and his associates.

III. Television Standards

By about 1935 electronic television had progressed to the point where it became obvious that the most pressing requirement was for a set of standards, in order that the development of systems could proceed in an orderly fashion. By that year there already existed experimental transmitters in the VHF band, camera tubes such as iconoscope and dissector, and cathode-ray display tubes. The necessary ancillary circuits, such as for deflection and video amplification, were already well known. It was obvious that vitally important factors to be decided were those of how many scanning lines should be used, whether they should be interlaced, and what field and frame rates were required. These were basic factors, but they reflected their influence widely throughout the whole system design, affecting video bandwidth (and in turn transmitter and receiver bandwidth), picture flicker, and deflection circuit frequencies. Along a parallel track lay the problems concerned with transmission and reception of the video signal and the accompanying sound. Items needing consideration here were method of modulation, polarity of transmission, polarization, the applicability of single or double sideband or other band saving means, the channel layout (spacing and relation between video and sound carrier and their locations in the channel), allocation of broadcast frequencies, and many others. Standardization of these items was achieved in the fall of 1941, and the FCC issued commercial television standards at that time. Some of the more important standards will now be discussed.

Amplitude modulation was finally selected for picture transmission. Partial suppression of the lower sideband was used to save bandwidth. Frequency modulation for picture transmission was tested early in 1941, but found to produce bad effects in the receiver in the presence of multipath; it was abandoned for broadcast use for this reason. The polarity of modulation was selected as negative (that is, more light at the scene, less carrier amplitude). On the other hand, frequency modulation was selected for sound transmission, since a sufficiently high deviation ratio could be provided to take full advantage of the FM system.

The spacing between sound and picture carriers was standardized at 4.5 Mc, with the sound carrier 0.25 Mc below the upper edge of the channel, and the total channel width 6 Mc.

It was realized quite early (1933) that the transmission of 60 fields per second, interlaced two to one, was desirable from the standpoint of elimination of flicker. The matter was briefly reexamined in 1941, but the original decision was reaffirmed. The number of scanning lines progressed from 343 in 1935, to 441 in 1937, and was finally standardized at 525 in 1941. The video bandwidth increased correspondingly from about 2 Mc in 1933 to 4.25 Mc in 1941. The horizontal deflection frequencies increased from 10.29 kc to 15.75 kc during this period.

The work of standardization was carried out by a number of industry groups. The final coordination was done by the National Television System Committee. Their work has been described in detail by Fink [10].

IV. Development of Circuits for Electronic Television Receivers

The development of television standards briefly referred to above was concurrent with a progressive development of television receiving techniques. During the period up to about 1945, superheterodyne receivers were developed capable of handling the necessary bandwidth of 6 Mc [11]–[14] and video amplifiers capable of amplifying adequate bandwidth with acceptable frequency and phase characteristics were developed and design formulas were obtained [15]–[19]. Basic factors having to do with receiver and system optimization were studied during this period, and papers were written by Kell, Bedford and Fredendall [20], Wilson [21], Wheeler [22], [23], Loughren [22] and Baldwin [24].

The design of intermediate frequency amplifiers was considered by many workers in the field, including Butterworth [11], Landon [13], Sziklai and Schroeder [14] and more recently by Avins [25], Ruth [26], Bridges [27], and Bradley [77].

Prior to 1945 the accepted method of sound reception had been to use a separate IF amplifier for this purpose. It required great stability of the local oscillator in order to maintain the sound within the narrow sound IF channel. A new method, utilizing the 4.5-Mc beat between sound and picture carriers, was developed and is now used exclusively [28]. Since the 4.5 Mc beat frequency is determined at the transmitter, receiver tuning is now not critical.

Synchronizing and deflection circuits and yokes have shown progressive development. In the 1930's deflection angles of 50° were considered large. At the present time the figure has risen to 110°, with corresponding decrease in cabinet depth [29]–[33]. This has been accomplished by increased efficiency in circuits and yokes. At the same time synchronizing circuits have been improved, notably by the introduction of AFC synchronization.

Transistors are now being used in television receivers, particularly in portable models and in IF amplifiers. Much work has been reported in this field [34]–[44].

Considerable development has been applied to picture tubes since 1930. This development has covered screens, guns and geometry. Prior to about 1935 engineers had been satisfied to do what came naturally with respect to phosphors—namely, to use the readily obtainable green fluorescing willemite as a screen phosphor on the receiver display. However, about that year it became obvious that such a color would be unacceptable for monochrome TV receivers, and development of efficient white phosphors began. By 1938 the present two component white-balanced screens began to appear [45]–[47].

In about the year 1935 it became recognized that the central regions of picture tube screens were exhibiting dead areas in the phosphor—that is to say, the picture contained a dark spot in the center. It was established that this was due to bombardment of the screen by negative ions, presumably emitted from the heated cathode. The ionic mass was so great, compared to that of the electrons in the beam, that the ions were practically undeflected by the magnetic deflection fields. They could be focused by an electrostatic field, but not by a magnetic field. Magnetically focused tubes showed a sharp shadow of the tube neck as a boundary of the ion blemish. Electrostatically focused tubes showed a smaller more intense black spot due to the electrostatic focusing action. Research using the picture tube as a mass spectrograph identified the various ions involved.

By devising an electron gun having a bent section (the axis of the bend being normal to the tube axis) the ions became trapped in the gun. By adding an external auxiliary magnetic field the electron beam in the gun was bent so it could pass down the axis of the second section of the gun, which was coaxial with the tube neck. However, the ions, which were not deflected by the auxiliary magnetic field, remained trapped in the gun and thus could not reach the screen to damage it. The development of the aluminized tube provided protection for the phosphor from ion bombardment, and ion traps are not required with such tubes.

Phosphor screens emit light on both sides. That which is emitted on the side facing the electron gun is wasted so far as the observer is concerned, if it is permitted to be absorbed and reflected by the inside surface of the picture tube. More than this, it destroys the over-all contrast of the picture. The operation of the aluminized tube depends on the fact that an electron beam of velocity such as used in a picture tube can penetrate a film of aluminum which is at the same time thick enough to reflect essentially all the incident light. Thus, light emitted from the rear surface of the phosphor screen is reflected and appears as a useful image illumination to the viewer. It is prevented from destroying image contrast. The aluminum film also protects the phosphor screen from injury due to ion bombardment [48].

Deflection of the cathode-ray beam may be accomplished either by electrostatic or electromagnetic fields; and the same is true of focusing. During the period 1930 to 1950, ES and EM focusing were used, with the general predominance on EM focusing since it was easier to get fine bright spots and the focusing element was outside the tube, giving a freedom to adjust and experiment. In the period 1950–1955 the disadvantage of the current drain of the EM focuser was overcome by designing permanent magnet focusers with adjustable airgaps. Since 1955, with the increase in knowledge of electron optics, it has been possible to design good picture tubes with electrostatic focus, and these are now widely used in television receivers. Regarding deflection, some of the early experimenters used crossed ES-EM systems, using the ES for the high-frequency deflection. This was probably done because of the lack of knowledge at that time concerning the design of magnetic deflection circuits operating at high frequency. By about 1935 magnetic deflection circuitry was well understood and has been used almost exclusively since then. There are new developments in ES deflection cathode-ray tube which may affect the present monopoly of EM deflection [49]–[53].

V. Color Television Receivers

A. Early Devlopment

Within the period of this review, the earliest color television work appears to have been carried out in England by Baird in about June 1928. This was an all mechanical field sequential receiver display, with 15 red lines and 15 cyan lines. Naturally the results were quite crude, and only demonstrated principles [54].

The next demonstration of which the author is aware was one given by Bell Laboratories some time in June 1929, at their Laboratories in West Street, New York [55]. The system was devised by Dr. Ives, the well-known expert in optics and colorimetry. Three separate channels were used, one each for red, green and blue. The receiver display had 50 lines and 18 pictures per second, using a 20-kc band for each channel.

What may be called semi-electronic color television was first broadcast by Dr. P. C. Goldmark of CBS in 1940 [56]. The system used a field sequential receiver display, with a disk having successive red, green and blue filters revolving in front of a cathode-ray picture tube. The disk was synchronized so that red, green and blue separation images were displayed in succession. This system received much attention, particularly in the post-war years. It suffered from the compromise between resolution and flicker, and from its noncompatibility with the existing monochrome service. Since it was a sequential system, there was no way in which the bandwidths required to satisfactorily transmit red, green and blue separation images could be tailored to fit the lesser demands of red and blue as compared to those of green. It also suffered from the disadvantage of not being entirely electronic, requiring (at least at the time it was under consideration) a synchronized filter disk. In spite of these disadvantages the system was adopted as U. S. standard by FCC [57] in October 1950, the scanning specifications being 405 lines, 48 color fields, interlaced, leading to a flicker rate of 48 cps. In order to partially offset the lack of detail, ingenious "crispening" circuits were devised [58] and applied to the receiver. It was demonstrated that a receiver could be built with changeover switches to show either monochrome or color pictures, thus rendering the lack of compatibility somewhat less objectionable. The standard was rescinded in December 1953.

Most of the development of color television receivers since 1950 has centered about the development of the display device. The progress will therefore be discussed mainly in terms of the display.

B. Line Sequential Color Display

An early approach to color display was that called "line sequential," in which successive lines were scanned in red, green and blue. This was first demonstrated in 1949; but, since the display involved a triple interlace, objectionable traveling patterns intruded upon the picture. No solution being found, the system was abandoned.

C. The Chromatron Tube

This tube, developed by Dr. Lawrence, is characterized by having a screen structure in which the phosphors are laid down in horizontal stripes [59]–[61]. The phosphor lines are of three separate kinds of phosphor, fluorescing in red, green, and blue, respectively. Close to the screen is a set of deflection plates running parallel to the phosphor lines. The plates cause the electron beam to be selectively deflected according to the potential between them. In this fashion the scanned raster can be caused to fall only on the green phosphor, or only on the red or blue phosphor. The selection is made at the instants when the video signal amplitude is appropriate to control the luminance of selected phosphor. This is accomplished by signal processing at the receiver.

D. The Shadow Mask Tube

This tube is of the three gun type—that is to say, separate electron guns are assigned to exciting the red, green and blue phosphors, respectively [62]–[65]. This is achieved by the parallax existing between the three guns due to their slightly differing space positions, as they beam electrons through an apertured plate on to a phosphor screen placed parallel to it. The plate has about 300,000 holes and is of the same size as the phosphor screen. The phosphor screen has 900,000 phosphor dots, arranged in groups of three, each group consisting of a red dot, a green dot, and a blue dot (the color named refers to the fluorescent radiation). Each group of three is arranged at the corners of a very small equilateral triangle of such a size and so oriented that the electrons from one gun (call it the red gun) can only fall on the red dots; and similarly for the green gun and green dots, and the blue gun and blue dots. This relation must hold over the whole screen no matter where the three beams are deflected by a deflection yoke through which they all pass. Naturally the size of phosphor dots must be carefully controlled, as must the relation between the dots and the aperture mask, which requires great precision of alignment.

The signal processing circuits effectively feed red, green and blue separation signals between grid and cathode of each respective gun.

E. The Apple Receiving System

The name "Apple" was applied by Philco to a specific type of single gun color tube developed in its Research Laboratories [66]–[70]. Considerable signal processing is involved in utilizing the tube. Basically the tube consists of a striped phosphor screen, with the stripes running vertically. The stripes fluoresce in the additive primary colors red, green, and blue. They are arranged in the repetitive order red, green, blue, such a triplet of stripes occupying a width of about $\frac{1}{16}$ inch on a 21 inch picture tube, and there being some 320 triplets across the face of the tube. Adjacent phosphor strips are separated by an inert black material. The rear of the aluminized surface is imprinted with a set of index stripes in register with the red phosphor stripes. The index stripes provide a beam position reporting signal, which is used in the signal processing circuits to cause the video signal applied to the control grid of the gun to assume values representative of the required luminance of a given phosphor at the moment the scanning beam crosses it.

A major advantage of the Apple system is that the screen white balance is automatically determined by the screen structure itself. Absence of color carrier (whether caused by a white section of a color picture being transmitted, or by the fact that the transmitted picture is monochrome) causes the screen to fluoresce in its characteristic white. The quality of this white depends only upon the balance achieved in laying down the phosphor screen during manufacture.

VI. Miscellaneous Approaches to Television Reception

A. *Supersonic Light Valve*

The most sophisticated embodiment of mechanical scanning was the system employing the supersonic light valve described by J. H. Jeffrees and developed as a receiving system by Scophony Ltd. of England [71], [72]. The light valve consisted of a liquid cell about 2 cm wide by about 4.5 cm long; at one end was placed a piezoelectric crystal having a natural frequency of about 20 Mc, and at the other an acoustic absorber. The crystal was highly damped (mechanically speaking) by the liquid, so that if excited by a video modulated 10-Mc signal, its amplitude of mechanical vibration would accurately follow the modulation envelope. The crystal induced a traveling compression wave at a frequency of 10 Mc in the liquid; the wave traveled down the length of the cell to the absorbing barrier at the far end with a velocity of 10^4 cm/sec, and a corresponding wavelength of 10^{-2} cm. As the wave was formed at the sending end, it bore the instantaneous amplitude of the video signal, which it preserved throughout its journey to the end of the cell. The compression waves caused the liquid to act as a diffraction grating to light passing transversely through the cell. The energy in the zeroth order was blocked out, and that in the remaining orders was permitted to pass, by means of an optical system including a slit focused upon a complementary barrier. The greater the amplitude of the video signal modulating the carrier applied to the crystal, the greater the intensity of the compressional wave at the corresponding instant of generation; and the greater the light passed by the light valve. It took about 45 μsec for the modulated wave to travel the length of the cell.

An optical system was arranged to focus an image of the cell upon the receiver screen, with the long axis of the image (corresponding to the direction of travel of the compressional wave) horizontal. A high-speed mirror drum scanner suitably synchronized was used to immobilize the image of the compressional wave on the screen; and the light coming through the immobilized wave structure was, at any point, proportional to the corresponding video signal amplitude. Thus, the scanning lines of the image were reproduced in corresponding detail at the receiver screen. A second synchronized mirror drum was used to provide the vertical scanning component.

The system employed a 20-facet mirror drum running at 30,000 rpm for horizontal immobilization. Trouble was experienced with electronic sync generators, since their rate of change of frequency was greater than could be followed by the immobilizer. This trouble would probably not exist today when crystal masters are used for sync generators instead of 60-cycle power mains.

Using the above principles, a home television receiver giving a highlight brightness of up to 10 ft lamberts, and theater projection equipment giving 3 ft lamberts on a 12×15 foot screen, were designed.

B. *Skiatron Display*

The skiatron depends upon the discoloration of certain metallic salts (such as potassium chloride) by bombardment with cathode rays. It has been proposed to use this effect by causing a transparency to appear on the face of the cathode-ray tube coated with such material instead of the usual phosphor, but with a scanned and modulated beam otherwise exactly like a normal picture tube. The transparency is then projected or viewed with a light source.

Such a device was used during the war for a bright radar display. For television use it has, at the moment, the disadvantage that the effect decays slowly so that it is difficult to depict motion.

C. *Eidophor Display*

This is a unit primarily designed for theater projection. Its basic principle is the disturbance of the surface of a thin film of liquid brought about by electric charge. The electric charge is deposited by a cathode-ray beam scanned and modulated in the normal television fashion. The beam is focused on to the surface of the liquid, and causes a charge pattern representation of the TV picture to appear. The liquid film is formed on the surface of a mirror, and the reflection, resulting from its surface irregularities, causes the direction of the rays reflected by the mirror to be affected accordingly. A Schlieren type of optical system correspondingly varies the amount of emergent light. The surface of the liquid is focused on the screen by a projection lens. The unit was developed by Dr. Fischer at the Federal Institute of Technology in Zurich [78]. It has been adapted to field sequential color television.

Conclusion

The story of the development of television receiver techniques is a long one, characterized by contributions to its progress made by many scientists and engineers, Only the contributions made by the very earliest workers occurred long enough ago that they have acquired historic perspective so that just credit can be given them.

An appreciation of the amount of work and of the number of participating engineers and scientists involved in the recent development of color standards may be had by consulting some of the more recent texts [73]–[76]. The author does not presume to have the ability properly to weigh and apportion the credit for the more recent work.

References

[1] C. F. Jenkins, "Radio Vision," Proc. IRE, vol. 15, pp. 958–964; November, 1927.
[2] H. E. Ives, "Television," *Bell Sys. Tech. J.*, vol. 6, pp. 551–559; October, 1927.
[3] H. H. Sheldon and E. N. Grisewold, "Television," D. Van Nostrand Co., Inc., New York, N. Y.; 1929.
[4] J. D. McGee, "Distant electric vision," Proc. IRE, vol. 38, pp. 596–608; June, 1950.
[5] A. G. Jensen, "The evolution of modern television," *J. SMPTE*, vol. 63, pp. 174–188; November, 1954.
[6] P. T. Farnsworth, U. S. Patent No. 1,773,980; August 26, 1930.
[7] P. T. Farnsworth, "Television by electron image scanning," *J. Franklin Inst.*, vol. 218, pp. 411–444; October, 1934.
[8] A. Dinsdale, "Television by cathode ray—the new Farnsworth system," *Wireless World*, vol. 28, pp. 286–288; March, 1931.
[9] V. K. Zworykin, "Television with cathode ray tube for receivers," *Radio Engrg.*, vol. 9, pp. 38–40; December, 1929.
[10] D. G. Fink, "Television Standards and Practice—NTSC," McGraw-Hill Book Co., Inc., New York, N. Y.; 1943.
[11] S. Butterworth, "On the theory of filter amplifiers," *Exptl. Wireless and Wireless Engr.* (London), vol. 7, pp. 536–540; October, 1930.
[12] E. W. Herold, "Superheterodyne converter system considerations in television receivers," *RCA Rev.*, vol. 4, pp. 324–337; January, 1940.
[13] V. D. Landon, "Cascade amplifiers with maximal flatness," *RCA Rev.*, vol. 5, pp. 347–362, 481–497; January, April, 1941.
[14] G. C. Sziklai and A. C. Schroeder, "Band pass bridged T network for television intermediate frequency amplifiers," Proc. IRE, vol. 33, pp. 709–711; October, 1945.
[15] V. D. Landon, "The band-pass—low-pass analogy," Proc. IRE, vol. 24, pp. 1582–1584; December, 1936.
[16] S. W. Seeley and N. Kimball, "Analysis and design of video amplifiers," *RCA Rev.*, vol. 3, pp. 290–308; January, 1939.
[17] H. A. Wheeler, "Wideband amplifiers for television," Proc. IRE, vol. 27, pp. 429–438; July, 1939.
[18] H. L. Donley and D. W. Epstein, "Low frequency characteristics of the coupling circuits of single and multi stage video amplifiers," *RCA Rev.*, vol. 6, pp. 416–433; April, 1942.
[19] H. E. Kallman, R. E. Spencer, and C. P. Singer, "Transient response," Proc. IRE, vol. 33, pp. 169–195; March, 1945.
[20] R. D. Kell, A. V. Bedford, and G. L. Fredendall, "Determination of optimum number of lines in a television system," *RCA Rev.*, vol. 5, pp. 8–30; July, 1940.
[21] J. C. Wilson, "Channel width and resolving power in television systems," *J. Television Soc.* (London), vol. 2, pt. 2, pp. 397–420; 1938.
[22] H. A. Wheeler and A. V. Loughren, "The fine structure of television images," Proc. IRE, vol. 26, pp. 540–575; May, 1938.
[23] H. A. Wheeler, "Interpretation of amplitude and phase distortion in terms of paired echoes," Proc. IRE, vol. 27, pp. 359–385; June, 1939.
[24] M. W. Baldwin, "Subjective sharpness of simulated television images," Proc. IRE, vol. 28, pp. 458–468; October, 1940.
[25] J. Avins, "IF amplifier design for color TV receivers," IRE Trans. on Broadcast and Television Receivers, vol. PGBTR-7, pp. 14–25; July, 1954.
[26] L. Ruth, "A printed circuit IF amplifier for color TV," IRE Trans. on Broadcast and Television Receivers, vol. BTR-2, pp. 50–52; July, 1956.
[27] J. E. Bridges, "Detection of television signals in thermal noise," Proc. IRE, vol. 42, pp. 1396–1405; September, 1936.
[28] R. B. Dome, "Carrier difference reception of television sound," *Electronics*, vol. 20, pp. 102–105; January, 1947.
[29] C. E. Torsch, "High efficiency, low-copper sweep yokes with balanced transient response," IRE Trans. on Broadcast and Television Receivers, vol. PGBTR-6, pp. 17–32; April, 1954.
[30] R. A. Bloomsburg, "The measurement of yoke astigmatism," IRE Trans. on Broadcast and Television Receivers, vol. PGBTR-7, pp. 26–33; July, 1954.
[31] C. G. Torsch, "Yoke development for standardization of 110° and 90° deflection angles," IRE Trans. on Broadcast and Television Receivers, vol. BTR-1, pp. 10–15; October, 1955.
[32] R. Gethmann, "Deflection distortions contributed by the principal field of a ring deflection yoke," IRE Trans. on Broadcast and Television Receivers, vol. BTR-4, pp. 24–34; February, 1958.
[33] C. Droppa, "Improvements in deflection amplifier design," 1958 IRE National Convention Record, pt. 7, pp. 147–153.
[34] W. Palmer and G. Scheiss, "Transistorized vertical deflection system," IRE Trans. on Broadcast and Television Receivers, vol. BTR-3, pp. 98–105; October, 1957.
[35] E. M. Creamer, Jr., L. H. DeZube, and J. P. McAllister, "Transistor circuit problems in TV receiver design," 1957 IRE National Convention Record, pt. 3, pp. 205–212.
[36] A. R. Curll, "A transistorized portable television receiver," IRE Trans. on Broadcast and Television Receivers, vol. BTR-6, pp. 9–16; May, 1960.
[37] J. G. Humphrey, "A transistor TV IF amplifier," IRE Trans. on Broadcast and Television Receivers, vol. BTR-6, pp. 17–20; May, 1960.
[38] W. H. Mead, "Performance of Nuvistor small-signal tetrodes in television video IF amplifiers," IRE Trans. on Broadcast and Television Receivers, vol. BTR-6, p. 50; May, 1960.
[39] C. D. Simmons and C. R. Gray, "The video processing circuits of an all transistor television receiver," IRE Trans. on Broadcast and Television Receivers, vol. BTR-6, pp. 25–32; May, 1960.
[40] F. L. Abboud, "Transistorized vertical scan system for magnetic deflection," IRE Trans. on Broadcast and Television Receivers, vol. BTR-6, pp. 33–38; May, 1960.
[41] R. B. Ashley, "Linearization of transistorized deflection system," IRE Trans. on Broadcast and Television Receivers, vol. BTR-6, pp. 39–47; May, 1960.
[42] Z. Wiencek, "Phase splitter video amplifier for transistor TV," IRE Trans. on Broadcast and Television Receivers, vol. BTR-6, pp. 18–24; November, 1960.
[43] G. C. Hermeling, "A Nuvistor low noise VHF tuner," IRE Trans. on Broadcast and Television Receivers, vol. BTR-6, pp. 52–56; November, 1960.
[44] R. W. Ahrons, "Factors in the circuit operation of transistorized horizontal deflection," IRE Trans. on Broadcast and Television Receivers, vol. BTR-6, pp. 57–59; November, 1960.
[45] R. R. Law, "Contrast in kinescopes," Proc. IRE, vol. 27, pp. 511–524; August, 1939.
[46] H. W. Leverenz, "Optimum efficiency conditions for white luminescent screens in kinescopes," *J. Opt. Soc. Am.*, vol. 30, pp. 309–315; July, 1940.
[47] M. Sadowsky, "The double layer projection tube screen for television," Proc. IRE, vol. 38, pp. 494–498; May, 1950.
[48] D. W. Epstein and L. Pensak, "Improved cathode-ray tubes with metal backed luminescent screens," *RCA Rev.*, vol. 7, pp. 5–10; March, 1946.
[49] R. R. Law, "Factors governing the performance of electron guns in television cathode-ray tubes," Proc. IRE, vol. 30, pp. 103–105; February, 1942.
[50] G. Liebmann, "The image formation in cathode-ray tubes and the relation of fluorescent spot size and final anode voltage," Proc. IRE, vol. 33, pp. 381–389; June, 1945.
[51] J. E. Rosenthal, "Correction of deflection defocusing in cathode-ray tubes," Proc. IRE, vol. 39, pp. 10–15; January, 1951.
[52] L. C. Wimpee and L. E. Swedlund, "The development of 110° television picture tubes having an ion trap gun," 1957 IRE National Convention Record, pt. 3, pp. 150–155.
[53] J. W. Schwartz, "The annular geometry electron gun," Proc. IRE, vol. 48, pp. 1864–1870; November, 1958.
[54] A. Russell, "Book Review on 'Practical Television' by E. T. Larner," *Nature*, vol. 122, pp. 232–234; August, 1928.
[55] H. E. Ives and A. L. Johnsrud, "Television in colors by a beam scanning method," *J. Opt. Soc. Am.*, vol. 20, pp. 11–22; January, 1930.
[56] P. C. Goldmark, J. N. Dyer, E. R. Piore, and J. M. Hollywood, "Color television," Proc. IRE, vol. 30, pp. 162–182; April, 1942.
[57] P. C. Goldmark, J. W. Christensen, and J. J. Reeves, "Color television—U.S.A. standard," Proc. IRE, vol. 39, pp. 1288–1313; October, 1951.
[58] P. C. Goldmark and J. M. Hollywood, "A new technique for improving the sharpness of television pictures," Proc. IRE, vol. 39, pp. 1314–1322; October, 1951.
[59] R. Dressler, "The PDF chromatron," Proc. IRE, vol. 41, pp. 851–858; July, 1953.
[60] J. D. Gow and R. Dorr, "Compatible color picture presentation with the single gun tricolor chromatron," Proc. IRE, vol. 42, pp. 308–315; January, 1954.
[61] R. Dressler and P. Neuworth, "Brightness enhancement techniques for the single gun chromatron," 1957 IRE National Convention Record, pt. 3, pp. 220–229.

[62] R. R. Law, "A three gun shadow mask color kinescope," Proc. IRE, vol. 39, pp. 1186–1194; October, 1951.
[63] M. J. Grimes, A. C. Grum, and J. F. Wilhelm, "Improvements in the RCA three-beam shadow mask color kinescope," Proc. IRE, vol. 42, pp. 315–326; January, 1954.
[64] N. F. Fyler, W. F. Rowe, and C. W. Cain, "The CBS-colortron: A color picture tube of advanced design," Proc. IRE, vol. 42, pp. 326–334; January, 1954.
[65] M. E. Amdursky, R. C. Pohl, and C. S. Szegho, "New high-efficiency parallax mask color tube," Proc. IRE, vol. 43, pp. 936–943; August, 1955.
[66] C. F. Barnett, F. J. Bingley, S. L. Parsons, C. W. Pratt, and M. Sadowsky, "A beam indexing color picture tube—the Apple tube," Proc. IRE, vol. 44, pp. 1115–1119; September, 1956.
[67] R. G. Clapp, E. M. Greamer, S. W. Moulton, M. E. Partin, and J. S. Bryan, "A new beam indexing color television display system," Proc. IRE, vol. 44, pp. 1108–1114; September, 1956.
[68] R. A. Bloomsburgh, W. P. Boothroyd, G. A. Fedde, and R. C. Moore, "Current status of Apple receiver circuits and components," Proc. IRE, vol. 44, pp. 1120–1124; September, 1956.
[69] R. A. Bloomsburgh, A. Hopengarten, R. C. Moore, and H. H. Wilson, Jr., "An advanced color television receiver using a beam indexing picture tube," 1957 IRE National Convention Record, pt. 3, pp. 243–249.
[70] H. Colgate, C. Comeau, D. Kelley, D. Payne, and S. Moulton, "Recent improvements in the Apple beam indexing color tube," 1957 National Convention Record, pt. 3, pp. 238–242.
[71] D. M. Robinson, "The supersonic light control and its application to television with special reference to the Scophony television receiver," Proc. IRE, vol. 27, pp. 483–486; August, 1939.
[72] J. Sieger, "The design and development of television receivers using the Scophony optical scanning system," Proc. IRE, vol. 27, pp. 487–492; August, 1939.
[73] D. G. Fink, Ed., "Color Television Standards," McGraw-Hill Book Co., Inc., New York, N. Y.; 1955.
[74] J. W. Wentworth, "Color Television Engineering," McGraw-Hill Book Co., Inc., New York, N. Y.; 1955.
[75] K. McIlwain, and C. E. Dean, Eds., "Principles of Color Television," John Wiley and Sons, Inc., New York, N. Y.; 1956.
[76] D. G. Fink, Ed., "Television Engineering Handbook," McGraw-Hill Book Co., Inc., New York, N. Y.; 1957.
[77] W. E. Bradley, "Wideband amplification," in "Television Engineering Handbook, McGraw-Hill Book Co., Inc., New York, N. Y., ch. 12; 1957.
[78] E. Baumann, "The Fischer large screen projecting system," *J. SMPTE*, vol. 60, pp. 344–356; April, 1953.

The Impact of Receiving Tubes on Broadcast and TV Receivers*

E. W. HEROLD†, FELLOW, IRE

Summary—**The receiving electron tube has been primarily responsible for the modern superheterodyne, with its high sensitivity, high selectivity, automatic gain control, and ease of tuning. In examining the history, one finds an early period of triodes and diodes (1907 to 1927), a second period (1927 to 1936) of indirectly heated cathodes and multigrid tubes, and a third period (1936 to 1960) of close-spaced tubes and VHF operation. The most significant tube concepts are the triode, the multigrid tube, and the indirectly heated cathode, all of which started in the first period. The second and third periods were marked by tremendous advances in the technology of production, and in extensive application of the early inventions to new receiver designs. At present, solid-state devices are gradually supplanting vacuum tubes in some receiver applications, and this trend is expected to continue.**

Introduction

The history of the receiving electron tube, and the history of the Institute of Radio Engineers are almost coincident in time. It was only eight years after the invention of the Fleming valve, and six years after the invention of the DeForest triode, that the Institute was founded in 1912. In the present survey, the highlights of tube development, as they affected broadcast and television reception, will be outlined. The five decades which have elapsed saw the early and rapid rise, a peak period extending through the introduction of television, and a saturation point with a gradual decline, as solid-state devices became more and more prominent. For convenience, the discussion will be largely chronological and is divided into three historical periods.[1] A concluding section will give a brief view of the present period, and of future trends.

The first historical period was from roughly 1907 to 1927, when only filament-type triodes were widely used. This period included the invention and application to receivers of cascade amplification, regeneration and oscillation, heterodyning, and the superheterodyne. All these basic circuit principles are still with us today, and all were made possible by the electron tube. Some other developments of the period, superregeneration and reflex amplification have never been widely used in broadcast or television reception.

The second major period is the shortest in time, and is possibly the most outstanding, so far as rate of receiving-tube development goes. It started about 1927 with the delveopment of indirectly heated cathode tubes (and a few years later of multigrid tubes), and ended about 1936 with the octal base metal tube and beam power tube. The tube developments of this period were pri-

* Received by the IRE, December 13, 1961.
† Varian Associates, Palo Alto, Calif.

[1] The author wishes to apologize in advance for the many important and excellent contributions, particularly those of foreign origin, which he has been forced to omit in a discussion of this length.

marily responsible for the application in receivers of ac operation, high-level detection, very selective and single-knob tuning, the built-in loudspeaker, automatic-gain-control, tuning indicators, and multiband operation This was also the period in which the superheterodyne became the universal receiving circuit. Once again, every one of these receiver improvements remains with us thirty years later.

The third major period, from 1936 to about 1960, started with close-spaced (high gain-bandwidth) tubes and miniature glass envelope types, ending with the color picture tube and tubes for hybrid transistor-tube combination receivers. This is the period of VHF, *i.e.*, FM and television; the tube developments made these possible, together with their special circuits such as limiters, discriminators, automatic-frequency-control, saw-tooth deflection, synchronization, inverse feedback, wide-band amplification, and multiplexing.

It is unlikely that there will again be a period in which the device we know as a receiving tube will have as revolutionary an impact as in the three periods summarized above. The present and future of broadcast and television receivers appears to depend more and more strongly on solid-state methods, micro-electronics and other newer technologies. There will always be some exceptions, possibly in color picture tubes, or in tubes for higher-frequency operation of broadcast services.

The Years of the Diode and Triode, 1907–1927

The Fleming diode [1] of 1904, and the DeForest triode [2], invented in 1906, were used in radio receiver circuits shortly thereafter. Prior to 1912, the triode (Audion) was used exclusively as a detector, and showed marked superiority over other radio detection methods. The first Audions were highly nonuniform, and had a very poor vacuum. In fact, DeForest as late as 1915 said he could not predict whether the anode current would go up or go down when a signal was impressed; each tube was different [3]. However, starting in 1912, both General Electric Company and Western Electric Company scientists learned to make well evacuated and uniform tubes [4]. By 1915 Langmuir described high-vacuum tungsten-filament diodes and triodes [5], and the Western Electric telephone repeater tubes with oxide-coated filaments were already in use in cascade audio amplifiers.

These early triodes first made practical the cascade amplifier [6], the triode oscillator [7] and the regenerative feedback circuit [8]. In turn, the triode oscillator made possible application of the very important but unused principle of heterodyne reception, invented by Fessenden years before [9] when only arc generators were available. During World War I, both the General Electric and the Western Electric Companies manufactured substantial numbers of receiving tubes, of which the Western Electric VT-1 tube was perhaps best known.

The stimulus of the war led to the invention of the single most important receiver circuit to this day, the superheterodyne [10], [11]. Triode RF amplifiers had been rather unsuccessful except at very low frequencies. Major Armstrong tried to devise a more sensitive receiver, to detect enemy aircraft by their short-wave ignition noise. It then occurred to him [10] to use the heterodyne principle to lower the incoming frequency to a value which could readily be amplified, but which was still high enough to use selective circuits, at what we now call the intermediate frequency. In Germany, the same invention was made, and at about the same time [11]. Although it is obvious that the triode tube, used as an oscillator, as a heterodyne mixer, and as an amplifier, was essential to the superheterodyne, another important impact of the triode was to make possible multiple tuned circuits [12], separated by nonreciprocal active elements.[2]

World War I also stimulated the invention of the second most important electron-tube principle, the use of another grid interposed between control grid and anode, as made by Schottky [13]. Schottky saw the significance of the reverse-phase ac swing of the anode in reducing the cathode current excursion, and he devised the screen grid, at a fixed positive potential, to screen the cathode region from anode fluctuations. At the same time, he also saw that a fixed-potential positive grid between cathode and control grid (space-charge grid) could be used to greatly increase the tube current, if low voltages only were available [13a]. Two- and three-grid tubes were made in many laboratories, and went into modest production in Europe, but the tremendous advantages of the screen-grid principle for both RF amplification and for increased audio output were not widely appreciated at first. About 1926 Hull and Williams described their highly shielded two-grid RF amplifier tube [14] and research laboratories also began to experiment with pentode (*i.e.*, 3-grid) tubes for audio use, and the suppressor grid was invented [15]. A few years later, as we shall see, the screen-grid principle was used in almost every new tube type. However, in this first period of our history, tube manufacturing techniques were only just about reaching the point at which a multigrid tube could be considered; receiver designers had to make do with the triode.

An important help to the receiver designer who had to use triodes for RF amplification was the analysis by Miller [16] and others, of the feedback due to grid-anode capacitance. One proposed solution was the use of a neutralizing or cancelling type of feedback [17], a principle which was later extensively used in the Hazeltine Neutrodyne circuit [18]. Another invention of a completely different kind, was thought to eliminate entirely the need for extensive use of RF cascade amplification,

[2] Only after the perfection of the screen-grid tube was almost complete unilateral amplification achieved.

namely superregeneration [19]. However, many triode RF amplifiers used simple resistors in the grid circuit to prevent oscillation; in superheterodynes, very low IF was used. In all RF cascade triode amplifiers, stage gain was quite low, perhaps 10 to 15 db.

Up to 1920, there were no receivers in the hands of the general public, but by 1923 the broadcast receiver business was flourishing [20]. Among the available tubes were the type 200, a "soft" detector tube, and types 201 and 202 which were high-vacuum amplifier types, all with tungsten filaments. Because electron tubes were expensive, crystal receivers, or one-tube regenerative sets, with earphones, were the rule. About this time, the General Electric Company started manufacture of the '99 and the '01A triodes [21], with thoriated filaments, and the Westinghouse Company introduced the WD-11 oxide-coated filament triode [22]. The reduced filament power of these types was a great advantage to receiver designs, which were battery operated. The reflex type of circuit was devised to hold down the number of tubes, since a single tube was used as both RF and audio amplifier. Gradual reduction of tube prices over the next decade eliminated the economic need for either superregeneration or the reflex circuit, and they soon become obsolete. The Neutrodyne [18] type of tuned RF receiver with '01A tubes became very popular in the early days of broadcasting in spite of the use of a separate tuning knob for each stage. Superheterodyne receivers with six tubes were finally put on sale [23] in 1924 and this type of receiver rapidly developed a reputation for sensitivity and ease of tuning.

As the first period came to a close, elimination of "B" battery operation by means of gas-filled or high-vacuum double-diode rectifiers became common. The demand for higher audio output led to use of higher power or low-mu triodes (types '12, '20, '71, and 210) in the output stage; even so, outputs of under 500 mw were common. At this time, also, a large number of receiving tube companies began to compete strenuously for the increasing volume of business, which attained 67 million dollars per year (1927), for receiving tubes [25].

AC Operation and the Multigrid Tube, 1927 to 1936

Although complete ac operation was feasible with filament-type tubes, further advance in circuit design required the equipotential or indirectly heated oxide-coated cathode. Such a cathode had been invented many years before [24], but high manufacturing cost and great difficulty in controlling heater-cathode leakage prevented its commercial use. The first reasonably reliable indirectly heated triode was the 2.5-volt heater type '27 which was announced in 1927 and adopted by most receiver manufacturers [26]. For a year or two, a companion tube for triode RF amplification, the type '26, which had a low-hum oxide-coated filament, was also used [26]. However, it quickly became apparent to tube manufacturers that all but output tubes would soon require indirectly heated cathodes. The first screen-grid tube, the filament type '22, was also introduced late in 1927, but received almost no use. By early 1929, the 2.5-volt heater, type '24, screen-grid tube became available [27] and began to be used extensively.

Complete ac operation was the key to tremendous improvement in receiver performance from 1927 to 1930. It became practical to use many tubes in a receiver without impractical battery drain. RF amplification and high audio output led to very selective and sensitive receivers [27], [28] capable of operating the newly introduced electrodynamic loudspeakers. Along with the indirectly-heated screen grid and triode tubes, the low-mu, oxide-filament type 45 was commonly used for audio output [29]. The typical 1929 high-performance set used two or three stages of tuned radio-frequency, with screen-grid tubes and a single tuning knob, followed by a high-level linear detector and a triode audio stage with one or more type 45 output tubes, delivering from 1 to 3 w to the speaker. It was at this time that some earlier ideas, the self-bias resistor, and the automatic-volume control principle [30], began to be widely applied [31].

By 1930, superheterodynes with type 24 tubes for the RF and IF amplification began to greatly out-perform the tuned RF sets, particularly for selectivity above 1 Mc. Broadcast stations increased power enough so that cross-modulation became a serious problem and the variable-mu tube (remote cutoff) was invented [32] and appeared in receivers as the screen-grid 2.5-volt heater type 35. The output pentode [33], already in use in Europe, was introduced in the U. S. by one of the smaller tube manufacturers, Arcturus [34], and the similar type 47 soon superseded the old low-mu triode type 45 as the most popular output tube. In the meantime, automobile receivers were introduced and a 6-volt heater line, in a small glass envelope, became available as the types 36, 37, 38 and 39. The type 38 was the first indirectly heated output pentode. In a few years, the 6-volt heater took over, and 2.5-volt tubes were discontinued.

Tube developments came so rapidly in 1931 to 1933 that it is impossible to list more than a fraction. Class B amplification was used in the highest-output audio stages [35], [36] and the types 46, 79 and 53 were developed; the latter two types were double triodes, and were the forerunner of a very popular type of tube ten years later. The type 55 and 85 were combination diode-triodes. The 41, 42, 59 and 89 were all indirectly heated screen-grid output tubes. For 110-volt operation ac/dc receivers, with series string heaters, the types 43 and 48 screen-grid output tubes were developed. The year 1932 also saw the introduction of the dome-top bulb and a whole series of new screen-grid RF amplifier tubes [37] in which a separately connected suppressor was used for the first time. This series included the 56 and 76 triodes, and the 57, 58, 77 and 78 RF amplifiers.

About 1933, a considerably more revolutionary tube development took place, one which once and for all consolidated the superheterodyne as the one and only receiver circuit, even in the lowest-cost sets. This was the pentagrid converter, type 2A7 and 6A7, which used electron coupling between an oscillator inner section and an outer, remote cutoff screen-grid mixer section [38]. Thus, in a single tube, high conversion gain was achieved together with freedom from oscillator interaction with the signal circuit. Within a year or two, this converter tube was almost universally used in the U. S., although in Europe a combination triode-hexode was also used in the same general way. From this time on, it became possible to make a 4-tube-plus-rectifier superheterodyne, with performance not too far from 8- and 10-tube receivers of a few years before. For battery portable receivers, a corresponding low drain filament type, the 1A6 was introduced and was used with other 2-volt types in such applications.

In 1935 a change was made in the base and envelope of almost every basic type of receiving tube. The octal base, with up to eight pins, and with a convenient locating lug, became almost a standard from this point in time. A line of octal-base metal-envelope tubes included the 6J5, 6J7, 6K7, 6A8 and 6F6, which were the counterparts of the earlier dome-top bulb line [39]. The electron ray type of tuning indicator tube continued to use glass, of course, and in a year or two there were many straight-sided glass envelope equivalents to the metal types, designated by the suffix GT. Over the years to follow, metal tubes gradually disappeared from use, but the octal base remained.

For a year or two prior to the metal tube, receiver manufacturers had been including from one to three short-wave bands, with frequency ranges up to 20 Mc or even to 60 Mc. This presented problems for the pentagrid converter, and a 5-grid mixer tube, the 6L7, was developed for the metal line [38]. The 6L7 required a separate oscillator, but had much less oscillator-signal circuit interaction at the short wavelengths. A few years later, the 6K8 and 6SA7 converter tubes were devised for the same purpose [38].

During the development of the metal tube, extensive laboratory work had been undertaken to improve the audio output pentode. These endeavors led to an oval cathode, an aligned control-grid and screen-grid to reduce the screen-grid current, and a minimal suppressor consisting of two side "beam-forming plates"; the remaining suppression of secondary emission came from space charge [40]. The first tube using the new beam principles was the 6L6, and within a few years it was followed by a slightly smaller version, the 6V6. These "beam-power" types had higher output, higher efficiency, and lower distortion than the pentode tubes which they largely replaced; the same beam principles are still in use today.

The close of this period of rapid expansion saw receiving tube production rise to nearly 100 million units per year (1936) and price competition, plus mass production techniques, led to cost reductions which were among the most remarkable in the history of the industrial revolution. Whereas the type '01A triode of 1923 had a list price of about $9.00, the five-grid converter tube of 1936 had a list price of only around $1.25. In 1936, about 9 million radio receivers were sold, with a retail value of around 500 million dollars.

Close-Spaced Tubes and VHF, 1936 to 1960

Just as the first period was chacterized by the triode, and the second period by the indirect heater and multigrid tube, the third period is characterized by the use of close grid-cathode spacing. For many years it had been appreciated that closer spacing between control grid and cathode led to higher transconductance. However, tube manufacturers prior to 1936 held firmly to a conviction that 0.015-in (0.4 mm) spacing was about as small as practicable for a low-cost mass-production tube. The improved understanding of transit-time effects [41], fluctuation noise [42], and the principles of wide-band amplification [43] all led to only one conclusion: Practical television and FM at VHF frequencies required that tube spacings be very much smaller.

Fortunately, acorn tube [44] manufacturing experience showed that short lengths of cathode and grid *did* permit reasonable yields, even with 0.005-in spacing. The additional feature of dc inverse feedback, through a cathode bias resistor, turned out to make such spacings practicable even for longer length cathodes, and the first wide-band amplifier tubes, the 1853/6AB7 and the 1852/6AC7 were born [45]. From this point on, almost every subsequent tube development of importance used spacings substantially less than had previously been believed practicable and, most remarkable of all, this was eventually done with very little increase in manufacturing cost. At the same time as this progress was made in receiving tubes, television picture tubes with up to 12-in diameter white screens were developed and the 1939 World's Fair saw the introduction of the first television receivers designed for home use [46]. These operated up to 90 Mc and had IF and video bandwidth of 4 Mc. Just a year before, the first commercial FM receivers [47] were put on sale; they operated in the 50 Mc region.

By 1939, some other chacteristics of future receiving tubes were evident. Metal tubes were getting much competition from glass types, particularly from the GT and a glass base type known as the "loctal." Tube types, essentially the same in principle, but with very minor deviations in basing, or internal capacitances, proliferated almost without end (for this reason, specific tube types will no longer be mentioned in the remainder of this section). A line of miniature glass tubes for battery-operated portable receivers was successfully manufactured [48], and this small sized glass envelope with a new integral 7-pin glass base was destined to become the standard for all except the highest power receiving tubes of the future. The structure was small, inexpensive, and the short leads were ideally suited for

VHF type operation. During World War II, this trend to the miniature glass envelope became consolidated [49], and the tubes which are in use today (1962) in almost every broadcast and TV receiver are essentially World War II types modified for commercial production.

After the war, just about every tube which had been known before the war had a redesigned counterpart in either a miniature, $\frac{3}{4}$-in diameter envelope, a 9-pin $\frac{7}{8}$-in envelope, or in the GT octal-base glass construction. A major change, however, was a greatly increased use of double-triodes in receiver circuitry, particularly in TV receivers. In fact, this type of tube essentially halted the more extensive use of multigrid structures, since it was found that a double triode was a less expensive way to solve a circuit problem than use of two higher-performance multigrid structures [50]. In the case of low-noise TV input amplifiers, the cascode circuit [51] for the double triode actually provided better performance than any known multigrid tube circuit.

In their impact on receivers, the post-war tubes permitted excellent FM performance in the newly assigned 100-Mc band, and television reception up to 216 Mc was most satisfactory. In the television receiver, the special problems of magnetic deflection [52], [53] were solved by special beam type deflection tubes, and low-impedance damping diodes. Picture tube anode voltages were raised to high values permitted by pulse transformers operating filamentary high-voltage rectifiers. The picture tubes themselves were greatly improved, with greater size and aluminized backing for the phosphor [54]. By 1950 it was clear that electronic color television was entirely feasible, at least from a technical point of view, and color-television picture tubes had already been successfully demonstrated [55], [56]. Extensive use of inverse feedback in high-gain audio amplifiers led to greatly improved sound quality, both in receivers and in separate hi-fi equipment.

From 1953 on, tube developments played a diminishing role in broadcast receivers, because of the major impact of the solid-state art, *i.e.*, the transistor and the junction rectifier. The transistor quickly dominated the portable receiver, and the silicon rectifier is rapidly replacing the vacuum tube types in ac receivers. By 1956, tube developments were already altering to match the transistor. In the automobile receiver, an old principle, the space-charge grid (invented in 1913, but never used), was revived, so that 12-volt operation became feasible in the so-called "hybrid" receiver, with tubes for the RF and IF and a tube driver for a transistor audio output [57]. A radically new line of tubes, called Nuvistors, was developed [58]. These were very small in size, had extremely low power consumption, and had superior high-frequency performance over other low-cost tubes and transistors [59].

At the close of this period, receiving tube sales were still high[3] because the transistor had not yet entered the TV or FM receiver, and low-frequency broadcast receiver sales were chiefly confined to portable and auto receivers. However, the trend is clear and the amplifying type of receiving tube, as we knew it in the past, has clearly seen its zenith.

[3] In 1959, over 400 million receiving tubes were sold.

The Present and the Future

At the present time, the major advantages which the receiving tubes have over the competing solid-state devices are cost and performance. The tube has had forty years of improving technology, mass production, and cost reduction, while solid-state devices have had only about ten years. If size and power consumption are disregarded, there is almost no function in which the receiving tube is not superior in performance to the semiconductor device, except perhaps for the power rectifier. Nevertheless, it is also clear that the receiving tube is at a time of minor refinements, and no major new inventions are likely. The solid-state device, on the other hand, is only just beginning to achieve low-cost production, and major performance-improving inventions are being made each year.

Perhaps the real clue to the future lies in integrated and functional electronics, or microelectronics as it is also called. Future receivers will surely include functional subsystems, replacing the arrays of active and passive components we now use. By far the greatest research effort in micro-electronics is going on in the solid-state area. However, there are several groups of competent investigators who are convinced that a form of vacuum-tube technology may yet be the answer. The outcome will probably lie in both areas, *i.e.*, there will be microelectronic systems of both types. Nevertheless, in the writer's opinion, the broadcast and TV receivers of the future will slowly and inevitably tend toward predominant use of solid-state devices, with the exception of the picture tube.

References

[1] J. A. Fleming, "Instrument for Converting Electric Currents into Continuous Currents," U. S. Patent No. 803684, filed April 19, 1905; issued Nov. 7, 1905.
[2] L. DeForest, "Space Telegraphy," U. S. Patent No. 879532, filed January 29, 1907; issued February 18, 1908.
[3] L. DeForest, "Discussion of E. H. Armstrong paper," Proc. IRE (Correspondence), vol. 3, p. 239, September, 1915.
[4] For a comprehensive review of tube developments of the period 1910–1918, see G. F. J. Tyne, "The saga of the vacuum tube," *Radio News*, in 22 parts, vol. 29, March, 1943 to vol. 35, April, 1946.
[5] I. Langmuir, "The pure electron discharge," Proc. IRE, vol. 3, pp. 266–269; September, 1915.
[6] W. R. Maclaurin, "Invention and Innovation in the Radio Industry," The Macmillan Co., New York, N. Y., p. 77; 1949.
[7] L. DeForest, "Radiosignaling System," U. S. Patent No. 1,507,016, filed September 23, 1915, and "Wireless Telegraph and Telephone System," U. S. Patent No. 1,507,017, filed March 20, 1914; issued September 2, 1924.
[8] E. H. Armstrong, "Some recent developments in the audion receiver," Proc. IRE, vol. 3, pp. 215–238; September 1915.
[9] R. Fessenden, "Wireless Signaling," U. S. Patent No. 706,740, 1902; and "Electric Signaling Apparatus," U. S. Patent No. 1,050,441; 1913.
[10] E. H. Armstrong, "A new system of short-wave amplification," Proc. IRE, vol. 9, pp. 3–11; February, 1921.

[11] W. Schottky, "On the origin of the superheterodyne method," Proc. IRE, vol. 14, pp. 695–698; October, 1926.
[12] E. F. W. Alexanderson, "Selective Tuning System", U. S. Patent No. 1,173,079, filed Oct. 29, 1913; issued February 22, 1916.
[13] W. Schottky, "High-vacuum amplifiers," *Arch. Elektrotech.*, vol. 8, pp. 1–41, 299–328; 1919.
[13a] Previously invented by I. Langmuir, "Electrical Discharge Apparatus," U. S. Patent No. 1,558,437, filed October 29, 1913; issued January 1, 1924.
[14] A. W. Hull and N. H. Williams, "Characteristics of shielded-grid pliotrons," *Phys. Rev.*, vol. 27, pp. 432–438; 1926.
[15] G. Holst and B. D. H. Tellegen, "Means for Amplifying Electric Oscillations," U. S. Patent No. 1,945,040, filed Nov. 25, 1927; issued January 30, 1934.
[16] J. M. Miller, "Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit," *Sci. Paper Bur. Stds.*, Washington, D. C., vol. 15, No. 351, pp. 367–385; 1919.
[17] C. W. Rice, "System for Amplification of Small Currents," U. S. Patent No. 1,334,118, filed July 31, 1917; issued March 16, 1920.
[18] L. A. Hazeltine, "Tuned R-F amplification with neutralization of capacity coupling," *Proc. Radio Club America*, vol. 2, no. 8, 1923.
[19] E. H. Armstrong, "Some recent developments of regenerative circuits," Proc. IRE, vol. 10, pp. 244–260; August, 1922.
[20] G. L. Archer, "Radio's great era of expansion," in "History of Radio," American Historical Society, New York, N. Y., ch. 15; 1938.
[21] J. C. Warner, "Recent developments in high vacuum receiving tubes," Proc. IRE, vol. 11, pp. 587–600; December, 1923.
[22] F. S. McCullough, "Thermionic tubes," Proc. IRE, vol. 10, pp. 468–484; December, 1922.
[23] E. H. Armstrong, "The superheterodyne—its origin, development, and some recent improvements," Proc. IRE, vol. 12, pp. 539–552; October, 1924.
[24] A. Nicolson, "Thermionic Translating Device," U. S. Patent No. 1,459,412, filed September 10, 1919; issued June 19, 1923.
[25] "Statistics of Radio Sales," *Electronics*, vol. 2, pp. 534–535; March, 1931.
[26] A. N. Goldsmith, "Progress in radio reception during 1927," *GE Rev.*, vol. 31, pp. 64–69; January, 1928.
[27] A. N. Goldsmith, "Progress in radio reception during 1929," *GE Rev.*, vol. 33, pp. 69–79; January, 1930.
[28] W. E. Holland and W. A. MacDonald, "The Philco 95 screen-grid plus," *Radio Broadcast*, vol. 16, pp. 111–112; December, 1929.
[29] F. H. Engel, "Engineering features of the UX-245," *Radio Broadcast*, vol. 15, pp. 167–168; July, 1929.
[30] L. Espenschied and R. Bown, "Radio Transmission Control System," U. S. Patent No. 1,447,773, filed 1921; issued 1923.
[31] H. A. Wheeler, "Automatic volume control for radio receiving sets," Proc. IRE, vol. 16, pp. 30–35, January, 1928.
[32] S. Ballantine and H. A. Snow, "Reduction of distortion and cross-talk in radio receivers by means of variable-mu tetrodes," Proc. IRE, vol. 18, pp. 2102–2127; December, 1931.
[33] S. Ballantine and H. L. Cobb, "Power output characteristics of the pentode," Proc. IRE, vol. 18, pp. 450–470; March, 1930.
[34] B. V. K. French, "Power pentodes for radio receivers," *Electronics*, vol. 2, pp. 576–577; 614; April, 1931.
[35] L. E. Barton, "High audio output from relatively small tubes," Proc. IRE, vol. 19, pp. 1131–1149; July, 1931.
[36] L. E. Barton, "Application of the class B audio amplifier to a-c operation," Proc. IRE, vol. 20, pp. 1085–1100; July, 1932.
[37] J. C. Warner, E. W. Ritter and D. F. Schmidt, "Recent trends in receiving tube design," Proc. IRE, vol. 20, pp. 1247–1262; August, 1932.
[38] For references to the history of converter tubes, see E. W. Herold, "The operation of frequency converters and mixers," Proc. IRE, vol. 30, pp. 84–103; February, 1942.
[39] R. H. Langley, "Review of broadcast reception in 1935," Proc. IRE, vol. 24, pp. 376–384; March, 1936.
[40] O. H. Schade, "Beam power tubes," Proc. IRE, vol. 26, pp. 137–181; February, 1938.
[41] W. R. Ferris, "The input resistance of vacuum tubes as ultra-high-frequency amplifiers," Proc. IRE, vol. 24, pp. 82–105; January, 1936.
[42] B. J. Thompson, D. O. North and W. A. Harris, "Fluctuations in space-charge-limited currents at moderately high frequencies," *RCA Rev.*, in 5 parts; vol. 4, January, 1940 to vol. 6, July, 1941.
[43] H. A. Wheeler, "Wide band amplifiers for television," Proc. IRE, vol. 27, pp. 429–437; July, 1939.
[44] B. Salzberg and D. G. Burnside, "Recent developments in miniature tubes," Proc. IRE, vol. 23, pp. 1142–1157; October, 1935.
[45] A. P. Kauzman, "New television amplifier receiving tubes," *RCA Rev.*, vol. 3, pp. 271–289; January, 1939.
[46] For example, see "*Practical Television*," RCA Manufacturing Co., Camden, N. J., Copyright 1939, which describes the RCA Victor TRK-12.
[47] G. W. Fyler and J. A. Worcester, "FM in radio broadcasting—a new Armstrong FM receiver," *Proc. Radio Club of America*, vol. 16, July, 1939.
[48] N. R. Smith and A. H. Schooley, "Development and production of the new miniature battery tubes," *RCA Rev.*, vol. 4, pp. 496–502; April, 1940.
[49] N. H. Green, "Miniature tubes in war and peace," *RCA Rev.*, vol. 8, pp. 331–341; June, 1947.
[50] G. C. Sziklai and A. C. Schroeder, "Cathode-coupled wide-band amplifiers," Proc. IRE, vol. 33, pp. 701–708; October, 1945.
[51] H. Wal man, A. MacNee and C. Gadsden, "A low-noise amplifier," Proc. IRE, vol. 36, pp. 700–708; June, 1948.
[52] O. H. Schade, "Magnetic-deflection circuits for cathode-ray tubes," *RCA Rev.*, vol. 8, pp. 506–538; September, 1947.
[53] O. H. Schade, "Characteristics of high-efficiency deflection and high-voltage supply systems for kinescopes," RCA Rev., vol. 11, pp. 5–37; March, 1950.
[54] For a d scussion of picture tube development see F. J. Bingley, "Forty years of television reception," Proc. IRE, this issue.
[55] E. W. Herold, "Methods suitable for color kinescopes," Proc. IRE, vol. 39, pp. 1177–1185; October, 1951.
[56] H. B. Law, "A three-gun shadow-mask color kinescope," Proc. IRE, vol. 29, pp. 1186–1194; October, 1951.
[57] C. E. Atkins, "Low plate potential tubes," *Radio and Television News*, vol. 57, pp. 47–49, 130; January, 1957.
[58] G. M. Rose, "The Nuvistor—a new look in tube design," *RCA Engineer*, vol. 4, pp. 42–44; April–May, 1959.
[59] G. C. Hermeling, "A Nuvistor low-noise TV tuner," IRE Trans. on Broadcast Television Receivers, vol. BTR-6, pp. 52–56; November, 1960.

Section 6

BROADCASTING

Organized with the assistance of the IRE Professional Group on Broadcasting

AM and FM Broadcasting*

RAYMOND F. GUY†, FELLOW, IRE

Summary—**Interesting early experiments and significant technical milestones are described in radiotelephony preceding the birth of the broadcasting industry at KDKA in 1920. The technical problems and difficulties which faced the burgeoning early growth and subsequent development are outlined and the most significant advances in the technology are outlined as they occurred in both AM and FM broadcasting. The descriptions are accompanied by comments concerning the past, present and possible future status and trends in both services and the manner in which they may be influenced by factors which are unique to each.**

USEFUL COMMUNICATION by radio had its humble beginnings as the 19th century drew to a close. By 1920 it had attained stature as an important component in the world's economies, and had become a great boon to mankind. Yet throughout this period commercial radio operations were conducted by technically trained professional people at specially built commercial, governmental or military installations. In 1920 a profound change took place. Radio also moved into the home, to become a vital force to be enjoyed by all members of the family. A new service was born.

In retrospect it seems inevitable that such a development should have taken place because of the advances in the technology of transmission and reception of radiotelephony. In particular, the mass production of simple and inexpensive radio receivers, the "little black box," which could be installed and operated by any member of the family, represented the fruition of the dreams of pioneers such as Fessenden, de Forest, and Sarnoff.

* Received by the IRE, November 14, 1961.

† 264 Franklin Street, Haworth, N. J.

Early Experiments in Broadcasting

The first documented successful broadcasting of speech and music was conducted by Dr. Reginald Fessenden at Brant Rock, Mass., on Christmas eve, 1906, utilizing a 50 kc radio-frequency alternator which produced about 1 kw of power and which was built by the General Electric Co. under the direction of Dr. E. F. W. Alexanderson. Modulation was accomplished by means of a microphone which is believed to have been water-cooled and which was connected in the antenna circuit. Clear reception was obtained at many locations including ships at sea [1].

Subsequently Dr. Lee de Forest conducted experimental broadcasting in 1907 from his laboratory in New York City, in 1908 from the Eiffel Tower in Paris, and in 1910 from the Metropolitan Opera House in New York City. These experiments, which were conducted with arc transmitters of about 500 watts power, modulated by microphones in the antenna-ground system, while successful, were handicapped by the high noise level inherent in arc transmitters. However, the Metropolitan broadcasts were particularly noteworthy in that they consisted of performances from the stage featuring Enrico Caruso and other great artists of the day.

The discovery that the de Forest triode could be made to produce continuous wave oscillations, and its adoption as a carrier generator, led to a succession of experimental transmissions of high-quality radiotelephony including those of de Forest in 1916 from High Bridge, N. Y., those at Tufts College in 1915–1916, and those of the ATT Co. in 1915 at the Naval Station, Arlington, Va. The ATT Co. equipment employed 500 tubes in parallel as Class A amplifiers and was heard in Paris and Honolulu.

Although it was demonstrated during this period that high-quality radiotelephony could be achieved with practical vacuum tubes and ancillary components, it was not until after World War I that it was shown that high plate operating efficiency could be obtained in combination with highly sophisticated and satisfactory methods of modulation. For the development of the class C radio-frequency amplifier and efficient methods of modulating it, including the constant current system, we are indebted to Dr. Raymond A. Heising. His work on modulating systems was described in the August, 1921, issue of the IRE PROCEEDINGS for which he was awarded the Morris Leibman prize [2], [3].

Experimental transmissions by Dr. Frank Conrad of the Westinghouse Electric and Manufacturing Co. in 1919 and 1920 from his amateur station 8 XK, in Pittsburgh, created such great interest that musical programs were transmitted at scheduled times. His rapidly growing audience led to the inescapable conclusion that the time was ripe for the creation of a public broadcasting service and a U. S. Government license was requested and received for the purpose. The world's first broadcasting station to be assigned special call letters and a special frequency for its purpose was KDKA, Pittsburgh, Pa., which inaugurated a public broadcasting service in November, 1920.

Early broadcasts from the studio soon were supplemented by broadcasts from churches and other remote points, using wire lines for conveying the program signals to the transmitter. Public acceptance and enthusiasm for this service was immediate and created a phenomenon which was not to be duplicated until many years later with the new television service after World War II. The tremendous rate of growth in broadcasting, particularly during the first few years, may be seen from the listing of station authorizations granted by the Department of Commerce as shown in Table I.

TABLE I

Year	New Authorizations Granted
1920	1
1921	30
1922	639
1923	239
1924	286

Allocation of Frequencies to AM Broadcasting

In 1920 and 1921, broadcasting stations were assigned the wavelength of 360 meters, or 833 kc. The creation of new stations quickly produced an allocations problem which in some areas led to the enforced sharing of time on the air. Recommendations of the First National Radio Conference in 1922 led to the recognition of the need for more adequate regulation of the new service. As a result the classification of "limited commercial" station was abandoned with respect to broadcasting stations, which became classified as such. In addition, the problem of crowding and time sharing was dealt with. More adequately to cope with the growing number of stations, the wavelength of 400 meters, or 750 kc, was also assigned to the broadcasting service, and the power limits of a minimum of 500 watts and a maximum of 1000 watts were established. But these regulations also quickly became inadequate as hundreds of new stations came into existence.

Based upon recommendations of the National Radio Conference of 1923 and 1924, the Department of Commerce allocated the whole band from 550 to 1500 kc to the rapidly growing service and authorized powers up to 5000 watts. This also quickly proved to be inadequate as the demands for frequencies and power increases grew inexorably. In 1925 a Fourth National Radio Conference requested that limits be established on power and broadcast time to relieve the severe and growing interference. However, the attempt to implement these limitations failed. While the Secretary of Commerce was empowered to grant authorizations, he found that he was unable to impose the desired limitations because court decisions held that under the Radio Act of 1912 he had no authority to do so. Thereafter, because of the recognized and admitted lack of regulatory authority, many broadcasters changed their operating frequencies and increased their powers at will. Bedlam resulted as Federal Regulation disintegrated.

In 1926 Congress was urgently requested by President Coolidge to take remedial action to restore control, which was accomplished under the Dill-White Act of 1927. The Federal Radio Commission was created with broad powers and one of its first tasks was to restore order and correct the chaotic conditions. It proved to be impossible to make satisfactory provisions for the 732 stations then in operation, as a result of which about 150 stations were required to surrender their licenses and cease operating, and many others to change frequencies and other operating conditions.

The final and most significant action in the regulatory field was the creation of the Federal Communications Commission under the Communications Act of 1934. This body was given broad authority and charged with responsibility for regulation of all interstate and foreign communications by radio and wire.

Under the FCC, growth has been orderly. New Rules and Regulations have taken advantage of technical developments through the years of its existence to increase the service to the public, and the number of stations which provide it. The broadcast band has been widened to encompass 540 to 1600 kc and the number of stations has grown to over 3500, with nearly 900 new applications currently awaiting processing.

International Interference Problems

The burgeoning broadcasting industry experienced other growing pains during its early years, as more stations were built and operating powers were increased in all of the countries of the North American Region. Each country has a sovereign right to use any part of the frequency spectrum at will. But the exercise of this right quickly led to interference between stations in the various countries in the North American Region, particularly during the night time hours when skywave propogation produced substantial interfering fields at distances of many hundreds of miles. International interference was becoming intolerable by 1935.

By mutual agreement the countries of the North American Region met in Havana in an effort to negotiate a satisfactory agreement on the sharing of the broadcast spectrum and the mutual protection of the services of the participants. An agreement was reached in 1937, signed by the United States, Canada, Cuba, Mexico, the Dominican Republic and Haiti. Unfortunately, this agreement was not made effective until 1941 and it expired in 1946. It was succeeded by an interim agreement, arrived at with some difficulty, which extended the original North American Regional Broadcasting Agreement for three years, until 1949, with concessions to some countries. Upon expiration, another difficult period of negotiations followed which led to a new agreement in 1950 among all of the countries except Mexico. This agreement was not ratified until about 8 years had elapsed, largely because, without the participation of Mexico, there was strong opposition on the grounds that protection to the participating countries was not assured. In 1957 a bilateral agreement between the United States and Mexico cleared the way for ratification which followed shortly thereafter.

Negotiation of such agreements was complicated by the conflict of interests among not only the countries which participated, but also by such conflicts between various interests in the individual countries. As an example, in the United States there were conflicts of interest between organized groups of stations limited to daytime operation, regional channel operation and clear channel operation. One of the most difficult issues was the demand for additional clear channel rights by Cuba and Mexico.

Insofar as technical performance was concerned, in all countries and to a large degree, these conflicts could have been relieved by the use of FM broadcasting by stations limited in power and service area by crowding in the standard broadcasting band. But the long established public habit of listening to AM stations, the cost of developing and building an FM audience, and the competitive dislocation and hardship involved left little choice for stations which would not have been able to survive a hazardous and uncertain transistion. The existence of a large, ready-made and established AM listening audience, compared with the far smaller FM audience, has been the incentive behind the very large increase in the number of new AM stations during the post-war period, despite the fact that for most small stations a larger area could be served, free of interference and with the other inherent advantages of the FM system.

Directional Antenna Systems

The very rapid growth of broadcasting during the period 1920–1926 required much sharing of carrier frequency assignments at what were deemed to be adequate geographical separations to maintain reasonable limits of mutual interference. The antennas employed at that time were omnidirectional. Under these conditions this technique for a time made it possible to find suitable frequency assignments for new applicants. But as the number of new applications continued to mount it was not long before difficulties arose. It became evident that if a technical solution were not found, either the interference standards would have to be relaxed or the granting of new applications would have to be increasingly limited until a static condition prevailed.

The technical solution which went far to relieve this problem was provided in 1927, early in broadcasting history, by two consulting engineers, Raymond Wilmotte and T. A. M. Craven, acting on behalf of a Florida applicant. Mr. Craven is now an FCC Commissioner.

Although directional antenna technology was not unknown, and a three-element medium frequency array had been described in a textbook around the turn of the century, none had been proposed for this service. Directional antennas used in other services were designed to increase the field intensity in a discrete direction. For the Florida application, Craven and Wilmotte proposed to utilize the directional properties to decrease it in the direction of a mid-western station, in order to protect it against interference.

Although directional antennas inevitably would have come into use at some time, the success of this project was particularly timely. It marked an important milestone in efficient use of the broadcast spectrum. The impact on the industry was substantial and beneficial, as evidenced by the hundreds of directional stations now operating which otherwise could not have been authorized.

Design Trends in AM Broadcasting Apparatus

Beginning with KDKA, the earliest broadcast transmitters generally consisted of self-excited oscillators modulated by class A constant-current Heising modulators. Because of the cost and design problems in developing large amounts of audio frequency power with class A modulators, most of the transmitters of that time were limited to 15 to 50 per cent modulation. One of the first 50 kw transmitters required 22 water-cooled modulator tubes to achieve 40 per cent modulation. The original KDKA transmitter produced 100 watts and others ranged from 5 to 500 watts. Lacking any special provisions for precise control of the carrier frequency, and a more precise measuring device than a wavemeter, stations operated on a rough approximation of the assigned frequency. It was not unusual in the earliest years to change the operating frequency 50,000 cycles to avoid, temporarily, interference from an unknown source.

Radiating systems consisted of *T* or *L* antennas with very poor ground systems and were characterized by radiation efficiencies which were very low by modern standards. Filament and plate supplies normally consisted of batteries or motor generators. One station, owned by a battery manufacturer, obtained plate power from a 10,000 volt battery. Safety devices were often crude or omitted entirely.

Improvements followed rapidly. The term "superpower" was loosely applied first to 5 kw in 1923, then to 50 kw in 1925, and finally to 500 kw in 1934 [4]. The one and only 500-kw transmitter licensed and used in the United States was at WLW, but this was destined to continue for only a limited time. A few years later the authorization was withdrawn and replaced with one for 50 kw, the highest power permitted since that time. In the absence of such a restriction, powers of 500 kw or more would have been used by many clear channel stations to provide new and improved service to millions of listeners. The wisdom of this restriction has been a highly controversial and debated question in the industry for over a quarter of a century.

It may be surprising to some that the achievement of satisfactory 100 per cent modulation in broadcast transmitters was, in general, delayed until new and improved techniques such as the class B modulator [5], [6], the Chireix outphasing system [7], and the Doherty high-efficiency system [8] were developed from 7 to 10 years after the birth of such a virile industry, but such was the case. All of these systems have been in use for about a quarter century. By far the most widely used is the class C radio frequency amplifier modulated by a class B modulator.

In the early years of broadcasting, during which self-excited oscillators were used to generate the carrier power, precise control of the operating frequency was impossible. Carrier frequencies varied from hundreds to thousands of cycles from the assigned frequency. In co-channel operation, nighttime interference between stations was characterized by interchannel beat notes which were usually in the most responsive range of loudspeakers, and extremely destructive to reception. Fortunately, this aggravation could be and was solved following the evolution of the crystal controlled oscillator about 1926. A first step toward a solution was taken when, shortly thereafter, the Regularity Authority imposed a 500-cycle tolerance on carrier frequency deviation, representing a reasonable recognition of the technology. A few years later, as the technology advanced, this tolerance was reduced to 20 cycles. Since this limits the frequency of a beat note between two stations to a maximum of 40 cycles, and stations actually maintain their frequencies to within one or two cycles, audible beat notes ceased to be a problem.

The development of efficient broadcasting antenna systems lagged somewhat in the early years but gained impetus with the work of Brown, Gihring and Ballantine [9]–[12] in 1934–1936. These published works quickly were reflected in new designs and also new FCC minimum standards of performance which stations are required to meet.

Viewed in the light of improvements which have taken place in apparatus and techniques, technical progress in the broadcasting industry has been gratifying. However, these improvements in general had been accomplished 20 or more years ago. The technology of AM broadcasting at that time had attained substantial maturity and stability. Although the number of stations has tripled in the ensuing years, the growth has taken place largely within the technical framework and level of competence established then. Subsequently, changes have been marked principally by the growing application of automated operation, the terminal development and widespread use of audio tape, growing application of remote control of apparatus, including transmitters, and gradual refinements and improvements in moderate degree of apparatus units and components, of which solid-state power rectifiers is an example.

Were a listener able to compare the transmissions of today with those of 20 years ago, the technical difference would be unnoticeable. Program-wise the difference would be very noticeable because AM programming has had to be adjusted to meet the impact of television.

Stereophonic Transmission in AM Broadcasting

During the past few years there has been a growing interest in stereophonic reproduction of music. Several proposals have been made to the FCC that it be authorized in AM broadcasting. The methods proposed, in general, contemplated the use of amplitude modulation of $A+B$, and the use of velocity modulation of the same carrier for $A-B$. After studying these proposals, the FCC ruled that it would not be authorized because of the lack of public interest and the technical problems attendent upon its use in the Standard Broadcast band.

Compatible Single Sideband Transmission

Another proposal has been under consideration, the use of compatible single sideband transmission. This system, developed and advocated by Kahn [13], may be applied to a conventional Standard Broadcast transmitter by the addition of equipment mounted on standard telephone-type racks, and a few simple changes in the transmitter input circuitry. In the Kahn system only the low audio frequency components have sufficient amplitude to produce appreciable second-order sideband components. The bandwidth of the CSSB wave is essentially that of a conventional SSB wave, occupying about half of the bandwidth of a conventional AM double sideband wave. This reduction of transmitted sideband width may be employed effectively to reduce co-channel or adjacent channel interference. A desired-to-undesired sideband ratio of about 30 db is achieved with normal modulation. By tuning the receiver it is possible to utilize the full RF-IF pass band to obtain from the SSB system a wider range of audio frequency reproduction.

The Kahn system has merit and could be applied to Standard Broadcasting without cost to the public, obsoleting existing receivers, or other major complications. It is the first system of its kind to be proposed in the 41-year history of broadcasting and has the advantage that it could be authorized as an optional method of modulation. A technical description of the Kahn system is given in another paper in this issue by Reed [25].

Technical Status of AM Broadcasting

Although changes may take place in the continuing adjustment of the program format and the economics of AM broadcasting, it must be viewed as a relatively mature industry. Because of limitations of channel width and near-saturation of the broadcast band, the tremendous investment in publicly owned receivers, the cost, and the absence of public demand, any sweeping changes in frequency allocations or technical methods seem impractical. The Kahn system could be integrated with existing techniques without such sweeping changes.

The excellent national television service, and the growing FM service will not, in the foreseeable future, supplant AM broadcasting, which continues to expand. AM broadcasting provides service which cannot be duplicated by TV or FM with respect to daytime service range, and nighttime skywave service to relatively great distances.

Frequency Modulation Broadcasting

The development of commercial FM broadcasting, as was the case with commercial television, encountered some formidable obstacles. Both services were born in 1941; within a period of months the United States was at war. Production of apparatus was halted and expansion of these services came to a halt. After the war ended, both services were required to readjust to new changes in the portions of the frequency spectrum assigned to them.

Not the least of the problems facing the infant FM service was the competition from the well-established, prosperous and reasonably satisfactory AM service. In urban and suburban areas the vast majority of listeners received good service with a wide choice of programs, and could not receive FM transmissions without a substantial investment in receivers. The principal advantages of FM to the public were:

1) the availability of a system in which there were no inherent limitations on the fidelity of transmission and reception which could be achieved,
2) the inherent capacity to reduce noise and interference,
3) the availability of new channels on which more stations could operate to provide new or greater service.

FM broadcasting was authorized on January 1, 1941. For this service the FCC assigned the 42-50-Mc frequency band, which encompassed 40 channels. The 42–43-Mc bloc was for noncommercial operation and the 43–50 for commercial. About 350,000 receivers were sold for this allocation format. Then, in 1945, the entire FM allocation plan was changed, the 42–50-Mc bloc was abandoned and the service was moved up to 88–108-Mc to encompass 100 channels. 88–92 Mc are assigned for noncommercial use and 92–108 Mc for commercial.

For FM broadcasting, which produced little or no revenue, this was a difficult adjustment to make. Despite the fact that the increased number of channels was very desirable, this move was not made without spirited controversy based upon the financial hardship and dislocation involved, and doubts about the severity of the potential interference claimed to exist because of long-distance F-layer propagation at 40–50 Mc during periods of maximum sunspot activity. To a lesser degree there were some misgivings about the adequacy of 100 channels for a fully developed national competitive FM service, with particular reference to the heavily populated northeastern states.

Much of the impetus to the growth of FM broadcasting arose from the concern of AM station owners that if they did not have an FM station they might not be able to obtain a frequency assignment for one when, possibly, FM broadcasting would supplant AM. Most of the early stations, and many of those now operating, duplicated or duplicate their programs on AM and FM in the same area. Possibly 10 per cent of the early FM operations were abandoned as these fears diminished.

Subsidiary FM Services

In order to stimulate the growth of the struggling FM industry, the FCC, in May, 1945, authorized a subsidiary form of service upon the application for and granting of a Subsidiary Communications Authorization. Weather information, time signals, background music for restaurants, stores, homes, etc., special programs for merchandising establishments, religious, professional, educational, agricultural groups, etc., were permitted under the condition that the service on the main channel be conducted simultaneously, for general public reception, with conventional receivers, without interference. Within two years about 100 stations took advantage of this option.

The growth of FM broadcasting in recent years has been more solidly based upon a desire to provide high-fidelity transmission of good music, cultural and educational programs, or other forms of service which could not be provided in the almost saturated AM band. A lack of public understanding of high-fidelity reproduction, and rather widespread indifference to it in the early years of FM operation, have gradually diminished. Growing appreciation of high-fidelity reproduction and stereophony are creating a growing interest and appreciation of FM broadcasting and there is every reason to believe that this interest, and the FM broadcasting service, will continue to expand.

FM growth already has reached the milestone where FCC competitive hearings have been necessary to determine which of many applicants should be granted the only remaining unassigned channels in some of our large cities in the northeastern states and elsewhere. In the spring of 1961 over 800 FM stations were licensed, construction permits had been requested for about 300 more, and there were on file over 100 other new applications.

An Interesting Sidelight on FM Development

Any author writing about the history of FM would be remiss in omitting reference to two early patent applications 60 years ago by Cornelius D. Ehret, a Philadelphia patent attorney. On February 10, 1902, he filed with the U. S. Patent Office disclosures covering the transmission and reception of code signals or speech by varying the frequency of wireless waves. So far as is known these were the first disclosures to describe any system of modulation by name. His method of modulation consisted of varying the resistance or reactance in the circuitry of an oscillator. Inasmuch as the de Forest triode was not to be invented until four years later, and an additional period of several years was to lapse before it was discovered that it could be made to oscillate or amplify, Ehret showed the spark type of impulse oscillator of the day. He described a demodulator employing slope detection and specified that the method did not depend upon increasing or diminishing the amount of energy transmitted.

The amount of interest in FM in the very early years of the industry may be surprising in view of the fact that it was not used for telephony to any significant extent until 39 years after Ehret's disclosures. Table II shows year by year the number of patent applications applying to frequency or phase modulation, of which roughly 90 per cent cover FM.

It is evident that there have been many workers in the field of FM, and many contributors to the technology, but two certainly warrant special recognition, Edwin H. Armstrong and Murray Crosby.

TABLE II

Year	Applications or grants
1905	2 (Ehret grants)
1916	1
1917	3
1919	2
1921	1
1922	1
1923	1
1925	3
1926	4
1927	6
1928	6
1929	6
1930	1
1931	9
1932	16
1933	13
1934	10
1935	9
1936	17
1937	16
1938	15
1939	18
1940	24
1941	61
	245

An early and enthusiastic advocate of FM broadcasting, Armstrong, through his tireless dedication and efforts, his technical papers and his innumerable promotional activities, brought to public attention the merits of the system and provided the leadership which led to the adoption of frequency allocations and technical standards, and the authorization of commercial service by the FCC on January 1, 1941. His name has been inseparably linked with FM since his historic paper in 1936 [14].

Crosby, over a period of many years, enriched the technical literature with the results of his tireless investigation of phase and frequency modulation methods and systems [15]–[21]. His contributions were particularly timely and valuable during the years prior to 1941, when the system was being developed and field tested [24].

Stereophonic Transmission in FM Broadcasting

Since World War II there has been mounting public interest in high-quality reproduction of music. In the home instrument field the techniques for accomplishing it with audio tape and grooved recordings have improved a great deal, and the sale of such equipment has grown to hundreds of millions of dollars each year. There has been a parallel interest also in stereophonic reproduction in both home record players and FM broadcasting. In recognition of this interest the FCC invited industry proposals for national stereophonic standards for FM. With the assistance of a specially appointed National Stereophonic Radio Committee, under the chairmanship of A. Prose Walker, the several systems which were proposed were field tested [22], one was adopted, and the FCC on June 1, 1961 authorized commercial stereophonic transmission in the FM broadcasting service. This system is described by Reed [25].

The introduction of stereophonic and subsidiary subcarrier transmission requires more rigid tolerances in the transfer characteristics of the FM system to minimize crosstalk between carrier and subcarriers [23]. In this respect the achievement of a low voltage-standing-wave ratio in the transmission line and antenna system assumes increased importance. Multiplex crosstalk increases with increasing nonlinear phase shift, which usually accompanies a reactive component in the termination. Transient changes caused by ice deposits, rain or snow on some types of antennas may become so severe that continuity of operation may be acheived only by a reduction in power. In a multiplexed system the added vulnerability of crosstalk exists unless the VSWR is low and remians stable.

Trends in FM Transmitting Apparatus

Improvements in FM transmitting apparatus since 1941 have followed advances in the technology and consist principally of

1) the introduction of solid-state components in power supply rectifiers and other circuits,
2) improvements in over-all efficiency through the use of improved tube types and circuitry,
3) improvements in FM modulators to achieve more precise maintenance of average carrier frequency, simplicity, elimination of moving parts, and superior operational stability,
4) the introduction of stereophonic transmission,
5) the introduction of remote control of transmitters.

With the exception of the modulating systems, these new components or techniques are well known and require no recounting even though their application has substantially improved over-all operating efficiency and enabled economies to be achieved in the cost of construction and operation. With respect to the stereophonic FM modulating system which has been adopted as the industry standard, it is described in detail in this issue by Reed [25].

References

[1] G. L. Archer, "History of Radio to 1926," The American Historical Society, New York, N. Y., 1938.
[2] R. A. Heising, "The audion oscillator," *J. AIEE*, April–May, 1920.
[3] ——, "Modulation in radiotelephony," Proc IRE, vol. 9, pp. 305–352; August, 1921.
[4] J. A. Chambers, L. F., Jones, G. W. Fyler, R. H. Williamson, E. A. Leach, and J. A. Hutcheson, "The WLW 500 kilowatt broadcast transmitter," Proc. IRE, vol. 22, pp. 1151–1180; October, 1934.
[5] L. E. Barton, "High power audio from relatively small tubes," Proc. IRE, vol. 19, pp. 1131–1149; July, 1931.
[6] ——, "Recent developments of the class B audio and radio frequency amplifiers," Proc. IRE, vol. 24, pp. 985–1006; July, 1936.
[7] H. Chireix, "High power outphasing modulation," Proc. IRE, vol. 23, pp. 1370–1392; November, 1935.
[8] W. H. Doherty, "A new high efficiency power amplifier for modulated waves," Proc. IRE, vol. 24, pp. 1163–1182; September, 1936.
[9] S. Ballantine, "High quality radio broadcast transmission and reception. Part I," Proc. IRE, vol. 22, pp. 564–629; May, 1934.
[10] *Ibid.*, part II, vol. 23, pp. 618–652; June, 1935.
[11] G. H. Brown and H. E. Gihring, "General considerations of tower antennas for broadcast use," Proc. IRE, vol. 23, pp. 311–356; April, 1935.
[12] G. H. Brown, " A critical study of the characteristics of broadcast antennas as effected by antenna current distribution," Proc. IRE, vol. 24, pp. 48–81; January, 1936.
[13] L. H. Kahn, "Compatible single sideband," Proc. IRE, vol. 49, pp. 1503–1527; October, 1961.
[14] E. H. Armstrong, "A method of reducing disturbances in radio signalling by a system of frequency modulation," Proc. IRE, vol. 24, pp. 689–740; May, 1936.
[15] M. G. Crosby, "Frequency modulation noise characteristics," Proc. IRE, vol. 25, pp. 472–514; April, 1937.
[16] ——, "Frequency modulation propagation characteristics," Proc. IRE, vol. 24, pp. 893–913; June, 1936.
[17] ——, "Carrier and sideband relations with multitone frequency or phase modulation," *RCA Rev.*, vol. 3, pp. 103–106; July, 1938.
[18] ——, "Communication by phase modulation," Proc. IRE, vol., 27, pp. 126–136; February, 1939.
[19] ——, "The service range of frequency modulation," *RCA Rev.*, vol. 4, pp. 349–371; January, 1940.
[20] —— , "Reactance tube frequency modulators," *RCA Rev.*, vol. 5, pp. 89–96; July, 1940.
[21] ——, "Bandwidth and readability in FM," *RCA Rev.*, pp. 363–370; January, 1941.
[22] A. P. Walker, "Engineering performance of six proposed stereo systems," *Electronics*, pp. 85–89; November, 1960.
[23] A. H. Bott, "The effect of SWR on cross modulation of FM multiplexed signals," *Broadcast Engrg.*, vol. 3, pp. 4–38; August, 1961.
[24] R. Guy and R. Morris, "NBC frequency modulation field test," *RCA Rev.*, vol. 5, pp. 190–225; October, 1940.
[25] O. Reed, Jr., "Broadcasting developments now taking place," this issue, p. 837.

Television Broadcasting*

CLURE H. OWEN†, MEMBER, IRE

Summary—The story of television broadcasting may be divided into three periods of time, *i.e.*, the historical period—to about 1930, the developmental period—to the end of the Second World War, and the commercial period.

One of the first public demonstrations was given by C. Francis Jenkins in Washington, D. C. in 1925. During the 1930's the mechanical scanning equipment was replaced with all electronic equipment. The first "network" or long-distance pickup was the 1940 Republican National Convention in Philadelphia which was broadcast by the NBC station in New York and the General Electric station in Schenectady.

Commercial television was authorized by the Federal Communications Commission, effective July 1, 1941; however it was not until shortly after the end of the Second World War when the electronics industry returned to peace-time conditions that the public became aware of the magic of television in the home. In a relatively short period of time the television broadcasting industry developed from a few stations with very limited programming to a national industry with more than 600 television stations and over 50,000,000 receivers in use by the public. The impact of television has affected the lives of nearly every citizen of this country and its effects are being felt in most of the other countries of the world.

THE STORY OF television broadcasting, even when limited to the technical aspects, cannot be told by one person or contained in a single volume. Accordingly only the highlights can be touched upon in this brief article. Television broadcasting can roughly be divided into three periods of time, *i.e.*, the historical period up to about 1930, the developmental period from the early 1930's to the end of World War II, and the postwar commercial period.

Historical Period

For ages man has dreamed of transmitting sight and sound from one place to another. One of the milestones along the road to television was the isolation of selenium by the Swedish scientist Baron Berzelius in 1817. While the light-sensitive properties of selenium were known for some years, little application of this property was made until 1892 when Elster and Geitel devised a photoelectric cell [1]. This cell is the basic principal upon which television, as we know it today, is based. Another milestone, that of scanning and reproducing an image, was made by Paul Nipkow, a German experimenter, when he invented the scanning disc in 1884. It was Nipkow's concept of changing the picture into electrical bits at the transmission end of a circuit, the sequential transmission of these bits and the reassembly at the receiving end into a complete picture, which makes it possible to transmit a television picture over a single telephone or radio frequency circuit.

The invention of the triode vacuum tube by Dr. Lee deForest in 1906 was one of the greatest achievements in the advancement of communications including, of course, both radio and television. In the early 1920's J. L. Baird in England and C. F. Jenkins in the United States carried on separate experiments using mechanical scanning. The mechanical system of scanning had very definite limitations on the amount of information that could be transmitted (picture quality) and the size of the reproduced image. The appeal of an electronic system without the above limitations was evident by the work of several of the early experimenters. M. Dieckman and G. Glace received a German patent in 1906 and B. Rosing of St. Petersburg received a patent in 1907; both for systems of television using a cathode-ray tube receiver [2]. The first suggestion of an all electronic television system was made by A. A. Campbell-Swinton in 1908 [2]. His suggestion of the camera tube included a multiplicity of photoelectric cells in parallel so that each cell could store a charge until scanned and would deliver only one electric pulse for each picture period. In 1923, V. K. Zworykin filed a U. S. patent for an electronic camera tube known as an "Iconoscope." This tube differed from earlier electronic camera tubes in that it employed beam intensity modulation by an axially symmetric grid. One of the first public demonstrations of television or radio vision, as it was then called, was given by Jenkins on June 13, 1925 when a live picture was transmitted from the Naval Air Station at Anacosta, Maryland, to the Jenkins Laboratories in Washington, D. C., a distance of some miles. A mechanical system of scanning was used in this demonstration [3].

In 1929 Zworykin attained success with an electrostatic cathode-ray picture display known as a "Kinescope." By the early 1930's most work on systems of scanning had changed over to one of the electronic systems. In the United States Philo T. Farnsworth and Vladimir Zworykin headed parallel investigations which by 1933 proposed equipment to make electronic television possible. Other persons making contributions to television about this time were Ives of AT&T, Alexanderson of GE, Goldmark of CBS, Engstrom and Goldsmith of RCA and DuMont [4]. Experimental work was also being carried on in Germany using cathode-ray tubes in receivers [5].

Developmental Period

It is difficult to draw a line of demarcation between the experimental period and the developmental period as there was, during the 1930's, a gradual transition from an occasional equipment test to more frequent,

* Received by the IRE, June 29, 1961; revised manuscript received, November 20, 1961.
† American Broadcasting Company, New York, N. Y.

and in some cases more or less regular, program operations. One of the earliest and most important stations in the development of television was the RCA station W2XBS in New York which went on the air with mechanical scanning equipment in 1928.

During 1931 and 1932 field tests were made with the transmitter located at the Empire State Building [6]. The transmitter used 120-line, 24 frames/sec mechanical scanning. The receivers for these tests used cathode-ray tubes instead of mechanical scanning.

During 1933 a system was built in Camden, New Jersey which used the "Iconoscope" as a pickup device. The system operated on 240 lines, 24 frames/sec. The picture carrier was at 49Mc and the sound carrier was at 50Mc.

In 1934 the system operated with 343 scanning lines and used interlaced scanning. The frame frequency was 30/sec and the field frequency was 60/sec. The last mechanical element of the system was removed by the development of an electronic synchronizing generator.

In 1936 provisions were made for live and film studios at Radio City to be used with the Empire State Building transmitter. The picture carrier was operated at 49.75 Mc and the sound carrier at 52Mc. The two transmitter outputs were fed through coupling filters so that a common transmission line and antenna could be used. The receivers used in these tests were superheterodynes capable of receiving both the picture and sound simultaneously.

During the latter part of the 1930's regular studio programs were instituted (Figs. 1 and 2). At first most of the receivers were company-owned receivers in the homes of RCA-NBC employees; however, by 1940 it was estimated that there were about 3000 receivers in the New York area. On April 30, 1940 a regular public series of television programs was inaugurated with the opening of the New York World's Fair by President Roosevelt.

The Philco Radio and Television station W3XE-WPTZ [7] in Philadelphia carried on extensive broadcasts with remote pickups of outdoor events at Franklin Field. It used a high-frequency radio relay to transmit the programs from Franklin Field to the television station.

The General Electric Company's television station WRGB in Schenectady [7] was one of the early television stations carrying on both equipment and program development.

The Columbia Broadcasting System's station WCBW in New York had its transmitter located in the Chrysler Building and studios in the Grand Central Terminal Building. While this station carried on public interest programming its niche in history is on the basis of the equipment development and color television investigations under the capable leadership of Dr. Peter Goldmark.

The DuMont Laboratories' station WABD in New York [7] carried on both program and equipment development. The resourceful DuMont engineers, whose

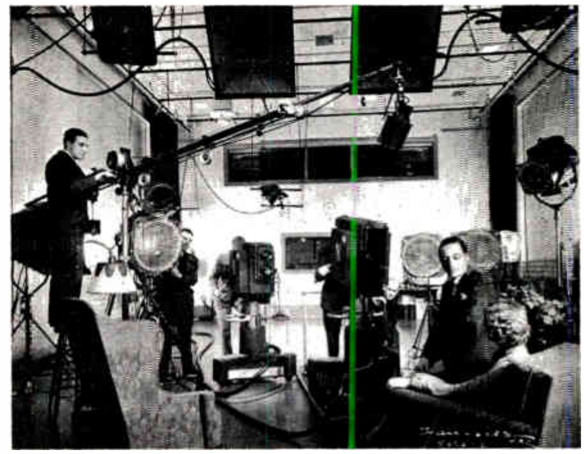


Fig. 1—NBC experimental television studio.

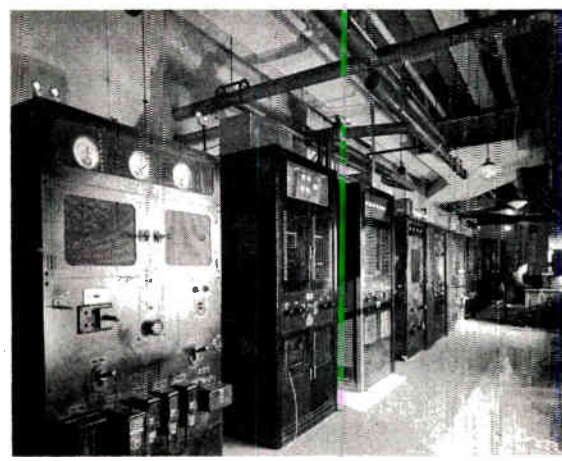


Fig. 2—NBC experimental television transmitter.

great experience in cathode-ray electronics made many contributions to television, produced some of the earliest large kinescope picture tubes.

The Don Lee Company's station W6XAO in Los Angeles [7] was instrumental in the development of interest in television on the West Coast. Station W6XAO carried on extensive tests of television reception in an airplane. Self-synchronized cathode-ray television receivers demonstrated that television reception was possible using power sources other than that common to the power source of the transmitter. A field strength contour map was prepared based upon reception of signals at 3000 feet [8].

The Paramount stations WBKB in Chicago and W6XYZ in Los Angeles [7] were more concerned with the development of programming methods and the production of entertainment than the development of equipment.

The Zenith station W9XZV-WTZR in Chicago [7] carried on small-scale but intensive development of equipment and field testing of television receivers.

The Farnsworth station W2XPF [7] originated on the West Coast but was later moved to Philadelphia. This organization under the guidance of the youthful electronics genius Philo T. Farnsworth made many contributions in the field of television, most noteworthy of which was the "image dissector" camera tube.

In 1936 the British Broadcasting Corporation opened a London transmitting station at Alexander Palace. Two systems were used for television transmission. The Baird system used 240 lines, 25 pictures/sec with sequential scanning. The Marconi-E.M.I. system used 405 lines, 25 pictures/sec, interlaced, to give 50 frames/sec. The Marconi-E.M.I. system used cathode-ray-type pickup for both direct and film transmission [9].

The first "network" or long-distance television pickup [10], [11] was the 1940 Republican National Convention held in Convention Hall in Philadelphia. It was broadcast by the NBC station W2XBS in New York and the GE station W2XB in Schenectady. The pickup was made by NBC field pickup equipment and transmitted by wire lines from the Convention Hall to the Bell System carrier terminal in Philadelphia, to the RCA Building in New York by coaxial cable, to the Empire State Building by wire lines and to the Helderberg mountain site of W2XB by an off-the-air pickup of W2XBS in New York. The wire lines used from the Convention Hall to the Bell System carrier terminal in Philadelphia and from the RCA Building to the Empire State Building in New York were carefully selected telephone circuits with wide-band repeaters and equalizers installed at the terminals, as well as at intervals along the circuits. The coaxial cable used was the experimental coaxial cable installed between New York and Philadelphia by the Bell System for the study of multiple carrier telephone transmission. The off-the-air pickup at the Helderberg mountain site near Schenectady was made possible by a specially constructed receiving antenna system.

Commercial Television

Commercial television was authorized by the Federal Communications Commission, effective July 1, 1941; however, the start of the Second World War shortly thereafter effectively curtailed any rapid development of commercial television. Following the conclusion of the war the television broadcasting industry developed at an unprecedented rate. With only a few low-power television broadcasting stations on the air and a small number of receivers in use the public suddenly became aware of the magic of television in the home. The television home in each neighborhood was a popular meeting place with neighbors, both adults and children, streaming in to watch the new medium of entertainment. Although the size of the television screens was small and the technical quality and program fare inferior by latter-day comparison, the public interest generated by this new medium far exceeded the most optimistic opinions of the "experts." The neighbors soon desired receivers in their own homes and the sale of receivers sky-rocketed. The popularity of television resulted in an increased number of stations until 1948 when the Federal Communications Commission stopped granting authority for additional stations pending a reallocation to minimize interference between stations. The television "freeze" imposed in 1948 was not lifted until 1952 when again new stations were permitted to start broadcasting and the total number of television stations rapidly increased. Table I shows the growth in the number of television stations authorized and the television receivers in use by the public from 1946 through 1960.

The impact of television upon the home, advertising, entertainment, etc., has been the subject of vast quantities of written material. For this article it is sufficient to say that the impact of television has affected the lives of nearly every citizen of this country and its effects are being felt in most of the other countries of the world.

TABLE I

	Commercial Television Stations Authorized* [12]	Television Receivers In Use [13]
1946	30	5,000
1947	66	150,000
1948	109	1,010,000
1949	117	3,660,000
1950	109	9,785,000
1951	109	15,590,000
1952	108	21,460,000
1953	483	26,920,000
1954	573	32,750,000
1955	582	37,360,000
1956	609	42,810,000
1957	651	46,690,000
1958	665	49,900,000
1959	667	53,290,000
1960	653	56,210,000

* At the close of each fiscal year, June 30.

Television Standards

The transmission and reception of television pictures require that the image be electronically disassembled at the transmitting end, transmitted bit-by-bit and reassembled in the correct order at the receiving end. For the receiver to correctly reassemble the picture it is necessary that the transmitter and receiver use the same set of transmission standards, *i.e.*, the same number of lines per picture, the same number of pictures per second, etc. To keep the received picture in step with the transmitted picture, synchronizing signals are transmitted.

During the experimental period of television various numbers of lines per picture and pictures per second were used, with as low as 30-line pictures scanned at a frequency of 5 pictures/sec. To obtain better definition of picture detail the number of lines per picture as well as the number of pictures per second were increased from time to time. During 1936 the RMA Television Committee recommended standards which included 441 lines, frame frequency of 30/sec and field frequency of 60/sec, interlaced. During 1941 the Federal Communications Commission adopted the 525-line standards proposed by the National Television Systems Committee.

Space does not permit the detailed discussion of all the standards of television transmission. The reader is referred to the Rules and Regulations of the Federal Communications Commission for details.

Unfortunately the various countries of the world have adopted different television standards. This, as might be expected, makes for some difficulty in transmitting television programs between countries. In transmitting live or taped television programs between countries where different television standards are used, various standards converters have been employed. In general these standards converters have been based upon the photographic principle, *i.e.*, the picture presentation on a kinescope monitor operating on one set of standards would be photographed by a television camera operating on a different set of standards. The resultant pictures suffered from loss of detail and often had some visible flicker. The American Broadcasting Company has developed an all electronic standards converter which has no optics in the system and has no perceptible picture degradation (Fig. 3). Table II gives the principal television systems of the world.

In 1953 the Federal Communications Commission adopted standards for color television. The standards adopted were the result of several years' work by the National Television Systems Committee. The color standards are compatible with the monochrome standards in that a monochrome receiver is able to receive a black and white picture when the picture is transmitted in color for color-equipped receivers. The color information is transmitted on an amplitude modulated subcarrier at about 3.58 Mc above the picture carrier.

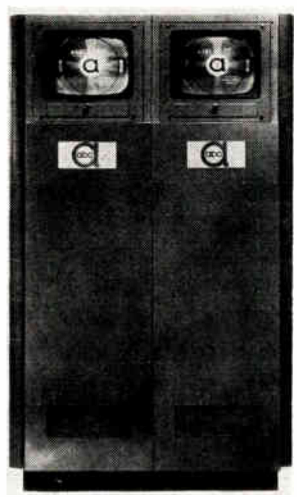


Fig. 3—All electronic standards converter.

TABLE II

PRINCIPAL TELEVISION SYSTEMS OF THE WORLD [14]

Description Used in Foreign Station Directory	British	World-Wide	Western Europe	Eastern Europe	French
Location of Principal Use	United Kingdom	Western Hemisphere, Far East	Western Europe	Soviet orbit	France and Possessions
Number of Lines Per Picture	405	525	625	625	819
Video Bandwidth (Mc)	3	4	5	6	10.4
Channel Width (Mc)	5	6	7	8	14
Sound Carrier Relative to Vision Carrier (Mc)	−3.5	+4.5	+5.5	+6.5	11.15*
Sound Carrier Relative to Edge of Channel (Mc)	+0.25	−0.25	−0.25	−0.25	0.10*
Interlace	2/1	2/1	2/1	2/1	2/1
Line Frequency (cps)	10,125	15,750	15,625	15,625	20,475
Field Frequency (cps)	50	60	50	50	50
Picture Frequency (cps)	25	30	25	25	25
Vision Modulation	+	−	−	−	+
Level of Black as Percentage of Peak Carrier	30	75	75	75	25
Sound Modulation	AM	FM	FM	FM	AM

* French standards invert the video and audio frequencies for certain channels.

Television Equipment

The experimental and developmental periods of television permitted a determination of the basic fundamentals of television equipment. However, commercial television resulted in demands for equipment which exceeded all expectations. These demands were not restricted to greater quantities of existing equipment but were also for specialized equipment to permit greater flexibility of operations, greater reliability, ease of maintenance, etc.

Competition between stations resulted in a necessity for technical quality control of broadcast signals which would permit the best possible pictures being received by the public. It is not the purpose of the paper to comment on all types of television equipment. However, comments follow on the basic types of equipment such as live cameras, projection, lighting, switching, video recording and transmitting equipment.

Live Television Camera

The live television camera, which can be considered the "eye" of television, is one of the most important single elements in the line-up of the equipment system. With the exception of film or tape reproduction of programs the live television camera serves as the "front door" for the entrance of all television signals into the broadcasting system.

The image orthicon camera is generally considered as the workhorse of the industry for studio use. Its ability to satisfactorily handle scenes with a wide range of light levels ranging from only a few foot-candles for mood scenes to some 300 foot-candles for high lights, is desirable for general studio use.

The vidicon camera has found universal use as a film camera, however, as a live camera its use has been limited. The target, being photoconductive rather than photoemissive as the image orthicon, has a higher sensitivity; as a result the size of the tube as well as the associated camera is much smaller.

Without the electron multiplier and shading control and with less critical beam current control, the operation of the vidicon camera is less critical than the image orthicon camera. The signal-to-noise ratio of the vidicon is more dependent upon light levels than the image orthicon. Accordingly, for general studio use, light levels above 200 foot-candles are required for satisfactory operation. The vidicon is being extensively used at present for unattended closed-circuit television, also for limited area scenes such as newscasts, flip-cards, etc. (Fig. 4).

Live color television cameras have three image orthicon cameras, one for each of the primary colors, red, green and blue. The light from the scene being televised is focused through a lens common to all of the tubes after which it is split three ways by means of dichroic mirrors and light filters. Each color channel has its own amplifiers and color correction circuitry.

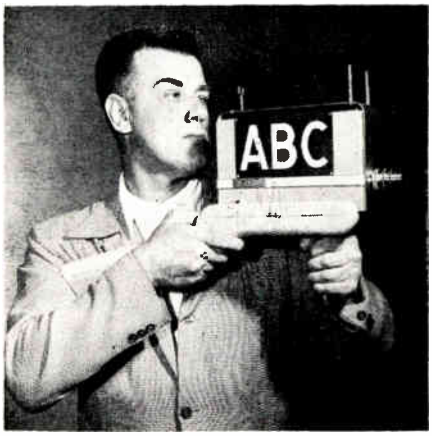


Fig. 4—Vidicon camera.

Film Projection Equipment

Film has always played a very important part in television programming. At first the film used was restricted to motion picture films. In later years more and more of the films used were made especially for television broadcasting. Commercial television started with the use of the iconoscope film cameras. Today the vidicon film cameras have almost completely replaced the old iconoscope cameras with a marked improvement in technical quality. Both 35-mm and 16-mm film projectors are used in the telecine rooms although many of the television stations in smaller markets use only 16-mm film projectors. The film projectors are sometimes used with an individual camera; however, most stations use multiplexers for selecting, by means of movable mirrors, any one of three or four light sources, *i.e.*, film or slide projectors, for use with one film camera chain.

Color television film projection requires the use of three vidicon cameras, one for each of the primary colors, red, green and blue. The light being projected through the color film is separated by appropriate dichroic mirrors and color filters. A separate lens is used for each vidicon film camera.

Lighting Equipment

Artificial lighting is necessary for practically all studio television program pickups. Outdoor scenes with adequate daylight can usually be picked up with an image orthicon camera without additional lighting.

In the early days of television the relatively insensitive iconoscope camera tube required very bright illumination of the subject to be photographed. To obtain adequate pictures light levels of 800 to 1200 foot-candles were needed. There were several disadvantages to these high light levels, such as actor discomfort, high power costs, the great number of lighting fixtures, and tremendous heat loads which made it essential to air-condition the studio.

With the greater sensitivity of the image orthicon camera tube the problem of studio lighting was greatly simplified. With the exception of mood scenes most studio illumination runs from 50 to 150 foot-candles. The lighting of television scenes is an art which involves not only the placement of the light fixtures but the regulation of the light output of the individual lights to obtain the desired artistic effect. The regulation of light output is accomplished by dimmers to control the voltage applied to the lamp filaments. The dimming circuits for the individual lamps vary from the old-fashioned series resistance, through the autotransformers, grid-controlled thyrathron rectifiers to the more modern solid-state silicon controlled rectifier dimmers with punch-card preset control (Fig. 5). The latter system permits almost instantaneous change from one lighting setup to the next.

Video Switching Equipment

In the early days of television the video switching requirements were relatively simple as it was only necessary to switch between cameras or from live to projection equipment. As great strides have been made in television programming great demands have been made upon the switching equipment. This has resulted in complex equipment for handling a large number of inputs and outputs as well as special effects and superimposition and montage of a portion of one scene as an insert within another scene.

Video Recording Equipment

The first type of video recording was accomplished by focusing a film camera on the face of a kinescope monitor. The film was then developed and the required number of prints made. This type of video recording has been used for network film distribution since the early days of commercial television. The principal users of kine recordings are stations which do not have live network connections or which desire to schedule the program at other than the time it is carried on the network.

Video tape recording equipment permits immediate playback of a program and is capable of higher technical quality than kine recordings. Video tape recording has added a third source of television programming to the previous live or filmed sources. The use of video tape equipment permits a more efficient use of studio space, as entire programs, selected scenes, or commercials can be prerecorded and inserted at the desired time.

Transmitting Equipment

Television transmitters are somewhat different from transmitters in other services as it is necessary to transmit both the aural and visual signals within the 6-Mc television channel. This is normally accomplished by two separate transmitters, one carrying the aural signals and the other carrying the visual signals. The outputs of both transmitters are generally fed into a diplexing circuit for exciting a single antenna with both aural and visual signals (Fig. 6).

The high-frequency propagation of television signals requires that the transmitting antennas be located as high as possible above the surrounding area to provide maximum service. Accordingly most transmitting antennas are located on mountain tops or tall buildings. Furthermore, due to the directional characteristics of television receiving antennas it is advantageous to have all of the television transmitters in the same area. Fig. 7 shows the transmitting antennas on the Empire State Building where all seven of the New York television stations are located.

Fig. 5—Solid-state silicon controlled rectifier dimmers with punch-card preset control.

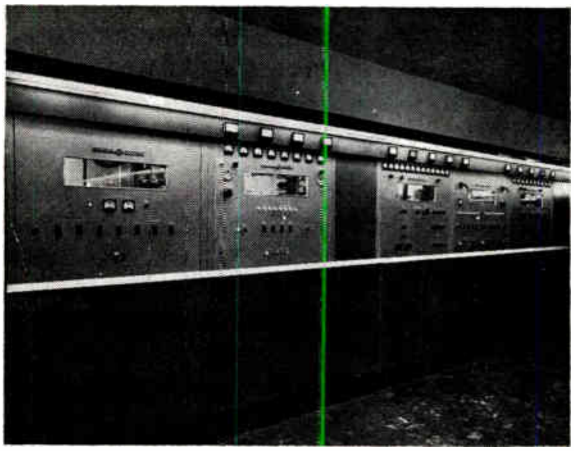


Fig. 6—50-kw television transmitter.

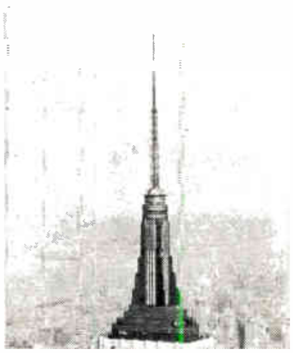


Fig. 7—The Empire State Building tower with seven television transmitting antennas.

Common Carrier Transmission of Television Programs

The common carrier facilities of the telephone companies are the arteries which carry the life-blood television programs from the network control centers to the affiliated stations across the country. The transmission of television programs, as distinguished from radio programs, involves the transmission of both the sound programs and the picture programs. Technically these could be transmitted together over the same circuits. In the United States, however, due to the existence of considerable audio facilities for more than two decades of radio network transmission, the aural and visual programs are transmitted over separate circuits.

Liaison between the television networks and the telephone companies is maintained by two special committees in addition to the normal operating contacts. The Video Transmission Engineering Advisory Committee (VITEAC) promotes a common understanding of the various factors of an engineering and policy nature as related to network service. The Network Transmission Committee (NTC) is the working subcommittee of the VITEAC.

It is the function of network facilities to deliver to numerous receiving locations a signal which is essentially unchanged from the signal which is applied at the originating point. To transmit audio signals satisfactorily, the facilities must transmit all frequencies with uniform loss and phase shift and must accommodate the complete range of loudness without distortion from overloading or interference from noise. For this reason they are gain and delay equalized, and maximum and minimum levels are established.

Similarly, video facilities are equalized for gain and delay and also for differential gain and phase, and operating levels are set. The frequency band required for video transmission extends from 0 to 4.5 Mc and the entire band is equalized essentially flat. Variations from flatness in the region from 0 to 250 kc result in streaking or smearing of the picture. In the high-frequency range different tolerances apply to color and monochrome transmission. For color television signals differential gain and phase require close control to avoid effects on the hue and saturation of the color signal. Color facilities in general are equipped with special equalizers which operate in the range of the color subcarrier and its sidebands. Circuits which are set up to handle color television transmission will also handle monochrome transmission.

In the early period of television network transmission coaxial cable was used for intercity circuits. New circuits are now using microwave radio facilities.

Community Antenna Systems

The location and service range of television stations in the United States and Canada were such that many of the people, particularly those residing in the mountainous areas, were unable to enjoy television reception. Due to the very great demand for television reception many community antenna systems were installed. In general these systems consisted of a high-gain television receiving system installed on a mountain top near the community, a radio frequency amplifier and a coaxial cable system for transmitting the television signals from the mountain-top location to the various homes in the community. Most of the installations have been able to receive and distribute the programs of several television stations. Some of the community antenna systems use a microwave circuit where off-the-air reception of television programs is not possible.

Television Booster and Translator Stations

Television booster and translator stations, in a manner comparable to the community antenna systems, are means of extending television service into underserved areas. In general these are low-power relay stations located on mountain tops or other areas of possible television reception. The signals are received and retransmitted, on the same television channel in the case of the booster stations, or on a different channel in the case of the translator stations, to the television receivers in the areas of desired reception.

Acknowledgment

The author wishes to express his appreciation to R. M. Morris of the American Broadcasting Company and to H. A. Ahnemann of the AT&T Long Lines Department for their assistance in the preparation of this manuscript.

Bibliography

[1] A. M. Glover, "A review of the development of sensitive phototubes," Proc. IRE, vol. 29, pp. 413–423; August, 1941.
[2] V. K. Zworykin, E. C. Ramberg, and L. E. Flory, "Television in Science and Industry," John Wiley and Sons, Inc., New York, N. Y., p. 9; 1958.
[3] C. F. Jenkins, "Radio vision," Proc. IRE, vol. 15, pp. 958–965; November, 1927.
[4] W. C. Eddy, "Television, The Eyes of Tomorrow," Prentice-Hall, Inc., New York, N. Y., p. 3; 1945.
[5] M. Von Ardenne, "An experimental television receiver using a cathode-ray tube," Proc. IRE, vol. 24, pp. 409–425; March, 1936.
[6] L. M. Clement and E. W. Engstrom, "RCA television field tests," *RCA Rev.*, vol. 1, pp. 32–40; July, 1936.
[7] W. C. Eddy, "Television, The Eyes of Tomorrow," Prentice-Hall, Inc., New York, N. Y., pp. 17–25; 1945.
[8] H. R. Lubcke, "Television image reception in an airplane," Proc. IRE, vol. 20, pp. 1732–1740; November, 1932.
[9] "Radio Progress During 1936: Television and Facsimile," Proc. IRE, (Rept. by Technical Committee), vol. 25, pp. 199–211; February, 1937.
[10] O. B. Hanson, "RCA-NBC television presents a political convention as first long distance pickup," *RCA Rev.* vol. 2, p. 267; January, 1941.
[11] M. E. Strieby and J. F. Wentz, "Television transmission over wire lines," *Bell Sys. Tech. J.*, vol. 20, pp. 62–81; January, 1941.
[12] Annual Reports of the Federal Communications Commission.
[13] "TV Factbook No. 32," Television Digest, Triangle Publications, Inc., Radnor, Pa., p. 24; 1961.
[14] *Ibid.*, p. 832.

Frequency Allocations for Broadcasting*

GEORGE R. TOWN†, FELLOW, IRE

Summary—**Any sound plan for frequency allocations for broadcasting must be based on engineering considerations. A broadcasting system—AM, FM or television—may be divided into its component elements, the characteristics of these elements pertinent to allocations may be analyzed, and a technically consistent allocation plan can be developed. The major components are the eye of the television viewer or the ear of the radio listener, the receiving installation, the existing electrical noise, the physical facts of wave propagation, and the transmitting installation. Historically, allocation plans were not developed solely through such engineering studies. AM broadcasting grew in a chaotic manner, and early attempts at governmental regulation were declared illegal by the courts. Remedial legislation resulted in the establishment of the Federal Radio Commission and its successor, the Federal Communications Commission (FCC) with legally delegated regulatory responsibilities and authority. Television and FM broadcasting, which came later, were always subject to governmental regulation. Television profited greatly from the high degree of cooperation at the engineering level which has existed between the FCC and the television industry. The practical administration of broadcast allocations depends upon other factors as well as engineering considerations. In analyzing the critical needs of the future, some of these non-engineering factors assume major importance.**

Introduction

THE Federal Communications Commission (FCC) has the responsibility for the allocation of radio frequencies in the radiated electromagnetic spectrum for all purposes except their use by the various agencies of the Federal government. Of the 762,886 stations authorized by the Commission on March 31, 1961, only 5699 were broadcasting stations—AM, FM and television. The allocation of frequencies for broadcasting is, however, one of the most troublesome of the many problems faced by the Commission.

The objectives of frequency allocation plans are "to maximize the service rendered, minimize interference, and ensure the fullest possible occupancy of the assigned channels."[1] To achieve these objectives, allocations must be based on sound engineering principles, which include both technological and economic factors. The first section of this paper is devoted to a statement of the general nature of the engineering factors which are basic to the development of a sound allocation plan. This is followed by a brief summary of the historical development of broadcasting allocations in this country. The third section of the paper will be concerned with some of the factors other than engineering which influence allocations for broadcasting. The paper will conclude with a brief discussion of the most critical problems now faced by those responsible for broadcasting allocations.

Engineering Basis for Allocations

Any sound allocation plan must be based on engineering considerations. Admittedly, factors other than engineering must also receive consideration; but unless adequate attention is paid to engineering factors, the plan will be far from optimum. The general engineering factors basic to allocations are the same for all types of broadcasting: AM, FM and television. These will now be discussed.

The ultimate purpose of a broadcasting system is to deliver a program to a viewer or listener. The final judge of the technical quality of the program is the eye or the ear of the consumer. Hence, any successful allocation plan must be based upon a study of the reaction of typical viewers and listeners to the controlled, varying technical quality of typical programs. In particular, it is necessary to determine on a sound statistical basis the minimum signal-to-interference level which is judged to provide satisfactory service. This determination must be made for all types of interference: thermal noise, impulse noise, and interfering carrier signals modulated in the same and in different manners as the desired signal. The interrelated nature of allocation studies and transmission standards is immediately evident, as the answers to the questions suggested above obviously depend upon the standards chosen, for example, the number of scanning lines per television picture, or the bandwidth of an AM channel.

The next step in an allocation study is to determine the critical characteristics of typical receiving installations, including antenna, lead-in and receiver. It is this equipment which serves to translate the signal at the receiving site into a picture or into sound. Here the most important factors include the gain of the antenna, the losses in the lead-in, the noise figure and sensitivity of the receiver, and the ability of the receiver to discriminate between desired and undesired signals. This latter factor includes such items as receiver selectivity, receiver rejection circuits (for example, adjacent channel traps and image rejection) and noise limiters. It also inherently includes a study of the noise-reducing effects permitted by the type of modulation employed. A major question is whether allocations should be based on the characteristics of new or of old receivers, or perhaps of hoped-for future designs; or on the characteristics of the best or of average receiving installations.

Information must next be obtained regarding typical electrical noise levels at the receiving antenna. Both

* Received by the IRE, June 12, 1961.
† Iowa State University, Ames, Iowa.
[1] D. G. Fink, "Television Engineering Handbook," McGraw-Hill Book Co., Inc., New York, N. Y., p. 2-2; 1957.

natural noise, such as static and galactic noise, and man-made noise, such as that from ignition systems and electric motors, must be considered. Typical noise levels are different in urban and rural areas, and these differences must be determined.

A knowledge of electromagnetic wave propagation phenomena is an essential ingredient in the development of an allocation plan. It is also the most complex of the technical factors basic to allocations. Until recently, there was not even general agreement as to how field strengths should be measured at VHF and UHF television frequencies and as to how the information regarding station coverage should be analyzed and portrayed.[2] Propagation studies must include a determination of the influence of such factors as transmitting- and receiving-antenna heights, terrain, climatic conditions, vegetation, time of day and of year and, of course, carrier frequency. Because of the many variables, results will usually be most useful when expressed in statistics which are probability functions both of location and of time. It is evident that two aspects of propagation must be considered. The first is propagation over the area served by a transmitter. The second is propagation over the greater distances at which a station provides a signal strength which is insufficient to give service but which is still sufficient to cause undesirably high levels of interference. Great differences exist between propagation at the medium frequencies used for AM broadcasting and at the VHF and UHF regions used for FM and television. The allocation objectives are, however, the same, namely, to provide an adequate ratio of desired to undesired field strength.

An allocation study must next consider the characteristics of practical transmitting installations, including transmitter, transmission line and antenna. It is here that economic, as well as technical, considerations become of real significance, although they are also of importance in studies of receiver performance.

There are, of course, many other engineering factors involved in addition to the major ones outlined above. These include such matters as the performance of directional transmitting antennas, the synchronization or the precise frequency offset of carriers, multipath interference (in television) and the manner of combining interferences from several sources. Finally, careful field studies, conducted in a statistically sound manner, must be made whenever possible to determine the over-all service provided by existing broadcasting facilities having known performance characteristics.

The synthesis of these various elements into a broadcasting system is easy to visualize in general terms. But the design of an allocation plan to maximize service and minimize interference is no simple matter. The basic engineering ingredients of such a design are, however, found in an analysis of the system components such as that which has been outlined, as the results obtained from this analysis permit the determination of such fundamental items as the minimum spacings of co-channel, adjacent channel and image-frequency stations of given power to produce a given statistical amount of interference in the customer's picture or sound program. Decisions on such basic matters as the quality of service which will be considered "normal" or "standard" are, however, not engineering in nature. They depend upon the judgment of the regulatory authorities. These, too, are essential ingredients in an allocation plan.

[2] H. T. Head and O. L. Prestholdt, "The measurement of television field strengths in the VHF and UHF bands," Proc. IRE, vol. 48, pp. 1000–1008; June, 1960.

Historical Development of Broadcasting Allocations

The dependence of an allocation plan on the characteristics of the individual components of a broadcasting system and the evaluation of existing or proposed plans by means of a study of these characteristics seem obvious today. This is the line of attack now taken by the FCC and by industry groups in allocation studies. Historically, however, allocations for broadcasting were not always made in this manner. A detailed study of the historical development of broadcasting allocations is most fascinating. Space permits only a brief outline which summarizes authoritative information available elsewhere.[3–10]

AM Broadcasting

The early days of radio broadcasting were marked by confusion, not only with respect to allocations themselves but also with respect to the question of whether or not the Federal government had authority to make allocations. In 1910, Congress passed the Wireless Ship Act, the first legislation pertaining to wireless telegraphy. Under the provisions of this act, which became effective on July 1, 1911, and which had as its purpose the protection of life at sea, the Radio Service of the

[3] L. F. Schmeckebier, "The Federal Radio Commission—Its History, Activities and Organization," Inst. for Govt. Res., Service Monographs of the U. S. Govt., No. 65, The Brookings Institution, Washington, D. C., pp. 1–38; 1932.

[4] C. M. Jansky, Jr., Address upon occasion of presentation of the Award for Distinguished Service to the Honorable Herbert Hoover, presented at the Annual Convention of the Natl. Assoc. of Radio and Television Broadcasters, Chicago, Ill.; April 9, 1957. Reprinted by Natl. Assoc. of Broadcasters, Washington, D. C., 1957.

[5] C. F. Lindsley, "Radio and Television Communication," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 23–28, 55–75; 1952.

[6] D. G. Fink, "Television Standards and Practice—Selected Papers from the Proceedings of The National Television System Committee and Its Panels," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 1–13; 1943.

[7] D. G. Fink, "Television broadcasting practice in America—1927–1944," *J. IEE*, vol. 92, pt. 3, pp. 145–164; September, 1945.

[8] D. G. Fink, "Color Television Standards—Selected Papers and Records of the National Television System Committee," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 1–40; 1955.

[9] "Allocation of TV Channels," Report of the Ad Hoc Advisory Committee on Allocations (E. L. Bowles, Chairman) to the Committee on Interstate and Foreign Commerce of the U. S. Senate, Govt. Printing Office, Washington, D. C.; March 14, 1958.

[10] E. W. Allen, "Television allocations problems," Proc. IRE, vol. 48, pp. 991–992; June, 1960.

Bureau of Navigation administered the laws pertaining to radio communication until 1927. An Act of Congress dated August 23, 1912, extended the authority of the Radio Service to regulate interstate radio communication and specified that such communication must take place at frequencies above 500 or below 187.5 kc. Broadcasting was first recognized by the Radio Service in a ruling dated December 1, 1921 (13 months after the broadcast of the Harding-Cox election returns by KDKA), in which the frequencies of 833 and 619 kc (wavelengths of 360 and 485 m respectively) were assigned to "limited commercial," or broadcasting, stations.

With the rapid increase in the number of broadcasting stations in 1921, it became evident that more effective governmental regulation was required to prevent intolerable interference. In 1922, Secretary of Commerce Herbert Hoover called the first of four annual industry-government conferences which "established the basic principles upon which our great American system of free (radio and television) broadcasting has been built."[4] This conference also set the pattern for the continuing cooperation which has since existed between governmental and nongovernmental organizations in developing allocation plans and transmission standards for broadcasting. As a result of these conferences, a third frequency, 750 kc, was allocated to broadcasting, and in 1923 the band of 550 to 1351 kc was established, with stations assigned to particular frequencies. Different classifications of stations were defined, based upon transmitter power, with different frequency bands allocated to the various classes. Rules were drawn up regulating the separation of stations according to geography and frequency, and a 10-kc channel was specified.

In the meantime, in 1923, the Court of Appeals of the District of Columbia had ruled that while the Secretary of Commerce had the right to assign an operating frequency to a station, he had no right to deny a license to any applicant. This was followed in 1926 by a ruling by the U. S. District Court for Northern Illinois that the Secretary of Commerce had no right "to establish regulations" for radio transmissions. This meant that, contrary to the previous ruling, the Secretary could not assign frequencies, and that a station could use any frequency it chose, within the limits established by the Act of 1912. This was confirmed on July 8, 1926, in a ruling by the Acting Attorney General, who recommended that remedial legislation be passed.

The effect of this chaotic situation was that Congress, which had failed to take action on a number of earlier pieces of remedial legislation, finally passed the Radio Act of 1927 which established the Federal Radio Commission (FRC). The act was signed by President Coolidge on February 23, 1927. Under the provisions of this act, and of subsequent extensions, the Commission was given authority to classify stations, to assign frequencies, to specify transmitter power and to make regulations to minimize interference. All types of interstate and foreign radio communication (except governmental) were placed under the purview of the Commission. The act stated that actions of the Commission should be taken in the interest of "public convenience, interest or necessity," a phrase which has been widely quoted in interpreting the powers of the FRC and its successor, the FCC. It should be noted that the act stated that it provided "for the use of channels, but not the ownership thereof," by station licensees.

One of the first actions of the FRC was a sweeping re-allocation of frequencies for broadcasting. In 1928, Congress passed the so-called Davis Amendment which directed the FRC to allocate broadcasting licenses, powers and frequencies equitably among the states, in proportion to population and area. This resulted in the development of a very complex allocation scheme which was doomed to ultimate failure by its very complexity. It did, however, permit the establishment of an orderly procedure for the growth of radio broadcasting and "by 1930, broadcasting had been brought to the level of functional public service."[5]

In 1934, after five years of discussion, Congress passed the Communications Act of 1934 establishing the Federal Communications Commission, which took over the functions of the FRC and other responsibilities relating to interstate and international wire and radio communications. With respect to broadcasting, the new act followed the general principles of the Act of 1927, particularly in its delegation to the FCC of complete authority over the licensing and allocation of frequencies.

In this account of the development of Federal authority over frequency allocation, only standard AM broadcasting has been considered. It was the only type of broadcasting of any significance at the time of the adoption of the Communications Act of 1934. From the standpoint of allocations, AM broadcasting has not changed greatly since that time, except for a slight extension of the broadcast band (now 535 to 1605 kc) and a trend toward an ever increasing number of stations with, as viewed by at least some engineers, an ever increasing amount of interference. The FCC, in contrast with the FRC, has been primarily concerned with FM and television broadcasting, rather than with AM broadcasting. These newer systems will also be the concern of the remainder of this paper.

It must be noted that during the first 13 years of government regulation (1921–1934), several basic principles were established which are now taken for granted. These include the private (as contrasted with public) ownership of our broadcasting system; government regulation through licensing, frequency allocation, and specified transmission standards; and the recognition of the frequency spectrum as a public resource, with portions allocated, but not given, to licensed stations. These principles, once established for AM broadcasting, have been extended to FM and television without question.

Television Broadcasting

Television was the second broadcast service to appear on the scene. The historical development of allocations for television has been traced so thoroughly, and at the same time succinctly, by Fink[7] and especially by Allen,[10] that repetition here is needless. It is sufficient to recall that experimental television broadcasting in the United States commenced in 1927; that the first stations had an almost unlimited choice of frequencies above 1500 kc; that in 1929, the FRC set aside for television five channels each 100-kc wide between 2000 and 2300 kc and between 2750 and 2950 kc; that licenses for experimental television in the region between 43 and 80 Mc were granted in 1931; that in 1936, permission for operation in the 2000-kc region was withdrawn; that in 1937, the FCC authorized 19 VHF channels each 6-Mc wide between 44 and 294 Mc; that in 1945, following World War II, the FCC revised the allocation plan to provide 13 (and later only 12) VHF channels plus an additional space 440-Mc wide for experimental operation in the UHF region; and that a comprehensive allocation plan providing the present 12 VHF channels (54–72, 76–88 and 174–216 Mc) and 70 UHF channels (470–890 Mc) was issued by the FCC in its Sixth Report and Order in 1952. Rather than discussing these developments in detail, it is more worthwhile to note one very significant aspect of the manner in which television allocation plans and transmission standards have been developed. This is the extensive participation, at the engineering level, of the television industry in this activity for which the Federal government has primary responsibility. The FCC and the industry have not always agreed. In fact, there have been several sharp disagreements. But the joint cooperative efforts have on the whole resulted in mutual understanding and the development of a reasonably satisfactory nation-wide television service.

In 1929, the Radio Manufacturers Association (now known as the Electronic Industries Association) formed its Television Standards Committee which in 1931 recommended the allocation of VHF channels to television broadcasting and which in 1936 and 1938 recommended transmission standards that form the basis for modern television broadcasting. Lack of complete unanimity in the industry led the FCC to refuse to give its expected approval to these standards. The Commission did, however, cooperate in the formation of the National Television System Committee (NTSC) composed of engineers representing all parts of the television industry. In a period of 12 months, this organization developed slightly modified standards which won the approval of the industry and the FCC.

In 1943, the electronics industry and the Institute of Radio Engineers, at the request of the FCC, formed the Radio Technical Planning Board (RTPB) to develop plans for post-war electronic activities. This Board included panels on AM, FM, and television broadcasting. One of the recommendations of the Television Panel in 1944 was that 30 channels be set aside in the VHF region for television broadcasting, starting at about 40 Mc and extending upward in a continuous band. The FCC actually allocated only 13 channels, because of the needs of other services. The RTPB recommendations were, however, realistic in terms of today's needs.

In June, 1948, the Joint Technical Advisory Committee (JTAC) was formed by the electronics industry and the IRE to make a study for the Commission of the use of UHF for television broadcasting. This was followed in 1952 by the publication of a comprehensive JTAC treatise on allocations.[11]

In the meantime, a second NTSC was formed in 1950 to develop acceptable standards for color television and an acceptable VHF-UHF allocation plan. A second group, known as the FCC Ad Hoc Committee on Propagation, composed of representatives of government and industry, advised the Commission on wave propagation problems. These groups played a large part in conducting the basic engineering studies leading to the Sixth Report and Order. This Order, however, disregarded an allocation plan of real merit, which minimized the intermixture of VHF and UHF stations, which had been submitted by the Allen B. DuMont Laboratories.

Finally, in 1956, the television industry, at the request of the FCC, established the Television Allocations Study Organization (TASO) to make a comprehensive study of the engineering factors underlying the allocation of frequencies for television broadcasting.[12] The two TASO reports[13,14] and a series of IRE papers[15] present the results of this contribution by the television industry to the work of the Commission.

This very incomplete summary of the cooperation between the television industry and the FCC on matters pertaining to television broadcasting indicates a most unusual and a most significant type of high-level joint activity between a governmental regulating agency and the industry which it regulates.

FM Broadcasting

Space does not permit a detailed discussion of FM allocations. Major Armstrong's pioneer FM station at Alpine, N. J., operated experimentally at a frequency of 42.8 Mc, although he also conducted experimentation at 110 Mc. As a result of hearings held in 1940, the FCC allocated channels 200-kc wide between 42 and 50 Mc to FM broadcasting. The RTPB recommended a mini-

[11] "Radio Spectrum Conservation," a Report of the Joint Technical Advisory Committee, McGraw-Hill Book Co., Inc., New York, N. Y.; 1952.

[12] G. R. Town, "The Television Allocations Study Organization—A summary of its objectives, organization and accomplishments," Proc. IRE, vol. 48, pp. 993–999; June, 1960.

[13] "Engineering Aspects of Television Allocations," Report of the Television Allocations Study Organization to the FCC, G. R. Town, Ed., TASO, Washington, D. C.; March 16, 1959.

[14] "Engineering Aspects of Television Allocations—II," Supplementary Report of the Television Allocations Study Organization to the FCC, G. R. Town, Ed., TASO, Washington, D. C.; June 13, 1960.

[15] Series of 16 papers in the "TASO Issue" of Proc. IRE, vol. 48, pt. I, pp. 991–1121; June, 1960.

mum of 80 to 100 channels (or a band of 16 to 20 Mc) in the vicinity of 40 Mc for FM broadcasting. The FCC, however, felt that space in this region was needed more urgently for television and other services and it developed a series of alternative FM allocation plans, finally choosing the present allocation of 100 channels extending from 88 to 108 Mc. This choice was made after lengthy discussions with the industry and in spite of vigorous opposition by a rather large segment of the FM industry.

Non-engineering Factors

It would be foolish to pretend that engineering factors are the only ones which influence the allocation of frequencies for broadcasting. For example, the practical economic fact that UHF television stations have difficulty in competing in an intermixed VHF-UHF area has led to a demand for more VHF television channels, while UHF channels are plentifully available but unwanted. Again, the increasing interest in educational television is expected by many to result, in the not far distant future, in a demand for more channels for this service, with a consequently greater scarcity of channels for commercial television. This demand may possibly, however, be met largely by UHF assignments.

The policy of the FCC of maintaining practically complete protection of the service area of an existing station from intrusion by an interfering signal from a new station can lead to a scarcity of channels. In some cases, a greater total population might be served under a different policy. Allied to this policy, is the practice of minimizing changes in frequency assignments to existing stations. The licensed stations do not legally have perpetual claim to a frequency assignment or to a license, but the FCC seldom makes major changes. In some instances, changes might be beneficial from an over-all point of view. These policy considerations are not mentioned as matters which necessarily should be changed, but merely as examples of non-engineering factors which definitely do affect allocations.

Finally, the Commission has a difficult task in maintaining close liaison with the industry which it regulates and still maintaining freedom from undue pressure from that industry. The Commission has a record of high integrity in dealing with this problem. It appears to an outsider that it must have an even more difficult task in maintaining its freedom from political pressures, especially from influential members of the government whose constituents wish to have favorable attention in allocation matters.

Critical Needs

If one were asked to name the most critical allocation need, the most common answer would be, more television channels, especially more VHF channels. The other broadcasting services have similar needs, but the question will be discussed in terms of television. And it may be that the above answer is not correct.

The first real need is to determine rather accurately just what service is being provided now. This cannot be determined in an office, but can be found only through extensive field work. The work of the TASO Field Test Panel[16] pointed the way toward the type of activity which is needed.

Second, there needs to be a serious study of just what type of service the public really needs. The purpose of broadcasting is to serve the public, not to serve the broadcaster. Should a broadcasting system serve the maximum area or the maximum number of people? Is it essential that every viewer receive television programs from all three (or more) networks? Does any viewer really need to receive as many as seven local television stations? What are the real needs of educational television? These, and similar, questions must be answered before a truly logical allocation plan can be developed.

Finally, after these first two studies have been made, an optimum allocation plan will need to be developed on a completely objective, and nonpolitical, basis. This would make use of all existing engineering techniques, such as directional transmitting antennas, to maximize service. With the use of modern computers, the development of an optimum plan, taking into account actual, existing geographic conditions, would not be an insurmountable task. This optimum plan, once developed, obviously would not be used, at least *in toto*. It would, however, serve as a standard against which other allocation plans could be compared and it would thus be of inestimable value. And it is only after the development of such an optimum plan that the extent of the real need for more channels can be determined.

[16] C. M. Braum and W. L. Hughes, "Studies of correlation between picture quality and field strength in the United States," Proc. IRE, vol. 48, pp. 1050–1058; June, 1960.

The Technology of Television Program Production and Recording*

JOHN W. WENTWORTH†, MEMBER, IRE

Summary—**The basic concepts for television origination equipment are determined by operational requirements, which have remained relatively constant throughout the history of television broadcasting. In brief, television cameras and related studio equipment are designed to permit the assembly of a smooth flow of picture and audio signals from multiple sources, including both "live" scenes and prerecorded material. The photography of kinescope images on motion-picture film was an early approach to practical television recording which still offers enough advantages that major engineering efforts continue to be placed on the development of improved film recording equipment. A great deal of the current television recording activity, however, involves the use of magnetic tape. The most common types of television tape recorders employ 2-in. wide magnetic tape, on which picture information is recorded by rotating video heads which produce tracks extending across the width of the tape. Four separate heads are mounted on the rotating headwheel, but switching circuits are employed to deliver a single continuous signal during the playback process. Through the use of FM techniques and precise servomechanisms, it is possible to accomplish television tape recording with extremely high picture quality.**

Introduction

THIS PAPER is offered as one span of a "bridge of understanding" between the broadcast engineering fraternity and the larger circle of IRE members whose professional activities have nothing to do with broadcasting. One of the major benefits of IRE membership is the opportunity to keep in touch with major developments in all branches of electronic engineering, although the ever-increasing complexity of the electronic arts tends to make communication between dissimilar specialists increasingly difficult. The author has assumed that most IRE members have at least a mild *interest* in the technology of the broadcast industry. Most engineers maintain a healthy professional curiosity about all electronic systems with which they come in contact, and nearly all American families are influenced by the broadcast industry. The specific objective of this paper is to convey to non-television specialists a basic understanding of the technical concepts involved in the production and recording of television programs.

Picture Origination

In the typical studio scene shown in Fig. 1, only a small fraction of the electronic equipment required for the production of television programs is visible from the studio floor. Some of the equipment is usually mounted in a nearby control booth, like that shown in Fig. 2, but still more equipment is mounted in unattended racks. The equipment may be subdivided into six basic categories: 1) camera chains for live pickup, 2) television film equipment, 3) synchronizing equipment, 4) video distribution and switching equipment, 5) audio origination equipment, and 6) intercommunications equipment. While there have been major innovations and improvements in the circuit details of the equipment in all categories, the basic concepts involved have remained relatively constant throughout the history of television. Before discussing the equip-

Fig. 1—View of a typical television studio in operation. (Photo courtesy of South Carolina ETV Center.)

Fig. 2—View of a typical control booth for a television studio, showing program director's position, video control center, and audio controls. (Photo courtesy of South Carolina ETV Center.)

* Received by the IRE, July 14, 1961.
† RCA, Camden, N. J.

ment itself, let us review briefly the production *objectives* and *problems* which influence the design of television studio equipment.

Basic Studio Operations

The technology of television program production has been greatly influenced by the "tyranny of the clock." Since *time* is the basic commodity broadcasters offer for sale, a major objective of television production facilities is the maintenance of a smooth, continuous flow of program material, usually accomplished by the *multiple-camera technique*. Except for the very simplest types of programs, television productions involve the use of two or more live pickup cameras, frequently supplemented by film cameras. The live cameras are so located that they command different views of the scene to be televised. While one camera is "on the air," another may be moved to a new position, or its objective lens may be changed (with the aid of a quick-change turret), to compose the next picture. Still shots from a slide projector or brief sections of motion-picture film can be used either to present pictorial information that cannot readily be set up in the studio, or to bridge the time intervals that may be required for more elaborate set or camera changes. The multiple-camera technique imposes special requirements on nonelectronic aspects of studio design, such as lighting and provision for rapid handling of sets.

The operator who stands behind a television camera has only two major responsibilities: 1) to compose the picture in accordance with instructions from the program director, and 2) to keep the picture in proper optical focus. (All other camera adjustments are of a setup nature only, or are made by remote control.) To aid the camerman, television cameras for live pickup are almost universally equipped with a viewfinder (a small television monitor displaying the camera's picture output), a lens turret with four lenses of different focal length (or a "zoom" lens), an optical focus knob, and a built-in intercom circuit permitting conversations with the program director. In addition, the camera is usually mounted on a dolly or pedestal which permits relatively free movement over the studio floor, as well as such basic movements as panning and tilting.

The Television Camera Chain

In addition to the camera proper on the studio floor, a complete television *camera chain* typically consists of a remote control panel, a signal processing amplifier, a "master monitor," and a regulated power supply. The remote control panel permits a video operator in a control booth to keep the circuits associated with the pickup tube in proper adjustment, to apply such electronic corrections to the signal as may be necessary, and to maintain a proper signal level at the output of the camera. The video operator is responsible for the technical quality of the picture in all respects except optical focus. The "master monitor" is an important tool for the video operator, providing both picture and waveform displays with special features for critical examination of picture quality. The *signal-processing amplifier* (sometimes called the *camera control unit*) serves two major functions: 1) it provides buffer amplification for the pulse signals which are transmitted outward along the camera cable to control the deflection circuits in the camera, and 2) it "cleans up" the camera output signal to make it conform to the desired signal specifications. In the special case of color television, the camera proper and the processing amplifier have *three* video channels for red, green, and blue signals, and the chain as a whole includes a *colorplexer* for encoding the three primary signals to form a single compatible signal.

Film Cameras

Television film cameras are specially designed to facilitate the generation of television signals from photographic materials. To preserve maximum quality, the photographic image is usually projected directly on the sensitive surface of the television pickup tube in the camera with no intermediate screen. Film cameras are frequently used with *optical multiplexers*, which employ movable mirrors or other means to combine the optical paths of several projectors. A very popular arrangement of television film equipment consists of a pair of 16-mm film projectors and a remotely-controlled slide projector mounted with an optical multiplexer as alternative inputs for a single film camera. Such a facility not only provides for film or slide inserts into live programs, but also permits smooth operation for prolonged periods with exclusively photographic material.

Synchronizing Equipment

All of the cameras, monitors, and other apparatus in a television studio system are kept in proper synchronism by pulses from a *sync generator*. One of the waveforms provided by the sync generator is known as *blanking*, and serves to define the retrace intervals during which no picture information is transmitted. The waveform commonly designated *sync* is a complex train of pulses with both horizontal-frequency and vertical-frequency components, timed to occur during the blanking intervals. This sync waveform is added to picture signals prior to transmission to provide timing control for the deflection circuits in distant television receivers. To assure continuity of service, sync generators are frequently installed in pairs with automatic changeover facilities.

Electronic Editing

The final assembly or editing of television programs involves the use of *video switching equipment*, controlled through pushbutton panels like those shown in Fig. 2. In very small-scale switchers, the push buttons actuate switches connected directly to video circuits, but in larger systems the push buttons are merely remote controls for relays or transistorized switching elements.

The switching operations are usually performed by the program director or his assistant with the aid of a bank of monitors displaying all the picture sources available.

A typical arrangmenet of switching equipment is shown in Fig. 3. Each vertical row of buttons corresponds to a single picture source, and each horizontal row corresponds to an output bus. One such output bus feeds the outgoing line directly, and is designated the "program bus." The other buses provide signals for several types of effects equipment or for preview monitors. All output buses are *interlocked*, so that only one picture signal can appear at a time, although a given input signal may be connected simultaneously to several output buses.

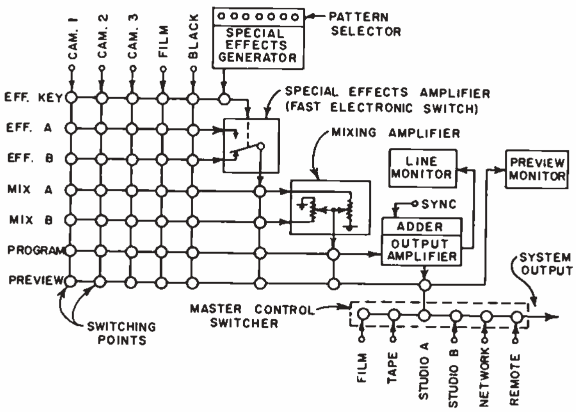


Fig. 3—Simplified sketch of a typical television switching panel and associated amplifiers.

Basic Editing Operations

The most basic electronic editing process is the direct "cut," achieved by simply pushing a button on the program bus to switch cleanly from one signal to another. The preview bus and the associated preview monitor permit the program director to check an upcoming picture before he actually switches to it.

The addition of two more buses plus a mixing amplifier provides facilities for fades, lap-dissolves, and superimpositions. In principle, a mixing amplifier is nothing more than a pair of remotely-controlled, variable-gain amplifiers. For a lap dissolve, the two picture signals involved are "punched up" on Bus A and Bus B, and the control levers for the two variable-gain stages are locked together in such a way that as the level of one signal decreases, the level of the other increases a corresponding amount. For a "fade to black," one may either punch up the "black" input on one of the two "mix" buses, or one may separate the two control levers and reduce the level of one signal to zero before advancing the gain control for the subsequent picture. For a superimposition, one may also separate the control levers to adjust the relative levels of the two picture signals to be combined. Because the sync pulses must always be transmitted at a standard level, they are normally added after the mixing amplifier at the output of the program bus. Remote signals which enter the studio with sync pulses already present are made available in a secondary portion of the program bus beyond the sync addition point, or in an entirely separate "master control" switcher, as indicated in Fig. 3.

Special Effects

Still more elaborate editing effects are possible with a *special effects amplifier* and a *special effects generator.* The special effects amplifier is essentially a very fast electronic switch (operating in the order of 0.1 μsec) carefully designed to minimize switching transients. The fast switching between the A and B picture inputs is controlled by a *keying signal*, which may be any camera signal or the output of a special effects generator. Typical camera signals used for keying purposes might consist of letters to be inserted in a background picture, or special shapes such as keyholes or trademarks. With a special effects generator, a wide range of transitional effects, including horizontal wipes, vertical wipes, diagonal wipes, and iris effects may be obtained. The desired transitional pattern is selected by means of a switch associated with the generator, and the actual movement of the pattern may be controlled by a pair of levers comparable to those used for a mixing amplifier. The two levers, normally locked together, may be separated to provide independent control of horizontal and vertical movement. Many of the transitional patterns may be held stationary for extended periods for interesting "split screen" effects. With careful control of lighting and signal levels, it is even possible to obtain "traveling mattes" or self-keyed insets with special effects equipment.

Audio Origination

Many electronic engineers are surprised to learn that audio problems are frequently more difficult in television program production than video problems. The capital investment in audio facilities is normally much less than for video equipment, but the number of input channels required and the complexity of operations during actual program time is usually much greater. It is not at all uncommon to employ ten or more microphones plus several turntables and tape recorders in conjunction with only two or three cameras. Even the acoustic environment is normally unfavorable. There is usually considerable ambient noise in a television studio resulting from "off-camera" activity, so directional microphones are used frequently and studio walls and ceilings are lined with sound-absorbing materials to make them as "dead" as possible. It is often necessary to employ artificial reverberation to achieve proper audio quality.

A second type of audio system in a television studio is the intercom facility which plays a vital role in coordinating the complex activity on the studio floor and in the control rooms. During rehearsals, the program director frequently employs a public-address system to give instructions to people on the studio floor, but during actual programs it is necessary to rely on headsets connected by wires or low-power radio links. While the investment in intercom facilities is usually even less than for program audio, and while the requirements for audio quality are not very high, the importance of such facilities as major production tools should not be minimized.

Since the problems in audio origination are primarily nonelectronic in nature, we shall not review them in greater detail here, but the writer hopes that he has aroused some sympathy for the often-neglected audio engineer.

Television Film Recording

One of the earliest and most obvious approaches to television recording involves the use of a motion-picture camera to photograph the image on a kinescope. Superficially, this approach seems to offer great advantages. Once a film negative is recorded, it is possible to produce, at reasonable cost, a large number of prints which can be "played back" either through a television film chain or on low-cost, direct-view projectors. In practice, however, the photography of television pictures involves so many practical problems that the phrase "kinescope recording" became, at one time, almost a synonym for "poor picture quality."

Some of the practical problems in television film recording are rather self-evident. Because of the required chemical processing, the final result of a film recording process cannot be monitored instantaneously to check the need for either major or touch-up adjustments. The intermediate picture on the kinescope necessarily introduces some image degradation, especially with respect to graininess and sharpness. Proper reproduction of the gray scale requires extremely good stability in the kinescope picture, precise adjustment of the camera exposure, and close control of the chemical developing process. Added to the other difficulties is the fact that kinescope phosphors with substantial ultraviolet output are commonly employed to gain maximum sensitivity with fine-grain emulsions, so even optical focus cannot be checked properly by simple visual means.

A special problem in television film recording is the "shutter bar" effect resulting primarily from the need to convert the 30-frame-per-second television signal to 24-frame-per-second motion pictures. The nature of this problem is illustrated in Fig. 4. Because of the difference in frame rates, the time interval corresponding to two motion-picture frames is equal to 2.5 television frame periods or 5 television *field* periods. In one sense, the presence of the extra television field period is an advantage, because it provides time for the film shutter to close while the film is "pulled down" from one frame to the next. Unfortunately, however, it is very difficult to assure identical exposure conditions for each motion-picture frame, because at least some of the film shutter operations must occur during active scanning time. In the case illustrated in Fig. 4, for example, film frame 1 (and all subsequent odd-numbered frames) has been made to coincide with a complete television frame period (two successive fields), so a proper exposure of all TV lines in the image can be obtained without much difficulty. The even-numbered film frames, however, must begin and end during the active scanning periods of two different fields. Unless the shutter timing is very precise, the exposure for lines near the center of the image is not quite correct, and a prominent 12-c flicker appears in the final pictures. Even with perfect shutter timing, the photographic effect known as "reciprocity law failure" prevents uniform exposure unless elaborate means of precompensation are employed.

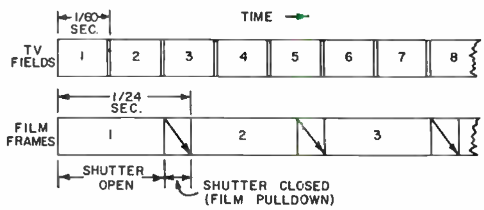


Fig. 4—Sketch illustrating the frame rate and shutter bar problems in television film recording.

In spite of these problems, the case for television film recording is far from hopeless. Interestingly enough, the advent of television tape recording has resulted in a strong rebirth of interest in television film recording, since a combination of the two techniques offers distinct advantages. At the time of this writing, serious development and design work on improved television film recording equipment is in progress in a number of laboratories, and it is quite possible that the term "kinescope recording" will become fully respectable within the next few years.

Television Tape Recording

The recent history of the television industry has been profoundly influenced by the development of practical equipment for recording television programs on magnetic tape. Signals recorded on tape can be played back immediately after recording, and costs are held within reasonable limits by the fact that tape can be erased and re-used a number of times. The lack of any intermediate image-forming process permits very high-quality picture reproduction.

The television tape recorders in most widespread use at the time of this writing employ a *lateral scanning technique* to obtain the very high writing speeds required to record wide-band video signals. Machines operating on this principle are manufactured in the U. S. by the Ampex Corporation and the Radio Corporation of America. The recorders produced by these two companies differ considerably in appearance and design details, but they involve similar basic concepts and produce recordings that can be played back readily on other machines of either make. A typical RCA television tape recorder is shown in Fig. 5.

The magnetic tape employed for television recording is 2-in wide, and information is recorded in the pattern of tracks shown in Fig. 6. The tape moves through the machine at 15 in/sec, but the picture signals are recorded by heads which move across the tape at a speed of nominally 1561 in/sec. Four such heads are spaced around the circumference of an approximately 2-in-diameter wheel, revolving at 240 revolutions per second (or 14,400 rpm). Each track contains approximately 16 lines of the television picture, with an overlap of slightly more than one line period in the information at the end of one track and the beginning of the next.

In addition to the transverse video tracks, the recording also contains three longitudinal tracks. One of these is used for the program audio signal, recorded by techniques equivalent to those in normal audio practice. A second, much narrower, audio track near the opposite edge of the tape is provided for cue information or special-purpose tone signals. The third track contains control information needed by the servo system which controls the capstan speed during playback.

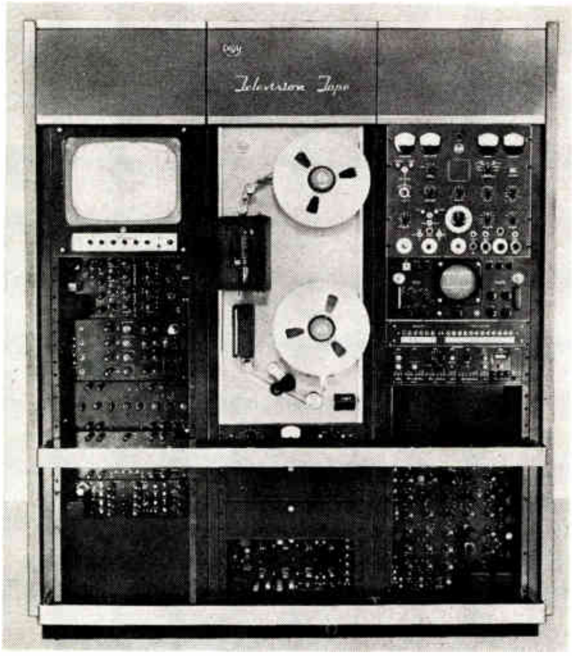


Fig. 5—Photograph of a typical television tape recorder.

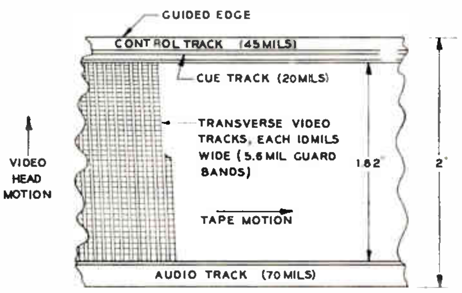


Fig. 6—Pattern of tracks on magnetic tape for television recording (not to scale).

Tape Transport Panel

A typical television tape transport panel is shown in Fig. 7. The driving force which pulls tape through the machine is imparted by a rotating capstan in association with a solenoid-operated pinch roller. In passing through the video head-wheel panel assembly, where the picture recording takes place, the tape is cupped to conform to the circumference of the rotating video head wheel by means of a guide shoe, which employs a vacuum to hold the tape firmly in place. Signal currents to and from the video heads are conducted through brushes and slip rings at one end of the motor shaft, and a tone wheel at the other end provides a feedback signal for the servo system which controls the speed and phase of the head-wheel motor. The control track is recorded by a small head mounted near the base of the guide shoe.

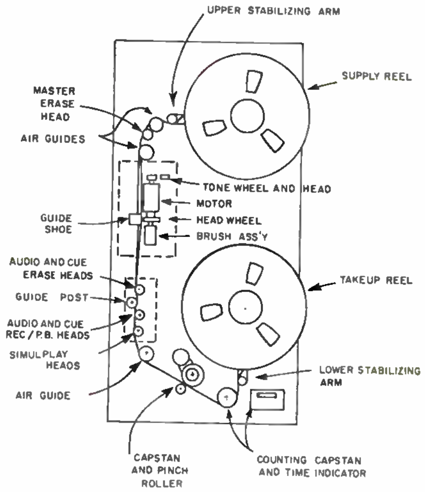


Fig. 7—A typical transport panel for a television tape recorder.

Three heads are used in sequence for the program audio track along the outer edge of the tape: 1) a narrow erase head, which removes the transverse video information previously recorded in the audio area, 2) the main record/playback head, and 3) a simultaneous playback head, which may be used while recording to give complete assurance that a proper audio signal has been recorded. For actual program playback, it is necessary to use the same head that made the recording (or one in the same relative position on a different machine) in order to preserve proper picture-sound synchronization. The cue channel employs similar (though somewhat narrower) heads mounted near the opposite edge of the tape path.

Video Recording and Playback Circuits

The circuits involved in handling the picture signals are shown in greatly simplified form in Fig. 8. To overcome the poor linearity of the tape medium, the video signal is converted to an FM wave before it is applied to the tape. The FM technique also provides a better SNR than would direct recording, and eliminates the need for extreme degrees of LF equalization by translating the entire signal upward to a somewhat higher frequency band. For monochrome recording, the normal deviation characteristic extends from 4.3 Mc for sync tips to 6.8 Mc for reference white. The band actually recovered from the playback of the tape extends from approximately 0.5 to 7.0 Mc, a range great enough to embrace the actual carrier deviation plus the first-order lower sideband only. Fortunately, the missing upper sideband can be restored by a limiting process, and the low deviation index serves to keep the significant signal energy within the first-order sideband spectrum.

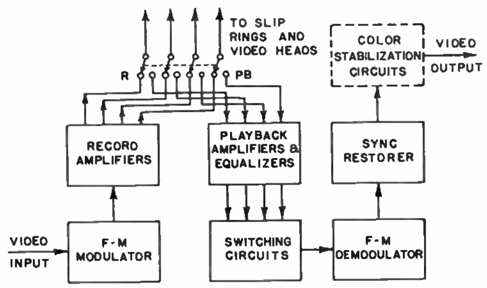


Fig. 8—Block diagram of video record and playback facilities.

Upon leaving the frequency modulator, the signal is passed through the record amplifiers, where the signal is split into four channels which are separately adjustable for optimum recording levels in the four heads. In some machines, the record amplifiers also include means for adjusting the relative *delays* of the four channels to compensate for minor dimensional variations in the physical spacing of the heads around the circumference of the video head wheel. The signals from the record amplifiers are conducted to the heads via a transfer relay, brushes, and slip rings.

In the playback process, the signals recovered from the heads are passed through a set of four-channel amplifiers and equalizers. These amplifiers not only raise the signal levels, but also provide adjustments for the frequency-response characteristics of the four channels and electronic quadrature adjustments in the form of variable delay lines. The signals are then passed through switching circuits which transform the original four head channels to a single channel. As noted earlier, there is a slight overlap in the information delivered by any two successive heads, so smooth switching is possible. To permit complete removal of switching transients, the transition from one head to the next is made to occur during a horizontal blanking interval. The FM signal delivered by the switching circuits is converted to a video signal by a frequency demodulator. This video signal may be somewhat distorted in the sync and blanking regions by switching transients and other imperfections in the many signal-handling operations, so a *sync restorer* or *processing amplifier* is employed to regenerate the sync and blanking portions of the signal. Finally, if the machine is to be used for color television signals, the output is passed through a set of *color stabilization circuits* which utilize the information in the color synchronizing bursts to remove the effects of jitter resulting from small-order imperfections in the servo systems.

Guide Servo System

Three different servo systems are employed in most television tape recorders. One of these is used during the playback mode to control the precise positioning of the vacuum guide shoe which pushes the tape against the rotating video head wheel.

As shown in exaggerated form in Fig. 9(a), the inner diameter of the guide shoe is deliberately made slightly smaller than the outer diameter of the video head wheel at the pole tips. Thus the pole pieces are forced to deflect the tape, placing it under rather severe local tension as the heads pass by. The *degree* of this tension may be adjusted by movement of the guide shoe to compensate for minor dimensional variations between different machines or in the same machine as a result of wear. If the center of curvature of the guide shoe is too high or too low relative to the center of the video head wheel, the tape tension varies across each track, yielding a characteristic "scalloping" effect along all vertical lines in the picture, as illustrated in Fig. 9(c). Fortunately, this effect can be avoided by careful mechanical positioning of the guide shoe. If the guide shoe is too close or too far away from the head wheel, the tension is not correct and the signals are not truly coincident at the instant a switch is made from one head to the next. Under these conditions, vertical lines in the picture break up into "venetian blinds," as shown in Fig. 9(b). To prevent this effect, the servo system shown in Fig. 10 is employed.

A servo system is, of course, an application of the negative feedback principle. In the case at hand, the feedback path comprises the entire video playback system which includes, among other things, a sync separator to recover the sync pulses from the composite video signal. These sync pulses are passed through an error detection circuit designed to sense the presence of timing variations at 960 cps, which is the head-to-head switching rate. If such timing errors are found, their polarity or direction is also sensed, and the error information is amplified and processed in suitable fashion to actuate a motor which moves the guide shoe to the proper operating position.

Head-Wheel Servo System

A second servo system is employed to control the speed and phase of the head-wheel motor. A simplified block diagram of a typical circuit arrangement is shown in Fig. 11. During the "record" mode of operation, the major objective of the head-wheel servo system is to assure the formation of a standardized pattern of tracks on the tape, in which each television field occupies exactly 16 tracks. (Such a standardized track pattern permits both interchanging and *intersplicing* of tapes.) The feedback signal employed in the head-wheel servo system consists of a pulse signal derived from the tone generator on the motor shaft. This pulse signal is compared in a phase detector circuit with a 60-cps pulse signal derived from the vertical synchronizing portion of the signal to be recorded. By means of this error detector and the subsequent motor drive circuits, the 240-cps output of the motor is kept "locked" to the fourth harmonic of the 60-c reference pulse.

In the playback mode, the 60-c reference pulse for the head-wheel servo system is derived from an external source, such as a local sync generator or a 60-c power line. In some more recent television tape recorder designs, the head-wheel servo operates with sufficient precision when locked to a sync generator that the tape output signal may be handled through electronic editing equipment as freely as a local camera signal.

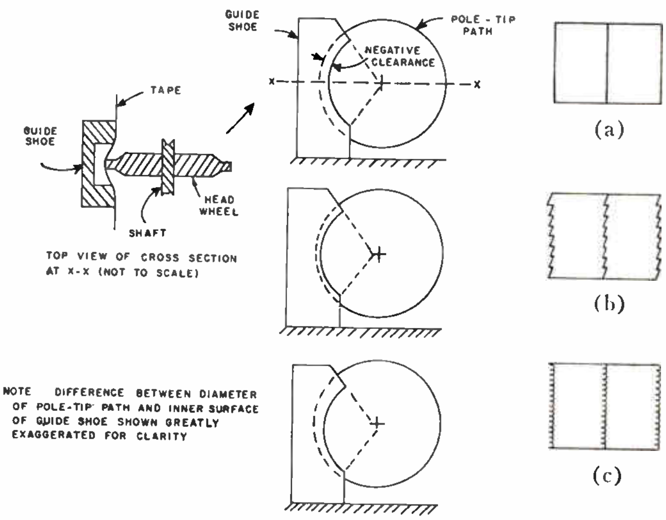


Fig. 9—Sketch illustrating the significance of guide shoe position adjustments. (a) Guide shoe in correct position. (b) Guide shoe displaced to left. (c) Guide shoe too low.

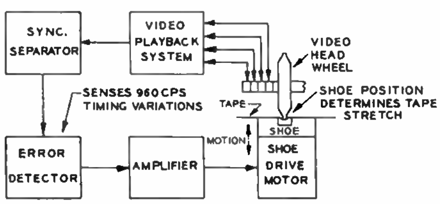


Fig. 10—Block diagram of guide shoe servo system.

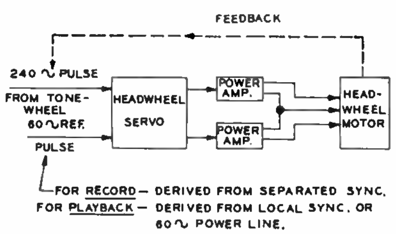


Fig. 11—Simplified block diagram of a typical head-wheel servo system.

Capstan Servo System

The third and final servo system controls the speed of the capstan through an arrangement like that shown in Fig. 12. When a recording is made, the capstan motor is driven synchronously by 60-c power derived from an oscillator kept locked to the 4th subharmonic of the 240-cps tone-wheel signal. Information about the relative speed of the capstan is recorded on the tape in the form of a 240-cps tone on the control track. Since this 240-cps tone is derived from the tone wheel on the head-wheel motor shaft, each cycle corresponds to exactly four video tracks. In the playback process, the error detectors in the capstan servo system compare the 240-cps tone from the control track with the 240-cps output of the tone wheel. Any error signal that may be present varies the speed of the capstan just enough to assure that exactly four video tracks are moved past the head wheel during each revolution. The phasing adjustment required to center each video head over a track on the tape is usually provided in manual form.

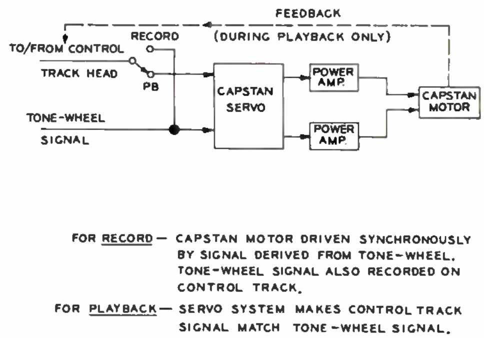


Fig. 12—Simplified block diagram of a typical capstan servo system.

Control and Monitoring Facilities

The operating controls for a television tape recorder are grouped for maximum operational convenience, and built-in monitoring instruments are provided to check significant aspects of the machine's performance during operation. Routine operation of a television tape recorder is surprisingly easy, and most basic operations can even be performed from a remote location.

Conclusion

In the brief space of this "overview" paper, it has been possible to review only the highlights of the technology of television program production and recording, but it is hoped that IRE members have found here a little insight into the problems that have already been solved and those still under exploration by their colleagues in the broadcast industry. During the 50 years of the IRE's history, broadcast engineers have provided for the American people the world's best facility for mass communications. There is every reason to assume that major creative contributions will continue to flow from this branch of the engineering profession throughout the next half-century.

Broadcasting Developments Now Taking Place*

OSCAR REED, JR.†, SENIOR MEMBER, IRE

Summary—**New developments in broadcasting in the past 20 years are treated from the general allocations development viewpoints including regulatory aspects and the evolution in the broadcast standard, frequency modulation, and television station industry. The prospects for compatible single sideband modulation techniques in standard broadcasting, developments in FM stereo regulations and experimental techniques in the color television field are explored. Brief treatment is accorded the horizon subject of extra terrestrial satellite relay usage in broadcasting as well as the currently developed backbone of international broadcasting.**

References are included to fundamental explorations of several government-industry engineering committees including the Television Allocations Study Organization, technical study papers issued under the auspices of the Chief Engineer of the Federal Communications Commission, and studies carried on by broadcasters, network organizations, and manufacturers.

NEW DEVELOPMENTS in broadcasting now taking place may be scanned from several points of view. The broadcasting industry is regulated by a government agency, the Federal Communications Commission. The interaction between technological progress and the official Rules and Regulations of this Commission, including engineering Standards, is continuous. Broadcast industry engineers and government agency engineers work toward the common goal of "public interest, convenience and necessity" set forth as a guideline in the Communications Act of 1934. The following general groupings will be used for consideration:

I. Regulatory Aspects—General Allocations Developments.

II. Developments in Systems and Techniques.

In each of these general groups it is pertinent to consider specific items in three major types of broadcasting. These are standard amplitude modulation (AM) broadcasting, frequency modulation (FM) broadcasting, and television (TV) broadcasting. These types of broadcasting are all comprehended in Part 3 of the Rules and Regulations of the Federal Communications Commission. In addition, Part 4 of the Rules and Regulations sets forth provisions relating to experimental, auxiliary, and special broadcast services.[1] Developments in this area will be outlined in some instances under the general groups breakdown. International broadcasting, also provided for in Part 3 of the Rules and Regulations will be treated from an over-all point of view without regard to the general categories. Satellite relays for broadcasting also will be considered.

* Received by the IRE, July 5, 1961; revised manuscript received, August 16, 1961.

† Jansky & Bailey, Washington, D. C.

[1] "Rules and Regulations," "Part 3—Radio Broadcast Services," and "Part 4—Experimental, Auxiliary, and Special Broadcast Services," FCC, Washington, D. C., vol. 3, as amended.

I. Regulatory Aspects—General Allocations Developments

A. Standard (AM) Broadcasting

At present there are 3543 licensed standard broadcast stations and construction permits have been issued for an additional 205 stations. Presently pending before the Federal Communications Commission are 711 applications for new AM facilities some of which are mutually exclusive. An additional number of 578 applications pend for changed facilities on the part of existing stations wishing to improve their operations. This picture represents a growth from 882 operating stations in 1941. Standard broadcast stations operate on frequencies ranging from 540 kc to 1600 kc inclusive, representing 107 ten kc channels.

1) Class IV "local" type stations are designed to render service to cities or towns and the rural area contiguous thereto. They formerly were authorized powers not to exceed 250 watts day and night. Under an Order adopted in Docket No. 12064, effective July 7, 1958, many stations have been able to increase their daytime power to 1000 watts. This trend still continues. Of the approximately 1100 stations of this class, approximately 300 have now increased their daytime operating power, in some cases installing daytime directional antennas to do so. No changes in night operating powers have been allowed on the six channels of this class.

2) With the entry into effect of the new United States-Mexican Broadcast Agreement on June 9, 1961, United States daytime Class II stations operating daytime only on the Mexican clear channels 730, 800, 900, 1050, 1220, and 1570 kc had their power ceiling lifted from 1000 watts to 5000 watts. With 22 applications on file for increased power as of that date the Federal Communications Commission began processing these requests. About 20 per cent of the approximately 1500 stations in the United States which operate daytime only utilize these channels.

3) The Commission on June 12, 1961 instructed its staff to prepare a report and order concluding the clear channel proceeding in Docket No. 6741 initiated in February, 1945. Rules amendments look toward opening the way to the assignment of one unlimited time Class II station on each of 13 selected Class I-A clear channels under controlled conditions to provide service to underserved regions of the country. The proposed action does not affect the present use of the remaining 12 Class I-A frequencies. Class I and II stations employing up to 50 kw power continue to provide secondary sky-wave coverage to much of the United States.

4) In June, 1960, the Commission instituted a Notice of Inquiry in Docket No. 13596 looking toward a possible amendment of its Rules to permit standard broadcast stations to operate with a "compatible single-sideband system of modulation" known as CSSB. The petitioner claimed that the proposed mode of operation would result in improved high frequency audio response, increased signal-to-noise ratio, reduction in cochannel and adjacent channel interference, reduction in television receiver radiation interference, and reduction of certain types of fading distortion. Several standard broadcast stations including KDKA, Pittsburgh, WSM, Nashville, WGBB, Freeport, New York, WABC, New York, N. Y. and WMGM, New York, N. Y. were authorized to conduct experiments to test the CSSB system and to gather data regarding this type of operation in the standard broadcast band. The record was closed on this proceeding in the Spring of 1961 with a number of proponents including Westinghouse Broadcasting Company supporting favorable action.

B. Frequency Modulation (FM) Broadcasting

There are presently 812 licensed FM broadcast stations. Construction permits have been granted for 274 more. Pending applications total 107. Since the authorization of Subsidiary Communications Authorizations in May, 1955, the number of FM stations has been on the increase, approximately doubling in number since that year. The Subsidiary Communications Authorizations (SCA) services are provided on a multiplex subchannel basis by the licensee of an FM broadcast station supplying the conventional type of programming on the main channel. Transmission of programs which are of a broadcast nature, but which are of interest primarily to limited segments of the public wishing to subscribe thereto is carried on under the (SCA) authority. Illustrative services include background music, storecasting, detailed weather forecasting, special time signals and other material of a broadcast nature expressly designed and intended for business, professional, educational, religious, trade, labor, agricultural or other groups engaged in any lawful activity. Such operations may be conducted without restriction as to time so long as the main channel is programmed simultaneously. Supersonic tones or other similar means may be employed in conjunction with the material transmitted during the SCA operation in order to promote or maintain its commercial marketability, with the station using appropriate actuating devices with the subscriber's receivers. Subchannel frequencies in common use include 67 and 41 kc frequency modulated subcarriers.

FM broadcasting has been assigned 100 channels each 200 kc wide between 88 and 108 Mc. The lower four megacycles of this band have been reserved exclusively for noncommercial educational broadcasting. Stations in the educational service may employ powers as low as 10 watts while stations in the commercial service employ effective radiated powers (transmitter power times antenna gain) up to several hundred kilowatts. Use of the higher effective radiated powers is quite common in the educational service also. Antenna heights utilized range from below 100 feet effective height up to 1000 feet and above in some instances.

1) Stereophonic FM broadcasting was authorized to commence June 1, 1961, by Order adopted April 19, 1961, in Docket No. 13506. FM stations operated by Zenith Radio Corporation in Chicago and the General Electric Company in Schenectady, New York, initiated operations on that date. After an extensive study of several proposed systems of FM stereocasting including field testing,[2] conducted with the help of the National Stereophonic Radio Committee (NSRC), the Commission adopted a system akin in all major respects with those techniques proposed by Zenith and GE. FM was given first priority of consideration with the indication that standards for stereo in the standard broadcast and television services would be the subject of further determinations.

Stereophonic FM broadcasting is defined as the transmission of a stereophonic program by a single FM broadcast station utilizing the main channel and a stereophonic subchannel. The stereophonic subchannel is that band of frequencies from 23 to 53 kc containing the stereophonic subcarrier (38 kc, suppressed), and its associated sidebands.

2) FM technical standards first established in 1941 were last extensively revised upon the adoption of a new frequency band for FM in 1945. Currently deliberations are under way in government-industry groups looking toward a revision and modernization of these standards. The Technical Standards specify engineering standards of allocation which relate to FM station coverage, determination of interference between stations, the location of transmitters, and other technical factors. Considerable incentive to revision of the FM Technical Standards has been occasioned by the availability of new field strength measurements made in connection with the program of the Television Allocations Study Organization. The results of this measurements program probably will yield new field strength vs distance curves for FM broadcasting.

On June 29, 1961 the Commission issued a Notice of Inquiry, Notice of Proposed Rule-Making and Memorandum Opinion and Order dealing with the following matters:

> Proposed an over-all plan of new FM station assignments based on minimum mileage separations with respect to existing stations, and also to some extent based on maximum separations.
>
> Proposed three classes of commercial FM stations instead of the present two classifications, and two classes of educational stations.
>
> Proposed particular channels for operation of the different classes of FM stations with protection to existing FM stations if they conform to at least the new minimum operating requirements; and

[2] A. P. Walker, "Engineering performance of six proposed stereo systems," *Electronics*, vol. 33, pp. 85–89; November, 1960.

Instituted an inquiry into numerous technical subjects related to FM broadcast, such as signal ratios, polarization, directionalizing, receiver characteristics, etc.

In substance, the proposals contemplated FM broadcast operation on the following basic pattern, giving each station a certain protected service area with respect to cochannel and adjacent-channel interference, and a somewhat smaller protected service area as to second and third adjacent-channel interference. This is achieved by providing for minimum cochannel and adjacent-channel mileage separations. The proposed maximum facilities, cochannel protected service radius, and minimum cochannel separation, are as follows:

Class	Maximum facilities (or equivalent)	Protected service area radius	Minimum co-channel spacing
Class A (low power)	1 kw ERP, 250 ft*	25 miles	115 miles
Class B (intermediate power)	20 kw ERP, 500 ft*	50 miles	190 miles
Class C (high power)	100 kw ERP, 2000 ft*	100 miles	300 miles
Class D (low power educational)	10 watts (transmitter power), 100 ft*	6 miles	25 miles
Class E (high power educational)	Same as for maximum commercial station at the same location		

* Above average terrain.

Provisions with respect to FM antenna systems also will be of importance. In recent years a proposal was filed with the Federal Communications Commission requesting that FM broadcast stations be allowed to operate with either horizontal or vertical polarization. This petition was subsequently denied. However, the present Standards indicate that while horizontal polarization shall be the standard for FM, circular or elliptical polarization may be employed if desired. Thus, within these present Standards a signal having some vertical component may be provided and certain experiments have been carried out along this line. Results indicate that this can be beneficial in consideration of the problem of providing adequate field strength to automobile receivers which conventionally operate using a vertical antenna. The present provision of the Standards in this respect requires merely that the supplemental vertically polarized effective radiated power shall in no event exceed the effective radiated horizontal power authorized to the licensee.

C. Television (TV) Broadcasting

Television stations operate on VHF and UHF frequencies and are assigned three bands of frequencies which are divided into six Mc channels. The frequency ranges are 54–88 Mc (except 72–76 Mc), 174–216 Mc, both VHF ranges and 470–890 Mc in the UHF range. A total of 585 stations is now operating of which 494 utilize VHF channels and 91 occupy UHF channels. This number includes 56 noncommercial educational stations of which 16 operate in the UHF portion of the range. Availability of channels is determined by a Table of Assignments contained in the Rules of the Federal Communications Commission listing the various cities and towns. The table lists approximately 2200 channels of which 270 are presently reserved for noncommercial educational television use.

There are about 140 construction permits outstanding which represent uncompleted construction or stations which have been on and gone off the air reverting to construction permit (CP) status for economic problem reasons. Many of the latter represent UHF facilities whose coverage has not measured up to that available to VHF facilities, which in many cases are assigned in the same market area. Thus with only 12 channels available for assignement in the VHF portion of the spectrum and 70 channels, many of which are unused in the UHF region, the Commission has been faced for some years with an allocation problem by reason of coverage and other technical inequities.

1) In January, 1960, a new Notice of Proposed Rule-Making was issued by the Federal Communications Commission identified as Docket No. 13340. The Commission therein invited comment on its proposal to:

a) Consider applications for waivers of minimum television station separation in exceptional, individual cases meeting certain criteria; and

b) Adopt revised rules and standards governing calculations of the service areas of television broadcast stations.

In April of 1959 the Commission indicated to the Senate Committee on Interstate and Foreign Commerce that it was pursuing studies and negotiations needed to ascertain the practicability of four alternative approaches to basic revision of the television allocations structure. These alternatives included further exploration of de-intermixture of VHF and UHF channels in the same market area, the possible ability to secure a continuous block of VHF channels in the VHF portion of the spectrum by trading with UHF channels, or the move of all television broadcast services to the UHF portion of the spectrum. In the interim period, developed in the Docket 13340 proceeding notice, the Federal Communications Commission proposed to consider "interim" courses of action utilizing short spaced channel separations under the following conditions:

c) The assignment would make possible a second or a third VHF television station in an important market.

d) The need for the additional service outweighs the need for any service lost as a result of additional interference to existing stations.

e) The new VHF service would not have substantial adverse effect on established UHF television services.

f) A new assignment would not require an excessive number of channel changes of existing stations.

It was proposed as a corollary to these provisions to require any new station assigned at a substandard cochannel separation to suppress radiation (presumably through the use of a directional antenna) in the direction of the existing station. Directionalization would be employed to the extent necessary to insure that such a new station created no more interference to the existing one than that which would be caused if the two stations were to be operating at the standard minimum mileage separations of 170, 190, and 220 miles, depending upon the locality of the United States involved. For this purpose tentative methods for calculating television broadcast interference also were proposed. The existing rules and regulations lack any such provision.

Revisions of the Television Engineering Standards also were set down in proposed form in this Notice. These revisions included new VHF propagation curves and definitions of signal strength required to provide various Grades of Service using in part recommendations of the Television Allocations Study Organization (TASO) which had been submitted in March, 1959.[3] The work of this organization made up of industry and government engineers and scientists previously had covered a period of two and one-half years. A supplementary TASO report was also issued on June 13, 1960, relating to television directional antenna tests, further propagation studies, and certain minority report matters.[4] In this effort extensive comparisons were made of VHF versus UHF capabilities to provide television services through the work of six major panels. The primary areas involved were transmitting equipment, receiving equipment, field tests, propagation data, analysis and theory, and levels of picture quality.

After the issuance of two supplementary FCC notices in May and July of 1960, comments in this proceeding were closed at the end of the year.

2) In mid 1960 the Federal Communications Commission received an authorization from the United States Congress to undertake a $2,000,000 study of the feasibility of providing UHF television service to the City of New York, admittedly one of the most difficult areas to provide with VHF and UHF service in the nation. This investigation is currently going forward. A UHF transmitter and high gain antenna capable of an effective radiated power of 1,000,000 watts is being installed at the same location as the seven VHF stations atop the Empire State Building in New York City. Measurements of field strength and observations of picture quality are to be made at 5000 statistically selected locations within a 25-mile radius of the transmitter. Comparisons will be made with existing VHF signals and prototype UHF receiver installations will be made at 1000 of the locations selected above for more intensive study. One hundred UHF receivers, 10 of them for color reception, representing the best the state of the art can provide, would be used for these investigations. Furthermore, RF preamplifier use would be investigated in those cases where very difficult reception conditions were encountered and where special receiving antennas still proved incapable of securing a satisfactory picture.

A second UHF transmitter had been planned for installation in the New Jersey area to test the possible advantages of multicasting.[5] This technique has been proposed to overcome the problems of shadowing and multipath transmission difficulties which are much more pronounced on the UHF portion of the spectrum than the VHF. In effect several transmitters would be used within a given market area to provide a signal source dispersion and facilitate reception of at least one good quality broadcast signal.

3) In June, 1961 the Federal Communications Commission through the Office of the Chief Engineer and the Associated UHF-TV project issued a UHF-TV Report TRR-5.2.2 containing a possible approach to a UHF assignment plan.[6] Based on studies conducted in Europe of the most efficient means of assigning channels, when a number of criteria must be met in addition to simple cochannel and adjacent channel spacings, a way was pointed theoretically which would permit the assignment of over 8000 stations in the UHF television band. The practical limits imposed by UHF receivers and their susceptibility to picture and sound image interference, oscillator radiation and intermodulation would still be a factor in any such utilization. These "taboos" which dictate certain mileage separation criteria would still impose severe restrictions on the flexibility of assignment if such an assignment efficiency is to be approached. Under the present Table of Assignments, however, which was based on a triangular lattice approach, the required cochannel separation factor is the major inhibiting consideration as more channels are sought in a given region. The rhombus configuration suggested in the report, while requiring a more arbitrary assignment grid, does offer considerable hope of increased channel availability for the UHF television broadcast service.

D. Experimental, Auxiliary, and Special Broadcast Services

This category of services includes the following types of operation, many of which are tied in with the day-to-day operations of AM, FM, and TV stations:

Conelrad provisions,
Experimental television broadcast stations,
Experimental facsimile broadcast stations,
Developmental broadcast stations,
Remote pickup broadcast stations,

[3] "Engineering Aspects of Television Allocations," TASO Rept. to the FCC, pp. 1–731; March, 1959. See also Proc. IRE, TASO issue, vol. 48; June, 1960.

[4] "Engineering Aspects of Television Allocations—II," TASO Rept. to the FCC, pp. 1–210; June, 1960. See also Proc. IRE, TASO issue, vol. 48; June, 1960.

[5] E. W. Allen, H. Fine, and J. Damelin, "A Preliminary Analysis of Multicasting," Tech. Res. Div., Office of Chief Engineer, TRR Rept. No. 5.1.1, pp. 1–8; February 4, 1958.

[6] A. G. Skrivseth, "UHF Assignment Plan," Office of Chief Engineer, UHF Project, FCC, UHF-TV Rept. No. 5.2.2, pp. 1–5; May 29, 1961.

Standard and FM broadcast STL and FM intercity relay stations,
Television auxiliary broadcast stations,
Television broadcast translator stations, and
Television broadcast booster stations.

Several years ago the Federal Communications Commission authorized the use of UHF translators. A translator generally is used to extend the service area of an existing station. A station in this service is operated for the purpose of retransmitting the signals of a television broadcast station or other television broadcast translator stations. Direct frequency conversion and amplification of the rebroadcast signals without significantly altering any characteristic of the incoming signal other than its frequency and amplitude are used to provide an extended area of television reception to the general public. UHF translators may operate on UHF Channels 70 through 83. Such translators have been authorized for some time and there are now 248 in operation and applications are pending for an additional 34 stations.

More recently VHF translators have been authorized and there are currently five of these on the air. However, construction permits have been granted for an additional 194 such operations and the backlog of pending applications currently totals 706. Present UHF translators can employ a maximum peak visual power output of 100 watts and there is no limitation on the effective radiated power which may be used within the limits of practical antenna gain. VHF translators are limited to a maximum peak visual power output of one watt and similarly there is no limitation as to effective radiated power.

UHF translator signal boosters are also provided for in the Rules. The transmitting apparatus consists of a simple RF amplifier, with one or more amplifying stages, which is capable of receiving, amplifying and transmitting the signals of the parent translator without significantly altering its electrical characteristics other than amplitude. The purpose of this type of operation is to fill a coverage gap in the nominal service area of the translator station. The maximum power input to the final radio frequency amplifier is not to exceed five watts. Generally, television broadcast translator and booster stations may be operated as remote controlled and unattended operations. The equipment is generally so arranged as to come on the air when the main transmitter comes on and goes off when the main station operation ceases.

An additional class of television booster station exists which may be employed to fill the gaps within the intended service area of a main UHF television broadcast station which may not be getting service directly from the main station. Whereas the translator type of station frequently extends the nominal service area of the main station, this latter class of booster station is intended to fill "holes" in the licensed service area. This type of booster station may employ any power needed to fill the gap in the main service area providing that it does not operate with an effective radiated power of more than 5000 watts peak visual power.

Several years ago in the mountainous regions of the west, in areas inaccessible to signals from the main commercial or educational stations, the enterprising public began to put in low power "boosters" which served to relay signals from the main station back into the hills and valleys. These were not licensed or regulated by the Federal Communications Commission and were generally condemned as potential sources of interference having been placed in operation illegally. More recently, however, under the impact of public and congressional pressure the Federal Communications Commission has taken steps to legitimatize these operations and they are currently referred to as repeaters. These repeater stations operate on the VHF channels, rebroadcasting programs on a different channel to avoid interference to the received signal being used as a program source. Several hundred of these stations in Arizona, Colorado, Idaho, Montana, Nebraska, Nevada, New Mexico, Oregon, South Dakota, Utah, Washington, and Wyoming as well as several other states have been afforded recognition by the Commission. For example, there are now over 150 such operations listed for the State of Colorado alone. Such stations provide service to the public by means of direct home reception in contrast to the many homes in such areas which may obtain their signal by direct cable means through the medium of a remote master receiving antenna and signal distribution amplifiers. This latter type of system is not regulated under the present Rules of the Commission and is generally known as a community antenna television system.

The Rules governing the class of stations known as television auxiliary stations now provide that a licensee of a television broadcast station may secure authority for television intercity relay stations. This provision is in contrast to an earlier period in the development of the television broadcast service when this intercity relay prerogative was the exclusive right of the common carrier operator. While television remote pickup facilities and television studio to transmitter links had long been authorized directly to the licensee of the broadcast facility, the intercity provision is a relatively new one. The license may be issued in any case where the relay will operate between television broadcast stations either by means of "off the air" pickup and relay or involve location of the initial relay station at the studio or transmitter of a television broadcast station. Such stations may be operated by remote control and be unattended provided that certain precautions are met.

II. Developments in Systems and Techniques

A. Standard (AM) Broadcasting

Remote control of standard broadcast stations has been allowed for the past several years. Stations utilizing directional antennas have been the most recent ones to come under such an allowance with new provisions in the Engineering Standards. In the spring of 1961 the industry put forward a proposal to the Federal Communications Commission to allow automatic logging of station operation.

In the realm of new modulation techniques the compatible single sideband system has been advanced for use by the AM broadcaster. In the programming field the industry has begun to make extensive use of cartridge tape systems with automatic cuing provisions. Various types of automation are being applied to studio operations utilizing punched tapes and cards geared to program department requirements.

1) Applications to operate a station by remote control may be made as a part of the application for construction permit for a new station provided that the proposal is for nondirectional operation with a power of 10 kw or less. If a station operates with a directional antenna, or higher powers, certain further evidences of stability of operation and suitability for remote control must be given over a period of time prior to application for such operation. Control and monitoring equipment has to be so installed at the remote control point to allow the licensed operator at such a distance to perform all the functions in a manner required by the Commission's rules.

The dc control system using two wire pairs is most commonly in use. Indications at the remote control point of the antenna current meter, or for directional antennas the common point current meter and remote base current meters, are read and entered in the operating log each half hour. If a directional antenna is used, the tower base currents, phase monitor sample loop currents and phase indications must also be read and entered in the operating log once each day for each pattern. These readings must be made by personal inspection at the transmitter within two hours after the commencement of operation for each pattern.

Several hundred AM and FM stations now operate employing remote control. Approximately 100 of these involve directional antenna operations. The Columbia Broadcasting System now handles six of its seven owned and operated AM radio stations by remote control. Five of these are 50 kw stations and one a 5 kw plant. One of the 50 kw operations at KCBS uses a directional antenna with different patterns for day and night operation. The stations involved operate on 24-hour daily schedules with a minimum time off for maintenance. In general experience with remote control has been good.[7] The two major problems encountered were personnel training to handle remote type operation and reliability of control lines. Training in handling failure procedures and instructions in correct logging helped overcome the former. Telephone lines providing dc paths are the lowest grade of circuit available and apparently are given only minimum consideration by the phone companies. Disruption from telephone company tampering has been given as the main difficulty.

2) On January 13, 1961, the National Association of Broadcasters filed a proposal with the Federal Communications Commission to amend its rules to permit the use of automatic logging devices for broadcast transmitter logs. Logs currently require entries every 30 minutes of the operating parameters of the transmitter including final plate current, final plate voltage and RF power output. The proposal filed contained extensive experimental material obtained from prototype operations using such techniques, while continuing manual logging, at several stations including WIP, Philadelphia, Pa.; KFI, Los Angeles, California; and WTOP, Washington, D. C.

The equipment at WTOP was designed and built by the Brown Instrument Industrial Division of the Minneapolis-Honeywell Company in Philadelphia, Pennsylvania.[8,9] Automatic logging is accomplished by means of a recording device driving a strip chart on which 15 different parameters are recorded in a continuous cycle. On a remote system this information may be transmitted by a telemetering servo via a leased metallic telephone line. The line need not be equalized nor are amplifiers required. Its maximum resistance may be as high as 10,000 loop ohms. For remote sampling a measuring unit and a telemetering unit must be placed at the transmitter, the equipment filling about 20 vertical inches of standard relay rack space. Where the transmitter and logging equipment are to be at the same location, a combination arrangement may be made incorporating the measuring system and the recording device in a single rack.

The recorder strip chart is provided with a scale of 0–100 units which can be expanded to any convenient value in multiples of 10. The chart is marked by a dot representing the exact value of measurement. The chart speed is two inches per hour and is calibrated in 10 minute segments. There is sufficient paper on one chart roll for one month's operation. All sampling of the various parameters is transduced to direct current or voltage of low level and close to ground reference. These are fed through a selector switch into the measuring servo.

3) A form of AM modulation developed for use in the standard broadcast band by Leonard R. Kahn of Kahn Research Laboratories has been termed compatible single sideband (CSSB). The system is said to be compatible in the sense that it may be received on existing broadcast receivers. Field tests have been made to determine its capabilities.

CSSB is generated through the use of an adaptor that

[7] O. L. Prestholdt, "Experience in Remote Control Operations of AM Plants," Natl. Assoc. of Broadcasters, Broadcast Engrg. Conf., CBS Engineering Dept., No. E1044-T; May, 1961.

[8] G. Klink, Jr., "Automatic Logging of Operating Parameters in lieu of a Manual Log for AM-FM and TV Stations," Natl. Assoc. of Broadcasters, Broadcast Engrg. Conf., WTOP Engrg. Dept., Wash., D. C.; April, 1960.

[9] G. Ehrenberg, "Automatic logging of AM, FM, and TV transmitting station parameters," *Elec. Engrg.*, vol. 79, pp. 304–308; April, 1960.

can be used with an existing broadcast transmitter.[10] The adaptor provides two outputs for connection to the transmitter. One output is a phase modulated RF signal at the transmitter's normal operating frequency. The phase modulation is in accordance with the audio information to be transmitted. This same audio is provided at the second output. The unique characteristic of the audio at the second output is that it has been subjected to the same delay as the audio used for the phase modulated output. The adaptor's RF output is applied to the transmitter's regular first RF amplifier stage. The second output (audio) is applied to the transmitter as a modulation source in the normal manner. When modulation takes place the phase modulated source causes the transmitter carrier to change frequency and at the same time output No. 2 amplitude modulates the moving carrier. The effect is a redistribution of the radiated energy from the usual symmetrical fashion to an asymmetrical one.

Measurements undertaken of the performance of the KDKA transmitter when using the CSSB adaptor have been reported.[11] Frequency response runs and distortion checks at all levels of modulation were made. The transmitter performance was very nearly the same as with the normal double sideband system. Wide band receivers respond about the same to the conventional modulation process and the CSSB system. On the other hand, narrow band receivers will have to be shifted 1 to 2 kc in the direction of the favored sideband. Generally such tuning will make the CSSB signal sound more treble and if the transmitter is not modulating heavily on either double or CSSB, make it sound louder. With CSSB and "off resonance" tuning the amplitude of the favored side frequencies is enhanced and the carrier somewhat attenuated giving the effect of increasing the depth of modulation. Results in noisy low signal areas favor the CSSB signal, especially if a narrow band receiver is used and the transmission is predominantly voice modulation. The chief gain of using CSSB to KDKA has been the reduction of adjacent channel interference between KDKA on 1020 kc in Pittsburgh and WBZ in Boston on 1030 kc. The suppression of KDKA's upper side frequencies has relieved interference to WBZ reception and the ability of receivers tuned to KDKA to tune slightly off resonance away from WBZ and toward the radiated lower sideband of KDKA has also given KDKA some additional immunity from WBZ.

4) Automatic tape control record and playback machines have been designed to accept tape packaged in hand size plastic cartridges in which the magnetic tape is both contained and recorded or played.[12] The plastic cartridges are loaded with an endless tape, generally supplied in any length required from a 30 second announcement to a 30 minute program. To play the tape, the cartridge is inserted in the playback machine, a button is pushed to activate the tape and the announcement goes on the air, on cue, with no manual threading or cuing. In addition to the basic tape playback and recording equipment, remote control panels and automatic master switchers are available. Automatic transfer can be arranged between two playback machines with one unit supplying the program material and the other machine the announcements or introductions. Magnetic disc units also have been introduced by one manufacturer which combine the advantages of magnetic recording and the conventional disc handling approach.

5) Present Federal Communications Commission rules and regulations require that program logs be kept showing the starting and ending time of each program and announcement as broadcast. One system of handling such record keeping has been suggested using a Productograph P131 unit manufactured by the Simplex Time Recorder Company. This time stamping device, when actuated, prints on a standard adding machine tape the month, day of the month, hour, minute, second, and an identifying letter code.[13] Systems also have been developed utilizing IBM punched cards to provide a combination program logging and billing record.

B. *Frequency Modulation (FM) Broadcasting*

The action of the Federal Communications Commission of April 20, 1961, concluded a proceeding initiated in March, 1959, looking toward the adoption of rules for FM stereophonic broadcasting. Eight proposed systems were put forward for study during this period and six were finally field tested. The Commission's decision in favor of the Zenith-GE system hinged in part on the desire to maintain an element of compatibility with existing multiplex SCA operations. As of January 31, 1961, more than 250 FM stations held SCA multiplex authorizations. Such stations would be able to utilize the stereophonic system adopted while retaining at least one SCA channel capability.

The instantaneous frequency of SCA subcarriers is required to be within the range of 20 to 75 kc at all times. When an FM station is engaged in stereophonic broadcasting the SCA subcarriers must lie within the range of 53 to 75 kc. The stereophonic components would require use of the range from 23 to 53 kc. Stereophonic transmission standards put forth in new Rule 3.322 are as follows:

a) The modulating signal for the main channel shall consist of the sum of the left and right signals.
b) A pilot subcarrier at 19,000 cycles plus or minus 2 cycles shall be transmitted that shall frequency modulate the main carrier between the limits of 8 and 10 per cent.

[10] R. D. Schneider, "Compatible single sideband for AM broadcasters," *Broadcast Engrg.*, vol. 1, pp. 4–7; August, 1959.

[11] R. N. Harmon, "Experience with CSSB at KDKA," presented at WESCON, Los Angeles, Calif.; August, 1958. Reprinted by Natl. Assoc. of Broadcasters with approval of the IRE.

[12] "New trends in broadcast equipment," *Broadcast Engrg.*, vol. 2, pp. 10–15; April, 1960.

[13] R. W. Flanders and R. M. Brockway, "An automatic program logging system," IRE Trans. on Broadcasting, vol. BC-7, pp. 11–14; January, 1961.

c) The stereophonic subcarrier shall be the second harmonic of the pilot subcarrier and shall cross the time axis with a positive slope simultaneously with each crossing of the time axis by the pilot subcarrier.
d) Amplitude modulation of the stereophonic subcarrier shall be used.
e) The stereophonic subcarrier shall be suppressed to a level less than one per cent modulation of the main carrier.
f) The stereophonic subcarrier shall be capable of accepting audio frequencies from 50 to 15,000 cycles.
g) The modulating signal for the stereophonic subcarrier shall be equal to the difference of the left and right signals.
h) The pre-emphasis characteristics of the stereophonic subchannel shall be identical with those of the main channel with respect to phase and amplitude at all frequencies.
i) The sum of the sidebands resulting from amplitude modulation of the stereophonic subcarrier shall not cause a peak deviation of the main carrier in excess of 45 per cent of total modulation (excluding SCA subcarriers) when only a left (or right) signal exists; simultaneously in the main channel, the deviation when only a left (or right) signal exists shall not exceed 45 per cent of total modulation (excluding SCA subcarriers).
j) Total modulation of the main carrier including pilot subcarrier and SCA subcarriers shall meet the requirements of Section 3.268 with maximum modulation of the main carrier by all SCA subcarriers limited to 10 per cent.
k) At the instant when only a positive left signal is applied the main channel modulation shall cause an upward deviation of the main carrier frequency; and the stereophonic subcarrier and its sidebands signal shall cross the time axis simultaneously and in the same direction.
l) The ratio of peak main channel deviation to peak stereophonic subchannel deviation when only a steady state left (or right) signal exists shall be within plus or minus 3.5 per cent of unity for all levels of this signal and all frequencies from 50 to 15,000 cycles.
m) The phase difference between the zero points of the main channel signal and the stereophonic subcarrier sidebands envelope, when only a steady state left (or right) signal exists, shall not exceed plus or minus 3 degrees for audio modulating frequencies from 50 to 15,000 cycles.
Note: If the stereophonic separation between left and right stereophonic channels is better than 29.7 db at audio modulating frequencies between 50 and 15,000 cycles, it will be assumed that paragraphs l) and m) of this section have been complied with.
n) Cross-talk into the main channel caused by a signal in the stereophonic subchannel shall be attenuated at least 40 db below 90 per cent modulation.
o) Cross-talk into the stereophonic subchannel caused by a signal in the main channel shall be attenuated at least 40 db below 90 per cent modulation.
p) For required transmitter performance, all of the requirements of Section 3.254 shall apply with the exception that the maximum modulation to be employed is 90 per cent (excluding pilot subcarrier) rather than 100 per cent.
q) For electrical performance standards of the transmitter and associated equipment, the requirements of Section 3.317 a) 2), 3), 4), and 5) shall apply to the main channel and stereophonic subchannel alike, except that where 100 per cent modulation is referred to, this figure shall include the pilot subcarrier.

When the main carrier is modulated by the sum of the left and right stereo signals a conventional FM receiver accepts the radiation as a full monophonic signal. The addition of a stereophonic multiplex adaptor to the conventional FM receiver allows reception of other information if stereophonic transmission is being employed. Another signal which is the difference of the left and right stereo signals amplitude modulates a supersonic 38 kc subcarrier. This subcarrier is suppressed and the resulting sidebands are used to frequency modulate the FM transmitter in accordance with the foregoing Standards. The transmitter also is modulated by a 19 kc pilot subcarrier. At the multiplex stereo adaptor attached to the receiver the second harmonic of this pilot subcarrier is used to synchronize the suppressed 38 kc reference with the sidebands received.[14] Several means are possible in the FM receiver stereo adaptors for recovering the multiplexed signals.[15] Adaptors may be devised that are relatively simple in circuitry and more complex units approaching the price of the conventional receiver itself have reached the market.[16]

C. *Television (TV) Broadcasting*

In the area of television broadcasting there are several potential developments of major importance and a number of existing ones. The use of directional antennas has been proposed by the industry and explored by TASO. Very precise frequency control of television carriers to reduce cochannel interference has been extensively studied. A system of color field redundancy for the simplification and cost reduction of color television relaying has been proposed and operated experimentally. A system involving airborne television is in use for educational television broadcasting purposes.

1) In the field of television broadcasting directional antennas have long been employed to achieve antenna gain through the concentration of horizontally polarized radiation into a narrow vertical plane in order to achieve relatively high effective radiated powers. Very little use, however, has been made of directivity in the horizontal plane. However, a few directional antennas have been employed to suppress radiation toward ocean areas or mountain regions and achieve the utmost efficiency in the desired service directions. As has been noted, the Federal Communications Commission in its considerations in Docket No. 13340 has now given some indication that directional antennas for allocation purposes may be authorized. The general advantage would be to secure additional VHF stations in an already fully occupied portion of this frequency spectrum.

During the latter phases of the TASO program, tests of directional antennas were undertaken at WBZ-TV in Boston, Massachusetts, and at WKY-TV in Oklahoma City. The installation at WBZ consisted of two three-section superturnstile antennas mounted one above the other on the 1100-foot tower at WBZ-TV. For the purpose of the test the upper antenna was used as an experimental directional antenna. Transmission lines were so arranged that north-south or east-west pairs of superturnstile elements could be used for directive operation.[17] During the period of field measurements a power divid-

[14] V. H. Pomper, "Engineering your hi-fi system," *Research-Development*, vol. 12, pp. 15–27; March, 1961.
[15] D. R. VonRecklinghausen, "An FM multiplex stereo adaptor," *Audio*, vol. 45, pp. 21–23; June, 1961.
[16] A. Csicsatka and R. M. Linz, "FM stereo—the General Electric system," *Audio*, vol. 45, pp. 24–28; June, 1961.
[17] F. G. Kear and S. W. Kershner, "Determining the operational patterns of directional TV antennas," PROC. IRE, vol. 48, pp. 1088–1097; June, 1960.

ing network was switched to provide alternate 15 minute periods of nondirectional and directional operation with the same power input from the aural transmitter.

Three types of field measurements were made in the evaluation program. The first of these involved measurements along four radial routes from the transmitter, at distances ranging from about 9 to 50 miles. At each measuring location a continuous mobile recording over a distance of approximately 100 feet was made using a half-wave dipole receiving antenna mounted 30 feet above the road. At each location recordings were obtained over the same path for both directional and nondirectional operation during the adjacent 15 minute periods. The second type of measurements consisted of similar field strength determination at locations traversing a "cross minimum" route at distances ranging from about 18 to 22 miles. A third type of measurement involved recordings made at a distance of about 120 miles at a fixed location. All types of measurements served to indicate that the expected directivity was obtained and that there was reasonably good correlation between close and far fields.

The measurements at WKY-TV in Oklahoma City again utilized a modified three-section superturnstile antenna for directional purposes. In this case a special reference antenna mounted about 20 feet above the upper superturnstile elements was used. Both the superturnstile antenna and the reference antenna were equipped with motor driven mechanisms so that the antennas could be independently positioned to give desired orientations. The antennas were installed on a 263-foot supporting tower located some distance from the main antenna in use at WKY-TV. In these tests also the measurements indicated the expected directivity within a fair degree of approximation. On the basis of these tests, recommendations were made concerning the preferred methods for proving the performance of television directional antennas and establishing the validity of suppression which might be expected.

The observations at Oklahoma City showed that for a theoretical indicated suppression of 20 db in the directional pattern, measured suppressions of 16 or 17 db could be counted upon at considerable distances from the transmitter. Furthermore, the field strength averages over a long period of time approximated the calculated 20 db value. The tests at WKY-TV also indicated that with suppressions of the order of 20 db reflections of the main beam from nearby objects could reach a magnitude equal to or greater than the direct signal. Caution was therefore indicated in siting a television station requiring a directional antenna with a high degree of directivity so that the suppressed direction is toward the area of lowest population density. In summary, TASO recommended that until further observations and experience were gained with directional antennas they should not be used to provide greater than 15 db suppression and that a 2 db "fill-in factor" should be used as a safety margin for required suppressions of this order of magnitude.

Various other types of directional VHF antennas have been suggested, some employing the slot radiator techniques similar to current UHF practices. Multiple slots around the surface of a cylindrical antenna structure may be used to provide a variety of patterns.[18]

2) About five years ago it was suggested that very precise control of television carrier frequencies could be employed to reduce cochannel interference between television broadcast stations.[19] In the early 1950's "offset carrier" operation was incorporated into the television Table of Channel Assignments to take advantage of the fact that when carriers were offset by about 10 kc the visibility of the carrier beat pattern from television stations on the same channel was at a minimum. If picture carriers are not offset, as was the case in the early years of operation, the visible beat pattern took the form of large horizontal bars moving across the picture. With 10 or 20 kc offset operation (to take care of a triad of stations) the beat pattern was reduced to a relatively fine horizontal line structure. The work of Behrend of RCA showed that if the optimum offset frequency values could be held within a few cycles a further and finer grain arrangement of the beat pattern would obtain resulting in a further reduction of interference. While it had been possible with the offset carrier technique to achieve an improvement from a required desired-to-undesired signal to ratio of 43 db for non-offset conditions to 28 db for the offset condition, studies showed that for very precise offset techniques this ratio might be reduced to the order of 20 db.[20,21]

Tests made between the NBC stations in Washington and New York City and between New York City and the Westinghouse station in Boston, all involving Channel 4 facilities, have indicated the validity of the demonstrated interference reduction previously explored in laboratory experiments and controlled subjective tests. The tests of this technique by the Federal Communications Commission at its Laurel, Maryland, Laboratory has given promise of utilizing this new scheme to improve television service in fringe areas. Conversely, this tool might be used for allocation purposes and implemented together with directional transmitting antennas to allow greater occupancy of the presently restricted VHF portion of the spectrum.

3) In an extensive research program at the Iowa State University Engineering Experimental station,

[18] G. H. Brown, "Directional antennas for television broadcasting," IRE Trans. on Broadcasting, vol. BC-6, pp. 13–15; August, 1960.

[19] W. L. Behrend, "Reduction of co-channel television interference by precise frequency control of television picture carriers," *RCA Rev.*, vol. 7, pp. 443–459; December, 1956.

[20] L. C. Middlekamp, "Reduction of co-channel television interference by very precise offset carrier frequency," IRE Trans. on Broadcast Transmission Systems, vol. PGBTS-12, pp. 5–10; December, 1958.

[21] E. W. Chapin, "FCC laboratory observations of precision frequency control of TV stations," 1961 IRE International Convention Record, pt. 7, pp. 36–37.

Dr. William L. Hughes developed a system of color television using redundancy for simplification of the color television system.[22] The technique appears to give promise of cost reduction in color television transmission by network means and certain simplifications in color television pickup equipment including live camera, film camera, film and video tape systems. This description of one feature of the Iowa System has been given:

> A special color television camera is required for the modified system. In this camera the image is focused on the Y pickup and the chroma pickup through the use of a lens and beamsplitter arrangement. The Y filter, together with the Y pickup, has a total light spectral response of the luminosity function y. This luminance signal goes to the NTSC encoder without modification except for amplification and gamma correction. This signal carries full video band width. The image on the chroma pickup is modified by the color filters in the color filter wheel. The chroma pickup is really a field sequential pickup, except that the field sequency is simply red, blue, red, blue, and so on. The color filter wheel is driven by a motor which is controlled by the vertical synchronizing pulse. Thus, at any one time, a luminance signal and one piece of color information, either red or blue, results. (Now comes the part which one has to see if he is to believe that it works, and it does work very nicely.) Each color field is delayed approximately 1/60 of a second. Then the result of any one time is a luminance signal, a red (or blue) signal, and a blue (or red) signal that is 1/60 of a second old. The three signals are sufficient to make a full three-color picture in the standard manner and they can be fed to an NTSC encoder as from three camera tubes.

4) Plans to transmit educational television from airplanes have developed to the operational phase. As an outgrowth of the Stratovision program engaged in by Westinghouse and the Glenn L. Martin Company both of Baltimore, Maryland, in 1948, DC-6 aircraft carrying UHF TV transmitters are broadcasting over Montpelier, Indiana. Flying at 23,000 feet, present indications are that service is being provided to participating schools within a radius of about 200 miles. Operation is on UHF Channels 72 and 76. Ultimately plans call for the use of special spectrum conservation techniques to have each transmitter broadcast two programs simultaneously. Special narrow band receivers under development by CBS Laboratories will be a factor in providing increased channel efficiency.[23]

Video tape recorders are used to provide the program material involved. Operations are under the auspicies of the Mid-West Council on Airborne Television Instruction. The Purdue Research Foundation and Purdue University are assisting in the project. The special antenna system, which is retractable for take-off and landing is maintained in a vertical plane gyroscopically. Initially programs were transmitted three hours a day for four days a week. It is anticipated that this schedule will be increased to six hours a day Monday through Friday after preliminary operational experiments are completed.

D. Experimental, Auxiliary, and Special Broadcast Services

FM transmitters are often located at remote points where high ground elevations are available and good effective antenna heights can be obtained. In many such locations wire line facilities are not available for either program material or remote control provisions. Arrangements may be made under Part 4 of the Rules to use radio relay for program or remote control purposes. There is also some networking between FM stations in adjacent areas on a direct pickup and rebroadcast basis. In the educational FM area of operation there is a system in operation utilizing radio relay frequencies which provide for talk back from receiving locations to the program source. Such facilities result in the ability to provide a "radio classroom."

1) FM transmitter sites which are beyond wire service may now be utilized employing radio facilities for interconnection with program sources. Radio remote control can also be accomplished in the 942–952 Mc band. Such an interconnection can provide for positive transmitter control circuits, a means of relaying monitoring information from the transmitter to the control point, program relaying and provision of means of communication between the studio and transmitter.[24] The use of simplex supersonic tones provides a reliable and relatively simple means of establishing transmitter control. Fail safe operation of the remote FM transmitter can be accomplished by a carrier operated relay in the link receiver so that in the event of link failure the transmitter would be removed from the air. Transmitter metering information also may be returned to the control point by multiplexing this information on the main transmitter FM carrier.

2) FM radio relay techniques have been demonstrated in the past and are currently being used in the WQXR-New York Times network. A suggestion has been put forward for an educational FM network to use this technique plus a multiplex adjunct to provide both radio remote control of isolated FM transmitters and a program relaying circuit.[25] Thus in addition to relaying the main channel program, a second audio program channel would be relayed via a multiplexed subcarrier at about 67 kc. The second audio program channel would be independent of the main program and would permit simultaneous two-way program relaying. Remote control is proposed by means of an audio channel multiplexed on each of the stations at about 41 kc. This subcarrier would carry coded tones to convey the control information to the remote stations. Other coded tone combinations would be used for the return of the remote transmitter monitoring information using this same subcarrier frequency multiplexed on the remote station.

[22] W. L. Hughes, "The Use of Color Field Redundancy in the Simplification and Cost Reduction of Color Television Transmission Systems," Iowa State Engrg. Experiment Station Rept., pp. 1–40; November, 1960.

[23] M. T. Decker, "Service area of an airborne television network" (Abstract), IRE Trans. on Broadcasting, vol. BC-6, p. 20; August, 1960.

[24] J. A. Moseley, "Design considerations for radio remote control operation," *Broadcast Engrg.*, vol. 3, pp. 10–12; March, 1961.

[25] R. E. Peterson, "The engineering aspects of an educational FM network," IRE Trans. on Broadcasting, vol. BC-6, pp. 5–11; September, 1960.

The remote control of the transmitters and the network switching would be broken down into four general functions. These would be audio switching control, transmitter function control, transmitter indicators, and supervisory indicators.

3) Considerable experience has been gained with two-way radio talk back for medical education in the Eastern New York and New England area.[26] WAMC located atop Mount Greylock, Massachusetts, utilizes an effective radiated power of 10 kw at 90.3 Mc. The effective antenna height is almost 2000 feet. Programs are sent to the transmitter site on a 940 Mc link from the Albany Medical College studios 40 miles away. Remote pickup transmitters in the 153 Mc band are used at the participating hospitals (some 30 in number) to transmit questions back to the Mount Greylock site which is on the highest mountain in Massachusetts. Thus, anything transmitted from either a remote hospital or from the originating studio at Albany may be heard by anyone with an FM receiver capable of receiving the broadcast signal. Specifically, this includes any medical men who can take time from their office duties to listen to these mid-day "round tables" from their homes.

A tone alerting system has been installed throughout the main conference network. This consists of tone sources at the remote participating locations and tone responsive circuits at the originating studio control room. When a participating hospital has a question they may use their tone device to actuate a signal which flashes a light to the medical program moderator in the main studio. Through a suitable control board function, the identity of the remote participating hospital may be recognized, and time cleared for the question. This technique need not be limited to medical education and it appears to afford greater flexibility than similar arrangements for talk back which have been made using wire circuits.

International Broadcasting

International broadcasting, a service of the United States Information Agency, is known as the Voice of America. Studios are located in the Health, Education and Welfare Building in Washington, D. C. The master control board is reputed to be the largest and most versatile in the world. Especially designed and built for the Voice of America facilities, it can select programs from 100 different sources and originate 26 programs simultaneously. Preset control is utilized to permit a 15 minute lead time set up and checkout in advance of program change. Thirty-five different languages are used plus a large number of dialects to reach a world-wide audience.[27]

Presently 30 short-wave transmitters at seven locations in the United States utilize powers ranging between 25 and 200 kw. Overseas the Voice of America has provided nine relay stations with 47 transmitters ranging in power from 35 to 1000 kw. Taken as a whole, the operation is on a 24-hour a day basis. Direct radio propagation from the United States to many important target areas of Eastern Europe and Asia is prevented by the shielding effect of the auroral zone. This necessitates the use of relay stations to overcome the vast distances involved and enable the "Voice" to reach the desired listener with strong signals competitive to local stations in the areas involved. The problems of overcoming Communist jamming are severe and the multiplicity of additional sources and flexibility in frequency utilization helps combat this problem.

Satellite relays for communication and broadcast relaying have been proposed.[28] To accomplish relaying by satellites wide band modulation must be used to secure the required information through noise. The expectancy is that a frequency channel of 250 Mc bandwidth would be required for each channel served. A channel could consist of 600 voice circuits or one television circuit. Some studies indicate that as many as 20 such channels might be needed requiring a spectrum of some 5000 Mc in width. In all probability because of this tremendous demand for frequency spectrum, many different satellites would be required so that many pairs of ground stations could reuse the same frequencies by orienting highly directive antennas at different satellites. The Federal Communications Commission is currently studying several proposals put forward for such purposes by common carrier and manufacturing interests. A satellite broadcasting station theoretically could put out a signal that could be received on home receivers. A synchronous satellite operating at about 22,000 miles height and having the same period of rotation as the earth could conceivably provide service to an area the size of the United States. The power requirements to provide such a system, however, would seem to be beyond the reach of immediately foreseeable techniques to solve. It does appear, however, that intercontinental television and communication by means of relaying by satellites will be feasible within the next few years.

[26] A. P. Fredette, "40,000-square mile classroom," and C. M. Jansky, Jr., "A comment," *Natl. Assoc. Educational Broadcasters J.*, vol. 20, pp. 1–5; May–June, 1961.

[27] G. Jacobs and E. T. Martin, "The international broadcasting system of the voice of America," 1961 IRE International Convention Record, pt. 7, pp. 25–35.

[28] J. H. Felker, "Satellite Relays and Broadcasting," Natl. Assoc. of Broadcasters, Engrg. Conf.; May 10, 1961.

Section 7

CIRCUIT THEORY

Organized with the assistance of the IRE Professional Group on Circuit Theory

Summary of the History of Circuit Theory*

V. BELEVITCH†, FELLOW, IRE

Summary—**After a brief survey of the state of circuit theory before World War I, the various directions of its development in the last 50 years are discussed, mainly in relation with applications to communication engineering. The early period of network design (1920–1925) was followed by the beginning of synthesis (1926–1935). The next steps were the development of feedback amplifier theory and insertion loss filter theory. The numerous new directions of research started during and after the second war are briefly mentioned. Finally recent progress is reported in formal realizability theory and in topological synthesis. A last section deals with nonlinear and linear variable circuits.**

Introduction

Although circuit theory is more than 100 years old (Ohm's law, 1827; Kirchoff's laws, 1845), it seems that no systematic account of its historical development has ever been written. The present essay attempts to cover the last 50 years, the fiftieth anniversary of the IRE being taken as an excuse to exclude the more distant past. This limitation is also justified by the development of circuit theory itself, which shifted from a steady to an accelerated progress a few years before World War I, simultaneously with the expansion of communication technology following the invention of the vacuum tube (de Forest's audion, 1906). This growth of circuit theory is directly testified by the number of articles published per year, which remained below unity till 1910 and increased from 5 to 25 in the period 1920–1940; after a drop during World War II, the increase continued, and a figure of 100 was exceeded in 1954.

We start with a brief survey of the state of the theory just before World War I and discuss in separate sections the various directions of its development up to the present days, mainly in relation with applications to communication engineering. Due to space restrictions, bibliographic references are omitted altogether: at least 200 important contributions out of a total of some 2000 would deserve mention. Important authors and dates are simply quoted in the main text; the dates generally refer to publications in regular scientific journals, for it was materially impossible to search through patents, theses and reports.

Circuit Analysis Before 1914

Long before 1914 circuit theory had emerged, from general electromagnetic theory, as an independent discipline with original concepts and methods. The conception of a circuit as a system of idealized lumped elements is already firmly established—drawings of Leyden jars and rheostats have gradually disappeared in favor of the now familiar graphical symbols. This assumes, at least implicitly, that a resistor is considered as a 2-terminal *black box* defined by the relation $v = Ri$, rather than as a physical device made of metal or carbon. This abstract point of view becomes

* Received by the IRE, May 22, 1961.
† Comité d'Etude des Calculateurs Electroniques, Brussels, Belgium.

more prominent in the modern developments, where the laws defining the elements, and the interconnection constraints (Kirchoff's laws) are explicitly taken as postulates. With this approach, *network theory*[1] only studies the properties of systems of elements and becomes a purely mathematical discipline dealing with abstract structures generated by various sets of postulates. Another consequence of this point of view is that the physical devices themselves are no longer studied by network theory but by what is sometimes called *device theory;* complicated physical devices are then naturally described by their *equivalent circuits.*

Conventional network elements are linear and time-independent; moreover, the stored electric and magnetic energies, and the dissipated power, are positive definite quadratic forms, and the constraints are instantaneously workless. This reduces the problem of network analysis to the classical theory of small vibrations in dissipative mechanical systems. The use of complex variables to combine the amplitudes and phases of harmonic steady states, and the separation of transients into normal modes, all familiar in analytical dynamics, were naturally taken over by circuit analysts. In particular, Steinmetz (1894) vulgarized the use of complex quantities in electrical engineering. Some specifically electric properties were, however, also established; we mention Kirchoff's topological rules (1847) and their extension by Feussner (1902–1904), the so-called Thévenin's theorem (Helmholtz, 1853), the star-delta transformation (Kennelly, 1899) and the concept of duality (Russel, 1904).

The impedance concept was fluently used by Heaviside, with p treated both as a differential operator and as an algebraic variable. The use of complex values for p is physically justified by Campbell in 1911. Although the relation between Heaviside's operational calculus and integration in the complex plane was only clarified after 1916, the mathematical tools allowing to derive transient behavior from harmonic response were available, since all is needed for lumped circuits is Heaviside's formula based on a partial fraction expansion. It took, however, 20 more years to introduce the terminology of poles and zeros into common engineering practice.

Although network analysis is, in principle, already separated from line theory, both branches remain in close contact in their later evolution, so that a few words about the latter are not out of place. Classical line theory has been worked out in detail by Heaviside, and such concepts as matching, reflection, iterative impedance, etc. began to be transferred from continuous to discrete structures, probably through the intermediary of artificial lines, a subject to which Kennelly devoted a book in 1917. Although the general concept of quadripole, or 2 port, as a black box characterized by its voltage equations, does not explicitly appear before 1921 (Breisig), attenuators and artificial lines calibrated in "equivalent length of standard line" were used earlier for telephone transmission measurements (the decibel was only introduced in 1924). Finally, Heaviside's invention of inductive loading successfully applied by Pupin led Campbell, through the theory of iterative structures (1903), to the invention of the electric filter to be discussed in the next section.

[1] Our restriction of the term *network* (as opposed to *circuit*) to situations where reference is made to an idealized model conforms with the IRE Standards ("IRE Standards on Circuits: Definitions of Terms for Linear Passive Reciprocal Time Invariant Networks, 1960" (60 IRE 4.S2), Proc. IRE, vol. 48, p. 1609; September, 1960). Both terms are in fact synonymous in common usage. Another distinction has sometimes been made which reserves *network* for systems with free terminals or terminal-pairs (Campbell and Foster, 1920); in this terminology, an n-port becomes a circuit when terminations are specified at all ports.

Network Design, 1920–1925

The progress in circuit theory during this period is connected with the development of long distance telephony in more than one respect. First, the theory of loaded lines is closely related with the invention of the electric filter, which found an immediate application in carrier telephony (first realiztion in 1918, Pittsburgh-Baltimore). Secondly, the design of bidirectional amplifiers (two-wire repeaters) immediately raised several problems of network synthesis (hybrid coils, balancing networks, filters, equalizers), not to mention the question of over-all stability. Although satisfactory engineering solutions were rapidly found to all these problems, their theoretical foundation was insufficiently systematic—the art of design, with its cut and try procedures had not yet matured into modern synthesis.

The electric filter was invented during World War I, independently by Campbell and Wagner. The first filters were iterative ladder stuctures, although lattice sections (in particular all-pass sections) are discussed by Campbell (1920–1922). After Zobel's invention of m-derivation (1923–1924), the catalogue of elementary sections was sufficiently extended to cover most practical requirements. In the spirit of line theory, a filter was designed as a cascade connection of sections with matched image-impedances, but the difference between image attenuation and insertion loss was thoroughly estimated by Zobel.

To Zobel are also due the understanding of the ideal transmission conditions (frequency independent attenuation and linear phase), the design of constant impedance equalizers (1928) including a method of approximation to prescribed frequency characteristics, the discovery of bandwidth conservation in the low-pass /band-pass transformation, and the first correction for mutual bridging effects in a parallel connection of two filters (x-derivation).

Simultaneously with the development of filter theory, the general concept of a quadripole, with its impedance, admittance and chain (ABCD) matrices was introduced in Germany and France, and the rules for computing the matrices of series, parallel and tandem combinations of quadripoles were discovered. One distinguished between 2-terminal pair networks and 3-terminal networks (grounded quadripoles), and the combination rules were correctly restricted to cases where no "longitudinal voltage" appears. Here again, the separation of a general disturbance into a transversal and a longitudinal wave (using both conductors as "go circuit" and the earth as "return") was familiar in line theory, and effectively used in long distance telephony (phantom circuits) since 1903. A similar separation into three modes (symmetrical components) for three-phase power circuits was introduced by Fortescue in 1918.

The concept of n-terminal-pair network, or n-port, does not seem to have appeared during this period. It occurs, however, implicitly in a 1920 paper by Campbell and Foster, which is probably the first publication on network synthesis in the true sense—the energy relations in matched nondissipative 4 ports are discussed, the biconjugacy of the networks is proved, and a complete enumeration of all realizations is given; moreover, the circuits (which include the familiar hybrid coil) are explicitly treated as composed of ideal transformers, a new network element whose theory is established.

Acoustical telephone repeaters (telephone+microphone) were used before the invention of the vacuum tube, and difficulties due to singing were experienced. It is not easy to trace the evolution during the first war, and the post war publications already show a well-established technique. This is testified by the Campbell-Foster paper of 1920, by the design formulas for balancing networks (Hoyt, 1923–1924), and by the recognition of the fact that singing is essentially limited by line irregularities (Crisson, 1925).

The Beginning of Synthesis, 1926–1935

Although Foster's proof of his reactance theorem (1924) is already a transition from the methods of analytical dynamics to those of modern network synthesis, the first paper dealing explicitly with the realization of a one-port whose impedance is a prescribed function of frequency is Cauer's 1926 contribution, based on continuous fraction expansions (also studied by Fry, 1926). With Cauer's and Foster's theorems, the synthesis problem for one ports containing two kinds of elements only was solved. The analogous problem for general one ports was solved by Brune (1931) and led to the concept of positive real function. Pomey (1928) proved that the real part of the impedance matrix of a general passive 2-port was positive definite at real frequencies. The remaining developments of this period deal mainly with nondissipative n-ports; although Gewertz (1933) found a synthesis method for general 2-ports (containing all three kinds of elements), the general problem for n-ports was only solved after World War II and will be discussed in another section. The concentration on LC networks is closely connected with filter theory and its engineering interest. Another aspect of network synthesis, the approximation problem, made also its appearance during this period; the maximally flat approximation was used by Butterworth (1930) in the design of multistage amplifiers; simultaneously and independently, Cauer realized the optimal character of the Chebyshev approximation and solved the approximation problem for an important class of image-parameter filters. Finally, it should be remarked that the canonical structures obtained as solutions of the various synthesis problems made a free use of ideal transformers; the much more difficult problem of synthesis without transformers was not of paramount interest for communication applications and has only been treated recently.

The simplest network after the one port is the symmetric 2-port, which involves two frequency functions only. Geometrically symmetric 2-ports were treated by Bartlett (1927) and Brune (1932), whereas Cauer (1927) and Jaumann (1932) found a number of canonical circuits for all symmetric 2-ports. Dissymetric, and, in particular, antimetric 2-ports were studied by Cauer, who also extended Foster's theorem to LC n-ports (1931) and showed (1932–1934) that all equivalent LC networks could be derived from each other by the linear transformations considered by Howitt (1931). Certain classes of symmetric n-ports were studied by Baerwald (1931–1932).

Cauer's first book on filter design (1931) contains tables and curves for the Chebyshev approximation to a constant attenuation in the stop-band of an image parameter low-pass filter, as well as frequency transformations for other filter classes. The solution of the approximation problem involved rational functions whose extremal properties were established by Zolotareff in 1877 and which reduce to ordinary Chebyshev polynomials when elliptic functions are replaced by trigonometric functions. Cauer's presentation of his design data was based on canonic structures, practically less convenient than ladder structures, and was not accepted in engineering circles before it was realized that the statement and the solution of the approximation problem were of interest in themselves, for most of Cauer's results could easily be transferred to the ladder structure. The systematic theory of image parameter filters was further developed by Bode (1934) and Piloty (1937–1938), thus placing Zobel's earlier results in a clearer perspective. The particular problems raised by crystal filters were studied by W. P. Mason (1934–1937).

Feedback Amplifiers, 1932–1945

The construction of oscillators was one of the first applications of vacuum tubes, and the amplification increase due to a positive feedback below the threshold of oscillation is mentioned in several early publications on radio-engineering. The elementary, but erroneous, physical reasoning based on the round and round circulation of the signal in the feedback loop yielded the geometric series $\mu(1+\mu\beta+\mu^2\beta^2+\cdots)$ for the effective amplification. This gave the impression that instability must occur for all $|\mu\beta|>1$, for the series then diverges.

Stable behavior with a negative feedback exceeding 6 db was thus apparently forbidden, and such low values have little practical interest. It was soon recognized that the theory based on the geometric series was contradicted by experience, and the expression $\mu/(1-\mu\beta)$ was used even for $|\mu\beta| > 1$ without theoretical justification until Nyquist (1932) proved his famous stability criterion and showed the error in the older theory—the physical reasoning based on the loop circulation is only correct when transients are taken into account, the terms in the geometric series then becoming convolution products; the steady-state formula is obtained as the limit after infinite time; the transient series is convergent, but not always uniformly, so that the limit operation and the summation cannot be interchanged. The case of conditional stability, predicted by Nyquist's theory, was experimentally confirmed (Peterson, *et al.* 1934) and the advantages of negative feedback were systematically discussed by Black (1934).

The domestication of negative feedback made it possible to design the wide-band highly linear amplifiers required for multichannel carrier telephony. Although various designs for interstage networks were proposed, for instance by Wheeler and Percival, the limitations due to parasitic capacities, and the way of overcoming them by an optimum over-all design, were only clarified by Bode in 1940, with the help of the integral relations between attenuation and phase.

The fact that the real and imaginary parts of physical network functions could not be specified independently from each other was first stated in connection with the ideal filter paradox. Küpfmüller's treatment (1926) of the transient response to a unit-step led to the well-known sine-integral embodying a response preceding the stimulus. Restrictions imposed by causality to physical response functions are mentioned by Y. W. Lee (1932) in connection with a method of synthesis for arbitrary transfer functions based on Fourier transforms, a method patented by Lee and Wiener (1938). Physical approximations to the ideal filter response, both in amplitude and phase, are discussed by Bode and Dietzold (1935). Explicit integral relations between real and imaginary parts of various network functions were studied independently by Cauer (1932–1940), Bayard (1935) and Leroy (1935–1937), but similar relations were known earlier in the theory of optical dispersion (Kramers-Kronig, 1926).

Bode extended the relations to the case where either component is specified in a partial frequency range, and worked out their consequences for input, output and interstage networks with prescribed parasitic capacities. He also computed the maximum obtainable over-all feedback in terms of band width and asymptotic loop transmission. The stability criterion for multi-loop amplifiers was obtained by Llewellyn, and the effect of feedback on impedances was discussed by Blackman (1943). Bode introduced the concepts of return reference and sensitivity in his classical book (1945).

Insertion Loss Filters and Related Problems

The limitations of image-parameter theory first appeared in connection with the design of filter groups, a problem frequently encountered in carrier telephony. Zobel's procedure of x-derivation, already mentioned, was first replaced by a more systematic method of impedance correction (Bode, 1930). An image-parameter theory of constant impedance filter pairs was developed by Brandt (1934–1936), Cauer (1934–1937) and Piloty (1937–1939), and it was recognized that this also yielded a solution to the equivalent problem of open-circuit filter design.

A completely new approach to the whole problem is contained in Norton's paper (1937) on constant impedance filter pairs, where the method of design starting from a prescribed insertion loss is established. The general synthesis problem for a reactance 2-port with prescribed insertion loss was solved independently by Cocci (1938–1940), Darlington (1939), Cauer (1939–1941) and Piloty (1939–1941). These contributions establish the canonic realization of a reactance 2-port as a ladder structure with mutual inductances restricted to adjacent or nearly adjacent arms and, as a consequence, the possibility of realizing an arbitrary passive one port by a network containing one resistance only. The approximation problem for insertion loss filters was reduced to the similar problem for image-parameter filters. Finally, Darlington also devised a method for precompensating the dissipative distortion.

As already mentioned, the prewar evolution of network theory was closely related with the development of wire communication, and the perfection reached by filter and feedback amplifier theory around 1940 made possible the design to strict tolerances which is required in long distance telephone equipment. On the contrary, the easier narrow-band problems of radio-engineering were treated by elementary circuit analysis, and it is only with the advent of video and pulse techniques that the theory of wide-band multi-stage amplifiers (without over-all feedback) underwent a systematic development, mainly during World War II. Another direction of war-time evolution was the extension of filter techniques to higher frequencies, leading to transmission line filters and microwave networks. Both directions have influenced classical filter theory and, as a consequence, common mathematical methods are now used in an extended field.

In microwave applications, the classical description of network performance in terms of voltages, currents, impedance and admittance matrices, was naturally replaced by a description based on transmitted and reflected wave amplitudes, leading to the concept of scattering matrix (Montgomery, Dicke and Purcell, 1948) taken over from general physics. This concept is also of interest in the field of lumped networks, where it was introduced independently and simultaneously by Belevitch. The scattering formalism allowed an easier presentation of insertion loss filter theory, and has been of great help in other applications to be discussed later.

The war-time progress in amplifier design is described in the book of Valley and Wallmann (1943), but similar work has been done independently in Germany, namely by Cauer (posthumous publications). Most input and inter-stage circuits are actually ladder filters, terminated or open-circuited, so that filter and amplifier problems are closely related. Explicit formulas for the element values of various important classes of ladder filters have been investigated by many workers, but it was recently realized that most results had been anticipated by Takahasi (paper in Japanese, 1951; English adaptation, 1960). The design of input and output circuits is also related with the broad-band matching problem, which consists in constructing a 2-port which transforms a given frequency-dependent output load (for instance a resistance shunted by a parasitic capacitance) into a pure input resistance. A rigorously constant resistance can be obtained by a lossy 2-port, but it is practically important to know the approximation to matching in a given frequency range obtainable from a lossless 2-port, and to see how the transmission loss varies with the degree of impedance equalization for lossy 2-ports. The broad-band matching problem with lossless 2-ports was solved by Fano (1950) for important classes of load impedances. The relations for lossy 2-ports were obtained by Carlin and La Rosa (1952–1955) with the help of the scattering formalism.

Distributed amplifiers, which overcame the limitations imposed by parasitic capacities to stagger-tuned circuits, had been invented by Percival in 1937, but were only practically exploited after 1948; the problem of their optimum design is still unsolved, although some progress has been recently achieved.

The scattering formalism proved useful in the design of various classes of n-ports of interest in telephone applications. Belevitch (1950–1955) discussed matched n-ports with equal losses between any couple of ports, and biconjugate n-ports for $n>4$. Dosoer (1958) and Oswald (1958) established design methods for a class of filter 4-ports (invented by Darlington in 1938) used in submerged repeaters.

The methods of best approximation used in filter and amplifier design were applied to other problems, namely to the design of various classes of delay networks. The maximally flat approximation for the group delay was obtained by Thomson (1949), and the Chebyshev approximation by Ulbricht and Piloty (1960). The Chebyshev approximation to a constant phase difference between two all-pass 2-ports (of interest for polyphase modulation) was obtained independently by Darlington, Orchard and Saraga (1950), and a mathematically equivalent problem on the phase angle of an impedance (of interest in feedback amplifier design) was solved by Baumann (1950). For problems having no exact analytic solution, Darlington (1952) described a procedure based on a series expansion in Chebyshev polynomials, whereas the potential analogy and the related electrolytic tank technique were practically applied at least since 1945 (Hansen and Lundstrom).

New Trends in Post-War Evolution

The field of application of circuit theory extended in so many new directions during and after World War II that only the major tendencies can be outlined. Several new developments lie on the borderline between circuit theory and other disciplines (information theory and noisy systems, electronic computers, automatic control), and will not be discussed. The present section attempts a brief classification of the new ideas, concepts and methods, in relation with their engineering applications; the next two sections review in detail the postwar developments of network theory *stricto sensu*. Nonlinear and linear variable circuits are treated separately in the last section.

Pulse techniques, already mentioned, raised various synthesis problems in the time domain. Although every problem stated in the time domain is, in principle, equivalent to a problem in the complex frequency domain, the approximation requirements are often difficult to translate from one domain into the other, and the convergence properties in the two domains may be quite different. Various mathematical methods, such as Laguerre functions (Lee, 1932), time series expansions (Lewis, 1952), complex integration (Guillemin and Cerillo, 1952), have been used, but there seems to have been little fundamental progress; for instance, the problem of the steepest monotonic response in presence of a given parasitic capacity is not completely solved, although important contributions have been published, namely by Zemanian (1954) and Papoulis (1958).

Microwave circuits have also been previously mentioned. The realization of a microwave gyrator (Hogan, 1952), based on the Faraday effect in ferrites, justified the interest of the ideal gyrator concept (Tellegen, 1948) and promoted theoretical research on nonreciprocal networks initiated by Tellegen (1948–1949). Theoretical work on the synthesis of passive n-ports (both reciprocal or not) is discussed in the next section. Its practical importance is due to the fact that, without the limitations of reciprocity, better performances can be obtained (interstage networks; Tellegen, 1951) or otherwise impossible behavior can be achieved (design of circulators: Carlin, 1955).

Progress has been slower in the theory of active networks because it was more difficult to represent the real physical devices by simple ideal elements; needless to say, even greater difficulties await the circuit theorists of the future, with the advent of integrated microelectronic devices. Classical amplifier theory dealt in fact with passive networks separated by ideal unilateral buffer stages, and amplifiers with large negative feedback were used to realize the active transfer function $-1/\beta$ involving only the transfer ratio β of the passive feedback network. With the advent of the transistor

(Bardeen and Brattain, 1948), practical design problems became much more difficult due to the internal feedback of the device. The fundamental limitations of active non-unilateral devices, at a fixed frequency, were progressively understood; for instance, the conditions for intrinsic stability were established by Llewellyn (1952), power invariants were found by S. J. Mason (1954), and noise invariants by Haus and Adler (1956–1959). The analysis of complex feedback structures was clarified by the use of the flow-graph notation (S. J. Mason, 1953). After the introduction of gyrators, it appeared more convenient to represent active devices by an equivalent circuit containing gyrators and (positive and negative) resistors, so that negative resistance remained the only new element in the theory of active networks. Although various negative resistance effects (arc discharge, dynatron, etc.) had been known for many years, practical and economic devices simulating linear negative resistors became available only with solid-state components (negative impedance converter: Linvill, 1953; tunnel diode: Esaki, 1958), and this favored the adoption of negative resistance as the basic element in theoretical work. Recent results in the theoretical field are summarized at the end of the next section.

Returning now to the synthesis of passive reciprocal networks, we briefly comment on the progressive separation between formal realizability theory (where ideal transformers are freely accepted) and topological synthesis (where even mutual inductances are excluded). The recent interest in topological methods originated from several distinct fields. First, mutual inductances, and even self-inductances, are difficult to realize at very low frequencies; this stimulated research on RC circuits, mainly for servomechanism applications, although similar problems arose earlier in the design of RC-oscillators (invented by van der Pol and van der Mark in 1934). Secondly, even at higher frequencies, it may be economical to replace inductances by capacitances combined with positive and negative resistances; this possibility was known theoretically before 1930, but became only practical with the availability of solid-state devices. Topological problems also arose in the design of contact networks, and some recent developments in the theory of switching are related with various fundamental problems of conventional network theory. Finally, the treatment of network problems on electronic computers asked for a more complete and detailed algebraization of all topological notions and raised various enumeration problems. Topological analysis was recently discussed by many workers, such as Bryant, Okada, Percival, Seshu and Watanabe; for combinatorial and enumeration problems, we refer to the book of Riordan (1958).

Formal Realizability

The synthesis of passive reciprocal n-ports has been achieved by three methods. The first one extended Gewertz's procedure for 2-ports and consisted in the successive extraction of reduced impedance matrices; the process is heavy and laborious, and only of historical interest; it enabled, however, Oono (1946) and Bayard (1949) independently to prove that any positive real impedance or admittance matrix is realizable, thus showing that RLC elements and the ideal transformer constitute a *complete* system of passive reciprocal elements. After a discussion of the 2-port case by Leroy and Belevitch (1949), it appeared that the first method was actually a disguised extension of Darlington's process for one ports, *i.e.*, the realization of an n-port as a reactive $2n$-port closed on n resistors. The second method consists in a direct application of this idea; it was used by Bayard (1950) and with the scattering formalism by Belevitch (1951). The third method extends Brune's process for one ports and arrives to a structure containing the minimum number of reactive elements; synthesis by this method was achieved independently by Oono (1948), Mac Millan (1948–1952) and Tellegen (1953) and contributed to a satisfactory definition of the degree of a rational matrix. In an important paper (1954), Oono and Yasuura rediscuss the synthesis by the second method, using the scattering matrix both for reciprocal and nonreciprocal n-ports and solve completely the equivalence problem.

In connection with the design of narrow-band unsymmetrical band-pass filters, Baum (1957–58) introduced a new fictitious passive element, the imaginary resistance (or frequency-independent reactance), and showed its interest in various synthesis problems. Using this concept, Belevitch obtained a simple derivation of Brune's for one-ports (1959) and extended Tellegen's process to nonreciprocal n-ports (1960).

Extensions to active networks were delayed due to difficulties arising with certain pathological n-ports admitting none of the conventional matrix descriptions (Tellegen, 1954; Carlin, 1955). These difficulties were circumscribed by Oono (1960), and by Carlin and Youla (1960). The latter showed that any active n-port is realizable as a passive nondissipative $3n$-port closed on n positive and n negative resistors. In a companion paper, Youla and Smilen (1960) establish the gain-bandwidth limitations due to parasitic capacitances in negative-resistance amplifiers and derive design formulas for optimum amplifiers.

Topological Synthesis

Logically, but not historically, the first problem arising in this field is the one of discriminating between constraints which are realizable without ideal transformers and general workless reciprocal constraints. In synthesis problems, constraints generally appear under the form of prescribed incidence relations between loops and branches, and the problem is to determine the necessary and sufficient conditions under which a prescribed incidence matrix corresponds to a graph. A necessary condition is well known (the matrix must be totally unimodular, *i.e.*, have all its minors of all orders equal to 0 or ±1), but is certainly not sufficient, a classical counter-example being the dual of the constraints in a nonplanar graph (Foster, 1952). Sufficient conditions were found by Tutte (1958–1959). Various algorithms have recently been devised which either prove a given matrix to be unrealizable, or lead to a unique realization, within trivial isomorphisms. The first such algorithm is due to Gould (1958) and was applied to the synthesis of contact networks.

The next problem is the synthesis of resistance n-ports, and of resistance networks having $n+1$ terminals (grounded n-ports). It was known for some time that dominancy (any diagonal element not smaller than the sum of the moduli of all elements in the same row) was sufficient, but not necessary, for the admittance matrix of an n-port. Necessary and sufficient conditions for a 3-port with prescribed impedance or admittance matrix were obtained by Tellegen (1952) by expressing that the network cannot yield current nor voltage gain under any set of open- or short-circuit conditions. A nontrivial extension of Tellegen's approach led Cederbaum (1958) to the condition of paramountcy (any principal minor not smaller than the modulus of every minor based on the same rows), which is weaker than dominancy. Paramountcy is necessary for admittance and impedance matrices, but its sufficiency is not established for impedance matrices with $n>3$. For $(n+1)$-terminal networks, the synthesis is equivalent to to a congruence transformation into a positive diagonal matrix by means of a realizable constraint matrix. An algorithm yielding a unique (except for trivial variants) transformation, if possible at all, has been found by Cederbaum, but this algorithm accepts totally unimodular transformation matrices, which are not necessarily realizable. Related synthesis procedures were described by Guillemin (1960) and Biorci (1961).

For two-element kind one ports, the canonical structures of Foster and Cauer contain no transformers. It has long been thought that the transformers appearing in Brune's process are unavoidable in the synthesis of general three-element kind one ports; a canonical realization without mutual inductance was, however, published by Bott and Duffin (1949). In spite of various small improvements (Pantell, 1954; Reza, 1954), the method is quite wasteful in elements; no further general improvement has been obtained, but procedures due to Miyata (1952) and Guillemin (1955) sometimes yield more economical realizations.

The synthesis of RC 2-ports was first treated by Guillemin (1949), who showed that the class of grounded RC 2-ports was practically not narrower than the class of general reciprocal 2-ports, in the sense that the modulus of the input-output voltage ratio of any 2-port of the second class can be arbitrarily approximated within the first class, except for a constant multiplier. Guillemin's synthesis was in terms of parallel ladders; Orchard (1951) established a synthesis by RC lattices and discussed the extraction of terminal resistances in order to realize prescribed insertion loss functions. The approximation problem was treated by Ozaki and Fujisawa (1953). Miscellaneous synthesis procedures were discussed by a number of authors, but the most general and complete results are due to Fialkow and Gerst in a half-dozen papers published between 1951 and 1955. These authors found the necessary and sufficient conditions for voltage ratios of grounded and nongrounded RC and RLC 2-ports, as well as for various restricted structures (ladder, lattice); they indicated canonic realizations and discussed the value of the constant multiplier which fixes the maximum available voltage gain. Articles by Cederbaum (1956) and Kuh and Paige (1959) bring certain additional precisions.

The synthesis of RC grounded 2-ports with prescribed admittance or impedance matrices is still an unsolved problem. A series-parallel synthesis procedure was described by Ozaki (1953), and sufficient conditions under which it succeeds are known, but have not been proved necessary. The problem progressed through contributions of Lucal (1955), Slepian and Weinberg (1958) and Adams (1958), but Darlington's conjecture (1955), stating that any RC grounded 2-port admits an equivalent series-parallel realization, is still unproved.

The problem of ladder synthesis of two-element kind 2-ports arose earlier in filter theory, and the important method of zero-shifting in cascade synthesis was introduced by Bader (1942). Necessary and sufficient conditions for ladder realizability, in the case of a prescribed impedance matrix, or prescribed insertion loss, are unknown in general, but have been established for important classes of networks (Fujisawa, 1955; Watanabe, 1958) and applied to filter design (Meinguet, 1958).

Active grounded RC networks have recently been discussed by Kinariwala (1959–1960) and Sandberg (1960).

Nonlinear and Linear Variable Circuits

In contrast with the maturity of linear network theory, it is often considered that the theory of nonlinear and of linear variable networks is still in its infancy: it is not yet completely separated from nonlinear mechanics and has not reached the stage of synthesis. The present brief review limited to purely electrical problems is intended to show, however, that there has been some systematic progress and that the theory in its present state is no longer a collection of odd results.

First, it is important to separate the problems raised by the unavoidable nonlinearities appearing in nominally linear systems, and the intentional nonlinear effects allowing performances unobtainable with linear systems. The calculation of nonlinear distortion is relatively elementary in the case of slightly nonlinear characteristics (vacuum tube amplifiers, carbon microphone, etc). A more difficult problem arose in connection with the cross-modulation due to hysteresis in loading coils (Kalb and Bennett, 1935). The computation of the modulation products generated by two harmonic signals turned out to be mathematically equivalent to the similar problem in an ideal rectifier, and the amplitudes were obtained as hypergeometric functions. Further progress is reported in later publications by Bennett and others, and the results proved useful in a number of problems involving sharp nonlinearities, such as overload effects in rectifier modulators and feedback amplifiers.

The typically useful properties of nonlinear, or linear variable systems are related with frequency conversion in a wide sense, including harmonic or subharmonic generation, and even simple oscillation, for an oscillator transforms dc power into power at its own frequency. The theory of the triode oscillator was developed by Appleton and van der Pol (1920–1926), and explained such effects as synchronization, amplitude hysteresis, etc. Subharmonic generations in the triode oscillator were analyzed by Mandelstam and Papalexi (1931). Although van der Pol had shown the continuity between quasi-harmonic and relaxation oscillations, these extreme types of behavior actually occur in different technical applications and continued to be treated in separated contexts. Starting from relaxation oscillators and multivibrators (Abraham and Block, 1919) a particular branch of electronic circuit technology developed in the direction of such applications as time bases, counters, logical circuits, wave form generators, etc. On the other hand, the main concern in the design of harmonic oscillators was frequency stability; the way it is affected by nonlinearities was deduced from the principle of harmonic balance (Groszkowski, 1933) and the analysis of linear effects led to the development of the numerous oscillator circuits bearing the names of their inventors. All the above problems are, however, treated by elementary circuit analysis and are more related with device technology than with circuit theory.

Frequency conversion by amplitude modulation raised only elementary problems as long as tubes were used. With the introduction of rectifier modulators, such as the Cowan and ring modulators, in the early thirties, more difficult problems appeared, due to the interaction of all products whose amplitudes thus finally depend on the load impedances at all frequences. In carrier systems, modulators normally operate between highly selective filters, and a corresponding small signal theory for rectifier modulators between selective terminations was developed after 1939 by such workers as Caruthers, Kruse, Stieltjes, Tucker and Belevitch. The conception of a linear variable network as a linear network with an infinite number of ports (treating the impedance presented to each modulation product as a separate termination) did not yield tractable solutions in the case of nonselective terminations, and the case of RC loads (Belevitch, 1950) was treated by a direct method. The general problem became recently of major importance for the optimum design of filters for pulse-time modulated systems; after contributions by Cattermole (1958) and Desoer (1958), an analysis leading to a finite set of linear equations was obtained by Fettweis (1959). Another application of rectifier modulators was treated by Miller (1959) who showed that better frequency dividers could be designed by separating the various nonlinear effects which occur simultaneously in the tube of ordinary oscillators—in the circuits of Miller, the tube is a linear amplifier and the necessary nonlinear effects are produced by rectifier circuits separated by selective filters.

Harmonic generation by rectifier circuits does not seem to have been treated until quite recently—Page (1956) published a theorem on the maximum harmonic efficiency, and Belevitch and Neirynck (1957) described optimal circuits. On the contrary, harmonic production by non-linear reactance is not subject to the same restrictions, and magnetic harmonic generators (Peterson, Manley and Wrathall, 1937) are widely used. The theory of frequency conversion in magnetic modulators is more recent—power relations in nonlinear reactances were discussed by Manley and Rowe (1956), Page (1957), Duinker (1957) and many others. The possibility of amplification and the related negative resistance effects (Hartley) were known however, around 1920 and applied quite early to frequency dividers (transformation of a 60 cps supply into a 20 cps telephone ringing current) and to magnetic amplifiers. The recent revival of interest in parametric amplification is due to the availability of solid-state nonlinear capacitances. On the other hand, resonant circuits with nonlinear inductances, and the phenomenon of ferroresonance were also discussed before World War I, and led to Duffing's equation (1918). New interest in related phenomena was raised by the invention of the parametron (Goto, 1955).

The analysis of complicated nonlinear oscillator circuits by the classical methods of nonlinear mechanics is practically impossible and various approximate engineering methods have been developed, the first one being the so-called equivalent linearization of Krylov and Bogoliubov (1937). During the Second World War, a method based on the describing function has been introduced separately in the U.S.A., the U.S.S.R. and in France. In this method, one derives, from the instantaneous response of the nonlinear element to a harmonic excitation or to a combination of such excitations, an amplitude response or a set of amplitude responses, which are functions of the incident amplitudes; the linear part of the circuit is characterized by its transfer function or matrix. With the additional hypothesis that the linear circuits are highly selective, a finite set of algebraic equations is obtained for the steady-state amplitudes, and it is no longer necessary to consider explicitly any differential equations. This also holds true for the stability analysis of the steady-states, which is performed by linear perturbation methods, thus using only the standard criteria of linear network theory (Hurwitz, Nyquist). This "filter method" (as it is called in Russia) allowed a much simpler derivation of the classical results of van der Pol and others, and is now being successfully applied to new problems.

From Circuit Theory to System Theory*

L. A. ZADEH†, FELLOW, IRE

Summary—**The past two decades have witnessed profound changes in the composition, functions and the level of complexity of electrical as well as electronic systems which are employed in modern technology. As a result, classical RLC network theory, which was the mainstay of electrical engineering at a time when RLC networks were the bread and butter of the electrical engineer, has been and is being increasingly relegated to the status of a specialized branch of a much broader discipline—system theory—which is concerned with systems of all types regardless of their physical identity and purpose.**

This paper presents a brief survey of the evolution of system theory, together with an exposition of some of its main concepts, techniques and problems. The discussion is centered on the notion of state and emphasizes the role played by state-space techniques. The paper concludes with a brief statement of some of the key problems of system theory.

I. Introduction

THE past two decades have witnessed an evolution of classical circuit theory into a field of science whose domain of application far transcends the analysis and synthesis of RLC networks. The invention of the transistor, followed by the development of a variety of other solid-state devices, the trend toward microminiaturization and integrated electronics, the problems arising out of the analysis and design of large-scale communication networks, the increasingly important role played by time-varying, nonlinear and probabilistic circuits, the development of theories of neuroelectric networks, automata and finite state machines, the progress in our understanding of the processes of learning and adaptation, the advent of information theory, game theory and dynamic programming, and the formulation of the maximum principle by Pontryagin, have all combined to relegate classical circuit theory to the status of a specialized branch of a much broader scientific discipline—system theory—which, as the name implies, is concerned with all types of systems and not just electrical networks.

What is system theory? What are its principal problems and areas? What is its relationship to such relatively well-established fields as circuit theory, control theory, information theory, and operations research and systems engineering? These are some of the questions which are discussed in this paper, with no claim that the answers presented are in any way definitive. Technological and scientific progress is so rapid these days that hardly any assertion concerning the boundaries, content and directions of such a new field as system theory can be expected to have long-term validity.

The obvious inference that system theory deals with systems does not shed much light on it, since all branches of science are concerned with systems of one kind or another. The distinguishing characteristic of system theory is its generality and abstractness, its concern with the mathematical properties of systems and not their physical form. Thus, whether a system is electrical, mechanical or chemical in nature does not matter to a system theorist. What matters are the mathematical relations between the variables in terms of which the behavior of the system is described.[1]

To understand this point clearly, we have to examine more closely the concept of a system. According to Webster's dictionary, a system is "... an aggregation or assemblage of objects united by some form of interaction or interdependence." In this sense, a set of particles exerting attraction on one another is a system; so is a group of human beings forming a society or an organization; so is a complex of interrelated industries; so is an electrical network; so is a large-scale digital computer, which represents the most advanced and sophisticated system devised by man; and so is practically any conceivable collection of interrelated entities of any nature. Indeed, there are few concepts in the domain of human knowledge that are as broad and all-pervasive as that of a system.

It has long been known that systems of widely different physical forms may in fact be governed by identical differential equations. For example, an electrical network may be characterized by the same equations as a mechanical system, in which case the two constitute analogs of one another. This, of course, is the principle behind analog computation.

While the analogies between certain types of systems have been exploited quite extensively in the past, the recognition that the same abstract "systems" notions are operating in various guises in many unrelated fields of science is a relatively recent development. It has been brought about, largely within the past two decades, by the great progress in our understanding of the behavior of both inanimate and animate systems—progress which resulted on the one hand from a vast expansion in the scientific and technological activities directed toward the development of highly complex systems for such purposes as automatic control, pattern recognition, data-processing, communication, and machine computation, and, on the other hand, by the attempts at

* Received by the IRE, August 21, 1961; revised manuscript received, August 25, 1961. The preparation of this paper was supported in part by the National Science Foundation.

† Dept. of Elec. Engrg., University of California, Berkeley, Calif.

[1] This and the following four paragraphs are reproduced with minor changes from an earlier article by the author: "System theory," *Columbia Engrg. Quart.*, vol. 8, pp. 16–19, 34; November, 1954.

quantitative analyses of the extremely complex animate and man-machine systems which are encountered in biology, neurophysiology, econometrics, operations research and other fields.

It is these converging developments that have led to the conception of system theory—a scientific discipline devoted to the study of general properties of systems, regardless of their physical nature. It is to its abstractness that system theory owes its wide applicability, for it is by the process of abstraction that we progress from the special to the general, from mere collections of data to general theories.[2]

If one were asked to name a single individual who above anyone else is responsible for the conception of system theory, the answer would undoubtedly be "Norbert Wiener," even though Wiener has not concerned himself with system theory as such, nor has he been using the term "system theory" in the sense employed in this paper. For it was Wiener who, starting in the twenties and thirties, has introduced a number of concepts, ideas and theories which collectively constitute a core of present-day system theory. Among these, to mention a few, are his representation of nonlinear systems in terms of a series of Laguerre polynomials and Hermite functions, his theory of prediction and filtering, his generalized harmonic analysis, his cinema-integraph, the Paley-Wiener criterion, and the Wiener process. It was Wiener who, in the late forties, laid the foundation for cybernetics—the science of communication and control in the animal and the machine—of which system theory is a part dealing specifically with systems and their properties. It should be noted, however, that some of the more important recent developments in system theory no longer bear Wiener's imprint. This is particularly true of the theory of discrete-state systems, the state-space techniques for continuous systems, and the theory of optimal control, which is associated mainly with the names of Bellman and Pontryagin. We shall touch upon these developments later on in the paper.

Among the scientists dealing with animate systems, it was a biologist—Ludwig von Bertalanffy—who long ago perceived the essential unity of systems concepts and techniques in various fields of science and who in writings and lectures sought to attain recognition for "general systems theory" as a distinct scientific discipline.[3] It is pertinent to note, however, that the work of Bertalanffy and his school, being motivated primarily by problems arising in the study of biological systems, is much more empirical and qualitative in spirit than the work of those system theorists who received their training in the exact sciences. In fact, there is a fairly wide gap between what might be regarded as "animate" system theorists and "inanimate" system theorists at the present time, and it is not at all certain that this gap will be narrowed, much less closed, in the near future. There are some who feel that this gap reflects the fundamental inadequacy of the conventional mathematics—the mathematics of precisely-defined points, functions, sets, probability measures, etc.—for coping with the analysis of biological systems, and that to deal effectively with such systems, which are generally orders of magnitude more complex than man-made systems, we need a radically different kind of mathematics, the mathematics of fuzzy or cloudy quantities which are not describable in terms of probability distributions. Indeed, the need for such mathematics is becoming increasingly apparent even in the realm of inanimate systems, for in most practical cases the *a priori* data as well as the criteria by which the performance of a man-made system is judged are far from being precisely specified or having accurately-known probability distributions.

System theory is not as yet a well-crystallized body of concepts and techniques which set it sharply apart from other better established fields of science. Indeed, there is considerable overlap and interrelation between system theory on the one hand and circuit theory, information theory, control theory, signal theory, operations research and systems engineering on the other. Yet, system theory has a distinct identity of its own which perhaps can be more clearly defined by listing its principal problems and areas. Such a list is presented below, without claims that it is definitive, complete and non-controversial. (To avoid misunderstanding, brief explanations of the meaning of various terms are given in parentheses.)

Principal Problems of System Theory

1) System characterization (representation of input-output relationships in mathematical form; transition from one mode of representation to another).
2) System classification (determination, on the basis of observation of input and output, of one among a specified class of systems to which the system under test belongs).
3) System identification (determination, on the basis of observation of input and output, of a system within a specified class of systems to which the system under test is equivalent; determination of the initial or terminal state of the system under test).
4) Signal representation (mathematical representation of a signal as a combination of elementary signals; mathematical description of signals).
5) Signal classification (determination of one among a specified set of classes of signals or patterns to which an observed signal belongs).

[2] To quote A. N. Whitehead, "To see what is general in what is particular, and what is permanent in what is transitory is the aim of modern science."

[3] Dr. Bertalanffy is a founder of the Society for General Systems Research, which publishes a yearbook and has headquarters at the Menninger Foundation, Topeka, Kans.

6) System analysis (determination of input-output relationships of a system from the knowledge of input-output relationships of each of its components).
7) System synthesis (specification of a system which has prescribed input-output relationships).
8) System control and programming (determination of an input to a given system which results in a specified or optimal performance).
9) System optimization (determination of a system within a prescribed class of systems which is best in terms of a specified performance criterion).
10) Learning and adaptation (problem of designing systems which can adapt to changes in environment and learn from experience).
11) Reliability (problem of synthesizing reliable systems out of less reliable components).
12) Stability and controllability (determination of whether a given system is stable or unstable, controllable—subject to specified constraints—or not controllable).

Principal Types of Systems

1) Linear, nonlinear.
2) Time-varying, time-invariant.
3) Discrete-time (sampled-data), continuous-time.
4) Finite-state, discrete-state, continuous-state.
5) Deterministic (nonrandom), probabilistic (stochastic).
6) Differential (characterized by integro-differential equations), nondifferential.
7) Small-scale, large-scale (large number of components).

Some Well-Established Fields Which May Be Regarded as Branches of System Theory

1) Circuit theory (linear and nonlinear).
2) Control theory.
3) Signal theory.
4) Theory of finite-state machines and automata.

Comment 1: Note that *approximation* is not listed as a separate problem, as it is usually regarded in classical circuit theory. Rather, it is regarded as something that permeates all the other problems.

Comment 2: Note that *information theory* and *communication theory* are not regarded as branches of system theory. System theory makes extensive use of the concepts and results of information theory, but this does not imply that information theory is a branch of system theory, or vice versa. The same comment applies to such theories as *decision theory* (in statistics), *dynamic programming*, *reliability theory*, etc.

Comment 3: Note that there is no mention in the list of *systems engineering* and *operations research.* We regard these fields as being concerned specifically with the operation and management of large-scale man-machine systems, whereas system theory deals on an abstract level with general properties of systems, regardless of their physical form or the domain of application. In this sense, system theory contributes an important source of concepts and mathematical techniques to systems engineering and operations research, but is not a part of these fields, nor does it have them as its own branches.

It would be futile to attempt to say something (of necessity brief and superficial) about each item in the above list. Instead, we shall confine our attention to just a few concepts and problems which play particularly important roles in system theory and, in a way, account for its distinct identity. Chiefly because of limitations of space, we shall not even touch upon a number of important topics such as the design of learning and adaptive systems, the analysis of large-scale and probabilistic systems, the notion of feedback and signal flow graph techniques, etc. In effect, the remainder of this paper is devoted to a discussion of the concept of state and state-space techniques, along with a brief exposition of systems characterization, classification and identification. We have singled out the concept of state for special emphasis largely because one can hardly acquire any sort of understanding of system theory without having a clear understanding of the notion of state and some of its main implications.

II. State and State-Space Techniques

It is beyond the scope of this presentation to trace the long history of the evolution of the concept of state in the physical sciences. For our purposes, it will suffice to observe that the notion of state in essentially the same form it is used today was employed by Turing [1] in his classical paper, "On computable numbers, with an application to the Entscheidungs problem," in which he introduced what is known today as the Turing machine.[4]

Roughly speaking, a Turing machine is a discrete-time ($t=0, 1, 2, \cdots$) system with a finite number of states or internal configurations, which is subjected to an input having the form of a sequence of symbols (drawn from a finite alphabet) printed on a tape which can move in both directions along its length. The output of the machine at time t is an instruction to print a particular symbol in the square scanned by the machine at time t and to move in one or the other direction by one square. A key feature of the machine is that the output at time t and the state at time $t+1$ are determined by the state and the input at time t. Thus, if the state, the input and the output at time t are denoted by s_t, u_t

[4] A thorough discussion of the Turing machine can be found in M. Davis, "Computability and Unsolvability," McGraw-Hill Book Co., Inc., New York, N. Y.; 1958. For a discussion of a variant of the Turing machine, see C. Y. Lee, "Automata and finite automata," *Bell Sys. Tech. J.*, vol. 39, pp. 1267–1296; September, 1960.

and y_t, respectively, then the operation of the machine is characterized by:

$$s_{t+1} = f(s_t, u_t), \qquad t = 0, 1, 2, \cdots, \tag{1}$$

$$y_t = g(s_t, u_t) \tag{2}$$

where f and g are functions defined on pairs of values of s_t and u_t. Note that (1) and (2) imply that the output symbols from any initial time t_0 on are determined by the state at time t_0 and the input symbols from time t_0 on.

An important point about this representation, which was not pointed out by Turing, is that it is applicable not only to the Turing machine, but more generally to any discrete-time system. Furthermore, we shall presently see that it is a simple matter to extend (1) and (2) to systems having a continuum of states (*i.e.*, continuous-state systems).

The characterization of a system by means of equations of the form (1) and (2) (to which we will refer as the *Turing representation* or, alternatively, as the *state equations* of a system) was subsequently employed by Shannon [2] in his epoch-making paper on the mathematical theory of communication. Specifically, Shannon used (1) and (2) to characterize finite-state noisy channels, which are probabilistic systems in the sense that s_t and u_t determine not s_{t+1} and y_t, but their joint probability distribution. This implies that the system is characterized by (1) and (2), with f and g being random functions, or, equivalently ,by the conditional distribution function $p(s_{t+1}, y_t | s_t, u_t)$, where for simplicity of notation the same letter is used to denote both a random variable and a particular value of the random variable (*e.g.*, instead of writing S_t for the random variable and s_t for its value, we use the same symbol, s_t, for both).

It was primarily Shannon's use of the Turing representation that triggered its wide application in the analysis and synthesis of discrete-state systems. Worthy of note in this connection is the important work of von Neumann [3] on probabilistic logics, which demonstrated that it is possible, at least in principle, to build systems of arbitrarily high degree of reliability from unrealiable (probabilistic) components. Also worthy of note is the not-too-well-known work of Singleton [4] on the theory of nonlinear transducers, in which techniques for optimizing the performance of a system with quantized state space are developed. It should be remarked that the problem of approximating to a system having a continuum of states with a system having a finite or countable number of states is a basic and as yet unsolved problem in system theory. Among the few papers which touch upon this problem, those by Kaplan [5], [6] contain significant results for the special case of a differential[5] system subjected to zero input. A qualitative discussion of the related problem of ϵ-approximation may be found in a paper by Stebakov [7].

There are two important notions that are missing or play minor roles in the papers cited above. These are the notions of equivalent states and equivalent machines which were introduced by Moore [8] and, independently and in a somewhat restricted form, by Huffman [9]. The theory developed by Moore constitutes a contribution of basic importance to the theory of discrete-state systems and, more particularly, the identification problem.

So far, our discussion of the notion of state has been conducted in the context of discrete-state systems. In the case of a differential system, the state equations (1) and (2) assume the form

$$\dot{\mathbf{s}}(t) = \mathbf{f}(\mathbf{s}(t), \mathbf{u}(t)) \tag{3}$$

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{s}(t), \mathbf{u}(t)) \tag{4}$$

where $\dot{\mathbf{s}}(t) = d/dt\ \mathbf{s}(t)$, and $\mathbf{s}(t)$, $\mathbf{y}(t)$ and $\mathbf{u}(t)$ are vectors representing the *state*, the *input* and *output* of the system at time t. Under various guises, such equations [particularly for the case where $\mathbf{u}(t) \equiv 0$] have long been used in the theory of ordinary differential equations, analytical dynamics, celestial mechanics, quantum mechanics, econometrics, and other fields. Their wide use in the field of automatic control was initiated largely in the late forties and early fifties by Soviet control theorists, notably A. I. Lur'e,[6] M. A. Aizerman, Ya. Z. Tsypkin, A. A. Fel'dbaum, A. Ya. Lerner, A. M. Letov, N. N. Krasovskii, I. G. Malkin, L. S. Pontryagin, and others. In the United States, the introduction of the notion of state and related techniques into the theory of optimization of linear as well as nonlinear systems is due primarily to Bellman, whose invention of dynamic programming [10] has contributed by far the most powerful tool since the inception of the variational calculus to the solution of a whole gamut of maximization and minimization problems. Effective use of and or important contributions to the state-space techniques in the field of automatic control have also been made by Kalman [11], Kalman and Bertram [12], LaSalle [13], Laning and Battin [14] (in connection with analog simulation), Friedland [15], and others. It is of interest to note, however, that it was not until 1957 that a general method for setting up the state equations of an RLC network was described by Bashkow [16]. An extension of Bashkow's technique to time-varying networks was recently presented by Kinarawala [17].

Despite the extensive use of the notion of state in the current literature, one would be hard put to find a satisfactory definition of it in textbooks or papers. A reason for this is that the notion of state is essentially a primitive concept, and as such is not susceptible to exact

[5] By a differential system, we mean a system which is characterized by one or more ordinary differential equations.

[6] We are not citing specific references to the work of Lur'e and others on state-space techniques, since their contributions are spread over such a large number of papers that it is difficult to isolate the most relevant ones.

definition. However, it is possible, as sketched below, to define it indirectly by starting with the notion of *complete characterization* of a system. Specifically, consider a black box B[7] and some initial time t_0. We assume that B is associated with three types of time-functions:

1) a controllable variable u [*i.e.*, a time-function whose values can be chosen at will from a specified set (input space) for all $t \geq t_0$];
2) an initially controllable variable s [*i.e.*, a time-function whose value can be chosen at will at $t=t_0$ from a specified set (*state space*), but not thereafter]; and
3) an observable variable y (*i.e.*, a time-function whose values can be observed for $t \geq t_0$, but over which no direct control can be exercised for $t \geq t_0$).

Furthermore, we assume that this holds for all values of t_0.

If these assumptions are satisfied, then we shall say that B is *completely characterized* if for every $t \geq t_0$ the value of the output at time t, $y(t)$, is uniquely determined by the value of s at time t_0 and the values of u over the closed interval $[t_0, t]$. Symbolically, this is expressed by

$$y(t) = B(\mathbf{s}(t_0); \mathbf{u}_{[t_0,t]}) \tag{5}$$

where $u_{[t_0,t]}$ denotes the segment of the time-function u extending between and including the end points t_0 and t; $\mathbf{s}(t_0)$ is the value assumed by $\mathbf{s}(t)$ at time t_0; and $B(\cdot\ ;\ \cdot)$ is a single-valued function of its arguments. [Note that B is a functional of $\mathbf{u}_{[t_0,t]}$ and an ordinary function of $\mathbf{s}(t_0)$; $\mathbf{s}(t)$ is usually a vector with a finite number of components.] It is understood that (5) must hold for all possible values of $\mathbf{s}(t_0)$ and $\mathbf{u}_{[t_0,t]}$ and that to every possible input-output pair $\mathbf{u}_{(t_0,t)}$, $\mathbf{y}_{(t_0,t)}$ there should correspond a state $\mathbf{s}(t_0)$ in the state space of B.

If B is completely characterized in the sense defined above, then $\mathbf{u}(t)$, $\mathbf{y}(t)$ and $\mathbf{s}(t)$ are, respectively, the values of the *input*, the *output* and the *state* of B at time t. [The range of values of $\mathbf{s}(t)$ constitutes the *state-space* of B. A particular value of $\mathbf{s}(t)$, *i.e.*, a particular state of B, will be denoted by $\mathbf{q}$.] In this way, the input, the output and the state of B are defined simultaneously as by-products of the definition of complete characterization of B.

The intuitive significance of the concept of state is hardly made clear by the somewhat artificial definition sketched above. Essentially, $\mathbf{s}(t)$ constitutes a description of the internal condition in B at time t. Eq. (5), then, signifies that, given the initial conditions at time t_0, and given the values of the input between and including t_0 and t, we should be able to find the output of B at time t if the system is completely characterized.

With (5) as the starting point, it is a simple matter to demonstrate that (5) can be replaced by an equivalent pair of equations:

$$\mathbf{s}(t) = \mathbf{f}(\mathbf{s}(t_0); \mathbf{u}_{[t_0,t]}), \qquad t \geq t_0 \tag{6}$$

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{s}(t); \mathbf{u}(t)), \tag{7}$$

the first of which expresses the state at time t in terms of the state at time t_0 and the values of the input between and including t_0 and t, while the second gives the output at time t in terms of the state at time t and the input at time t. Note that these relations are in effect continuous analogs of the Turing representation

$$s_{t+k} = f(s_t, u_t, \cdots, u_{t+k-1}) \tag{8}$$

$$y_t = g(s_t, u_t). \tag{9}$$

It will be helpful at this point to consider a simple illustrative example. Let B be the network shown in Fig. 1, in which u is the input voltage and y is the output voltage.

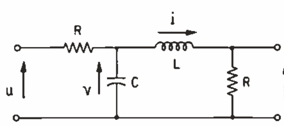


Fig. 1—Network for illustrative example.

It can easily be shown that the state of this network is a 2-vector, $\mathbf{s}(t)$, whose components can be taken to be $v(t)$ (voltage across C) and $i(t)$ (current through L). With this choice of $\mathbf{s}(t)$, the state equations can readily be set up by inspection. In matrix form, they read:

$$\dot{\mathbf{s}} = A\mathbf{s} + Bu \tag{10}$$

$$y = \alpha \mathbf{s} \tag{11}$$

where

$$A = \begin{bmatrix} -\frac{1}{CR} & -\frac{1}{C} \\ \frac{1}{L} & -\frac{R}{L} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{RC} \\ 0 \end{bmatrix}, \quad \alpha = [0 \quad R].$$

On integrating (10), we obtain

$$s(t) = \mathbf{G}(t - t_0)s(t_0) + \int_{t_0}^{t} \mathbf{G}(t - \xi)Bu(\xi)d\xi \tag{12}$$

$$y(t) = \alpha s \tag{13}$$

where the matrix G, called the *transition matrix* of B, is given by the inverse Laplace transform of the matrix $(sI - A)^{-1}$. Specifically,

$$G(t) = \mathcal{L}^{-1}\{(sI - A)^{-1}\}$$

$$= \mathcal{L}^{-1} \begin{bmatrix} \dfrac{s + \frac{R}{L}}{\det A} & \dfrac{-\frac{1}{C}}{\det A} \\ \dfrac{\frac{1}{L}}{\det A} & \dfrac{s + \frac{1}{RC}}{\det A} \end{bmatrix} \tag{14}$$

[7] For simplicity, it is assumed that B is (1) deterministic (*i.e.*, nonrandom), (2) time-invariant, (3) nonanticipative (not acting on the future values of the input), and (4) continuous-time (*i.e.*, with t ranging over the real time-axis).

where

$$\det A = \left(s + \frac{1}{RC}\right)\left(s + \frac{R}{L}\right) + \frac{1}{LC}.$$

Note that (12) expresses $s(t)$ as a function of $s(t_0)$ and $u_{[t_0,t]}$. Thus, (12) and (13) represent, respectively, the state equations (6) and (7) for B.

III. State and System Equivalence

One cannot proceed much further with the discussion of state-space techniques without introducing the twin notions of equivalent states and equivalent systems.

Suppose that we have two systems B_1 and B_2, with q_1 being a state of B_1 and q_2 being a state of B_2. As the term implies, q_1 and q_2 are *equivalent states* if, for all possible input time-functions u, the response of B_1 to u starting in state q_1 is the same as the response of B_2 to u starting in state q_2. Following Moore, B_1 and B_2 are said to be *equivalent systems* if, and only if, to every state in B_1 there is an equivalent state in B_2, and vice versa. What is the significance of this definition? Roughly speaking, it means that if B_1 and B_2 are equivalent, then it is impossible to distinguish B_1 from B_2 by observing the responses of B_1 and B_2 to all possible inputs u, if the initial states of B_1 and B_2 are not known to the experimenter.

To illustrate, let us consider the two simple networks shown in Fig. 2.

The state of B_1 is a vector with two components (which may be identified with the currents flowing through the two inductances); the state of B_2 is a scalar (the current through $2L$). Nevertheless, by writing the state equations for B_1 and B_2, it is easy to verify that B_1 and B_2 are equivalent in the sense defined above, as well as in the more conventional sense of circuit theory.

On the other hand, consider the constant-resistance network shown in Fig. 3. Its input impedance is equal to unity at all frequencies. Does this mean that B_1 is equivalent to a unit resistor?

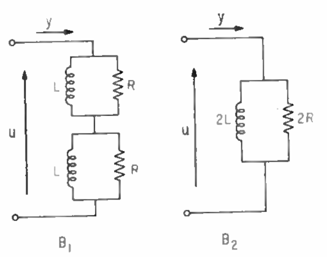


Fig. 2—Equivalent networks.

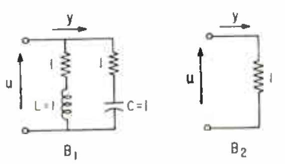


Fig. 3—Ground-state equivalent networks.

The circuit theorist's answer to this question would be "yes," since in circuit theory two networks are defined to be equivalent if their respective terminal impedances (or admittances) have the same values at all frequencies. By contrast, the system theorist would say "no," since there are states in B_1, say the state (0, 1) (where the first component is the voltage across C and the second component is the current through L), to which no equivalent state in the unit resistor can be found.[8] However, we note that the unexcited state (ground state) of B_1, which is (0,0), is equivalent to the ground state of the resistor. Thus, we can assert that, although B_1 is not equivalent to the unit resistor, it is ground-state equivalent to the unit resistor. In plain words, this means that if B is initially unexcited, then it will behave like a unit resistor. On the other hand, if B is initially excited, then it may not behave like a unit resistor.

This simple example shows that the notion of equivalence in circuit theory has a narrower meaning than it does in system theory. More specifically, the equivalence of two networks in the sense of system theory implies, but is not implied by, their equivalence in the sense of circuit theory. In effect, the notion of equivalence in circuit theory corresponds to the notion of ground-state equivalence in system theory.

IV. The Notion of Policy

Another important difference between circuit theory and system theory manifests itself in the way in which the input to a system (circuit) is represented. Thus, in circuit theory it is customary to specify the desired input to a network as a function of time. By contrast, in system theory it is a common practice—particularly in dealing with control problems—to express the input as a function of the state of the system rather than as a function of time.[9] In many ways, this is a more effective representation, since it is natural to base the decision on what input to apply at time t on the knowledge of the state of the system at time t. Furthermore, in the latter representation (input in terms of state) we have feedback, whereas in the former (input in terms of time) we have not.

To say that the input depends on the state of the system means, more precisely, that the input at time t is a function of the state at time t, *i.e.*,

$$u(t) = \pi(s(t)) \tag{15}$$

where π is a function defined on the state space with values in the input space. This function is referred to as a *policy function*, or simply a *policy*. In effect, a policy is

[8] The resistor represents a degenerate case in which the output (say voltage) depends only on the input (say current) and not on the state of the system. In such cases, the choice of state is immaterial. For example, it may be taken to be the current through R.

[9] For simplicity, it is tacitly assumed here that the system as well as its performance criteria are time-invariant; otherwise, the input would be expressed as a function of both the state of the system and time.

a function which associates a particular input with each state of the system.

The notion of policy plays a key role in system theory and, especially, in control theory. Thus, a typical problem in control theory involves the determination of a policy for a given system B which is optimal in terms of a specified performance criterion for B. More specifically, the performance criterion associates with each policy π a number $Q(\pi)$, which is a measure of the "goodness" of π.[10] The problem, then, is to find a policy π which maximizes $Q(\pi)$. Such a policy is said to be *optimal*.

As was stated previously, the most effective general technique for solving problems of this nature is provided by Bellman's dynamic programming. The basis for dynamic programming is the so-called *principle of optimality* which in Bellman's words reads: "An optimal policy has the property that whatever the initial state and the initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

Needless to say, one can always resort to brute force methods to find a policy π which maximizes $Q(\pi)$. The great advantage of dynamic programming over direct methods is that it reduces the determination of optimal π to the solution of a succession of relatively simple maximization or minimization problems. In mathematical terms, if the payoff resulting from the use of an optimal policy when the system is initially (say at $t=0$) in state $s(0)$ is denoted by $R(s(0))$, then by employing the principle of optimality one can derive a functional equation satisfied by R. In general, such equations cannot be solved in closed form. However, if the dimension of the state vector is fairly small, say less than four or five, then it is usually feasible to obtain the solution through the use of a moderately-sized digital computer. The main limitation on the applicability of dynamic programming is imposed by the inability of even large-scale computers to handle problems in which the dimensionality of the state vector is fairly high, say of the order of 20. A number of special techniques for getting around the problem of dimensionality have recently been described by Bellman [18].

A very basic problem in system theory which has been attacked both by the techniques of dynamic programming [19], [20] and by extensions of classical methods of the calculus of variations [21] is the so-called *minimal-time* or *optimal-time* problem. This problem has attracted a great deal of attention since the formulation of the so-called maximum principle by Pontryagin [20] in 1956, and at present is the object of numerous investigations both in the United States and the Soviet Union. Stated in general terms (for a time-invariant, single-input, continuous-time system) the problem reads as follows.

[10] It is tacitly assumed that B is a deterministic system. Otherwise, $Q(\pi)$ would be a random variable, and the performance of B would generally be measured by the expected value of $Q(\pi)$.

Given: 1) A system B characterized by the vector differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), u(t)) \tag{16}$$

where $\mathbf{x}(t)$ and $u(t)$ represent, respectively, the state and the input of B at time t. ($\mathbf{x}$ is a vector; u is a scalar, and $\mathbf{f}$ is a function satisfying certain smoothness conditions.)

2) A set of constraints on u,[11] *e.g.*, $|u(t)| \leq 1$ for all t or $|u(t)| \leq 1$ and $|\dot{u}(t)| \leq 1$ for all t.

3) A specified intial state $\mathbf{x}(0) = \mathbf{q}_0$ and a specified terminal state $\mathbf{q}_1$.

Find an input u (satisfying the prescribed constraints) which would take B from $\mathbf{q}_1$ to $\mathbf{q}_2$ in the shortest possible time. This, in essence, is the minimal-time problem.

In a slightly more general formulation of the problem, the quantity to be minimized is taken to be the cost of taking the system from $\mathbf{q}_0$ to $\mathbf{q}_1$, where the cost is expressed by an integral of the form

$$C(u; \mathbf{q}_0, \mathbf{q}_1) = \int_0^{t_1} f_0(\mathbf{x}(t), u(t))dt. \tag{17}$$

In this expression, f_0 is a prescribed function, t_1 is the time at which B reaches the state $\mathbf{q}_1$, and $C(u, \mathbf{q}_0, \mathbf{q}_1)$ denotes the cost of taking B from $\mathbf{q}_0$ to $\mathbf{q}_1$, when the input u is used.

It is not hard to see why the minimal-time (or, more generally, the minimal-cost) problem plays such an important role in system and, more particularly, control theory. Almost every control problem encountered in practice involves taking a given system from one specified state to another. The minimal time-problem merely poses the question of how this can be done in an optimal fashion.

Various special cases of the minimal time problem were considered by many investigators in the late forties and early fifties. What was lacking was a general theory. Such a theory was developed in a series of papers by Pontryagin, Boltyanskii, and Gamkrelidze [21], [22].

The maximum principle of Pontryagin is essentially a set of necessary conditions satisfied by an optimal input u. Briefly stated, let ψ be a solution of the variational system

$$\dot{\psi} = -\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)' \psi \tag{18}$$

where $[\partial \mathbf{f}/\partial \mathbf{x}]'$ is the transpose of the matrix $[\partial f_i/\partial x_j]$, in which f_i and x_j are, respectively, the ith and jth components of $\mathbf{f}$ and $\mathbf{x}$ in the equation $\dot{x} = \mathbf{f}(\mathbf{x}, u)$. Construct a Hamiltonian function $H(\mathbf{x}, \psi, u) = \psi \cdot \dot{x}$ (dot product of ψ and $\dot{x}$), with the initial values of ψ in (18) constrained by the inequality $H(\mathbf{x}(0), \psi(0), u(0)) \geq 0$. The

[11] More generally, there might be additional constraints imposed on $\mathbf{x}$. This case is not considered here.

maximum principle asserts that if $\hat{u}(t)$ is an optimal input, then $\hat{u}(t)$ maximizes the Hamiltonian $H(\mathbf{x}, \psi, u)$, with x and ψ held fixed for exact t.

An application of the maximum principle to a linear system characterized by the vector equation

$$\dot{\mathbf{x}} = \boldsymbol{A}\mathbf{x} + \boldsymbol{B}u, \tag{19}$$

where $\boldsymbol{A}$ is a constant matrix and $\boldsymbol{B}$ is a constant vector, leads to the result that an optimal input is "bang-bang," that is, at all times the input is as large amplitude-wise as the limits permit. More specifically, an optimal input is of the form

$$\hat{u}(t) = \text{sgn}\ (\psi(t) \cdot \boldsymbol{B}) \tag{20}$$

where sgn stands for the function sgn $x=1$ if $x>0$, sgn $x=-1$ if $x<0$, and sgn $x=0$ if $x=0$, and ψ is a solution of the adjoint equation

$$\dot{\psi} = -\boldsymbol{A}'\psi \tag{21}$$

satisfying the inequality

$$\psi(0) \cdot [\boldsymbol{A}\mathbf{x}(0) + \boldsymbol{B}u(0)] \geq 0. \tag{22}$$

This and other general results for the linear case were first obtained by Pontryagin and his co-workers. Somewhat more specialized results had been derived independently by Bellman, Glicksberg, and Gross [25].

The main trouble with the maximum principle is that it yields only necessary conditions. The expression for $\hat{u}(t)$ given by (20) is deceptively simple; in fact, in order to determine $\hat{u}(t)$, one must first find a $\psi(t)$ which satisfies the differential equation (21), the inequality (22), and, furthermore, is such that $\boldsymbol{B}$ reaches q_1 when subjected to the $\hat{u}$ given by (20). Even then, there is no guarantee that $\hat{u}$ is optimal, except when either the initial state or the terminal state coincides with the origin. Still another shortcoming of the method is that the optimal input is obtained as a function of time rather than the state of the system.

One could hardly expect the maximum or any other principle to yield complete and simple solutions to a problem as difficult as the minimal-time problem for nonlinear, continuous-state, continuous-time systems. Actually, complete solutions can be and have been obtained for simpler types of systems. Particularly worthy of note is the solution for the case of a linear discrete-time system which was recently obtained by Desoer and Wing [26]. Quite promising for linear continuous-time systems is the successive approximation technique of Bryson and Ho [27]. In the case of systems having a finite number of states, the minimal-time problem can be solved quite easily even when the system is probabilistic and the terminal state changes in a random (Markovian) fashion [28].

Closely related to the minimal-time problem are the problems of reachability and controllability, which involve the existence and construction of inputs which take a given system for one specified state to another, not necessarily in minimum time. Important contributions to the formulation and solution of these problems for unconstrained linear systems were made by Kalman [29]. It appears difficult, however, to obtain explicit necessary and sufficient conditions for reachability in the case of constrained, linear, much less nonlinear, systems.

V. Characterization, Classification and Identification

Our placement of system characterization, classification and identification at the top of the list of principal problems of system theory (see Section I) reflects their intrinsic importance rather than the extent of the research effort that has been or is being devoted to their solution. Indeed, it is only within the past decade that significant contributions to the formulation as well as the solution of these problems, particularly in the context of finite-state systems, have been made. Nevertheless, it is certain that problems centering on the characterization, classification and, especially, identification of systems as well as signals and patterns, will play an increasingly important role in system theory in the years to come.

The problem of characterization is concerned primarily with the representation of input-output relationships. More specifically, it is concerned both with the alternative ways in which the input-output relationship of a particular system can be represented (*e.g.*, in terms of differential equations, integral operators, frequency response functions, characteristic functions, state equations, etc.), and the forms which these representations assume for various types of systems (*e.g.*, continuous-time, discrete-time, finite-state, probabilistic, finite-memory,[12] nonanticipative,[13] etc.). Generally, the input-output relationship is expressed in terms of a finite or at most countable set of linear operations (both with and without memory) and nonlinear operations (without memory). For example, Cameron and Martin [30] and Wiener [31] have shown that a broad class of nonlinear systems can be characterized by (ground-state) input-output relationships of the form

$$y(t) = \sum_{n=0}^{\infty} A_n X_n(t) \tag{23}$$

where the $X_n(t)$ represent products of Hermite functions of various orders in the variables $z_1, z_2, \cdots$, which in turn are linearly related to u (input) through Laguerre functions. Note that the operations involved in this representation are 1) linear with memory, *viz.*, the rela-

[12] A system has finite memory of length T if its output at time t is determined by $u_{[t-T, t]}$.

[13] A system is nonanticipative if, roughly speaking, its output depends only on the past and present, but not future, values of the input.

tions between the z's and u; 2) memoryless nonlinear, *viz.*, the relations between the X_n and $z_1, z_2, \cdots$; and 3) linear with no memory, *viz.*, the summations. In this connection, it should be pointed out that the basic idea of representing a nonlinear input-output relationship as a composition of an infinite number of 1) memoryless nonlinear operations and 2) linear operations with memory, is by no means a new one. It had been employed quite extensively by Volterra [32] and Frechet [33] near the turn of the century.[14]

A key feature of the Wiener representation is its orthogonality [meaning uncorrelatedness of distinct terms in (23)] for white noise inputs. This implies that a coefficient A_n in (23) can be equated to the average (expected) value of the product of $X_n(t)$ and the response of the system to white noise. In this way, a nonlinear system can be identified by subjecting it to a white noise input, generating the $X_n(t)$ functions, and measuring the average values of the products $y(t)X_n(t)$. However, for a variety of technical reasons this method of identification is not of much practical value.

The problem of system classification may be stated as follows. Given a black box B and a family (not necessarily discrete) of classes of systems $C_1, C_2 \cdots$, such that B belongs to one of these classes, say C_λ, the problem is to determine C_λ by observing the responses of B to various inputs. Generally, the inputs in question are assumed to be controllable by the experimenter. Needless to say, it is more difficult to classify a system when this is not the case.

A rather important special problem in classification is the following. Suppose it is known that B is characterized by a differential equation, and the question is: What is its order? Here, C can be taken to represent the class of systems characterized by a single differential equation of order n. An interesting solution to this problem was described by Bellman [34].

Another practical problem arises in the experimental study of propagation media. Suppose that B is a randomly-varying stationary channel, and the problem is to determine if B is linear or nonlinear. Here, we have but two classes: C_1 = class of linear systems, and C_2 = class of nonlinear systems. No systematic procedures for the solution of problems of this type have been developed so far.

Finally, *the problem of system identification* is one of the most basic and, paradoxically, least-studied problems in system theory. Broadly speaking, the identification of a system B involves the determination of its characteristics through the observation of responses of B to test inputs. More precisely, given a class of systems C (with each member of C completely characterized), the problem is to determine a system in C which is equivalent to B. Clearly, the identification problem may be regarded as a special case of the classification problem in which each of the classes $C_1, C_2, \cdots$, has just one member. This, however, is not a very useful viewpoint.

It is obvious that such commonplace problems as the measurement of a transfer function, impulsive response, the A_n coefficients in the Wiener representation (23), etc., may be regarded as special instances of the identification problem. So is the problem of location of malfunctioning components in a given system B, in which case C is the class of all malfunctioning versions of B.

A complicating feature of many identification problems is the lack of knowledge of the initial state of the system under test. Another source of difficulties is the presence of noise in observations of the input and output. For obvious reasons, the identification of continuous-state continuous-time systems is a far less tractable problem than the identification of finite-state discrete-time systems. For the latter, the basic theory developed by Moore [8] provides very effective algorithms in the case of small-scale systems, that is, systems in which the number of states as well as the number of input and output levels is fairly small. The identification of large-scale systems calls for, among other things, the development of algorithms which minimize the duration of (or the number of steps in) the identifying input sequence. With the exception of an interesting method suggested by Bellman [35], which combines dynamic programming with the minimax principle, little work along these lines has been done so far.

Another important area which is beginning to draw increasing attention is that of the identification of randomly-varying systems. Of particular interest in this connection is the work of Kailath [36] on randomly-varying linear systems, the work of Hofstetter [37] on finite-state channels, the work of Gilbert [38] on functions of a Markov process, and the generalization by Carlyle [39] of some aspects of Moore's theory to probabilistic machines. All in all, however, the sum total of what we know about the identification problem is far from constituting a body of effective techniques for the solution of realistic identification problems for deterministic, much less probabilistic, systems.

VI. Concluding Remarks

It is difficult to do justice to a subject as complex as system theory in a compass of a few printed pages. It should be emphasized that our discussion was concerned with just a few of the many facets of this rapidly-developing scientific discipline. Will it grow and acquire a distinct identity, or will it fragment and become submerged in other better-established branches of science? This writer believes that system theory is here to stay, and that the coming years will witness its evolution into a respectable and active area of scientific endeavor.

[14] A more detailed expository account of the representation of nonlinear operators may be found in a paper by the author, "On the representation of nonlinear operators," 1957 IRE WESCON Convention Record, pt. 2, pp. 105–113. A thorough mathematical treatment of related problems is available in the books by M. M. Weinberg, "Variational Methods of Investigation of Nonlinear Operators," Moscow, USSR, 1956; and M. A. Krasnoselski, "Topological Methods in the Theory of Nonlinear Integral Equations," Moscow, USSR, 1956.

References

[1] A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem," *Proc. London Math. Soc.*, ser. 2, vol. 42, pp. 230–265; 1936.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379–423, 623–656; 1948.

[3] J. von Neumann, "Probabilisitic logics and the synthesis of reliable organisms from unreliable components," in "Automata Studies," Princeton University Press, Princeton, N. J., pp. 43–98; 1956.

[4] H. E. Singleton, "Theory of Nonlinear Transducers," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, RLE Rept. No. 160; August, 1950.

[5] W. Kaplan, "Dynamical systems with indeterminacy," *Am. J. Math.*, vol. 72, pp. 575–594; 1950.

[6] ——, "Discrete Models for Nonlinear Systems," presented at AIEE Winter General Mtg., New York, N. Y., February, 1960; Conference Paper 60–109.

[7] S. A. Stebakov, "Synthesis of systems with prescribed ϵ-behavior," *Proc. Conf. on Basic Problems in Automatic Control and Regulation*, Acad. Sci. USSR, Moscow, 1957; pp. 101–106.

[8] E. F. Moore, "Gedanken-experiments on sequential machines," in "Automata Studies," Princeton University Press, Princeton, N. J., pp. 129–153; 1956.

[9] D. Huffman, "The synthesis of sequential switching circuits," *J. Franklin Inst.*, vol. 257, pp. 161–190; April, 1954.

[10] R. E. Bellman, "Dynamic Programming," Princeton University Press, Princeton, N. J.; 1957.

[11] R. E. Kalman, "Analysis and synthesis of linear systems operating on randomly sampled data," Ph.D. dissertation, Columbia University, New York, N. Y.; August, 1957.

[12] —— and J. E. Bertram, "General synthesis procedure for computer control of single and multi-loop nonlinear systems," *Trans. AIEE*, vol. 77 (*Application and Industry*, pp. 602–609; 1958.

[13] J. P. LaSalle, "The time-optimal control problem," in "Contributions to the Theory of Nonlinear Oscillations," Princeton University Press, Princeton, N. J., vol. 5, pp. 1–24; 1960.

[14] J. H. Laning, Jr., and R. H. Battin, "Random Processes in Automatic Control," McGraw-Hill Book Co., Inc., New York, N. Y.; 1956.

[15] B. Friedland, "Linear modular sequential circuits," IRE Trans. on Circuit Theory, vol. CT-6, pp. 61–68; March, 1959.

[16] T. Bashkow, "The *A* matrix, new network description," IRE Trans. on Circuit Theory, vol. CT-4, pp. 117–119; September, 1957.

[17] B. Kinarawala, "Analysis of time-varying networks," 1961 IRE International Convention Record, pt. 4, pp. 268–276.

[18] R. E. Bellman, "On the Reduction of Dimensionality for Classses of Dynamic Programming Processes," RAND Corp., Santa Monica, Calif., Paper P-2243; March, 1961.

[19] C. A. Desoer, "Pontragin's maximum principle and the principle of optimality," *J. Franklin Inst.*, vol. 271, pp. 361–367; May, 1961.

[20] L. I. Rozonoer, "The maximum principle of L. S. Pontryagin in the theory of optimal systems," *Avtomat. i Telemekh.*, vol. 20, nos. 10–12; October–December, 1959.

[21] L. Berkovitz, "Variational Methods in Problems of Control and Programming," RAND Corp., Santa Monica, Calif., Paper P-2306; May, 1961.

[22] L. S. Pontryagin, "Some mathematical problems arising in connection with the theory of optimal automatic control systems," *Proc. Conf. on Basic Problems in Automatic Control and Regulation*, Acad. Sci. USSR, Moscow, 1957.

[23] V. G. Boltyanskii, R. V. Gamkrelidze, and L. S. Pontryagin, "On the theory of optimal processes," *Dokl. Akad. Nauk SSSR*, vol. 110, no. 5, pp. 7–10; 1956.

[24] ——, ——, and ——, "On the theory of optimal processes," *Izvest. Akad. Nauk SSSR*, vol. 24, pp. 3–42; 1960.

[25] R. E. Bellman, I. Glicksberg, and O. A. Gross, "Some Aspects of the Mathematical Theory of Control Processes," RAND Corp., Santa Monica, Calif., Rept. R-313; January, 1958.

[26] C. A. Desoer and J. Wing, "A minimal time discrete system," IRE Trans. on Automatic Control, vol. AC-6, pp. 111–125; May, 1961.

[27] Y. C. Ho, "A Successive Approximation Technique for Optimal Control Systems Subject to Input Constraints," presented at JACC, Boulder, Colo., June, 1961; ASME Paper No. 61-JAC-10.

[28] J. H. Eaton and L. A. Zadeh, "Optimal Pursuit Strategies in Discrete-State Problablistic Systems," presented at JACC, Boulder, Colo., June, 1961; ASME Paper No. 61-JAC-11.

[29] R. E. Kalman, "On the general theory of control systems," *Proc. 1st Internatl. Congress on Automatic Control*, Moscow, USSR, 1960.

[30] R. H. Cameron and W. T. Martin, "The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals," *Ann. Math.*, vol. 48, pp. 385–392; April, 1947.

[31] N. Wiener, "Nonlinear Problems in Random Theory," John Wiley and Sons, Inc., New York, N. Y.; 1958.

[32] V. Volterra, "Sopra le funzioni che dipendono da altre funzioni," Rend. *accad. sci. Lincei*, vol. 3, 1887.

[33] M. Frechet, "Sur les fonctionnelles continues," *Ann. école norm. sup.*, vol. 27, pp. 193–216; 1910.

[34] R. E. Bellman, "On the separation of exponentials," *Boll. d'unione math.*, vol. 15, pp. 38–39; 1960.

[35] ——, "Sequential machines, ambiguity and dynamic programming," *J. Assoc. Comp. Mach.*, vol. 7, pp. 24–28; January, 1960.

[36] T. Kailath, "Correlation detection of signals perturbed by a random channel," IRE Trans. on Information Theory, vol. IT-6, pp. 361–366; June, 1960.

[37] E. M. Hofstetter, "Large Sample Sequential Decision Theory," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, RLE Rept. No. 359; December, 1959.

[38] E. J. Gilbert, "On the identifiability problem for functions of finite Markov chains," *Ann. Math. Statistics*, vol. 30, pp. 688–697; 1959.

[39] J. W. Carlyle, "Equivalent Stochastic Sequential Machines," Ph.D. dissertation, Dept. of Elec. Engrg., University of California, Berkeley; August, 1961.

Academic and Theoretical Aspects of Circuit Theory*

RONALD M. FOSTER,† FELLOW, IRE

Summary—**This paper discusses various types of realizability problems in electric circuit theory, with particular attention to the role of necessary and sufficient conditions. The first example is the general realization of an arbitrary driving-point impedance function discovered by Brune, together with his introduction of the use of the positive real function. Next comes the remarkable synthesis of the same problem by Bott and Duffin, with their complete elimination of all transformers. A brief survey is then given of subsequent efforts to simplify the original networks prescribed by Bott and Duffin, together with various alternative conditions which are sufficient but not necessary. All of this is then illustrated by application to the relatively simple case of the biquadratic impedance function.**

The other chief example discussed is the current work on the realization of a given real matrix as the open-circuit impedance matrix, or closed-circuit admittance matrix, of a network composed entirely of ordinary resistors. Particular attention is given to the work of Cederbaum in this field. It would seem that, although complete necessary and sufficient conditions have not yet been obtained, this problem is almost at the point of being solved.

IN this fiftieth year of the Institute, it may well be said that circuit theory has come of age. Perhaps this is best indicated by the fact that circuit theory by now has become almost a mathematical discipline. This is true, not so much because it is founded upon a generally accepted and well-understood system of assumptions or postulates, nor because various abstruse branches of mathematics have been brought to bear upon its problems, but rather because of the mathematical point of view now in vogue.

The principal question now in almost any branch of circuit theory relates to the possibilities of physical performance, and the essential limitations on what can be done with given means—in other words, the determination of necessary and sufficient conditions for a given type of network performance. Furthermore, we now demand that these conditions be stated in precise form, with no exceptions, or if one prefers, that the exceptions to our statements be listed explicitly, with a statement of the modified conditions holding for each such exception. Exact and precise statements are now demanded.

Perhaps we should note that what are sometimes called "exceptions" to the stated rule may ultimately prove to be among the most useful cases. The familiar constant-resistance network (two resistors with one inductor and one capacitor suitably adjusted) is just the equivalent of a single resistor; therefore, it would tend to be ignored in any census of biquadratic driving-point impedances as a highly degenerate or pathological exception. Nevertheless, this exceptional network was the starting point for a whole series of networks [42] designed for groups of filters quite a number of years ago.

In asking for a set of necessary and sufficient conditions for a given type of network, we are now perhaps even more demanding than is usual in mathematics. We ask not only that a set of conditions be sufficient, but also that it be operational, in that the conditions actually prescribe a means of synthesis for the network in question; this synthesis includes both the determination of the particular structure involved and the determination of the element values. Furthermore, we ask that the sufficient conditions be stated in a form readily tested before proceeding to the actual synthesis of the network.

In any specified problem the answer will depend on the variety of network envisaged, that is, both on the structural form and on the kinds of elements permitted.

So far as structural form is concerned, there are two broad classes of networks: 1) the n-terminal network, and 2) the n-port network (that is, a $2n$-terminal network, but with the terminals always paired in a designated way—in other words, a n-terminal-pair network). In this category of structure, we may also encounter specifications of particular types of linear graphs such as series-parallel or bridge structures, planar or nonplanar, or other special forms.

Perhaps the most important type of structure is the 2-terminal, or 1-port, network, largely because it is a principal building block in all other types of structure. The problem here is often thought of as the synthesis of a prescribed driving-point impedance (or admittance).

Possibly next in importance is the 2-port network (2-terminal pair, quadripole, transducer). Certain special requirements may be added: ladder, lattice, bridged-T, or some other similar specialized structure; they may be possibly balanced or unbalanced, grounded or not grounded.

Recently, another structural form has received a considerable amount of attention, namely, the true 3-terminal network. This has proved to be of some interest in its own right, instead of being regarded as a grounded 2-port network.

The other question concerns the kinds of elements allowed. In the present discussion we are limiting ourselves to ordinary lumped, passive elements, finite in number, with certain specified additional types of ele-

* Received by the IRE, June 30, 1961.
† Dept. of Mathematics, Polytechnic Institute of Brooklyn, Brooklyn, N. Y.

ment, to be considered presently. Thus, in the main we are concerned with resistors, inductors, and capacitors.

The first general class of networks to be considered is that made up of a single kind of element. Usually, the language used implies that resistors are being used. Actually, however, this is without any loss of generality. Indeed, we may go further and say the network is made up of possible multiples of any physically realizable building block of impedance Z_1.

From this class, we naturally go to the class of networks composed of two kinds of elements, usually taken to be inductors and capacitors, or resistors and capacitors, again without any loss in generality. Here again, we may say we are really dealing with the case of networks made up of positive multiples of any two physically realizable building blocks, Z_1 and Z_2.

Finally, we have the class of networks composed of all three kinds of fundamental elements: resistors, inductors, and capacitors. Here, however, we do not have an immediate extension of the theory to three arbitrary kinds of elements: Z_1, Z_2, and Z_3. Very little seems to have been done in this particular field of three perfectly arbitrary kinds of elements. The extension of the theory from all three ordinary kinds of elements to three arbitrary kinds of elements would appear to hold only for the case when

$$Z_1 Z_3 = Z_2^2 \tag{1}$$

[assuming that the units of measurement have been so adjusted as to eliminate a positive multiplier in (1)].

The first extension to the above categories occurs with the addition of mutual inductance in those cases when inductors are included among the kinds of elements.

Then we come to the possible inclusion of various sorts of ideal elements, perhaps the best known being the ideal transformer. This has been discussed in the circuit literature [8], [14] since as early as 1920. Next, we have the gyrator [56], [57], discussed since 1948. Quite recently, we have had two new such ideal elements [9]: the "nullator," a 2-terminal device which acts simultaneously like a short-circuit and an open-circuit, that is, it imposes the requirement that the current through it and the potential drop across it are both equal to zero; and its companion, the "norator," which permits a perfectly arbitrary choice of current and potential drop.

Current interest today seems to be directed in two quite diverse directions: 1) toward a complete investigation into what can be accomplished with no more than the three ordinary kinds of elements (resistors, inductors, and capacitors) without the presence of any transformers, ideal or ordinary, and with no other ideal elements; and 2) toward developing a complete theory of the perfectly general finite, linear network.

To some extent, the present paper is a summary and an extension of four survey articles [20], [29], [30], [63] which have appeared in the last decade, and to which the reader is referred for further details.

With the above two-fold classification in mind, the first problem to be dealt with in some detail is that of the 2-terminal network, or 1-port—in other words, the physical realization of a specified driving-point impedance.

For a single kind of element, there is no problem. For two kinds of elements, the problem was given a virtually complete solution [10], [11], [12], [28] in the 1920's, that is, necessary and sufficient conditions given on the impedance function for it to be realizable by networks made up of the specified two kinds of elements, including a detailed set of rules for synthesis in various prescribed structural forms.

It might be noted in passing that the addition of mutual inductance adds nothing to the problem if inductors comprise one of the two kinds of elements. Networks composed of inductors and capacitors, for example, with mutual inductance linking any set of inductors, yield impedances which can be realized by inductors and capacitors alone.

It might be of some slight historical interest to note that the case of networks composed of resistors and capacitors was the first for which the general solution was obtained, in the very early 1920's, in connection with the design of balancing networks for submarine telephone cables. But, when written up for publication in 1924, it was presented in terms of inductors and capacitors.

It might also be noted that, before 1930, it was known that the most general driving-point impedance obtainable with the three ordinary kinds of elements (together with mutual inductance) could be realized in a certain canonical form [10], [13] by means of ideal transformers and the three kinds of elements. However, no conditions on the nature of the impedance itself were apparently known.

For the realization of the general driving-point impedance, the modern era may be said to have started in 1931 with the publication by Brune [6] of his celebrated theorem. The importance of this step rests essentially in the giving of necessary and sufficient conditions on the impedance function in order that it could be realized by means of the three kinds of ordinary elements, together with mutual impedances. Thus, we were given the general concept of a positive real function. Any driving-point impedance of the sort of network in question was a positive real rational function; any positive real rational function could be realized by that sort of network. In brief, a positive real rational function is a rational function of a complex variable with the following two properties: 1) if the real part of the variable is positive, the real part of the value of the function is positive; and 2) if the variable is real, the value of the function is real.

Brune's work covered a variety of cases, showing that in certain instances mutual inductances were not needed, but that in general his method of synthesis called for mutual inductance, introduced through the medium of perfectly coupled coils.

Thus the matter rested until 1949. I dare say that in the intervening years, many specialists in the subject may have ventured the opinion that mutual inductance could not be entirely eliminated. But in that year, Bott and Duffin [4] showed that any positive real rational function could be realized as the impedance of a network composed of just the three kinds of ordinary elements. This was really an amazing result; and it was presented to the public, proof and all, as a letter-to-the-editor that took only one-half a printed page!

This synthesis was accomplished at a price—the price of many superfluous elements, superfluous in the sense that they far exceeded in number the number of independent coefficients in the rational impedance function.

It was natural that there should be attempts to simplify the original structure as called for by the Bott and Duffin synthesis. Within a half-dozen years, several different papers [26], [44], [46]–[49], [55] presented the same essential simplification, reducing the number of reactive elements called for in each step of the synthesis from six to five. For a biquadratic impedance function, this simplification has recently been shown [51] to be the best possible, that is, so far as the number of elements is concerned. However, the general question still remains unanswered as to further possible simplifications in the case of impedances of higher order.

In the meantime, other methods of synthesis [2], [35], [37], [41] have been developed, which, though without the complete generality of the simplified Bott and Duffin method, nevertheless provide methods for realizing certain restricted classes of positive real rational functions with considerably fewer elements. Much remains to be done in this direction. The answer to the following question has not yet been given in all completeness: For a given numerical positive real function, what is the least number of elements with which this particular impedance function can be synthesized? And, having answered that question, what is the complete set of equivalent structures, each composed of this minimum number of elements, realizing the given impedance function?

Even in the comparatively simple case of the biquadratic impedance function, these questions have not been completely answered, although some attempts have been partially successful [37], and considerable progress has been made to date. Perhaps it may be worthwhile to consider this case of the biquadratic impedance function in some detail, in order to illustrate many of the matters mentioned above.

The general biquadratic impedance function is

$$Z(s) = \frac{As^2 + Bs + C}{Ds^2 + Es + F}. \tag{2}$$

There is no loss in generality in assuming that the six coefficients A, B, C, D, E, F are all non-negative real numbers. Furthermore, in the interest of complete generality, it is advantageous to retain all six coefficients, although there are just five parameters involved (the ratios of the six coefficients). It is perhaps a temptation to reduce some one of them to unity, but then a degree of generality would be sacrifice (the possibility of reducing that particular coefficient to zero).

A necessary and sufficient condition that this function $Z(s)$, with the assumptions as to its coefficients, be a positive real function is that

$$(\sqrt{AF} - \sqrt{CD})^2 \leqq BE. \tag{3}$$

Furthermore, we will assume that the biquadratic function (2) is a true biquadratic in the sense that at least one of the two polynomials (numerator and denominator) is of the second degree, with no common factors which might cancel out. Accordingly, the discriminant

$$\begin{aligned} K &= (AF - CD)^2 - (AE - BD)(BF - CE) \\ &= A^2F^2 - ABEF - 2ACDF + ACE^2 + B^2DF \\ &\quad - BCDE + C^2D^2 \end{aligned} \tag{4}$$

cannot vanish. For, if $K=0$, there is a linear factor common to both numerator and denominator, and so the function would reduce immediately to no more than a bilinear function.

If, however, $K<0$, then the roots of both the numerator and the denominator are real and separate each other. Under these circumstances, the impedance function is realizable with two kinds of elements: resistors and inductors if, in addition, $AF-CD>0$; resistors and capacitors if $AF-CD<0$. Note in passing that $AF-CD$ cannot equal zero if K is less than zero

In the case of most interest, therefore, we may assume that the discriminant K is actually positive, and so all three kinds of elements are required. The problem is therefore the synthesis of a positive real biquadratic function meeting this condition on K.

Now the modified Bott-Duffin synthesis in the case of the biquadratic function calls for no more than eight elements. We would expect that, in general, five elements would be needed in view of the five independent parameters in the function specified. We realize, of course, that for special choices of the coefficients, four, three, or even two elements may be sufficient.

The major questions to be answered are these: What additional restrictions must be imposed upon A, B, C, D, E, F in order that the impedance fucntion be realizable by a five-element network? By a six-element network? By a seven-element network? Then, those questions having been answered, it will be desirable, in the case of any specified impedance function meeting the requirements for an n-element network, to know the totality of equivalent n-element structures realizing this particular function.

Partial answers have been given to some of these questions, for example, in the extreme case when the equality sign holds in condition (3) for a positive real function—that is, when

$$\sqrt{AF} - \sqrt{CD} = \sqrt{BE}, \tag{5}$$

or

$$\sqrt{CD} - \sqrt{AF} = \sqrt{BE}. \tag{6}$$

In the first of these cases (5), the modified Bott-Duffin method calls for seven elements: two resistors, three inductors, and two capacitors. There are two different structures each using the same number of elements. In the second of these cases (6), the numbers of inductors and capacitors required are simply interchanged.

This seems to be a complete answer for these extreme cases, except for the obvious 2- and 3-element special cases, and for the rather unusual pair of 5-element bridge networks [35], [51], [59] corresponding to the extreme cases

$$AF = 4BE = 4CD, \tag{7}$$

and

$$4AF = 4BE = CD. \tag{8}$$

These bridge cases illustrate one of the few theorems dealing with three arbitrary kinds of elements. If one pair of opposite bridge arms consists of Z_1 elements, the other pair of Z_2 elements, and the cross arm of Z_3 element (a Z_1, Z_2 lattice terminated in Z_3), the driving-point impedance across the terminals of the bridge is

$$Z = \frac{2Z_1^2Z_2 + Z_1^2Z_3 + 2Z_1Z_2^2 + 2Z_1Z_2Z_3 + Z_2^2Z_3}{Z_1^2 + 2Z_1Z_2 + 2Z_1Z_3 + Z_2^2 + 2Z_2Z_3}. \tag{9}$$

Upon eliminating the common factor (Z_1+Z_2) from numerator and denominator, we obtain

$$Z = \frac{2Z_1Z_2 + Z_1Z_3 + Z_2Z_3}{Z_1 + Z_2 + 2Z_3}. \tag{10}$$

This can be made to yield a biquadratic in a variety of ways.

A somewhat different approach [36], [41] yields a synthesis in the biquadratic case when

$$|AF - CD| \leqq BE. \tag{11}$$

This yields a seven-element network, with a six-element network in the extreme case when the equality sign holds, once again with the exception of obvious special cases.

Still another approach [38] is to make a census of all five-element networks consisting of three resistors, one inductor, and one capacitor, covering all such structures realizing a biquadratic impedance (provided the particular impedance cannot be realized with even fewer elements). The result of such a census yields 25 distinct networks, arranged in 11 groups, the networks in any one group being completely equivalent, one to the other. Thus, there are 11 sets of special conditions on the coefficients, one set for each group.

A further elaboration of the above study includes a census of four-element structures (two resistors), leading to 18 networks in 12 groups, each group again with a special set of conditions on the coefficients. To complete this study we find eight networks in the three-element class, each with its own special set of conditions, and finally the two extreme two-element structures (inductor and capacitor in series, and in parallel).

Thus, we find a total of 53 distinct networks, arranged in 33 groups each irreducible in the sense that there exists no equivalent network containing fewer elements; each group has its own distinctive set of conditions on the coefficients A, B, C, D, E, F.

In addition, it may be pointed out that nothing is gained by adding resistors in excess of the three called for in the original census. Thus, in one sense, the problem has been completely solved for a biquadratic function which does not essentially have common factors added to both numerator and denominator, a case that perhaps might be termed the strict biquadratic.

An obvious extension would then be to consider, in much the same exhaustive way, the case of the biquadratic with one added linear factor in both numerator and denominator, thus in general entailing three reactive elements. Then consider the case with two added linear factors, calling in general for four reactive elements, and so on.

Another avenue of investigation might involve a comprehensive study of cases involving more than the minimum number of elements, especially those cases where the function form is simplified, somewhat similar to the case of the constant-resistance network mentioned above.

Let us now turn to a consideration of networks composed of just one kind of element, and since no loss in generality is involved, let us deal specifically with resistance networks.

With regard to the true n-terminal network, not much seems to have been done since the days of Campbell [7]. He showed that the network consisting of $n(n-1)/2$ elements, directly connecting the n terminals in pairs, was the most general possible n-terminal network. Furthermore, he gave formulas for obtaining the values of the elements in this directly connected network from the set of $n(n-1)/2$ impedance measurements—the impedance measured across each pair of terminals with all other terminals left free. Quite recently, a somewhat different line of attack [52] on this problem transformed the problem into an n-port problem, but an n-port problem of a very special type, where special methods can be used.

However, when we consider the n-port resistance network, we find a problem under active investigation. Here, the problem is usually stated as follows: Given a symmetric square matrix of order n composed of real numbers, what are necessary and sufficient conditions that this matrix be realizable as the open-circuit impedance matrix (or closed-circuit admittance matrix) of a resistance network?

It has long been known that, if ideal transformers are allowed, a necessary and sufficient condition for realizability is simply that the matrix correspond to a positive definite quadratic form, either singular or nonsingular (often referred to as semi-definite).

At present, efforts are being directed toward obtaining necessary and sufficient conditions for realizability by means of a network composed entirely of ordinary resistors, with no ideal transformers whatsoever. Important advances in this attempt have been made by Cederbaum [15]–[19] in the last few years, and by other investigators [53], [64]. One of the major points made is the discovery of a new necessary condition on such a matrix, namely, that it must be a paramount matrix. A paramount matrix, by definition, is a symmetric real matrix with the property that the determinant of any principal minor (of any order from 1 to n) is not less than the absolute value of the determinant of any other minor selected from the same rows of the matrix as the principal minor involved.

For several years, it had been known [58] that paramountcy was also a sufficient condition for realizability provided n did not exceed 3. Cederbaum [19], however, has shown quite recently that paramountcy is not sufficient, in general.

On the other hand, Guillemin [31], [32] has developed some sufficient conditions for realizability; but it has not been shown that these are necessary.

Thus, this general question is still open. It seems to present a very puzzling state of affairs: a seemingly simple question, easily stated, but so far defying a complete solution. Also, it represents a peculiar situation, where the necessary and sufficient conditions for realizability as an open-circuit impedance matrix must differ (in the light of Cederbaum's examples) from the conditions for realizability as a closed-circuit admittance matrix. This seems to present an extraordinary state of affairs when viewed in the light of the usual concepts of the principle of duality. Nevertheless, this problem would appear to be on the verge of being solved; all indications point that way.

When we come to networks composed of two kinds of elements, that is, beyond the 2-terminal (or 1-port) stage, we find much less progress. Perhaps the best progress has been made with the 3-terminal network composed of two kinds of elements.

One way of handling the 3-terminal network is to introduce the equivalent delta composed of three admittance-type elements (usually not all physically realizable), or, alternatively, the star composed of three impedance-type elements (usually not all physically realizable). Then, by passing alternately from the one representation to the other, extracting a suitably selected physically realizable portion at each step before proceeding to the next step, and also by a possible splitting of the delta representation into deltas in parallel, a 3-terminal series-parallel synthesis is obtained.

Adams [1] has given necessary and sufficient conditions for such a series-parallel synthesis for reactance networks, valid when the equivalent delta admittances (or star impedances) are of no higher than the sixth degree. These conditions are necessary for synthesis in the prescribed series-parallel form; but it is not known whether they are necessary for synthesis in arbitrary form.

Another approach has centered around the number of interior nodes involved in the network employed to connect the three terminals [20], [63]. Other authors have developed special sufficient conditions [39], [43].

Much the same state of affairs prevails with respect to the 2-port network composed of two kinds of elements, that is, with respect to a complete specification of the network. If only one network function is to be realized (usually a transfer function), then the problem has been almost completely solved [21]–[25]—but that is not the question under consideration here.

A somewhat similar situation holds with respect to the realization of a single transfer function, with all three kinds of elements allowed [27], [34], [45], [54], [60]–[63].

In addition to the search for necessary and sufficient conditions for specific kinds of network performance, there has developed, in recent years, a great deal of interest in the study of the abstract linear graphs corresponding to the physical electrical circuits. Witness the interest in circuit topology, so-called, as evidenced by at least two special issues of the IRE TRANSACTIONS ON CIRCUIT THEORY [65], [66], and elsewhere. The state of linear graph theory, especially in its applications to electric circuit theory, has been well summarized in a recent paper [33], to which the reader is referred for further details. Enumerative problems concerning graphs have also received comprehensive treatment in a recent book [50].

BIBLIOGRAPHY

[1] K. M. Adams, "On the synthesis of three-terminal networks composed of two kinds of elements," *Philips Res. Repts.*, vol. 13, pp. 201–264; 1958.

[2] R. F. Baum, "Synthesis of driving point impedances with geometric symmetry," IRE TRANS. ON CIRCUIT THEORY, vol. CT-5, pp. 359–363; December, 1958.

[3] G. Biorci and P. P. Civalleri, "On the synthesis of resistive N-port networks," IRE TRANS. ON CIRCUIT THEORY, vol. CT-8, pp. 22–28; March, 1961.

[4] R. Bott and R. J. Duffin, "Impedance synthesis without the use of transformers," *J. Appl. Phys.*, vol. 20, p. 816 (letter); 1949.

[5] D. P. Brown and Y. Tokad, "On the synthesis of R networks," IRE TRANS. ON CIRCUIT THEORY, vol. CT-8, pp. 31–39; March, 1961.

[6] O. Brune, "Synthesis of a finite two-terminal network whose driving-point impedance is a prescribed function of frequency," *J. Math. and Phys.*, vol. 10, pp. 191–236; 1931.

[7] G. A. Campbell, "Direct capacity measurement," *Bell Sys. Tech. J.*, vol. 1, no. 1, pp. 18–38; 1922. ("Collected Papers," pp. 169–189.)

[8] —— and R. M. Foster, "Maximum output networks for telephone substation and repeater circuits," *Trans. AIEE*, vol. 39, pp. 231–280; 1920. (Campbell, "Collected Papers," pp. 119–168.) Ideal transformers are treated in Appendix VI.

[9] H. J. Carlin and D. C. Youla, "Network synthesis with negative resistors," PROC. IRE, vol. 49, pp. 907–920; May, 1961.

[10] W. Cauer, "Die Verwirklichung von Wechselstromwiderständen vorgeschriebener Frequenzabhängigkeit," *Arch. Elektrotech.*, vol. 17, pp. 355–388; 1926.

[11] ——, "Über eine Klasse von Funktionen, die die Stieltjesschen Kettenbrüche also Sonderfall enthalt," *Jahresber. deutsch. Math.-Ver.*, vol. 38, pp. 63–72; 1929.

[12] ——, "Ein Reaktanztheorem," *Sitzber. preuss. Akad. Wiss., Phys.-math. Klasse*, vol. 30, pp. 673–681; 1931.

[13] ——, "Untersuchungen über ein Problem, das drei positiv definite quadratische Formen mit Streckenkomplexen in Beziehung setzt," *Math. Ann.*, vol. 105, pp. 86–132; 1931.

[14] ——, "Ideale Transformatoren und lineare Transformationen," *Elek. Nachr.-Tech.*, vol. 9, pp. 157–174; 1932.

[15] I. Cederbaum, "On networks without ideal transformers," IRE TRANS. ON CIRCUIT THEORY, vol. CT-3, pp. 179–182; September, 1956.

[16] ——, "Conditions for the impedance and admittance matrices of *n*-ports without ideal transformers," *Proc. IEE*, vol. 105, pt. C, pp. 245–251; 1958. (Monograph 276R.)

[17] ——, "Applications of matrix algebra to network theory," IRE TRANS. ON CIRCUIT THEORY, special suppl., vol. CT-6, pp. 127–137; May, 1959.

[18] ——, "Paramount matrices and synthesis of resistive *N*-ports," IRE TRANS. ON CIRCUIT THEORY, vol. CT-8, pp. 28–31; March, 1961.

[19] ——, "Topological considerations in the realization of resistive *N*-port networks," IRE TRANS. ON CIRCUIT THEORY, vol. CT-8, pp. 324–329; September, 1961.

[20] S. Darlington, "A survey of network realization techniques," IRE TRANS. ON CIRCUIT THEORY, vol. CT-2, pp. 291–297; December, 1955.

[21] A. D. Fialkow, "Two terminal-pair networks containing two kinds of elements only," *Proc. Symp. on Modern Network Synthesis*, Polytechnic Inst. of Brooklyn, N. Y., pp. 50–65; 1952.

[22] —— and I. Gerst. "The transfer function of an R-C ladder network," *J. Math. and Phys.*, vol. 30, pp. 49–72; 1951.

[23] —— and ——, "The transfer function of general two terminal-pair RC networks," *Quart. Appl. Math.*, vol. 10, pp. 113–127; 1952.

[24] —— and ——, "The transfer function of networks without mutual reactance," *Quart. Appl. Math.*, vol. 12, pp. 117–131; 1954.

[25] —— and ——, "Impedance synthesis without minimization," *J. Math. and Phys.*, vol. 34, pp. 160–168; 1955.

[25] —— and ——, "Impedance synthesis without minimization, *J. Math. and Phys.*, vol. 34, pp. 160–168; 1955.

[26] —— and ——, "Impedance synthesis without mutual coupling," *Quart. Appl. Math.*, vol. 12, pp. 420–422; 1955.

[27] —— and ——, "RLC lattice transfer functions," PROC. IRE, vol. 43, pp. 462–469; April, 1955.

[28] R. M. Foster, "A reactance theorem," *Bell Sys. Tech. J.*, vol. 3, pp. 259–267; 1924.

[29] ——, "Topologic and algebraic considerations in network synthesis," *Proc. Symp. on Modern Network Synthesis*, Polytechnic Inst. of Brooklyn, N. Y., pp. 8–18; 1952.

[30] ——, "Passive network synthesis," *Proc. Symp. on Modern Network Synthesis II*, Polytechnic Inst. of Brooklyn, N. Y., pp. 3–9; 1955.

[31] E. A. Guillemin, "How to grow your own trees from given cut-set or tie-set matrices," IRE TRANS. ON CIRCUIT THEORY, special suppl., vol. CT-6, pp. 110–126; May, 1959.

[32] ——, "On the analysis and synthesis of single-element-kind networks," IRE TRANS. ON CIRCUIT THEORY, vol. CT-7, pp. 303–312; September, 1960.

[33] F. Harary, "Graph theory and electric networks," IRE TRANS. ON CIRCUIT THEORY, special suppl., vol. CT-6, pp. 95–109; May, 1959.

[34] E. C. Ho, "RLC transfer function synthesis," IRE TRANS. ON CIRCUIT THEORY, vol. CT-3, pp. 188–190; September, 1956.

[35] W. H. Kim, "On non-series-parallel realization of driving-point function," 1958 IRE NATIONAL CONVENTION RECORD, pt. 2, pp. 76–81.

[36] N. Kowalewski, "Über die kopplungsfreie Realisierung von Zweipolen deren Scheinwiderstand eine vorgegebene Funktion der Frequenz ist," Ph.D. dissertation, University of Munich, Germany, 110 pp.; 1947.

[37] E. S. Kuh, "Special synthesis techniques for driving point impedance functions," IRE TRANS. ON CIRCUIT THEORY, vol. CT-2, pp. 302–308; December, 1955.

[38] E. L. Ladenheim, "A Synthesis of Biquadratic Impedances," Master's thesis, Polytechnic Inst. of Brooklyn, N. Y., 6+69 pp.; 1948.

[39] H. M. Lucal, "Synthesis of three-terminal RC networks," IRE TRANS. ON CIRCUIT THEORY, vol. CT-2, pp. 308–316; December, 1955.

[40] G. L. Matthaei, "Some techniques for network synthesis," PROC. IRE, vol. 42, pp. 1126–1137; July, 1954.

[41] F. Miyata, "A new system of two-terminal synthesis," IRE TRANS. ON CIRCUIT THEORY, vol. CT-2, pp. 297–302; December, 1955.

[42] E. L. Norton, "Constant resistance networks with applications to filter groups," *Bell Sys. Tech. J.*, vol. 16, pp. 178–193; 1937.

[43] H. Ozaki, "Synthesis of RC-3-terminal network without ideal transformer," *Tech. Repts. Osaka University*, vol. 3, pp. 57–77; 1953.

[44] R. H. Pantell, "A new method of driving-point impedance synthesis," PROC. IRE (Correspondence), vol. 42, p. 861; May, 1954.

[45] ——, "Minimum-phase transfer-function synthesis," IRE TRANS. ON CIRCUIT THEORY, vol. CT-2, pp. 133–137; June, 1955.

[46] F. M. Reza, "A bridge equivalent for a Brune cycle terminated in a resistor," PROC. IRE (Correspondence), vol. 42, p. 1321; August, 1954.

[47] ——, "Conversion of a Brune cycle with an ideal transformer into a cycle without an ideal transformer," *J. Math. and Phys.*, vol. 33, pp. 194–198, 1954; IRE TRANS. ON CIRCUIT THEORY, vol. CT-1, No. 1, pp. 71–75, March, 1954.

[48] ——, "Synthesis of one terminal-pair passive networks without ideal transformers," PROC. IRE (Correspondence), vol. 42, p. 349; January, 1954.

[49] ——, "Synthesis without ideal transformers," *J. Appl. Phys.*, vol. 25, p. 807–808 (letter); 1954.

[50] J. Riordan, "An Introduction to Combinatorial Analysis," John Wiley and Sons, Inc., New York, N. Y., and Chapman and Hall, Ltd., London, England, 11+244 pp.; 1958. Network problems are treated in Chapter 6.

[51] S. Seshu, "Minimal realizations of the biquadratic minimum function," IRE TRANS. ON CIRCUIT THEORY, vol. CT-6, pp. 345–350; December, 1959.

[52] G. E. Sharpe and B. Spain, "On the solution of networks by means of the equicofactor matrix," IRE TRANS. ON CIRCUIT THEORY, vol. CT-7, pp. 230–239; September, 1960.

[53] P. Slepian and L. Weinberg, "Synthesis applications of paramount and dominant matrices," *Proc. Natl. Electronics Conf.*, vol. 14, pp. 1–20; 1958.

[54] L. Storch, "Realization of minimum-phase and nonminimum-phase transfer functions by RLC ladder-type networks," IRE TRANS. ON CIRCUIT THEORY, vol. CT-7, pp. 137–150; June, 1960.

[55] J. E. Storer, "Relationship between the Bott-Duffin and Pantell impedance synthesis," PROC. IRE (Correspondence), vol. 42, p. 1451; September, 1954.

[56] B. D. H. Tellegen, "Gyrator, a new electric network element," *Philips Res. Repts.*, vol. 3, pp. 81–101; 1948.

[57] ——, "Synthesis of passive, resistanceless four-pole that may violate the reciprocity relation," *Philips Res. Repts.*, vol. 3, pp. 321–337; 1948.

[58] ——, "Theorie der wisselstromen," Noordhoff, Groningen, The Netherlands; 1952. See the treatment on pp. 166–168.

[59] M. E. Van Valkenburg, "Special case of a bridge equivalent of Brune networks," PROC. IRE (Correspondence), vol. 44, p. 1621; November, 1956.

[60] L. Weinberg, "Synthesis of unbalanced RLC networks," *J. Appl. Phys.*, vol. 24, pp. 300–306; 1953.

[61] ——, "RLC lattice networks," PROC. IRE, vol. 41, pp. 1139–1144; September, 1953.

[62] ——, "Unbalanced *RLC* networks containing only one resistance and one real transformer," IRE TRANS. ON CIRCUIT THEORY, no. PGCT-2, pp. 55–70; December, 1953.

[63] ——, "Sub-Commission VI-2 Report on Circuit Theory," presented at 13th Triennial General Assembly of URSI (Union Radio Scientifique Internationale), London, England; 1960.

[64] —— and P. Slepian, "Realizability conditions on *n*-port networks," IRE TRANS. ON CIRCUIT THEORY, vol. CT-5, pp. 217–221; September, 1958.

[65] Network topology papers, IRE TRANS. ON CIRCUIT THEORY vol. CT-5, pp. 4–54; March, 1958.

[66] Special issue on topology, IRE TRANS. ON CIRCUIT THEORY, vol. CT-8, pp. 4–58; March, 1961.

Teaching of Circuit Theory and Its Impact on Other Disciplines*

E. A. GUILLEMIN†, FELLOW, IRE

Summary—**In order to place the teaching of circuit theory in proper perspective, it is necessary first to formulate rather explicitly the principles and the philosophy of educational methods and objectives in an engineering curriculum. Subject content and subject order, as well as difficulties experienced because of existing academic policies, provide the central theme which is concerned with developments throughout the past 35 years in which circuit theory experienced its major growth.**

Introduction

THE TWO things that have interested me most throughout my professional career are teaching and circuit theory. I look with considerable pride and satisfaction upon the fact that many of the attitudes and principles that emerged out of these years of activity are now being ably carried on by some of my former students who, in their own teaching careers, are scattered throughout engineering schools both here and abroad. While I am tempted to dwell and elaborate upon this pleasant theme, my conscience and my deep-rooted interest in engineering motivates me rather to make comments and criticisms that may be helpful in guiding the future teaching of circuit theory.

I have been very fortunate in having lived through and taken part in an era during which circuit theory experienced its major growth and development. I have seen the effect of this growth upon all of the electrical engineering curriculum in which circuit theory, in reality linear system theory, has always played a dominant role. I have had an opportunity to compare the peculiarities of this evolutionary process in the conservative years preceding World War II with its strikingly accelerated character in the post-war years. I have seen many schemes tried; I have seen many successes and many failures.

Throughout these years I have consistently clung to defending and promoting those methods and procedures that provide the basic values in an engineering education. Strange as it may seem, the struggle to achieve these goals is never ending. One cannot rest upon one's laurels, for memories, unfortunately, are short and new recruits, eager to try their own wings and unwilling to pay heed to lessons of the past, are constantly changing the attitudes of engineering faculties.

So, once again, I feel compelled to make a speech that I have made many times before, but never to so wide an audience as the entire membership of the Institute of Radio Engineers. I can only hope that my message will fall upon a sufficient number of sympathetic ears to make this effort worth while.

* Received by the IRE, July 31, 1961; revised manuscript received, October 27, 1961.
† Massachusetts Institute of Technology, Cambridge, Mass.

Objectives and Basic Philosophy

Before any topic can be discussed intelligently, one must clearly understand what is meant by its title, which brings us to the question: what is circuit theory? Unlike other topics in electrical engineering, almost everyone in the profession is confident not only that he knows what circuit theory is, but also that he knows how to solve problems in circuit theory. Since this attitude is largely motivated by practical problems involving circuits of one sort or another, and since the variety of such circuit problems and their pertinent aspects is almost endless, there is a corresponding divergence of opinion as to what constitutes the essentials of circuit theory or what should be taught in a subject bearing that title.

It is not difficult to see that one of the first things a teacher must do here is to bring some order into this rambling situation through appropriate classification, and try, if possible, to find a "common denominator" or theoretical basis that he can regard as the "fundamentals" of circuit theory. On these he can, then, at first concentrate all of his attention and later expand logically by broadening his horizon so as to encompass as much more of the total picture as his available time permits, always stressing the basic character of fundamental principles in providing the essential means for constructing whatever partial or total solution is achievable.

While simple in principle, this statement of objectives is not simple in its implementation; neither is there a widespread agreement on its appropriateness for the training of the engineer who, unlike the pure science student, must be taught to cope with practical as well as the purely theoretical aspects of problems in his field. Can or should these practical aspects be taught in the classroom, and if so, what apportionment of time to such things is proper? This and a host of similar questions enter the picture and confuse the real issues. Industrial leaders are divided on the answers and so are those in the teaching profession. Attitudes today are even more divided than they were 35 years ago, in spite of the fact that experiences during World War II had a tremendous effect in pointing up the value of basic science in the engineering curriculum.

This is an era of gadgets and gadgetry; it is also an era in which the mania for changing things is rampant. Everything must be changed, whether for better or worse; and anyone who resists this trend, be it for stubbornness or sound reason, becomes an obstructionist. An idiotic idea is retitled a bold innovation and is seized upon by those who have little originality or sound judgement of their own but are ever eager to be found among that group enjoying the popular spotlight. This trend of the times penetrates even the academic environment of our learned institutions and causes confusion in the minds of those whose primary responsibility it is to distinguish clearly between the relevant and the irrelevant from a basic point of view.

It is a common inconsistency for educators, on the one hand, to cast aspersions upon the electrical engineering curriculum of yesteryear because it gave significant time and attention to rotating machine devices such as the synchronous converter instead of emphasizing basic energy conversion principles; while on the other hand, they find it quite appropriate and in keeping with the modern trend to spend half the available time in a fundamental circuits subject upon such gadgets as the transistor and the tunnel diode. Thirty years hence, these devices will be as obsolete as the rotary converter is today; and to give prominence to their discussion in the electrical engineering curriculum now is as proper or improper as it was to give similar prominence to rotating machines in days gone by.

Today's devices, like those of past decades, are merely current examples of the application of basic scientific principles. The devices live only long enough to be superseded by better ones, while the principles upon which their discovery rests endure and give birth to new devices in a never-ending sequence. It should be rather clear, therefore, that the emphasis should be placed upon basic principles and that devices should enter only by way of illustrative examples, except insofar as methods for their utilization involve techniques that, in themselves, are basic in character. In that case an appropriate degree of emphasis should be placed upon the techniques and their usefulness in dealing with a broad class of devices having certain common characteristics.

Since most of what has been said here is rather self evident, one may wonder why it needs to be said at all. That fundamental principles are important and enduring while their applications are essential only because it is through a discussion of them that the fundamental principles become thoroughly understood, has been said so often that it begins to sound like a platitude. In spite of this fact, and particularly in today's complex pattern of activity where relative values are often badly distorted, it is ever more important not to lose sight of our objectives, for it is only through a clear understanding of what we wish to accomplish that we can decide between a proper and an improper course of action.

Science and its pursuit have suddenly become glamorous; and in the minds of the general public, science and its achievements are measured solely by the performance of the gadgets that it produces. We should not be tempted to use the same yardstick in measuring the appropriateness of an engineering curriculum. To teach engineers how to build moon rockets would be a pedagogical blunder, even if we knew how to do it.

I know that many of my colleagues will disagree with this view. They will say: "What better way is there to teach engineers than to pick some comprehensive practical problem whose solution requires a combination of theoretical disciplines as well as engineering judgement and knowhow, and apply these in the proper manner? This really motivates and gets ideas across. In one package the student learns about theoretical dynamics, circuit theory, automatic control, celestial mechanics, radio transmission and reception, ballistics, etc. etc.!" The fallacy in this viewpoint is that the student gets only a smattering of each of these disciplines; and, even if many such problems are discussed, he still gets only a random collection of smatterings which add up to nothing but confusion. The discussion of such comprehensive problems is appropriate only *after* the separate disciplines have been adequately presented. Then such discussion serves as a useful integration and orientation medium, which is all that it can be expected to do. The disciplines themselves cannot properly be taught in this way.

Some argue that the comprehensive discussion should come *before* not *after* the introduction of separate disciplines, that it motivates interest in them and makes their subsequent separate consideration more meaningful to the student. This sounds like a good argument, but it doesn't work out that way, as those of us who have taught long enough and have tried this method know very well. Instead of becoming eager for a thorough understanding of the separate disciplines, the student becomes misled into thinking that he already knows all that he needs to know about them. Each subject is regarded as "old hat" and its further study as a waste of time and effort which could preferably be devoted to other specialized applications. Instead of being motivated, his attitude is one of intolerance and disinterest toward any systematic consideration of the topic in question.

Except for a few exceptional individuals, any attempt to fill in a proper background at this stage falls very flat. The students will not listen to it. Their attitude is apathetic. Trying to arouse them or reawaken their interest is like trying to stir an animal out of its state of hibernation. Their minds are already focused upon other things, and they refuse to become excited about what they believe to be unessential. They are wrong, of course, and will find this out when it is too late—when they have left behind those of us who were ready and eager to help them.

In the teaching of circuit theory we run into the same situation on a smaller scale; the conflict between those who would start with a discussion that includes literally everything but the kitchen stove, and those of us who would limit our introductory discussion to a minimum of essential topics and make thoroughness, scholarliness and depth rather than breadth our guiding theme.

The contention that we must acquaint the student with many different kinds of circuit elements that he will be expected to be familiar with when he assumes a post in industry, is in most cases a palliative with which the instructor tries to satisfy his conscience and rationalize his procedure. Students in electrical engineering, for the most part, have played with these gadgets in their spare time and in the associated laboratory. What they do not already know about them when they arrive at their jobs, they can readily acquire, especially in an environment that abounds in practical things. What they cannot acquire in such an environment is a habit of thoroughness and the ability to apply fundamental principles to create solutions to new problems.

A collateral argument put forth to justify the survey-type and gadget-filled presentation in an introductory circuits subject is that it provides its own motivation by keeping students amused with the manifold functions of devices, much as one would amuse a child by turning the pages in a picture book. This attitude literally degrades the dignity and scholarliness of a college atmosphere to the kindergarten level. Students who need this kind of motivation are misplaced in an academic environment; they are either too immature or too poorly prepared for it. To "nurse them through," so to speak, is a serious mistake. What the entering college student needs more than anything else is to be exposed to the discipline of a scholarly approach. If he cannot survive this, then he is not collegiate material, at least not in the best tradition of university standards. He must find another environment more suitable to his intellectual capacity.

All of these thoughts have a direct bearing upon the initial question raised here, for we cannot define the fundamentals of circuit theory without first clarifying what we mean by the term "fundamental" in connection with educational principles and methods. Whether or not a presentation is fundamental depends not only upon the material dealt with, but even more importantly, it depends upon the scholarliness with which it is considered, and the thoroughness with which its study is pursued. Unless the attitude is one in which thoroughness and depth of understanding are paramount, no claim regarding the fundamental nature of topics presented is justified. This point is of utmost importance, since it clearly establishes the fact that a syllabus or outline alone is nothing more than a collection of words; no criticism or discussion of its possible fundamental character can be based solely upon it. The all important factor concerns the manner in which the topics contained in such a syllabus are presented. Thoroughness and scholarliness are the indispensable attributes by which the fundamental character of a classroom presentation can be measured.

In this kind of a presentation we do not introduce a topic unless we are prepared to follow it to its ultimate conclusion, or at least to a point where the breadth of its implications, its limitations, and the subtleties of its detailed character reveal themselves. A quick look at many things, as is done in a survey type of presentation, can hardly be regarded as hitting at fundamentals regardless of content, since the student never learns what being fundamental means. Subject matter is important, of course, but primarily the question of being fundamental in an approach to a problem has to do with one's attitude or state of mind. This is why it is not a clear-cut matter to set down or define the fundamental principles pertinent to a subject whether it be circuit theory or anything else.

We touch here upon an old and much debated question regarding methods and objectives in engineering education. In view of the fact that the practice of engineering is largely practical rather than theoretical in nature, should not the practical aspects of the subjects taught receive major emphasis? Unless we regard engineering as limited to the practice of well established procedures and exclude creativeness and originality, the answer should be definitely negative. This does not mean that practical matters be excluded from our discussions. We need these to serve as illustrations and applications of basic theory and principles. But we should encourage, rather than discourage, students to consider the purely academic aspects of their studies. Once they leave the college campus behind them, they will have to be practical all the rest of their lives. We should not deny them the privilege and the luxury of being impractical during their few student years. The leaders in our profession are almost without exception those who spend at least some of their time in the role of an impractical dreamer.

Being thorough and teaching students how to think through a topic to its logical conclusion has collateral value also, in that it provides the basis upon which one exercises that elusive and much talked of quality called "engineering judgement" which, after all, is a decision making process. One cannot teach students how to make decisions, but one can teach them how to think straight, be thorough, and weigh carefully all the factors relevant to a given situation. Whether the pertinent problem acting as a vehicle for this process is practical or theoretical in nature is immaterial.

In reflecting upon these things one cannot help but be amazed at how short are the memories of many engineers and engineering faculties, and how often the battle to achieve these objectives needs to be refought. At the beginning of World War II, when engineers were presented the problem of developing radar, they were (except in very few cases) found woefully lacking in an ability to cope with such an unconventional situation; and physicists, both theoretical and experimental, had to be called in to do what essentially was an engineering job.[1] After the war was over, engineering educators told

[1] "Longhairs and short waves," *Fortune*, vol. 32, pp. 163–169, 206, 208; November, 1945.

themselves that never again should engineers be found wanting in the ability to solve problems for which no set pattern or method of solution exists; and thus was born the present-day clamor for more basic science in the engineering curriculum.

This move was in the right direction, but, unfortunately, clothes do not make the man, or rather, subject matter alone does not make a curriculum; and now after fifteen post-war years, many of our younger engineering teachers who professionally did not live through the pre- and post-war periods, are in favor of making all the old mistakes over again. The industry, with some notable exceptions, appears to be back to all of its old attitudes, and favors the kind of curriculum that makes men more immediately useful but leaves them essentially uneducated. This is why we need to be reminded constantly of the things that constitute an education in the true sense of the word.

Subject Matter and Subject Order

If, for the moment, we limit ourselves to content, it should be rather clear that the fundamentals of circuit theory are first of all concerned only with situations in which the behavior is essentially linear, and we may as well add the adjectives *passive* and *bilateral*. We can take it for granted that finiteness in element numbers or physical dimensions is an implicit restriction and that nothing essentially is lost by considering only lumped elements, since the transcendental functions characterizing so-called "distributed-parameter" systems such as lines and wave guides are conceptually rather effectively thought of as limiting forms of rational functions. In fact, if we present the subject of circuit theory as an approximation to field theory, then the so-called "lumped" network and its characterization in terms of rational functions is already playing the role of a discrete approximant to a distributed system, and any further distinction between general properties of these two embodiments is not only trivial but redundant as well.

Any first course or introductory treatment of circuit theory should limit its considerations to lumped, linear, passive and bilateral elements, for the methods of analysis applicable to this restricted class of circuits and a knowledge of their properties forms an indispensable basis for the clear understanding of methods designed to deal with problems involving more general types of circuit elements. For example, properties injected by the presence of active elements cannot be clearly understood without a background provided by properties of passive circuits. Circuits whose behavior patterns do not conform with the reciprocity theorem because they contain nonbilateral elements are presented preferably as a departure from the normal situation; and the existing methods of dealing with nonlinear circuits, such as the incrementally-linear method for small signal applications and the piecewise linear method for the determination of large-signal behavior, tacitly assume familiarity with methods of linear system analysis and, consequently, cannot be properly discussed without a solid foundation in the linear theory.

Linear circuit theory is linear system theory regardless of whether that system is electrical, mechanical, acoustic or a combination of these. Even problems in economics and business procedures, or any others in which the controlling relations are linear, may be included in this broad class. All of the theory of classical dynamics is applicable here, as is the theory of linear algebra and of functions of a complex variable with its collateral overlap in potential theory and conformal mapping. Last, but not least, is the pertinence of combinatorial topology through which the properties of circuits dealing with invariance to changes in geometrical form or topological structure, as well as questions relating to the introduction and transformation of variables, can be studied effectively. Linear system theory and the collateral branches of physics and mathematics upon which it rests is indeed the essential foundation. Without it, no engineer can get beyond the conventional things he has learned to do by rote; with it, his horizon is practically unlimited, for he can adapt and devise and, thus, create methods of solution where none have existed before.

The sequence in which various topics in circuit theory are presented is extremely important in providing a proper perspective and an understanding of the basic definitions of various pertinent quantities, especially those whose existence is solely mathematical in character. The most important of these is the so-called impedance function. It has no physical significance. It is merely a mathematical convenience. Nevertheless, its role is an important one and its usefulness can be moderate or great depending upon one's awareness of all the implications that stem from the way in which it is derived or defined.

The ideal arrangement of topic order would be to have the students equipped with the essential physics and mathematics background before presenting them with any circuit theory principles. Unfortunately, other curriculum constraints force us to begin our circuit theory discussions before the above mentioned mathematical topics can be introduced and even before, or simultaneously with, a presentation of the physics of electricity and magnetism. The difficulty presented by these circumstances can be met by beginning with the purely topological aspects of circuit theory concerning the proper introduction of sets of independent variables. This discussion requires as yet no background in the physics of electricity (beyond the concepts of voltage and current with which most students are familiar anyway) and settles an issue which is logically, at least, Number 1 on the agenda anyway. By the time the discussion gets around to inductance and capacitance elements, these have been introduced by the contemporary physics subject.

The point at issue here is not crucial. One can introduce these elements in the circuit theory discussions before students have become acquainted with the pertinent physics background by tentatively defining them in terms of their volt-ampere relations in the same way that Ohm's law defines a resistance. This expedient is a compromise, of course, but the many constraints which present themselves in the construction of a curriculum leave us little choice.

Although, as mentioned above, a discussion of network topology logically belongs at the beginning, there are those who feel it is better to postpone this topic of topology until some circuit theory motivation for it has been provided. In this arrangement there is, however, the danger that it will be postponed indefinitely and thus be crowded out entirely. Considering its growing importance and usefulness in various branches of electrical engineering as well as its indispensability in a thorough consideration of basic circuit theory, this possible end result should not be considered too lightly.

A second item of prime importance is the early consideration of transient response or the natural behavior of circuits. There was a time, not too long ago, when this topic was considered as the "frosting on the cake," a dispensable item that could be added as a final flourish if one wanted to dress things up. The sinusoidal steady-state behavior of circuits in terms of a very restricted interpretation of the impedance function encompassed what was then considered to be the essential aspects of circuit theory for electrical engineers.

The increase in importance of the communication and control art put an entirely different light upon this state of things. Transient response became of primary importance; the impedance concept took on a much more general aspect, and the sinusoidal steady state became a very minor part of the total picture.

Classical dynamics correctly places great emphasis upon the natural behavior of a system. Natural frequencies or natural modes and normal coordinates play a dominant role here; and so they should in the teaching of circuit theory, especially if one wants the student to acquire an understanding of the physical picture that accompanies network response. Although the concept of normal coordinates is a mathematical fiction, it provides a tremendously useful interpretation for the mechanism whereby an excitation spreads its influence throughout a network, and how the subtle results peculiar to various degeneracies come about. In fact, one may go so far as to say that a clear understanding of such peculiarities cannot be had without these concepts. Even the sinusoidal steady-state behavior is more effectively portrayed in terms of this same mechanism which provides an insight into the current or voltage distribution pattern throughout a network and shows how this pattern becomes frozen in the vicinity of resonance.

From this point of view, it is impossible to discuss steady-state response without first presenting the story about natural behavior. Transient theory must precede steady-state theory. The concept of complex frequency, so indispensable to the interpretation of the impedance function and to network response generally, is thus logically introduced by pointing out that natural frequencies have this character. The impedance concept is most effectively introduced by forming the total solution as a sum of forced and force-free components which mathematically are, respectively, the particular integral and the complementary function.

For a suddenly applied sinusoid, this approach introduces the impedance function as characterizing the particular integral or steady-state response and yields it in the form of a rational function of the complex-frequency variable. In the case of a driving-point impedance, its zeros and poles are revealed as the natural frequencies on short- and open-circuit, and thus the fact that the impedance function characterizes the transient as well as the steady-state response becomes obvious. In simple cases where the open- and short-circuit natural frequencies are determined by inspection, one can write down the pertinent impedance at once.

In terms of pole-zero plots in the complex frequency plane, and with the help of some principles of conformal mapping, one can readily study the properties of impedance or more general response functions; and one can classify these functions and the pertinent networks according to particular pole-zero patterns.

The advantages that flow from this approach to linear circuit theory are admittedly so great that any objections that may be raised against its adoption quickly vanish once the results are recognized. Although some missionary work still remains to be done, the teaching of linear circuit theory from this point of view is now quite generally established in our engineering schools.

A Remaining Difficulty

There is, however, an important item on which much confusion remains and where the tendency to crowd too much into too little time and, hence, jeopardize clear thinking and understanding still holds sway. This item concerns the introduction of Fourier (and Laplace) theory into the circuit-theory picture. It is a mistake to inject Fourier methods into the presentation of circuit theory at too early a stage. The purely mathematical subtleties of Fourier representation offer a challenge in themselves; superimposing them upon the problem of formulating network response at a time when that process is still only moderately clear, produces a state of confusion in the student's mind that precludes the possibility that he will gain a proper understanding of either topic. Moreover, the tendency is to keep the presentation so shallow that subtleties or difficulties of any kind are avoided, in which case the student is duped, for he is given a tool without proper instructions regarding its use. He is given a set of inadequate rules with little or no explanatory justification.

Fourier representation is a mathematical trick. It deals with the subject of how one constructs arbitrary functions out of sinusoids. The given function emerges as the result of an interference of sinusoidal components and, hence, is also referred to as an interference pattern. One can build the same interference pattern out of sinusoids with steady amplitudes or out of sinusoids with exponentially growing or decaying amplitudes, or out of endless mixtures of these; and there are all sorts of rules as to what can or cannot be built in this way and what restrictions must be placed upon the mixtures, and so forth. This is Fourier theory. It has nothing whatever to do with circuit theory, but when we apply Fourier theory to the problem of computing network response, some additional restrictions and rules enter the picture. It is important to be able to distinguish between conditions imposed by Fourier theory and those imposed by circuit theory. In the usual presentation of this topic, especially in the watered-down variety that is injected into an introductory circuits subject, these restrictions become gloriously mixed up. The poor student is either hopelessly confused or not even aware of the fact that he should be confused.

This state of confusion is not restricted to students. There is, for example, a general misunderstanding about what is Fourier theory and what is Laplace transform theory and how the two are related. Since many books on Laplace transform theory introduce the unit-step and unit-impulse functions, the mere use of these is sometimes mistaken for Laplace transform methods. Expanding a transfer impedance in partial fractions and writing down the time-domain equivalent of each term as a means for finding the impulse response of a given network has nothing whatever to do with either Fourier or Laplace theory, yet there is a general impression among engineers and engineering teachers that the use of this process is synonymous with the use of Laplace transform methods.

It is important to be aware of the fact that the same compact form of solution usually gotten by Fourier or Laplace transform methods is also obtainable by the differential equation approach with far less mathematical fanfare since the convergence of infinite integrals and the evaluation of complex integrals is not involved. From the point of view of computing network response for an applied impulse, step, or sinusoid, nothing whatever is gained by the use of Laplace transform theory, and much time is consumed by its introduction if a worthwhile understanding is to be achieved.

This statement does not imply that Fourier theory is not an essential part of the electrical engineering curriculum, especially for students interested in communications and automatic control systems. However, its introduction into the curriculum should be postponed until a fairly complete and thorough understanding of basic circuit theory principles has been achieved, so that the student does not confuse these with the principles of Fourier theory and can distinguish clearly the part played by each with reference to their practical application. Thoroughness and depth of understanding are here more than ever essential.

On this point, there seems to be a wide divergence of opinion and the confusion and misconceptions caused thereby are not inconsiderable. First of all, the early introduction of Laplace transform methods is pedagogically bad because it provides the student with a batch of rules and a procedure which he can learn to apply by rote. Many students like this sort of thing; it saves them from having to think, a painful process which they are eager to avoid. It is essentially an answer-getting process of fairly general capabilities. Once this crank-turning process is learned, circuit theory for them is "in the bag." This attitude might even be tolerated if it meant that a real depth of understanding of Fourier and Laplace theory lay behind it, but such is seldom the case.

The introduction of Fourier and Laplace methods early in the curriculum has even more fundamentally undesirable consequences. These stem from a tendency to make Fourier and Laplace methods, instead of classical dynamics, the basis for circuit theory or linear system theory. According to this attitude, the frequency-domain response function of a network (for example, the impedance function) is defined as the transform of the unit impulse response in the time domain. The difficulty with this definition is that, in calculating network response, the Fourier or Laplace integral yields only the forced part of the total solution and, hence, represents that solution only if conditions are such that the force-free part is zero. However, since this approach to circuit theory is intended to be independent of any classical background, other ways must be sought in which to express this condition. It turns out in some cases that the specification of a particular path of integration is a mathematically equivalent condition, and the pertinent response function (impedance) is then said to "exist" only in corresponding regions of the s plane. This makeshift method of dealing with the difficulty not only substitutes a false reason for the correct one and, hence, misleads the student, but it puts restrictions upon the impedance function which, from a fundamental point of view, do not belong there and make no sense at all.

This method of defining a response function implies by dictum that the total response is given by the forced part alone. It even goes so far as to deny the existence of a force-free behavior; it denies the existence of natural frequencies; it even denies that the system have an equilibrium condition of any kind. For all this it substitutes an infinite integral definition and a set of rules ostensibly based only upon a convergence argument for the integral, which is not always valid. The student who does not have a classical background in circuit theory cannot possibly salvage anything useful from such a presentation.

Summary

Although topic-wise the teaching of circuit theory has pretty well kept pace with that rapidly developing subject during the past 35 years, the struggle to see that students get the right kind of training in basic principles is never ending. To understand the problem of teaching circuit theory fully, requires a continued long-lasting interest that, alone, enables one to see clearly the whole picture and the relative aspects of its various parts. It takes an interest that absorbs one's entire time and energy on the single topic of network theory.

The young teacher of today, with very few exceptions, does not devote such wholehearted effort to a single subject, especially to its teaching aspects. As a result, his attitudes are usually based upon a very limited appreciation of the total picture and, moreover, are biased by those applications which his particular research interest finds useful. This situation has been aggravated in recent years by a strong trend in academic circles to emphasize research and minimize the importance of teaching effort as regards staff promotions.

One hears much comment to the effect that more weight is laid these days upon excellence in teaching. Yet, contrary to the statement that we encourage only those who have demonstrated both teaching and research ability, there are numerous instances where excellent research ability alone has sufficed to assure promotion, but never one in which excellent teaching ability alone was regarded as sufficient. Until this attitude among academic administrators changes, we shall have the curriculum and teaching difficulties pointed out above.

There is a growing tendency to regard as undesirable a continued interest on the part of a staff member in teaching the same subject year after year. It is tacitly assumed that this situation must lead to stagnation. Toward a teacher who is really interested in his subject, this is an unfair attitude. Excellence in teaching implies a continued freshness of viewpoint and a perpetual change that keeps pace with and participates in the growth of the subject. Frequent rotation of staff, on the other hand, tends toward a repetition of mistakes and offers little opportunity to benefit from past experiences.

Without bias, it may be said that the growth and development of modern network theory throughout the past 35 years had a marked influence upon the upgrading of teaching methods and attitudes in other subjects in the electrical engineering curriculum, especially those in which circuit theory or the principles of linear system theory are essential tools. This upgrading, although stimulated primarily by the war and significantly evident since then, had its roots definitely established for at least a decade before this period.

The modern approach to network theory and, in particular, the methods relevant to its synthesis aspects, with their strong tendency toward rigor and generality, establish a habit of thought for the engineer that cannot help but spill over into his other areas of interest and there awaken the desire to achieve a similar degree of formulative compactness and mathematical elegance. In the engineering curriculum revision process, circuit theory has a 35-year head start over the other subjects. Eventually these will surely face all of the trials and struggles that it has been our pleasure to cope with and to write about here.

Section 8
COMMUNICATIONS SYSTEMS

Organized with the assistance of the IRE Professional Group on Communications Systems

Antennas and Transmission Lines*

HAROLD H. BEVERAGE†, FELLOW, IRE

Summary—**Antennas and transmission lines used in long-distance point-to-point radio communication are described. Long wavelengths (3000 to 23,000 meters) were used exclusively in the early days of transoceanic communication, and are still used for specific services. The design parameters of long-wave transmitting antennas are briefly mentioned. This is followed by a description of some of the directive antennas used for reception of long waves to reduce the effects of atmospherics and interference.**

The discovery in 1924 that portions of the short-wave spectrum (10 to 100 meters) could be used for long-distance transmission during daylight as well as at night, completely revolutionized long-distance radio communication. The development of arrays and long wire types of antennas for short-wave transmission and reception are described. The radiation of undesired sidelobes from long wire types of antennas is discussed and means which have been used to reduce these side lobes are described.

The parameters which enter into the design of open-wire and coaxial transmission lines are briefly presented. Finally, trends which may tend to reduce the extreme congestion of the short-wave spectrum are mentioned.

THIS paper is limited to a discussion of antennas and transmission-line developments used for long-distance radio communication. The early experiments of Hertz, Marconi, and others were performed with short-wave dipole antennas, including cylindrical reflectors in some cases. Marconi found that as he increased the height and capacity of his antenna, he was able to increase the distance to which he could communicate. Hence, the wavelengths used for long-distance communication gradually increased until by 1914, wavelengths of 10,000 meters (30 kc) or more were in use with large top-loaded antennas of various types [1]. During and after the war, the wavelengths used continued to increase to upwards of 23,000 meters (13 kc).

Long-Wave Transmitting Antennas

In the design of long-wave transmitting antennas, it is of paramount importance to reduce the losses to the greatest extent economically feasible [2], [35].

Blondel pointed out in 1898 that the effect of the earth on a vertical antenna was to produce an image of the antenna, an approximation that was good for the long waves [3]. It can be shown that the radiation resistance R_r of a short vertical antenna is as follows [4], [5], [35]:

$$R_r = 160\pi^2\left(\frac{\alpha h}{\lambda}\right)^2 = 1576\left(\frac{H_{\text{eff.}}}{\lambda}\right)^2.$$

Where α is the form factor which takes into account the distribution of the current in the antenna,

h = height of the antenna in meters,
λ = wavelength in meters,
$H_{\text{eff.}}$ = effective height in meters $=\alpha h$.

$H_{\text{eff.}}$ is on the order of 0.60 to 0.80 of the antenna height h for long-wave top-loaded antennas. It is obvious that the ratio of effective height to wavelength for the long

* Received by the IRE, August 16, 1961.
† RCA Labs., Stony Brook, N. Y.

waves is a very small number, so the radiation resistance is also very small.

In 1917, the General Electric Company in cooperation with the U. S. Navy, installed Alexanderson alternators at the Marconi station at New Brunswick, N. J.; one of the outstanding features of this installation was the great improvement in the radiation efficiency of the antenna by the application of multiple tuning [6]. The Marconi antenna was a flat-top 5000 ft (1530 meters) long, 600 ft (183 meters) wide, supported by 13 guyed tubular masts 400 ft (122 meters) high. At the operating wavelength of 13,600 meters (22 kc) the effective height of the antenna was about 90 meters (295 ft) making the radiation resistance 0.07 ohm. Originally the antenna was tuned at the transmitter end only, and the total resistance was 3.8 ohms, making the radiation efficiency about 1.85 per cent. By multiple tuning the antenna at six points along its length, the ground resistance was reduced to such an extent that the total resistance of the antenna was 0.5 ohm, thereby increasing the efficiency from 1.85 per cent to 14 per cent. With multiple tuning, the total antenna resistance was made up of the following components [6]:

Radiation resistance	0.07 ohm
Tuning coils and insulation	0.10 ohm
Ground resistance	0.33 ohm
Total Resistance	0.50 ohm.

At Rocky Point, N. Y., two multiple-tuned antennas were erected in 1921. Each antenna consisted of twelve wires about 1.25 miles (2 km) long, hung from cross-arms 150 ft (46 meters) long at the top of six self-supporting towers 400 ft (122 meters) high. Because of the low conductivity of the soil at Rocky Point, a very extensive ground system was required. The total length of the buried ground wires was 284 miles (460 km) [7]. It is obvious that to obtain high radiation efficiency, it is necessary to reduce the ground resistance to the greatest extent possible. VLF antennas have been constructed with radiation efficiencies in excess of 50 per cent at 15 kc.

Receiving Antennas for Long Waves

In the early days of wireless communication, the same antenna was usually used alternately for both transmitting and receiving. Later, for duplex operation, a separate receiving antenna was used at a location several miles distance from the transmitting station.

For bidirectional reception, loops were sometimes used. Pickard combined a loop with an open antenna to obtain a unidirectional cardioid pattern. He applied for a patent on this combination on June 10, 1907 [8]. During World War I, several of Pickard's unidirectional antennas were used for transatlantic reception at the Naval Receiving Station, Otter Cliffs, Me. [8].

Ground wires were also used during World War I for directive reception by Taylor, Rogers, and others. Taylor utilized a balancing arrangement between loops and ground wires [9].

Weagant developed a receiving system utilizing large loops spaced several thousand feet apart along the great circle path to the desired transmitting station. His system was based on the theory that certain types of atmospheric disturbances originated from directly overhead. He stated that, "These results showed that static disturbances of the grinders type behaved as though due to heterogeneously polarized, electromagnetic, highly damped waves propagated in a direction perpendicular to the earth's surface" [10].

In an effort to solve a wartime interference problem, Alexanderson devised a receiving system which consisted of insulated wires laid on the ground, extending two miles in each direction from the receiver along the great circle bearing to the desired transmitting station. Means were provided for adjusting the phase and intensity from each wire to produce nulls in the antenna pattern in a direction more or less opposite to the direction of the desired signal. This arrangement was termed the "Barrage Receiver" [11].

The barrage receiver was found to be very effective in reducing atmospheric disturbances coming from the southwest while receiving European signals from the northeast. A barrage receiver was installed at the Naval Radio Station, Otter Cliffs, Me. in 1918. Experience with the long ground wires there indicated that each wire had a unidirectional characteristic, receiving best from the direction in which it was pointing. In the fall of 1919, ground wires several miles long were laid out on Long Island and it was confirmed that the unidirectional characteristic was due to the high attenuation of the ground wire which absorbed reflections from the distant end. This led to the development of the "Wave Antenna" [12]. In the spring of 1920, a two-wire wave antenna was erected on poles with a damping resistance equal to the surge resistance installed at the northeast end. Since the wave antenna was aperiodic, it was usable over a wide band of frequencies without adjustment of any kind. The receiving station at Riverhead, N. Y., was equipped with nine receivers and all long-wave reception was concentrated there.

The wave antenna depends for its operation upon the wave tilt produced by losses in the ground. This produces a horizontal component of the wave passing along the antenna, which induces a voltage in the horizontal wire. The ground parameters are the dielectric constant and the conductance. The wave antenna works best over poorly conducting ground, and will not work at all over a perfect conductor. For poorly conducting ground, the dielectric constant will range from about 6 to 15, which may be neglected in comparison with the conductance term. The wave tilt equation then reduces to the following simple relation [13], [14]:

$$T = 0.235 \times 10^{-10} \sqrt{\frac{f}{\sigma}}$$

where T is the wave tilt, or the ratio between the horizontal and the vertical component of the wave passing over the ground, and

f = frequency in cps,
σ = conductance in emu.

As an example the ground conductance at Riverhead is about 10^{-14} emu, at 20,000 cycles the wave tilt $T=0.033$. Hence, for the 9 mile (14,600 meters) long antenna, the "effective length" is approximately 1580 ft (428 meters).

The wave antenna has also been used as a transmitting antenna at locations where the ground is a poor conductor, which is a handicap to a conventional antenna. While the losses with the wave antenna are high, the directional gain more than makes up for these losses in situations where directional transmission is desired or can be tolerated.

Antennas for Short Waves—10 to 100 Meters

In 1924 Marconi reduced the wavelength of the Poldhu, Cornwall Station to 32 meters and made the important discovery that signals at this wavelength could be transmitted over distances of thousands of miles during daylight [15]. This opened the way for transmitting over great distances with equipment costing a fraction of the cost of the long-wave facilities required for a comparative grade of service. However, the atmospheric noise problem on the long waves was exchanged for deep fading on the short waves. One obvious approach to the fading problem was to transmit a signal of sufficient intensity such that at the receiver, the signal would seldom fade below the noise level.

On the short waves, it is possible to utilize antennas with gains of 20 db or more as compared with a single omnidirectional vertical antenna. As early as 1899, Brown [16] proposed the use of two vertical antennas, separated in space by an appreciable portion of a wavelength and excited at a half-period phase difference, as a means of directional transmission and reception. Foster calculated and plotted the directive pattern for 90 combinations of two antennas with spacings varying from zero to four wavelengths with phase differences varying from zero to 180 degrees. He also plotted 612 patterns for an array of 16 antennas with spacing and phase relations as the parameters [17], [35].

The early highly-directive short-wave antennas generally consisted of arrays of a number of half-wave radiators combined in proper phase for maximum directivity. The first extensive array was the Marconi beam antenna developed by C. S. Franklin. A typical Marconi beam array was supported from crossarms on towers 300 ft (92 meters) high spaced 650 ft (198 meters) apart. Vertical radiators consisting of several stacked vertical half-wave elements were suspended from the cross-arms with an arrangement of vertical wires functioning as a reflector located a multiple of a quarter wavelength behind the radiating elements. The radiators consisted of two or more half-wave elements stacked vertically, separated by nonradiating coils which reversed the phase 180 degrees, thereby causing the antenna elements to radiate in phase. A typical array designed for 22 meters utilized 16 vertical wires consisting of three half-wave elements stacked vertically. The reflector consisted of 53 wires spaced three-fourths of a wavelength behind the radiating array [18].

Several other arrays were developed using similar techniques. Sterba used half-wave balanced nonradiating horizontal sections to reverse the phase for the stacked vertical radiators [19]. Other arrays utilizing wires excited as half-wave or multiple half-wave antennas were developed by Koomans (Dutch), Chireix-Mesny (French), Tannenbaum antenna (Germany), and others. In some cases the radiators were vertical, and, in other cases, the radiators were horizontal [20], [35].

These arrays provided excellent performance, but were relatively expensive and were difficult to maintain in areas having severe icing conditions.

Some experiments were performed using the wave antenna [12] for short-wave reception. The simple horizontal wire was too short to be effective in comparison with the vertical antenna effect. Dr. Peterson overcame this difficulty by coupling horizontal doublets to a balanced transmission line through small capacitors in order to maintain the velocity in the transmission line near the velocity of light. This array was in the form of a fishbone and is often referred to as the "Fishbone Antenna." The fishbone antenna was useful over an extended range of frequencies, and since it was symmetrical with respect to the ground it could conveniently be used with long balanced open-wire transmission lines. These long transmission lines were required for the space diversity system where the antennas were located approximately 1000 ft (300 meters) apart [21].

In England, an array was developed which is similar in appearance to the fishbone antenna, but the antenna elements are approximately a half-wave long on each side of the balanced transmission line. This antenna is known as the HAD antenna (Horizontal Array of Dipoles) [22].

Carter, Hansell and Lindenblad attempted to devise an antenna similar to the fishbone antenna for transmission. They had difficulty in obtaining the proper phase relations for an end-on array, so it ended up as a broadside array, known as the "Type A" antenna [23]. This group then pioneered in the development of antennas utilizing simple long wires as radiators. Several different arrangements were developed, termed "Type B," "Type C," and "Type D" antennas [23].

A long wire in free space has a radiation pattern in the form of two hollow cones 180 degrees apart. The angle which the cones make with respect to the wire depends upon the length of the wire in wavelengths. For example, if the wire is 8 wavelengths long, the maximum radiation in the cones occurs at an angle of 17.5 degrees with respect to the wire. In the Model D antenna, two horizontal wires are utilized in the form of a "V" having an included angle of 35 degrees. The radiated energy in the two cones add along the bisector of the V in the forward and rear sectors, and tend to cancel elsewhere. By erecting another "V" antenna spaced an odd multiple of a quarter wavelength, the radiation in the rearward direction is canceled. Reflection from the ground tends to add the energy from the lower portion of the forward cone to the desired maximum main lobe. The angle which the main lobe makes with respect to the ground depends mainly on the height of the wires above ground. The long-wire type of antenna can be constructed at a small fraction of the cost of an array having similar gain.

Lindenblad proposed constructing a long-wire type of antenna in the shape of a diamond or rhombic. However, it remained for Bruce to propose the use of a damping resistance at the far end of the rhombic, thereby absorbing the energy which would otherwise be reflected to produce standing waves [24], [26]. In effect, the rhombic antenna with the damping resistor becomes a substantially aperiodic traveling-wave antenna readily matched to a transmission line over a relatively wide range of frequencies. Thus the terminated rhombic antenna has become the most widely used short-wave antenna all over the world because of its simplicity, wide bandwidth, and low cost.

An appreciable portion of the power is lost in the terminating resistance. This loss can be minimized by feeding energy back to the input end of the rhombic in the proper phase. However, when this is done, the rhombic is no longer aperiodic and can be operated only at a fixed frequency.

In spite of its many advantages, the rhombic antenna, or any simple long wire antenna, has some disadvantages which become increasingly important as the radio spectrum becomes more and more crowded. Dr. Brueckmann points out that rhombic antennas poorly reject interfering signals from the forward sector over the entire frequency range. At all frequencies, some side lobes exist which are only 6 to 10 db down from the main lobe. The back lobe varies between -6 and -20 db, decreasing gradually in strength with increasing frequency, as do the other side lobes in the rear sector. Brueckmann has described methods for suppressing the side lobes, finally resulting in the TAHA Antenna (Tapered Aperture Horn Antenna). From model measurements, the sidelobe level is -20 db or less over the entire frequency range; the bandwidth is equal to or better than the rhombic, and the sidelobe suppression is superior by a factor of 10 to 15 db. A full scale TAHA horn antenna has an over-all length of approximately 850 ft (259 meters), height 250 ft (75 meters), and width 560 ft (170 meters) [25].

The corner reflector antenna, developed primarily for ionospheric forward scatter, is another effective method for reducing the side lobes on the order of 40 db or more below the main lobe. For this service, which requires relatively high power, it is necessary to reduce the side lobes to prevent interference from back scatter and to or from other VHF services. This is particularly important during periods of high sunspot numbers, since there is occasional long distance VHF transmission due to *F*-layer propagation. The corner reflector antenna main beam can also be slewed to either side by grading the phases of the energy fed to the beam elements. Off-path transmission is sometimes advantageous with VHF ionospheric scatter [27].

In general, the side lobes in any type of multiple-element broadside array can be reduced by radiating the major portion of the energy from the central elements and tapering the energy radiated from the end elements [28], [29].

A method for reducing the undesired side lobes radiated from a rhombic antenna has been described by Christianson. In his arrangement, two or more rhombics are overlapped with the proper spacing and phase relations to reduce the side lobes [30].

Laport and Veldhuis have developed antennas that provide greater sidelobe suppression than conventional rhombic antennas, yet retain the desirable features of the rhombic. Their arrangement consists of two or more rhomboid antennas with a common apex wherein the length of the sides and the apex angles are adjusted to produce zeros in the array pattern to inhibit the growth of the side lobes [31].

Friis and Feldman devised a steerable antenna system consisting of six rhombics spaced end-on. This array had sufficient resolution to separate signals arriving at different vertical angles. Two or more discrete waves, or bundles of waves, may be separated and added in phase, thereby providing a system that might be considered "Vertical Angle Diversity." This arrangement not only increases the SNR, but, more importantly, it simultaneously reduces the distortion due to selective fading [32].

Ehrenspeck and Kearns have investigated another method for reducing the side lobes. Their Type II arrangement consists of a Yagi type array, six wavelengths long, designed to reduce the side lobes to a maximum level of 30 db. This is accomplished by utilizing parasitic elements spaced 0.66 wavelength each side of the main array. The gain of the Type II antenna was 30 per cent above that of a single Yagi of the same length [33].

Isbell has described dipole arrays which make use of the principles of the log periodic antenna design. Antennas are described which provide unidirectional radiation patterns of constant beamwidth and nearly constant input inpedances over any desired bandwidth [34], [35].

Transmission Lines

Standard two-conductor copper telephone lines, several miles long, with special transpositions, were used in connection with broadside VLF wave antennas [12].

Four-wire balanced lines were used with space-diversity antennas [21]. Both applications were for reception.

The writer's earliest knowledge of the application of balanced transmission lines for transmitters was based on a suggestion by Lindenblad about 1921. Lindenblad and the writer used a two-wire transmission line 1400 ft (427 meters) in length to energize a vertical antenna tuned to the amateur wavelength of 200 meters (1500 kc) at Belmar, N. J.

In 1924, two marine transmitters at Marion, Mass., were equipped with open-wire transmission lines to energize antennas which were respectively 300 ft (92 meters) and 600 ft (184 meters) from the transmitting building [36].

When the short waves became widely used, starting about 1925, transmission lines came into universal use for energizing antennas hundreds or even thousands of feet distant from the transmitter. Both balanced open-wire transmission lines and unbalanced coaxial transmission lines have been used extensively. Coaxial transmission lines have been used more widely in Europe than in the United States [20]. Open-wire transmission lines of the balanced type continue to be used extensively in the United States with high-frequency transmitters. Coaxial cable is used in the majority of receiving installations, although carefully balanced four-wire lines continue to be used in some large receiving centers [21].

Great care is required in the installation and maintenance of open-wire transmission lines for transmitters, since relatively slight unbalance may make the transmission line act like a wave antenna, causing it to radiate considerable energy in the direction in which it is pointed. Furthermore, unbalanced excitation of an open-wire line can occur through coupling to an adjacent perfectly balanced line, particularly at the transmitter building where the transmission lines are close together. In cases where special arrangements are made to reduce the intensity of the antenna side lobes, the improvement may be nullified by radiation from an unbalance in the transmission line [25].

In many respects, it is convenient to use open-wire transmission lines. For example, a rhombic or fishbone antenna is inherently balanced with respect to ground. The use of a transformer can be avoided by means of a tapered section to match the impedance of the antenna to a balanced transmission line [37]. Alternatively, broadside antennas have low enough impedance to enable matching the antenna directly to the transmission line [21]. The impedance of an open-wire transmission line is primarily a function of the ratio between the diameter of the conductors and the spacing of the conductors. Likewise, the impedance of a coaxial transmission line depends upon the relative sizes of the inner and outer conductors, and the dielectric constant of the insulation. Accordingly, the impedance of transmission lines can be varied over rather wide limits by design [35], [38].

Future Trends

For many years, the high-frequency spectrum has been close to saturation. Increases in traffic-handling capacity have been made through improved frequency stability, receivers with superior selectivity, multiplex printer operation, etc., but any improvement in spectrum utilization has quickly resulted in additional stations coming on the air, thus continuing the saturation of available channels. The congestion becomes particularly acute during periods of sunspot minimum when the band of frequencies suitable for long-distance high-frequency communication is sometimes squeezed down to a few megacycles by propagation conditions. Obviously, the high-frequency spectrum should be reserved for services that cannot be provided in any other manner [39].

There is great promise for the future in the use of satellites to provide wide bands of frequencies suitable for relaying television and hundreds of channels for telephony and telegraphy. Due to the high initial costs of wide-band relays by means of satellites, it seems apparent that such relays will be used initially only between cities having very large potential traffic requirements. The assignments in the high-frequency spectrum eventually released by satellite relays will be needed by aviation and other mobile services, as well as by recently independent countries which will need frequencies for international communication and broadcasting.

The antennas for use with satellite relays will probably be large parabolic or horn antennas on the ground cooperating with antennas on the satellite having relatively little gain. It will be essential to use antennas on the ground with minimum sidelobes to avoid mutual interference with terrestrial radio services. Most of the methods for suppressing the sidelobes described in this paper are applicable to the higher frequencies, plus other arrangements which are practical for microwave but not for the high-frequency spectrum. The characteristics of antennas for radio relays are covered elsewhere [40].

References

[1] P. S. Carter and H. H. Beverage, "Early history of the antennas and propagation field until the end of World War I, part I—antennas," this issue, p. 679.

[2] N. E. Lindenblad and W. W. Brown, "Main considerations in antenna design," Proc. IRE, vol. 14, pp. 291–323; June, 1926.

[3] A. Blondel, "Sur la théorie des antennes dans la télégraphie sans Fil" presented at Assoc. Franc. Avancement de Sci., Congres de Nantes, 1898, *Compt. Rend.*, pp. 212 *et seq.*; 1898.

[4] J. Zenneck, "Wireless Telegraphy," McGraw-Hill Book Co., Inc., New York, N. Y., p. 170; 1915.

[5] E. A. Laport, "Radio Antenna Engineering," McGraw-Hill Book Co., Inc., New York, N. Y., ch. 1; 1952.

[6] E. F. W. Alexanderson, "Trans-oceanic radio communication," Proc. IRE, vol. 8, pp. 263–285; August, 1920.

[7] E. F. W. Alexanderson, A. E. Reoch and C. H. Taylor, "The electrical plant of trans-ocean telegraphy," *Trans. AIEE*, vol. 42, pp. 707–717; June, 1923.

[8] G. W. Pickard, "Static elimination by directive reception," Proc. IRE, vol. 8, pp. 358–415; October, 1920.

[9] A. H. Taylor, "Long wave reception and the elimination of strays on ground wires," Proc. IRE, vol. 7, pp. 559–583; December, 1919.

[10] R. A. Weagant, "Reception through static and interference," PROC. IRE, vol. 7, pp. 207–256; June, 1919.
[11] E. F. W. Alexanderson, "Simultaneous sending and receiving," PROC. IRE, vol. 7, pp. 363–390; August, 1919.
[12] H. H. Beverage, C. W. Rice and E. W. Kellogg, "The wave antenna—a new type of highly directive antenna," *Trans. AIEE*, vol. 42, pp. 215–266; 1923.
[13] F. E. Terman, "Radio Engineers Handbook," McGraw-Hill Book Co., Inc., New York, N. Y., p. 697; 1943.
[14] K. A. Norton, "The propagation of radio waves over the surface of the earth and in the upper atmosphere," PROC IRE, vol. 25, pp. 1203–1236; September, 1937.
[15] G. Marconi, "Radio communication," PROC. IRE, vol. 16, pp. 40–49; January, 1928. Also E. H. Armstrong, "The spirit of discovery—an appreciation of the work of Marconi," *Elec. Engrg.*, vol. 72, pp. 670–676; August, 1953.
[16] S. G. Brown, Brit. Patent No. 14,449; 1899.
[17] R. M. Foster, "Directive diagrams of antenna arrays," *Bell Sys. Tech. J.*, pp. 292–307; April, 1926.
[18] A. W. Lardner and C. R. Stoner, "Short Wave Wireless Communication," Chapman and Hall, Ltd., London, England; 1936.
[19] E. J. Sterba, "Theoretical and practical aspects of directional transmitting systems," PROC. IRE, vol. 19, pp. 1184–1215; July, 1931.
[20] T. Walmsley, "Beam arrays and transmission lines," *J. IEE, Proc. Wireless Section*, vol. 6, pp. 531–561; March, 1931.
[21] H. H. Beverage and H. O. Peterson, "Diversity receiving system of RCA Communications, Inc., for radiotelegraphy," PROC. IRE, vol. 19, pp. 531–561; April, 1931.
[22] Lapport, *op. cit.*, p. 341.
[23] P. S. Carter, C. W. Hansell and N. E. Lindenblad, "Development of directive transmitting antennas at RCA Communications, Inc." PROC. IRE, vol. 19, pp. 1773–1842; October, 1931.
[24] E. Bruce, "Developments in short wave antennas," PROC. IRE, vol. 19, pp. 1406–1433; August, 1931.
[25] H. Brueckmann, "Suppression of undesired radiation of directional HF antennas and associated feed lines," PROC. IRE, vol. 46, pp. 1510–1533; August, 1958.
[26] S. A. Shelkunoff and H. T. Friis, "Antennas: Theory and Practice," John Wiley and Sons, Inc., New York, N. Y., ch. 14; 1952.
[27] R. C. Kirby, "1958 Critique of VHF ionospheric scatter communication," *Record of Natl. Symp. on Extended Range and Space Communications*, session V, October 6–7, 1958.
[28] C. L. Dolph, "A current distribution for broadside arrays which optimize the relationship between the beamwidth and the side-lobe level," PROC. IRE, vol. 34, pp. 335–348; June, 1946.
[29] R. J. Stegen, "Gain of Tchebycheff arrays," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-8, pp. 629–631; November, 1960.
[30] W. N. Christianson, "Rhombic antenna arrays," *AWA Tech. Rev.*, vol. 8, pp., 361–383; 1947.
[31] E. A. Laport and A. C. Veldhuis, "Improved antennas of the rhombic class," *RCA Rev.*, vol. 21, pp. 117–123; March, 1960.
[32] H. T. Friis and C. B. Feldman, "A multiple unit steerable antenna for short-wave reception," PROC. IRE, vol. 25, pp. 841–917; July, 1937.
[33] H. W. Ehrenspeck and W. L. Kearns, "Two-dimensional endfire array withi ncreaesd gain and side-lobe reduction," 1957 IRE WESCON CONVENTION RECORD, pt. 1, pp. 217–221.
[34] D. E. Isbell, "Log periodic dipole arrays," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-8, pp. 260–267; May, 1960.
[35] Henry Jasik, "Antenna Engineering Handbook," McGraw-Hill Book Co., Inc., New York, N. Y.; 1961.
[36] F. H. Kroger, "The Cape Cod Marine Station of the Radio Corporation of America," PROC. IRE, vol. 12, pp. 243–253; June, 1924.
[37] Laport, *op. cit.*, pp. 422, *et seq.*
[38] Laport, *Ibid.*, ch. 4, pp. 366 *et seq.*
[39] "Radio Spectrum Conservation," Joint Technical Advisory Committee, IRE and RTMA, McGraw-Hill Book Co., Inc., New York, N. Y.; 1952.
[40] E. C. Jordan and R. P. King, "Advances in the field of antennas and propagation since World War II. Part I—antennas," this issue, page 704.

Radio Receivers—Past and Present*

CHRISTOPHER BUFF†, SENIOR MEMBER, IRE

***Summary*—This paper reviews the origin and evolution of some of the more important techniques and instrumentalities of radio receiver development. A completely comprehensive story would require many more pages than are available in this anniversary issue. Highlights of development between the early era of the magnetic detector, or coherer, and the present time are discussed.**

INTRODUCTION

THE fundamental laws governing electromagnetic radiation were first stated by James Clerk Maxwell in 1865. Twenty years later, Heinrich Hertz first detected the presence of electromagnetic radiation from his spark-gap oscillator. His receiver was simply a piece of wire bent into a circle with a ball on one end and a sharp point on the other. When the circle of wire was tuned to the same frequency as the oscillator by adjustment of its circumference, the oscillator spark-gap discharge caused a spark across the air-gap of the receiving circuit.

Subsequent experimenters tried to improve receiver sensitivity by depending on luminous discharges through rarified gas, heat produced by HF currents, contraction of frogs' legs, explosion of oxygen and hydrogen by passage of sparks, mechanical attraction between oppositely charged conductors, demagnetization of iron, and depolarization of electrolytic cells. None of these schemes achieved significant results.

The first real advance in detection was the magnetic detector, or coherer, invented by Prof. Edward Branly. He did not, however, conceive its use for radio reception. The coherer operates on the principle that, while most powdered metals are poor dc conductors, metallic powder becomes conductive when HF current is applied. Shades of semiconductors! The iron filings that were used were continually vibrated to recondition them to their sensitive condition. The evolution of coherers resolved itself mainly into a study of the most sensitive metals and best mechanical tapping device permitting continuous use. Oliver Lodge coined the name coherer and first used the device for detecting

* Received by the IRE, June 6, 1961.
† American Cable and Radio Corporation, New York, N. Y.

radiotelegraph signals. Further improvements were made by Prof. A. S. Popoff and Guglielmo Marconi. For about 10 years, the coherer was the receiving detector used in the Marconi radio stations.

The mechanically unstable coherer gave way to the crystal detector using the semiconductors silicon or galena with a cat's-whisker contact, as invented by Pickard and Dunwoody.

At about the same time as the crystal detector, there appeared the first vacuum-tube-diode detector. Using the Edison Effect as a basis, Prof. Ambrose Fleming patented the Fleming valve in 1904. In its original form, the vacuum-tube diode was not as good a detector as the crystal and was therefore never used in practical radiotelegraph reception. Historically, however, the Fleming valve was greatly important as a forerunner of things to come. De Forest added the grid control, endowing what was previously a simple rectifier with the magical property of amplification. Arnold and Langmuir showed the necessity of a very high vacuum, which, in turn, was made possible by Wolfgang Gaede's molecular pump in 1913. The electron tube thus became the keystone of modern radio communications.

The early history of radio is an international one and it so continues today. Fascinating accounts of the early days were given by Espenschied, Hammond, and others.[1,2] Six decades of radio-receiver art have taken us from the iron-filing coherer to the ruby maser, from mile to multimillion-mile reception, from bandwidths of cycles to megacycles. That a few watts of man-made signals from millions-of-miles in outer space could be picked up on earth testifies to the great improvements embodied in the modern radio receiver. This is a startling example of the homily that a decibel is a decibel, whether costfully gained through more power at the transmitter or by reducing noise bandwidth at the receiver.

The basic obstructions to perfect reception are in the propagation medium, interference, and the receiver itself. Experiments conducted from the depths of the ocean into the cosmos are now revealing the origins of natural noise in relation to the frequencies employed. The International Radio Consultative Committee (CCIR) is contributing greatly in organizing and coordinating the impressive mass of data already obtained.[3]

A common and insidious source of noise is the receiver itself. Receiver noise is explained as originating in the unorganized motion of electrons in the early receiver stages. In vacuum-tube receivers, notable improvement has been achieved with traveling-wave tubes where the noise factor has been reduced from 10 or 12 to 5 db or less at about 3000 Mc. We are now witnessing staggering developments in parametric-amplifier, low-temperature, and atomic-excitation techniques that will extend the range and reliability of electromagnetic-wave reception from the VLF bands through bands probably far above those of visible radiation.

At a time when truly reliable earthbound communications are becoming achieved by improved types of ocean cable, it is apparent that these new developments will pay off in future space communications via orbiting satellites. At the very least, it seems safe to say that radio is here to stay when it comes to interplanetary communications; cable to the moon would be awkward indeed!

What is a radio receiver? A radio receiver might be defined as a device converting to useful form the information contained in the radio waves applied to its input. Basically it is a frequency converter with tuned circuits and amplifiers that specifically adapt it to a particular type of signal reception. In the early days, long-wave transoceanic telegraph signals were transmitted by very-high-powered alternators in the range 10 to 20 kc. Radio folklore tells us that a man on Long Island could simply connect his headphones to a good-sized antenna and read directly the telegraphic messages originating in Europe. This was about as simple a receiver of electromagnetic waves as has ever been achieved, comparing favorably with the first Hertzian detector.

Basic Receiver Circuits

Broadly speaking, all radio receivers might be classified by operating principle as follows:

1) tuned radiofrequency,
2) regnerative,
3) superregenerative, and
4) superheterodyne.

1) Tuned Radio-Frequency (TRF) Receivers

TRF receivers obtain selectivity through circuits tuned to the incoming signal frequency. Through hindsight, it might seem that receiver evolution could have progressed better from the TRF circuit to the superheterodyne, but this was not the case. Although Prof. Reginald Fessenden touched on a basic element by invention of the heterodyne oscillator in 1912, the contemporary invention of the self-oscillating detector, or Ultrauidon, by De Forest immediately focussed attention on this simpler means of increasing sensitivity.

Although the superheterodyne was invented by 1920, many TRF receivers were used in the following decade. This was in great part due to the invention and successful commercial development of the Neutrodyne circuit by Prof. L. A. Hazeltine. The Neutrodyne provided control of the regenerative feedback used by De Forest and others. Prior to this, the regenerative receivers used were unstable; the difficulty of tuning, their squeals,

[1] J. H. Hammond, Jr., and E. S. Purington, "A history of some foundations of modern radio-electronic technology," Proc. IRE, vol. 45, pp. 1191–1208; September, 1957.

[2] L. Espenschied, H. J. Hammond, Jr., and E. S. Purington, "Discussion of 'A history of some foundations of modern radio-electronic technology,'" Proc. IRE, vol. 47, pp. 1253–1268; July, 1959.

[3] Ponte, "Telecommunication in scientific research," *J. Union Internatl. Telecommun.*; June, 1960.

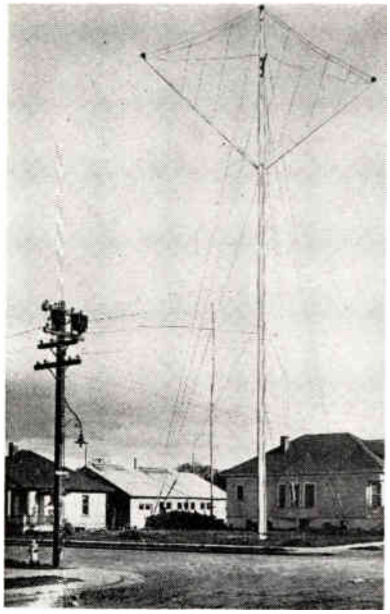


Fig. 1—The corrugated-iron shed behind the cottage was used as a factory and laboratory by the old Federal Telegraph Company in Palo Alto, Calif. It was here that Dr. Lee De Forest developed the first audion amplifier and oscillating detector. (*Courtesy of ITT.*)

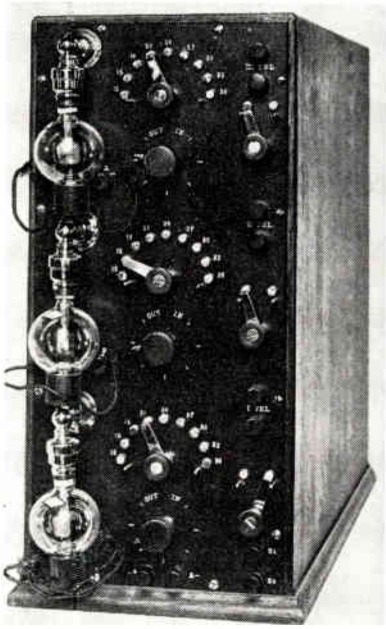


Fig. 2—Three-stage De Forest audion amplifier of the type first built in 1912 by the Federal Telegraph Company. This is the earliest known cascade AF amplifier. Tests made at that time indicated that the amplifier increased the received signal intensity by 120 times. (*Courtesy of ITT.*)

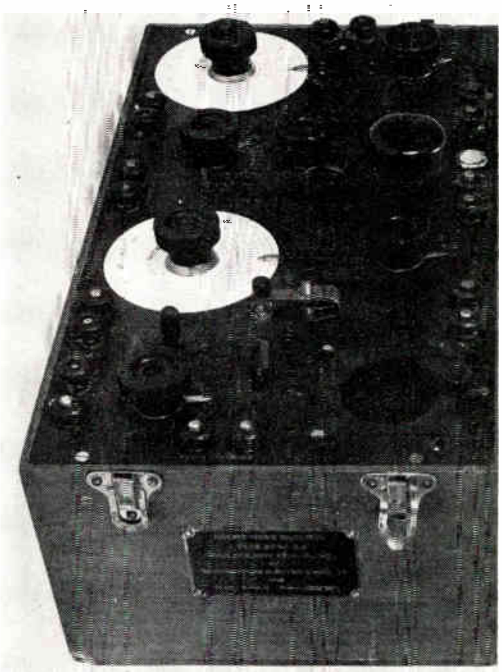


Fig. 3—Front view of the SE-1012A receiver manufactured for the Navy Dept., Bureau of Steam Engineering, by the Westinghouse Electric and Manufacturing Co., in the year 1918. It followed a design originally made by the Radio Test Shop, U. S. Navy Yard, Washington, D. C., and designated as type SE-1012. Signals could be received over wavelengths from 50 to 1100 m. The design included both a crystal-diode and a vacuum-tube detector, either of which could be used as desired. With the vacuum tube in use, the receiver could be operated with either a regenerative or an oscillating detector, and was usable with either damped or undamped signals, to use the terminology of that day. Either a Navy "G" or a Navy "J" tube could be used. (*Courtesy of USNRL.*)

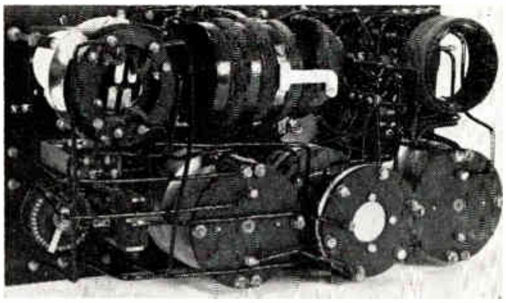


Fig. 4—Rear view of the SE-1012A receiver. (*Courtesy of USNRL.*)

howls, and the interference they caused were soul-racking. But now several stages of RF amplification could be used for isolation and for greater sensitivity; stations always came in at the same dial reading. Two factors contributed greatly to the demise of the TRF receiver during this period. These were the increasing signal congestion in the broadcast band caused by the ever-larger number of stations and the increasing use of what were then termed the short waves; those frequencies lying above 1.5 Mc (200 m) where the amplification and selectivity of the superheterodyne became manifest. By the early 1930's the superheterodyne gained dominance at all frequencies above 500 kc and TRF receivers were on the way out.

2) Regenerative Receivers

In a regenerative receiver, the received voltage is applied to the input of an amplifying stage and a portion of the amplified output of that stage is fed back to the input circuit in the phase that causes an increase in the input. This reduces the effective resistance (raises the q) of the input resonant circuit and thereby provides greater amplification of the signal. If the feedback is increased beyond a certain point sustained oscillations occur. The regenerative receiver is quite sensitive with even a single stage. Four inventors (De Forest, Armstrong, and Langmuir in the United States, and Meissner in Germany) independently discovered the increased receiver sensitivity produced by regenerative feedback around a triode. This feedback circuit made possible, in the heart of a city, the reception of transatlantic signals with comparatively small antennas, a theretofore impossible feat. It was then quickly recognized that the feedback circuit was to be of great commercial importance.

The patent litigation that followed was the most controversial in the annals of electronics. The case dragged on for 20 years. It was finally settled in De Forest's favor by the Supreme Court in 1934. By this time legal expenses had reached huge totals and improved receiver techniques had by then made the regenerative receiver obsolete except for some amateur and marine work.

Regeneration is still used occasionally in q-multiplier circuits for improving tuned-circuit selectivity.

3) Superregenerative Receivers

Superregeneration was discovered by E. H. Armstrong during the defense of his patent case for the regenerative receiver. In a superregenerative receiver, sustained oscillations are squelched by periodic variation of the effective resistance of the input resonant circuit. Oscillations periodically build up in a circuit resonant at the signal frequency. Sustained oscillations are prevented by periodic application to the grid of the superregenerative tube of a signal that damps the oscillations. The quenching frequency is usually between 20,000 and 100,000 cps. The superregenerative detector, because of its broad tuning and considerable sensitivity, was practical and popular earlier when unstable modulated oscillators were used as transmitters at frequencies above 30 Mc.

More recently, superregeneration has been applied to parametric amplifier detectors. Their relatively broad bandwidth with stable and high gain are desirable attributes in S-band experiments. The theory of superregeneration has generally not been well understood in the past and is only now being clarified.

4) Superheterodyne Receivers

In superheterodyne receivers, the incoming signal is combined with a local oscillator voltage to derive an IF signal, marked by the information-bearing modulation, which is then amplified and detected to reproduce the information. Underlying the superheterodyne principle is the basic modulation or mixing process that generates the sum-and-difference sideband pairs that make frequency translation possible.

Prof. Fessenden, in his search for an improved receiver, invented the heterodyne system in 1912. Previous receivers had merely acted as valves, detecting by turning a direct current on and off in amounts proportional to the received signal. In contrast, the heterodyne system operated through the joint action of the received signal and a local wave generated at the receiving station. Combination of these two alternating currents resulted in an audio beat-note, the difference frequency between the two waves. Although Fessenden's local oscillator was an arc source, very bulky and troublesome, it was nevertheless the forerunner of superheterodyne and single-sideband reception.

The next advance in double-detection technique involved amplification of the beat-note or intermediate frequency. Several parallel developments took place in the United States and in Europe. It is difficult to name an inventor since the superheterodyne system as a basic idea seemed to appear from several sources at about the same time. The works of J. H. Hammond, A. Meissner, Lucian Levy, E. F. Alexanderson, and E. H. Armstrong stand out. Armstrong fully appreciated the problem and obtained a patent in 1920 that was of major importance in the practical application of the superheterodyne system. Superheterodyne technique today has universal application in transmission as well as reception.

Superheterodyne receiver technique derives its main advantage in that almost all the receiver selectivity and most of the receiver gain can be obtained in stable IF-amplifier stages. Recently, extremely small and efficient megacycle-range quartz-crystal IF band-pass filters have been developed. These improved selective elements greatly reduce the number of conventional tuned circuits required in IF amplifiers. Where triple IF conversion was sometimes formerly required, it is now feasible to use a single conversion. Indeed, the carrier frequencies of yesterday have become the intermediate frequencies of today.

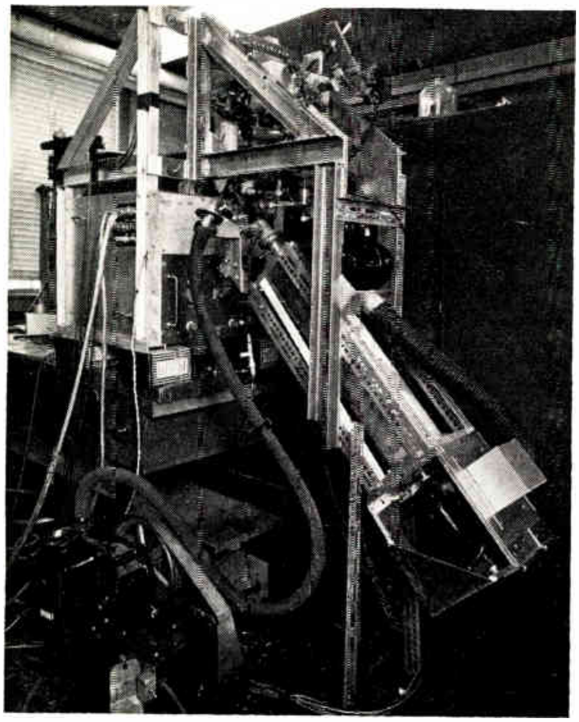


Fig. 5—9500-Mc radiometer with maser preamplifier shown in laboratory before mounting on the 50-foot radiotelescope at the USNRL. The maser was designed and constructed by the Columbia University Radiation Laboratory. (*Courtesy of USNRL.*)

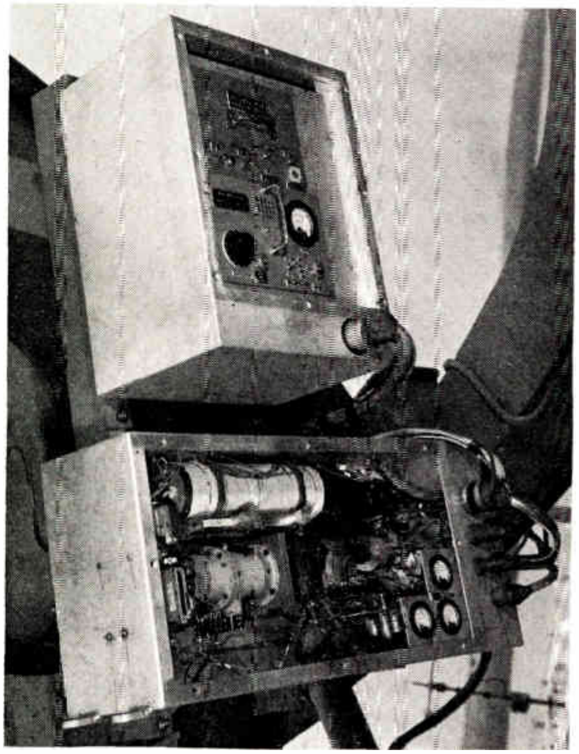


Fig. 6—35-kMc radio receiver for mounting on back of 10-foot radiotelescope at the USNRL. The waveguide from the focal point of the reflector can be seen entering the receiver on the left of the assembly. (*Courtesy of USNRL.*)

In communications receivers there is still, however, a serious fault. This is the wide-open door at the input to a noise bandwidth that is usually many times greater than that actually required to accommodate the desired information. What we must develop is an adjustable "window" between the antenna and the first RF stage that can be tuned over a wide frequency range and widened or narrowed as required to accept only the information content represented by the signal bandwidth.

Although present receivers have narrow IF bandwidth, much unnecessary noise and cross-modulation can result from the wide-open front end.

It is suggested that development of sharp, tunable receiver-input filters would hold great possibilities for improving radio-receiver performance.

Diversity Reception

As a result of some elegant experimental work by radio amateurs in this country and in Europe, the short-wave band was opened for all in the early 1920's. The variable nature of HF propagation was quickly recognized and led to many ingenious innovations in receiving technique. On the lower frequencies, below 500 kc, improved communications were attained through use of higher transmitter power and larger antennas. Because of the turbulence and abrupt changes encountered in the HF medium, special attention had to be given to improved means of reception as it became apparent that transmitter power increases were not alone sufficient.

Among the significant techniques developed, one of the most important is diversity reception, wherein a considerable improvement is obtained due to the statistical independence in the fading characteristics of two or more paths. Diversity reception is basic, improving the reception of any type of modulation at any frequency. H. H. Beverage, H. O. Peterson, and J. B. Moore described and developed a triple-space-diversity system for on-off telegraph reception around 1928.[4] This employed three antennas spaced about 1000 feet apart. The rectified outputs of three separate receivers were combined across a common load resistor and the voltage across the resistor keyed a local tone generator. As long as the voltage was above a certain minimum from any receiver, a properly keyed tone signal was reproduced. Prior to this, separate "raw" signals from two receivers were connected to separate earphones on the operator's head. Through his mental processes a diversity improvement resulted. This method was dubbed by some, "Scotch Diversity." Statistics on the improvement obtained or on the longevity of the operators do not seem to be available.

Also about 1927 or 1928, A. de Haas of the Netherlands Post, Telephones, and Telegraphs Administration described a complete diversity system used in the

[4] H. H. Beverage and H. O. Peterson, "Diversity receiving system of R.C.A. Communications Inc., for radiotelegraphy," Proc. IRE, vol. 19, pp. 531–561; April, 1931.

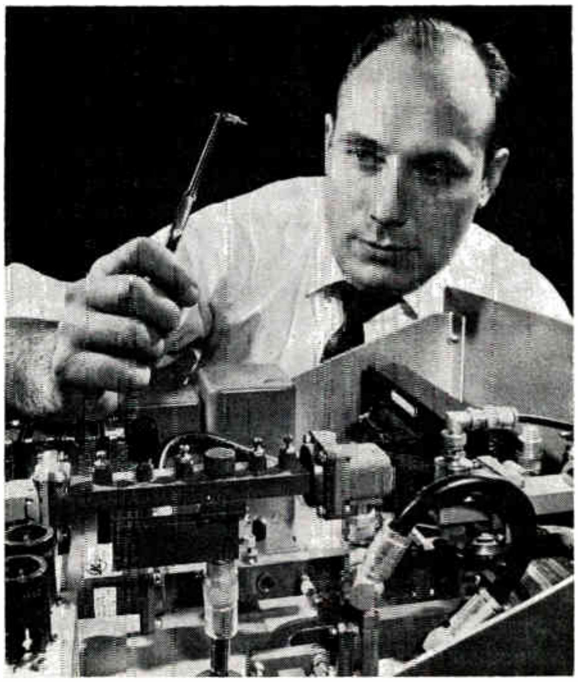


Fig. 7—Front end of a Tacan ground-beacon receiver employing a parametric amplifier. The parametric diode held in the tweezers is the heart of the amplifier. Lower-sideband up-converter is used to receive signals in the 1025–1150-Mc range. The pump frequency is 10 kMc. Another oscillator beats the lower sideband down to the 63-Mc IF. Noise figure of the amplifier is 2.5 db. (*Courtesy of ITT.*)

Dutch East Indies to improve reception of the Kootwijk, Holland, transmitters.[5] De Haas had all the essentials of an equal-gain postdetection combiner. This is another instance where a basic idea germinated at two sources at the same time. Confronted by similar problems, different people sometimes arrive at the same solution!

In recent years more has probably been written about space, frequency, and polarization diversity and combining methods than about any other receiving technique. With notable exceptions, much of this concerns refinement rather than true innovation, plus mathematical studies of the statistical performance of the various combining methods. The excellent analysis of D. G. Brennan is an example of this work.[6]

L. R. Kahn's important analysis showed that the optimum diversity advantage for frequency-modulation and frequency-shift reception could be obtained only by prelimiter evaluation of the signals.[7] Individual diversity channel contribution is then made proportional to the square of the signal level at the limiter input. Actual postdetector combining is by suitable discriminator cross-connections. This method is widely used on HF for individual telegraph channel combining and also at VHF and higher where the principle is used in baseband combining systems.

Another and more elegant solution of the diversity problem is that provided by F. J. Altman and W. Sichak in their predetector IF phase combiner.[8] Essentially, this system phases the signals in two diversity channels by phase control of one of the receiver local oscillators. This predetector combining system provides simplified circuits and improved distortion ratio as compared to baseband postdetector combiners using out-of-band noise as a criterion. This scheme has been mainly used in over-the-horizon scatter systems.

Assuming proper design, there seems to be very little practical difference in performance of the several basic diversity combining systems. However, an adequate diversity arrangement of some type is considered a necessity for reliability in any radio system involving even a minor degree of multipath propagation. This is especially true for telegraph and data transmission which inherently can tolerate less signal mutilation than voice transmission.

Frequency-Shift Telegraph Reception

By increasing the reliability of HF radio circuits in the international record traffic services, frequency-shift telegraphy has certainly become a classic receiving method deserving some mention. Its origin is somewhat nebulous although credit for the underlying modulation principles properly belong to E. H. Armstrong, the father of FM.

As modulation technique, frequency-shift keying (FSK) dated back to the Poulsen-arc days of the old Federal Telegraph Company in Palo Alto, Calif., circa 1915. Because their high-power arcs could not be on-off keyed, the transmitted frequency was instead shifted a few hundred cycles by means of resonant circuits switched in and out by the telegraph key. At the receiver, the key-up or spacing frequency was disregarded in the detection process. Obviously, there was no improvement obtained as far as reception was concerned.

In 1927 Armstrong is reported to have proposed an FSK system for VLF telegraph circuits to combat atmospherics. After this there seems to have been a lapse of several years. The early 1930's saw the utilization of subcarrier FM (SCFM) methods by Press Wireless, Inc., for receiving radiophotos from the Byrd south-polar expedition. The improvements in picture quality as compared to the former AM method were immediately apparent. By the late 1930's the significance of frequency-shift technique for teleprinter as well as facsimile purposes was widely recognized. The work of J. A. Smale in England on two-tone keying around this time is worthy of note. With the continuing improvement in transmitter and receiver oscillator stability,

[5] B. B. Barrow, "Translation of a historic paper on diversity reception," Proc. IRE (Correspondence), vol. 49, pp. 367–369; January, 1961.

[6] D. G. Brennan, "Linear diversity combining techniques," Proc. IRE, vol. 47, pp. 1075–1102; June, 1959.

[7] L. R. Kahn, "Ratio squarer," Proc. IRE (Correspondence), vol. 42, p. 1704; November, 1954.

[8] F. J. Altman and W. Sichak, "A simplified diversity communication system for beyond-the-horizon links," IRE Trans. on Communications Systems, vol. CS-4, pp. 50–55; March, 1956.

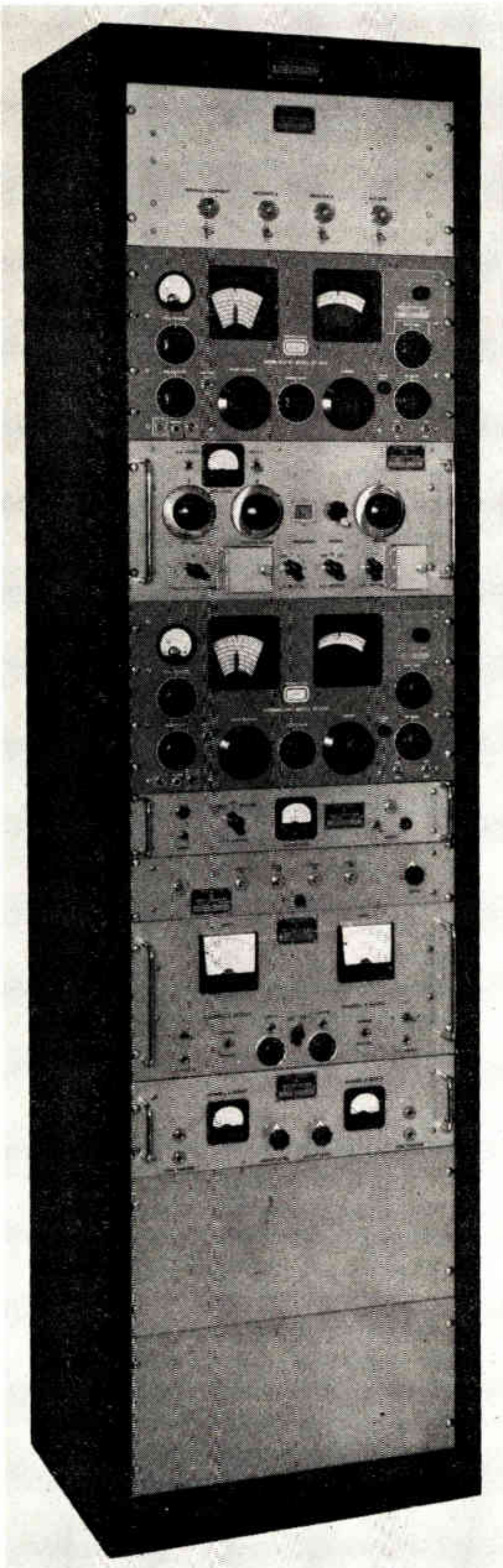


Fig. 8—Dual-diversity FSK Twinplex radio receiving terminal, as used in international point-to-point HF communications, features automatic frequency control and narrow-band Twinplex reception using 200-cps shift intervals. Eight 50-baud channels can be received with a total frequency-shift spread of 600 c through the combination of synchronized time-division multiplex and narrow-band twinplex techniques. (*Courtesy of Mackay Radio and Telegraph Company.*)

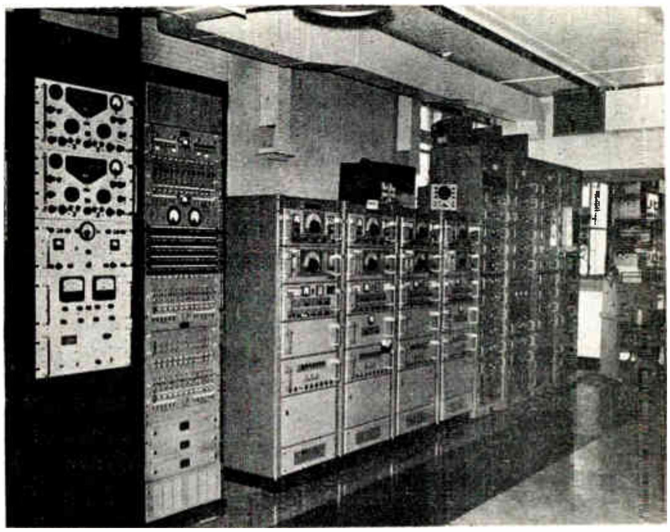


Fig. 9—Two types of ISB radiotelegraph receivers for HF are shown. At the left is a new type of dual diversity ISB receiver working into a twelve-channel transistorized FM voice-frequency-telegraph terminal directly adjacent. The lower bays in the center are older types of ISB receivers, two bays for dual-diversity, operating into tube-type FM voice-frequency-telegraph terminals at the far right. The differences in floor space requirements are readily apparent. (*Courtesy of Mackay Radio and Telegraph Company.*)

direct RF-carrier shift became a practical reality around 1941. As is well known, the ensuing war gave great impetus to FSK because of the wide application of teleprinters to radio circuits, which, in turn, was partly caused by the shortage of skilled Morse-code operators.

Essentially, the greatest single improvement produced in FSK reception is obtained by the action of a high-speed transient-free amplitude limiter inserted between the radio receiver proper and the telegraph-signal detector. The radio signals can then assume a very wide range of fading levels without altering the mark-to-space ratio of the telegraph signal. It is this effect more than the noise-reducing properties of FM that has made automatic (teleprinter) reception a practicality.

Over-all performance during fades is impaired by circuit elements, including band-pass filters, that have a poor transient response. These elements may, of necessity, precede the limiter for selectivity purposes or may, through poor design, be unintentionally incorporated in the limiter. Because narrow-band filters inherently have poorer transient response, systems using such filters, for example those for 85-cps shift, must place greater reliance on optimum diversity-path combination than those using relatively wide prelimiter band-pass filters. H. B. Law in England and others have developed FSK receivers that dispense with predetection limiting and through an "assessor" circuit evaluate both mark and space frequencies to obtain a frequency diversity effect.[9] Improvements are claimed, but it can probably be shown that a narrow-band independent-

[9] "Six papers on radiotelegraphy and fading," reprint from the *Proc. IEE*, vol. 104, pt. B, no. 14; March, 1957

sideband (ISB) multitone system with conventional limiting and optimally executed ratio-squared diversity combination performs at least as well and with better bandwidth utilization.

The entire subject of frequency-shift telegraphy, particularly reception, has been rather continuously reexamined during the past fifteen years; the practicing communications engineer may now feel the need for a really basic evaluation of the various FSK systems. Until now, there have been under study phase-shift, multiple-frequency-shift, four-frequency diplex (Twinplex), and coherent detection systems. Probably the most widely used variations of FSK are the Twinplex and ISB multitone methods. Significant reductions in transmitted bandwidth have been achieved through improved detection schemes at the receiver. Frequency-shift deviation at HF has been reduced from 850 to as low as 60 cps in ISB multitone systems. Twinplex bandwidth, by frequency-multiplication techniques at the receiver, has been reduced from 1200 to 600 cps while retaining its speed of 200 bauds per channel. Kineplex, which may be thought of as a phase-shifted Twinplex system, reduces the deviation to but a fraction of a cycle.

After much backing-and-filling, it has been established that the actual frequency deviation is not as all-important as it was once thought to be. The bandwidth employed can mainly be considered to be a function of the signaling speed and the deviation adjusted accordingly within this prescribed bandwidth. Through the use of extremely stable oscillators coupled with automatic frequency control and optimum diversity combination at the receivers, a 50-baud HF teleprinter circuit need occupy only one-tenth the bandwidth required for early FSK circuits.

Conclusion

The advent of space-probing efforts has overtaxed radio-receiving techniques once again. Until that day when high-powered ground transmitters become a reality and higher-gain antennas are available at the distant end, be it a spaceship or another planet, improvements must seem to come mainly from improved receivers. We find used in these receivers most of the basic developments of the past 50 years, such as the superheterodyne, sharp IF filtering, diversity reception, automatic gain control, FSK, and automatic frequency control. The major innovations in space-age receivers are mainly those relating to improvement of the front-end noise figure: the traveling-wave tubes, parametric amplifiers, and masers.

In earth-bound communications, receiver front-end technique has, it seems, always lagged behind. Now a supreme effort must be made to eliminate internal receiver noise, for we seek to use signals that are but flickering ghosts. No matter how low the noise figure of the input amplifier, however, the accepted bandwidths, being extremely wide compared with the desired signal, leave something to be desired. Interference from other signal sources, for example, cannot be wholly negated by IF filtering. What can be hoped for from our filter network discipline toward developing tunable filters that will eliminate everything but a desired signal right at the front end of the receiver? Radio astronomy is purely a receiving technique. The signal sources are not man-made. We have to do the best we can with these cosmic "noises." The pioneering work of Karl Jansky led the way to the exploration of space. He was followed by many others, including Grote Reber and John Kraus. At present, major efforts in radio astronomy are being carried forward in many countries. The magnitude and world-wide scope of the work being done on masers and parametric amplifiers can best be evaluated by referring to the literature.[10]

It has been half-seriously observed by some practical earth communicators that had it not been for the untimely intervention of HF radio, we would have been blessed by a highly-developed intercontinental cable network much sooner. Before the advent of space technology, this attitude may have had a degree of validity, albeit small. We realize now, however, the neatness with which the lower-frequency techniques are assuming their place in the pattern of requirements for space communications. Because of the past research into the vagaries of radio propagation and development of means to combat these, the new problems before us become much more tractable. The basic tools are available; still further refinement is needed.

Bibliography

[1] F. J. Mann, "Federal Telephone and Radio Corporation—a historical review: 1909–1946," *Elec. Commun.*, vol. 23, pp. 377–403; December, 1946.

[2] H. Mosler, "Einfuhrung in die moderne drahtlose Telegraphie und ihre praktische Verwendung," F. Vieweg und Sohn, Braunschweig, Germany; 1920. (In German.)

[3] McIlwain and Brainerd, "High Frequency Alternating Current," John Wiley and Sons, Inc., New York, N. Y.; 1931.

[4] W. R. MacLaurin, "Invention and Innovation in the Radio Industry," The Macmillan Company, New York, N. Y.; 1949.

[10] E. Mount and B. Begg, "Parametric Devices and Masers, An Annotated Bibliography," ITT Labs., Nutley, N. J., Biblog. Ser. No. 105; October, 1959.

The Compatible Technologies of Wire and Radio*

IVAN S. COGGESHALL†, FELLOW, IRE

Summary—**During the 50-year history of the Institute of Radio Engineers the respective technologies of wire and submarine cable telegraphy and telephony, on the one hand, and mobile and point-to-point radio, broadcasting, and television, on the other, have been commingled in an ever increasing degree. Initial contributions leading to the mutuality of the wire and radio branches of applied science came from pioneers who worked in both media. Subsequent discoveries were eagerly put to work in both, prompted by users' requirements for direct, end-to-end service. Opportunities for interchange of general communication engineering ideas and experience through the IRE were purposely fostered by its directorate from 1912 to 1962. An impressive list of exchanged benefits is attributable in part to the Institute's hospitable climate. An Appendix lists sources of information on the present status of wire communication, technical and organizational.**

Introduction

THROUGHOUT its 50-year history, the IRE has provided the men, terrain, and climate for ideas to undergo cross-pollination between the fields of wire and radio communication. Fenced-in frequencies protected the purity of their separate strains when the Institute was founded in 1912. But in the ensuing years, through the IRE's meetings, its adaptive organization, and the pages of its PROCEEDINGS, engineers, busy as bees in their deep probings of a flowering science, not only have caused the distinctions to be erased which classified the early genera of wire and wireless, but within the genera have transmutted their species: telegraphy, telephony, television, and data processing and control.

In 1962, raw bandwidth is being sold for end-use; and end-users, beholden to the miracle of their modulations only by the buttons they push, know not the successive channeling of their signals through courses of wire and space nor the modes nor the ducts whereof.

In 1912, telegraphy and telephony were clearly differentiated. In the 1920's and 1930's a first integration was occasioned by the vacuum tube; the two arts adopted identical techniques of ac modulating higher frequencies, in the form of code, voice, and photographs. A second integration followed World War II. The miniscular subdivision of time brought about by multiplexing and television scanning, the entry of binary dc computers into the field of megacycle communication, the time refinements of pulsed radar, and the realization that even speech and music could be converted into yes-no code by a sampling process, all conspired to reveal the digital pulse as the ultimate communication building block of Nature. Once the informational bit began to preoccupy engineering thought, telegraphy, telephony, and television could never again appear otherwise than as separate mutations of a single species.

So, after 50 years of the IRE, sight and sound, ideas and images, languages of men and stimuli of machines, together ride the carriers, stack upon stack, and are transferred willy-nilly between copper and air. Collectively they may find themselves poured through the same sieve and transmitted 0,1.

A Pioneering Spirit

Contemporaneously with and even antedating Marconi, there were a few physicists, acquainted with the mathematical findings of Faraday and Maxwell, the radiation experiments of Hertz, and the demonstrations of sound-resonance by Helmholtz, who conceived a pattern of unity in R, L, and C bridging the great gulf fixed between the frequencies of wire and wireless transmission.

Three of them—Stone, Kennelly, and Pupin—became the fourth, fifth, and sixth Presidents of the IRE. To their mathematical minds the wide separation in frequencies between radio and audio (using the terminology of today) offered no obstacles to a unified approach and solutions before the turn of the century; regardless of the fact that, at that time, radio was strictly a laboratory experiment holding great interest, while audio was a matter of field importance involving cross-country configurations.

Stone worked in the research laboratory of the American Bell Telephone Company, Boston, from 1890 to 1900. His early work on circuit resonance was with voice-frequencies on tone channels. His inductive and capacitative components in resonant circuits were substituted for the vibrating reeds unsuccessfully used by Bell in his search for a harmonic telegraph, out of which emerged the telephone. In 1894, Stone was using a voice-diaphragm-operated mechanical modulation of a high-frequency dc arc to generate a voice-bearing composite transmission signal over wires.

Pupin was adjunct Professor of Mathematical Physics at Columbia University, where he assembled units of L and C in circuits resonant to the harmonics of ac power. With them he demonstrated in 1893 what he named "electrical tuning" to his former professor, Helmholtz. When he applied for a patent that year, he found himself in interference with Stone.

* Received by the IRE, June 19, 1961.
† 670 Maywood Ave., Maywood, N. J.

Stone and Pupin were now working so close to the radio borderland that when the news broke in 1895 that Marconi was sending Morse messages by wireless over field distances, both were swept up in an endeavor to make additional contributions to the success of the new mode. Stone's researches led him to his loose-coupled "4-circuits" patent of 1900, which went into litigation against claims of Marconi, to whose American company Pupin sold all his wireless inventions outright in 1903.

In 1896, Pupin set about solving the problems of loading telephone lines inductively to decrease transmission losses. He deduced, from a shot-loaded, weightless, stretched-string mathematical model after La Grange, that on telephone spaced pairs on pole-lines and in multipair cable, the optimum relationship should be: $L>C>R$; hence that several lumped loadings should appear in each wavelength representing a frequency of interest in transmission. By attention to this detail he succeeded where predecessors had failed.

The 5th IRE President-to-be, A. E. Kennelly, worked as a submarine cable operator and ship's electrician with the Eastern Telegraph Company, London, before he came to the United States and became principal electrical assistant to Edison. From 1902 on, he was Professor of Electrical Engineering at Harvard and Massachusetts Institute of Technology. His nonradio specialties of value to radio included the mathematical treatment of oscillatory phenomena, acoustics, and the introduction of hyperbolic functions applied to ac.

Of such wide-ranging pioneers, was the IRE in its formative years.

A Benign Climate

From its beginnings, the Institute's terrain was what its members would have it. They expanded its field of influence year by year as accumulating knowledge made it expedient to cultivate new territories. But the unvarying, benign, hospitable climate with which it surrounded those engaged in "arts allied with radio," was purposefully created by the men who firmly directed its continuing policies and operations.

The founding fathers were radio men of broadest gauge. That they did not look askance at engineers in allied fields is witnessed from Volume 1, Part 1, paper No. 1 in the 1913 Proceedings; discussion:

> Pupin: I regard the one hundred thousand cycle alternators which General Electric Company produced for Mr. Fessenden, to deliver two kilowatts, as a triumph of mechanical and engineering skill.
>
> Goldsmith (*Editor*): . . . an excellent illustration of the courage needed to transfer ideas from one field of research to allied fields.

After the first three IRE Presidents had been drawn from eminently highly-placed wireless men, the next three, as related above, were chosen with full knowledge that before their conversion they had offered sacrifices to Moloch. Wire communication could not have been more highly honored. On the other hand, the versatility of such leaders to shift readily from laboratory to radio to open-wire and cable lines, and to find their way by the light of mathematical physics, cannot but have had an intellectually buoyant effect in a new organization which more than anything else wished to demonstrate engineering maturity.

Goldsmith led the Proceedings to unassailably high standards. They were read and the IRE meetings were regularly attended by many wire engineers who joined because of the "pull" of radio; or because of the stimulation, or direct usefulness in their work, of what they picked up; or because they liked the other people who were at the Institute or Section meetings or running the show.

The IRE poured the honor of scores of Fellowships upon men whose technical achievements spanned the wire and radio arts, bestowed upon 21 of them its most preciously held medals, entrusted 20 of them to help guide the Institute as Managers and Directors, and elected 9 of them Presidents. In the IRE's 50-year history, there was no period of five years when either the Medal of Honor or the Morris Liebman Memorial Prize, or both, or a number of these awards, did not go to recipients the value of whose achievements was shared by wire and radio communication. There were only four years in the 50 when the Board of Managers or Directors did not have in its membership at least one engineer whose reputation was established in both arts; and in 24 of the 50 years there were from three to five such men on the Board.

The mid-century Professional Group movement within the IRE supplied the finishing touches of radio's hospitality to its allied arts. Wire-oriented members were quick to avail themselves of alliance with radio men in groupings that both helped to create. At this writing the process of mutation goes on in the fields of microwave theory and techniques, vehicular communication, information and circuit theory, communication systems, and as many more areas of Professional Group interest.

A Lush Terrain

Radiocommunication had no need to invent either telegraphy or telephony. Marconi received from Morse a heritage which needed only adaptations, some of them difficult. By the same token, to Fessenden, Poulsen, De Forest, and other pioneers of radiotelephony around 1900–1907, Bell's invention had sprung full-blown from the head of Zeus.

By 1912, radiotelegraphy was well established as a maritime adjunct and was making valiant but feeble inroads on the traffic of wire telegraphs and the submarine cable. De Forest had invented, merchandised, but not tamed the vacuum tube. Within the realm of technology in the life of the people, his invention was to vie with the internal combustion engine as the most important socio-cultural, civilizing influence of the 20th Century. In expanding its utility, the IRE was to play a vital part. But before 1912 it had not come into its own.

In 1912, radio was suffering from want of amplification. Only the Rutherford-Marconi magnetic detector and Fessenden's arc-oscillated heterodyne receiver had received any at all, and in the latter, most favorable, case only the beat-energy to which a weak signal could contribute a component. Wire telegraphy was more fortunate, in dealing with frequencies so low that steady-state transmissions could provide enough received power to actuate relays; as a result the repeater art was well established. Submarine telegraphy had an amplifier of sorts in the hot-wire relay, but its gain was small and the technique applicable only to slow, galvanometer frequencies. But in long-distance telephony, power increments at repeaters were limited to those produced by microphones in battery circuits. Bell's harmonic telegraph had failed and Stone's carrier telephone channels had languished—either the received currents were too weak to vibrate the reeds or to survive the attenuation of tuned circuits.

Nineteen-twelve, the year of the IRE's birth, was in at least four other respects a significant wire-radio landmark. Arnold of Bell and Langmuir of General Electric pumped life into De Forest's vacuum tube. De Forest sold his wire rights in it to the telephone company and he began using tube circuits in cascade. He and others learned how to control them as oscillators and with feedback or regeneration.

In retrospect, it was "the year the dam broke." Released were the pent-up frustrations which hitherto had limited radio to the trickle of energy rectified from its antennas, and wires and cables to bandwidths and distances achievable within the strictures of signal attenuation. New found flood waters of local power distantly valved would thereafter course through prepared channels, bringing fresh life to a lush terrain.

By the time of the broadcasting and shortwave achievements of the 1920's, remote pick-ups and end-to-end sequences of wire and radio preoccupied their engineers domestically and internationally in shared problems and shared triumphs of quality improvements, traffic bandwidth stretching, and distance records. It was like the confluence of two streams of technology, both fed from headwaters in Europe and North America, mingling before being joined downstream by electronic tributaries rising in noncommunication sources. This figure of speech can be documented:

Wire Telegraphy's Assistance to Radio

Basic were the audible Morse and Continental code systems. Siphon recorders and undulators were adopted by radio from landline ocean cable arts. High-speed telegraphy, much favored in radio, stemmed from wire Wheatstone.

Wire telegraphy furnished the high-speed mechanical relay. Thyratrons, as substitutes for relays, were first used on telegraph cables. Repeaters, both telegraph and telephone, were developed for wire use. Wire rotary repeaters first utilized the principle of regeneration of unit signal pulses.

The uniform 5-unit printer code and its 7-unit start-stop version came from the wires, as indeed did machine telegraphy generally with its perforators, tape transmitters, page and tape printers and reperforators. Storage of signals, now universal in radio and computers, originally was employed on wires in holes-in-tape embodiment; so were machine encryption, multiple-storage, sequential, flip-flop, and programmed release, and demand read-out. The idea of interrupting and holding the flow of transmission enroute for the interpolation of pulses, characters, printer-control and switching functions, and stored information, came from wire telegraph and cable arts.

Simultaneous two-way operation, synchronism and isochronism of sending and receiving apparatus, and facsimile transmission and telephotography were worked out on landlines, together with the Nipkow-disk application of scanning as an unfolder of time-spatial relationships utilized in television. Time-division multiplexing by the interlacing of channels was a wire-art extension of synchronism. The Varioplex and Tasi systems of switching, making available to other users the line capacity of temporarily idle channels, were contributions of landline telegraphy and cable telephony, respectively.

Telegraph wire systems produced teletypewriter exchanges, reperforator-switching offices, automatically switched telex exchanges (from Europe), and customers' private wire systems.

From the telegraph wires and cables came initial steps in circuit network theory, in their artificial lines and employment of sending-end curbing and receiving-end shaping of signals; and the selenium cell, forerunner of the phototube.

Wire-Telephony's Assistance to Radio

Basic were the microphone, the telephone receiver, and amplitude modulation. Plug-and-jack circuits and multicontact Craft relays provided the essentials of manual and remote switching. Wire telephony also furnished the essentials of machine switching, dial controlled.

The wire-telephone art pioneered the frequency fundamentals of speech analysis, stereophonic sound, the theory of sidebands, a succession of "laws" in information theory, and the decibel system of measurement.

Out of telephone carrier technics came filtering, suppressed-carrier, and single-sideband practices; also speech inversion, compression, and scrambling; a number of types of pulse modulation, and practical negative-feedback amplifiers.

Wire telephony is credited with pioneer use of the "hard" vacuum tube. From telephone sources came nonlinear circuit elements: the varistor, thermistor, and transistor—the latter, in its versatility, constituting something of a writing off of indebtedness to radio for the vacuum tube and its circuitry.

The telephone system was first in communication to adapt the relay and vacuum tube to digital arithmetic of radix 2, and to the logic processes of Boolean algebra. Coaxial cables and waveguides, useful in radio installations, also came from wire telephone art.

Radio's Assistance to the Wire Arts

By far the most valuable single gift in the entire two-way exchange was De Forest's vacuum tube with control-grid and B-battery. Radio also contributed C-battery grid bias, tubes in cascade, the stabilizing screen-grid and other multielement tubes. The older and less spectacular radio crystal- and VT-diodes also served the wire arts.

Radio adopted, improved upon, and handed back to the wires better components like microphones and loudspeakers; or speeded them up, as in the case of telegraph recorders. Public demand, created by the wide acceptance of higher-fidelity radio broadcasting, caused the broadening of the commercially acceptable speech band; and radio produced frequency modulation. The radio public also brought into being the ac-dc power pack and the ac heated cathode. Automatic gain controls evolved from radio's meeting its AM transmission problems; piezoelectric frequency stabilizers were originally used to prevent radio transmitter drift. Error detection and correction in radiotelegraphy also were inventions to overcome transmission vagaries. All have been adopted to advantage by the wire companies.

The heterodyne reached perfection in radio; the earliest form of magnetic amplifier appeared there. Proximity fuses, rockets and satellites spurred radio to initial supremacy in printed circuits, modular construction, and microminiaturization.

It is not far-fetched to classify the IRE as one of radio's gifts to wire communication. Certainly the list under this subheading would be much longer had the fraternizing of engineers under the IRE not been so complete that inventions were adopted as hybrids on wire lines and in radio simultaneously and thus lost their identity as "gifts" of one to the other. American innovations were the more quickly put to use in tandem wire and radio circuit sections because wire, radio, and, later, submarine cables came together in one principal corporate organization. British inventions, particularly in globe-girdling shortwave telegraphy, had an advantage over American developments by virtue of corporate unity of cables and wireless, to which the Post Office landlines later were added.

Under such consolidations, determination of the realm of first application is of secondary interest and leads to omitting in the listing certain major items, like television and color TV, where landline demonstrations may have preceded first use on radio (or vice versa) but where neither branch of the art could thereby be held to be in obligation to the other.

Conclusion

It would be trite to say the trend outlined in the foregoing will continue. The thing has been done. More indicative of the future is the close involvement of communication *in toto*, as a transmitter of electrical intelligence, with electronic, nonlanguage, thought simulators like computers and servomechanisms, applied to machine sensing, memory, logic, and control.

This narrative perforce has had nothing to say about communication's pioneering spirit in the discovery of electronic bypaths. Discoveries there have been many and constant, but none more significant to the future than the route uncovered, by way of radar and video, to the speechless realm which lies between engineering and psychology. Here, where the techniques of mark-space find analogies in the yes-no elemental reactions of brain and nervous system, electricity bids fair to take over the repetitive tasks of man's mind and his directing hand much as in the past electric power helped emancipate him from muscular drudgery. For telecommunication engineers happy to accept challenges, what a wonderful approach to tomorrow's far horizons!

Appendix

The space allotted to this article featuring wire communication precluded adequate survey of the present state of the wire art or estimates of its future—both of which are desirable in a commemorative issue likely to be referred to, say, 25 and 50 years hence. Fortunately, reference may be made to several recent articles written expressly for those purposes, and the author's responsibility discharged in that way.

1) *Present technical status.* Two new chapters were added by Keith Henney, Editor, in the 5th edition of his handbook:

I. S. Coggeshall, "Wire telegraphy," ch. 27 (66 pp);
H. A. Affel, "Telephony," in "Radio Engineering Handbook," McGraw-Hill Book Co., New York, N. Y., ch. 28 (51 pp.); 1959. (Bibliographies.)

2) *Looking 25 years ahead.* To celebrate its 75th anniversary, the American Institute of Electrical Engineers had four articles prepared:

E. I. Green, "The evolving technology of communication," "*Elec. Engrg.*," vol. 78, pp. 470–480; May, 1959.
H. I. Romnes, "Advances in communication," *Elect. Engrg.*, vol. 78, pp. 481–492; May, 1959.
I. S. Coggeshall, "Telegraphy's next 25 years," *Elect. Engrg.*, vol. 78, pp. 493–499; May, 1959.
W. P. Marshall, "Electrical intelligence and power—A trusteeship," *Elect. Engrg.*, vol. 78, pp. 808–810; August, 1959.

3) *Encyclopedic-type articles.* The article by Green was prepared for "Encyclopaedia Britannica;" the other reference contains many short articles on facets of wire communication:

E. I. Green, "Telephone," *Bell Sys. Tech. J.*, vol. 37, pp. 289–338; March, 1958.
"Encyclopedia of Science and Technology," McGraw-Hill Book Co., Inc., New York, N. Y.; 1st ed.; 1960.

4) *Business features.*

Federal Communications Commission, "Annual Reports;" also "Statistics of the Communications Industry in the United States," Supt. of Documents, Washington, D. C., annually from 1935. (The "Statistics" includes the current corporate structure of the holding and operating companies, and thus is basic to use of the next reference.)

"Public Utility Manual," Moody's Investor Service, New York, N. Y.; annually; 1960. (The physical aspects of plant and services are usually included, from facts supplied by the companies.)

International Telecommunication Union, "General Telegraph Statistics"; also "General Telephone Statistics," General Secretariat I.T.U., Geneva; annually from 1947. (Statistics cover most of the world. Lists of international physical wire facilities and maps are also obtainable from this source.)

5) *Representative current developments.* Following list has been selected on topics of general present interest:

A. E. Bachelet, C. A. Collins, and E. R. Taylor, "New television network switching facilities," *Trans. AIEE*, vol. 79 (*Commun. and Electronics*, no. 60-81, pp. 625-634; 1961.

H. C. A. Van Duuren, F. L. H. M. Stumpers, and B. B. Barrow, "The international symposium on data transmission," IRE TRANS. ON COMMUNICATIONS SYSTEMS, vol. CS-9, pp. 4-96; March, 1961. (With 13 related papers.)

A. Boggs and J. E. Boughtwood, "Application of telegraph techniques in data transmission," *Trans. AIEE*, vol. 78 (*Commun. and Electronics*, no. 59-107), pp. 336-340; 1959.

M. J. Kelly, Sir G. Radley, *et al.*, "A transatlantic telephone cable," *Trans. AIEE*, vol. 74 (*Commun. and Electronics*, no. 55-230), pp. 124-139, 1955; also published with 11 related papers in *Bell Sys. Tech. J.*, vol. 36, pp. 1-326, January, 1957.

K. Bullington and J. M. Fraser, "Engineering aspects of Tasi," *Trans. AIEE*, vol. 78 (*Commun. and Electronics*, no. 59-70), pp. 256-260; 1959.

I. S. Coggeshall and P. Holcomb, Jr., "A character-metered transatlantic switching system," *Trans. AIEE*, vol. 79 (*Commun. and Electronics*, no. 58-1143), pp. 56-64; 1960.

F. W. Smith, "European teleprinter developments," *Western Union Tech. Rev.*, vol. 14, pp. 162-175; October, 1960.

E. D. Anderson, "Facsimile telegraph network for weather maps," *Western Union Tech. Rev.*, vol. 14, pp. 41-51; April, 1960.

W. D. Cannon and T. F. Cofer, "Nationwide facsimile network for weathermap service," *Western Union Tech. Rev.*, vol. 14, pp. 104-116; July, 1960.

Modulation Methods*

RAYMOND A. HEISING†, FELLOW, IRE

***Summary*—This paper covers briefly the early radio telephone experiments with efforts by the various inventors to modulate spark, arc and alternator outputs with microphones of various kinds, and also with variable reactances controlled by voice signals. The development of the high-vacuum triode provided a basis for a completely new art, both in sustained wave generation, and in modulation. A few of the methods of modulation by the triode, as well as the development of the first practical radio telephone which led the way to early broadcast transmitters, are outlined. The discovery of sidebands with its early influence on transmitter circuits, which led to channel spacing, and also to single sideband use for ether space conservation, is mentioned. The early fumbling with frequency modulation and then the eventual discovery of how it could be used to advantage is outlined. Pulse-modulation systems, which have not yet reached their full development, are covered briefly. Nothing is given about modulation methods for the infrared optical waves for which methods of generation have lately been discovered, but for which modulation methods have not yet been published.**

Introduction

THE word "modulate" did not originate in the radio field. Like many other words, it was adopted in radio because its prior meaning appeared to be expressive of a certain action that was a necessary part of radio. Basically it means: "to proportion; to adjust; to vary; to alter; to deviate; to diverge; or to change." In speech it meant to vary or inflect the sound in such a manner as to give expression to what is uttered, or to vary (the voice) in tone. In music it meant to change the key or mode of, in the course of composition.

When adapted to radio, it originally meant to vary in some suitable manner the character of a radio wave so as to cause the latter to act as a carrier of speech or music. Shortly after radio and carrier systems came into extensive use, some engineers began to use the word to mean the interaction in a circuit of one wave upon another without regard to communication. In the early standardization by the Institute, the definition restricted its use to signalling. However, some engineers continued to use it in the broader circuital sense, and so the definition has been broadened.

The earliest use of the word in radio was found by the author in an article by Fessenden in *The Electrical Review*, February 15, 1907. It next appears in books by J. A. Fleming and Prof. G. W. Pierce, both in 1910. Prior to 1910, most radio experimenters used *modify*, *control*, *vary*, *mould*, with *control* used in the majority of cases. In Telegraphy, *keying* was used.

Early History

The early experimenters appeared interested only in AM. Fessenden, in 1902, proposed using a condenser type of microphone in circuits that appear to produce FM, but in his published remarks about tests he seemed interested in using the frequency variation to produce amplitude variation by throwing the frequency in and

* Received by the IRE, June 6, 1961.
† 232 Oak Ridge Ave., Summit, N. J.

out of resonance with the antenna or other tuned circuit to modulate the amplitude. Fessenden used the Elihu Thomson (Duddell) arc in the early 1900's and keyed the waves by shunting part of the inductance, thus using FM in telegraphy. There is no clear idea that he was interested in, or tested, FM telephony as such. The earliest discussion on FM in telephony is a short item in a book by Dr. A. N. Goldsmith in 1918.

It is interesting to note that none of the early experimenters discovered sidebands. Several of them have mentioned a phenomenon they called the "fly-wheel effect" in that they were aware that a tuned circuit or antenna might not permit the radio wave to build up, or die down fast enough to handle the higher rates of variation required by the higher frequencies in the voice. Solutions were not provided, but the discussions appear to put the blame on the inductance in the circuit, and it was thought that careful design would overcome the trouble. However, no suggestions were made as to just how to do this. Others believed that the only limit to the number of channels to be secured depended upon how low the resistance of the antenna circuit could be made. No equation for a modulated wave is found in the literature prior to World War I. So far as the author has any direct knowledge, the equation was first set up and side frequencies discovered by Carl Englund in 1914. The disclosure to the armed forces during World War I and to the public afterward stemmed from Englund's discovery. It is true that some engineers of the American Telephone and Telegraph Company seem to have had earlier knowledge of the sidebands from their study of carrier-system theory for wire lines, but the perusal of their memoranda and other company writings has not turned up a prior discoverer. Dr. G. A. Campbell (AT&T Company) believed their knowledge was derived from mathematical expressions in Rayleigh's *Theory of Sound* involving periodically interrupted sound tones.

Knowledge of sidebands led to departures in antenna adjustments from the practices in radio telegraphy. In the AT&T Company tests of 1915, closer coupling of the antenna circuit to the power tubes than the adjustment for maximum power was resorted to in order to bring up the side frequencies. In the General Electric Company work of early 1919 providing a ship-shore circuit for President Wilson while traveling to Europe, the shore antenna was tuned to one sideband with an arrangement to accentuate the carrier.

The early radio telephone tests were made using spark systems, where the spark rate was pushed up as high as 10,000 cps. The noise convinced the experimenters that sustained waves were preferable, and efforts were made to devise sustained wave generators. Fessenden used arcs in 1902, as did others about that time, while Alexanderson and associates developed HF generators, one of which was used by Fessenden in 1904. Fessenden appears to have dropped his spark and arc work with the appearance of generators, but the development of these devices was continued both here and in Europe for a decade. Poulsen had developed a usable arc by 1904, and various forms of arcs were used by a large number of experimenters in Europe, and a few in the United States. Attempts were made on both continents to devise usable spark systems and some were put on the market with little success.

As generators of suitable HF waves were developed, circuits and devices were invented for modulating the waves. The carbon button microphone would not handle sufficient power to modulate the RF power that was to be used. There began, then, efforts to try to find circuits that would utilize the microphone more efficiently. The microphone was tried in the antenna circuit, then in shunt to the antenna coil, or with part of the coil, in a circuit coupled to the antenna and in numerous other positions. It was evident that a new problem had arisen.

In 1907 Fessenden said, "The difficult problem in wireless telephone is, of course, the modulation of the large amount of energy used for transmission." In 1910 J. A. Fleming said in, *The Principles of Electric Wave Telegraphy and Telephony*, "The difficulties still outstanding in radio telephone are largely those of modulating large high-frequency currents by means of some form of speaking microphone . . . " and again "The chief difficulty which presents itself in connection with radio telephony is that of modulating the oscillations of sufficient intensity in the sending antenna by means of some suitable form of microphone."

All experimenters, most of whom were European, began by using a single simple microphone. After overheating appeared, they began making efforts to surmount or avoid that trouble. It is interesting to see the extent to which they went to devise suitable modulators of the variable resistance type.

Poulsen used six microphones in series connected to one mouthpiece. Colin and Jeance used two microphones, switching from one to the other when one heated. Ditcham used two microphones at a time with several pairs of microphones on a turntable so he could quickly rotate it and switch from one set to the next. Lorenz used 25 microphones in parallel connected to one mouthpiece. De Forest used a microphone shaker that was periodically operated. Goldschmidt used several microphones at a time in parallel, with transformers arranged to divide the load between them and prevent one from taking all the load upon its resistance dropping with heating. Fessenden devised a special microphone that was water jacketed as well as having other advantages of design. Dubilier devised a form of wire-to-radio modulating relay using water cooled carbon-containing elements. Jahnke and Tate passed alcohol through their granulated carbon in their microphone to keep it cool. Berliner devised a heavy-current microphone with fan cooling. Egner-Holmstrom used 16 small carbon buttons, oil cooled, and connected to a single diaphragm. Marzi devised a microphone as part of a telephone relay in which powdered carbon passed from a hopper through a specially shaped carbon button, during operation. Majorana devised a liquid microphone in which he vibrated one side of an opening from which a fine even stream of acidulated water issued so as to transfer the speech wave form to the falling stream of water, and the stream passed onto two electrodes in the RF circuit. Good results were secured. Chambers devised a liquid microphone in which an electrolyte emerged vertically from a flat ended nozzle against the vibrating diaphragm and the thickness of the film of water was varied by the speech sounds so as to provide the modulating resistance. Vanni devised a liquid microphone in which a stream of acidulated water fell between two specially shaped electrodes, one of which was vibrated by a connection to a diaphragm.

Fessenden tried variable reactances. He made the tests mentioned earlier with a condenser microphone. He also planned to secure modulation by means of a variable inductance in which the inductance contained iron, and its permeability was varied by means of a variable current from a microphone. This appears to be the first "magnetic amplifier," but it is not known if he made use of it or not.[1]

Alexanderson and Nixdorff in 1916 published their paper describing their magnetic amplifier and its use in wireless telephony. They used it for modulating the output of Alexanderson's HF alternators.

During 1911 and 1912, Fritz Lowenstein, Lee De Forest, E. H. Armstrong, and Alexander Meissner (with gas tube) observed that the thermionic triode could generate oscillations. The last three filed patent applications in the United States and the litigation following was not ended until 1930. Though De Forest prevailed in the litigation, it was Armstrong who educated the engineering profession in the audion's possibilities. With the development of high-power vacuum tubes a few years later, the efforts of using arcs or alternating current generators of high frequency for radio telephony came to an end, though both devices continued in use in telegraphy until recently.

In 1912 De Forest's audion was called to the attention of both I. Langmuir of the General Electric Company and H. D. Arnold of the Western Electric Company. They both conceived of improving the vacuum and securing more power than De Forest's original tubes would give. With the respective assistance of W. C. White and H. J. van der Bijl, they converted the audion into a modulation device as well as a good amplifier. Langmuir and Arnold both filed patent applications on the improvement, and the litigation that followed went to the Supreme Court before being settled.

Alexanderson appears to have been the first to conceive of using the audion as a variable resistance to use as a modulator in modulating the output of his alternators. In the Western Electric Company, Colpitts conceived of using it as an oscillator that could be modulated by impressing a signal voltage upon the grid. Colpitts knew from telephone experience with mechanical audio amplifiers that oscillations could be generated in an amplifier if some of the output energy got back into the input. Apparently he didn't know just how to secure this action with the vacuum tube so he devised a circuit with many elements whose actual mode of operation is uncertain, and secured a patent.[2] Logwood, working with De Forest, devised a simpler oscillator circuit with means for impressing the audio signal upon the grid, and this circuit was used in 1916 by De Forest in broadcasting around New York.

H. J. Round in England devised a circuit in 1914 with a vacuum tube oscillator, but he used an ordinary microphone for modulating the output. In the earlier form, the output was fed to an antenna and the microphone was inserted in the antenna. In a later form, the output was modulated by a microphone and fed to a second tube which acted as amplifier and delivered its output to the antenna. The author does not know to what extent these sets might have gone into actual use.

Practical Applications

In the Western Electric Company, van der Bijl in 1913 observed that the improved triode had a square law characteristic curve. He conceived the idea of utilizing the variable slope of the curve to produce a modulated wave, such wave to be amplified to a desired power in a later stage. A small HF voltage, and a larger signal voltage were to be applied to the grid of the tube, and the HF output was to be amplified as much as needed. This system of modulation was used in the Bell System tests of 1915 in which speech was transmitted various distances, eventually reaching San Francisco, Panama, Hawaii and Paris, with the transmitter located at the U. S. Naval Radio Station in Washington, D. C. (June to October, 1915).

The author, then in the Research Department of the Western Electric Company (which later became part of the Bell Telephone Laboratories), devised several transmitter circuits using a vacuum tube as an oscillator and a second one as a modulator arranged in various ways. Two of the systems were plate-modulating systems, and two were grid-modulating systems. In the plate-modulating systems, the oscillator and modulator were in series in one system with a constant potential power supply, and in the second were in parallel on a constant current supply. In the grid-modulating systems, one utilized a large RF voltage, and a medium signal voltage on the grid, and the second system utilized a modulator tube with two grids, on one of which the HF voltage was impressed and on the other the signal voltage

[1] Fuller descriptions of these devices may be found in J. Zanneck, "Wireless Telegraphy," Dr. A. Enke, Stuttgart, 1912. (In German.) Translated by A. E. Seelig in 1915, McGraw-Hill Book Co., Inc., New York, N. Y. Also see A. N. Goldsmith, "Radio Telephony," Wireless Press, New York, N. Y.; 1918.

[2] Not to be confused with the simple oscillator circuit known as the "Colpitts oscillator."

was impressed. The HF output was amplified to the desired power. Also a circuit was devised using a rectifier as a modulator, where both the HF and the signal waves were impressed in series with the rectifier and its output circuit. The output was amplified when necessary. It might be noted that the first wire carrier systems used van der Bijl's modulating system, with a single sideband as proposed by Carson. Later, the triode modulators were replaced by rectifiers as modulators.

During this period, W. C. White of the General Electric Company came up with some modulating systems resulting in a complicated patent situation on radio transmitters such that no one could build equipment for commercial use until arrangements were made to exchange licenses.

In 1917, upon our entry into the war, the author was appointed a member of the subcommittee on Wireless Communication between Aircraft, of the National Research Council, and was called upon to develop a radio telephone transmitter for aircraft use. He chose one of his own inventions, a circuit that was simple and easy to adjust, and which is known as the "constant current modulating system." The radio sets known as the SCR 67 and 68 were the first ones developed for the Army, and the CW 936 for the Navy. Thousands of these sets were manufactured for the government before the war was over. Some of the CW 936 sets were used as transmitters in the early days of broadcasting. These two developments were probably the first successful practical radio telephone sets made.

The use of double modulation in radio telephony, involving a frequency intermediate between the carrier and signal frequencies, was conceived by Hammond as evidenced by a patent[3] filed in 1912. The author independently devised an operable double-modulation system in 1915[4] in which one or more speech signals individually modulated one of a series of spaced superaudible intermediate carriers, and all of the latter, then modulated a higher radio carrier. Experimental equipment utilizing the system was lent the Navy in January 1917. A modified form involving single sideband[5] was employed in the transoceanic long-wave system, demonstrated for one channel, one way to England on January 5, 1923, and it has been employed in multiplex form[6] on wire carrier and UHF radio systems since to shift an entire group of single-sideband channels from a frequency region where the single sideband is easily separable, to much higher regions in the frequency spectrum. Double detection is employed in reception.

A closely similar type of modulation is that devised by Kendall[7] for use at a radio repeater station in which a radio signal coming in on one frequency is shifted to a different outgoing frequency in a single step similar to the converter step in the superheterodyne receiver. It is widely used in radio relay stations.

HF carrier systems on wire lines were proposed before radio appeared on the scene. They began as "harmonic telegraph" systems, and are associated with the names of Gray, Bell, Edison, Mercadier and others. In the telephone field they appeared with Pupin, and Hutin and Leblanc in 1892, and with others following, such as Gibboney, J. Stone Stone, and Ruhmer. Arcs and high-frequency alternators were proposed as generators. Stone mechanically moved one electrode of the arc in modulation while the others were used microphones. None of the systems showed much promise because of the arc and high-frequency alternator required, together with modulation difficulties. About 1910 General Squier began tests in this field. Bell System work with an all-audion system, and the van der Bijl modulation system began in 1914, utilized Carson's idea of single sideband without carrier. In 1916 the first equipment was installed on wire lines between South Bend, Ind., and Toledo, Ohio. Since that time numerous advances have been made, one of which, from the modulating standpoint, was the use of rectifiers instead of vacuum tubes as modulators.

The carrier systems on wire are almost all of the single-sideband type without carrier. A device necessary for their original success was the band-pass filter. About 1909, G. A. Campbell discovered how to construct networks with substantially flat transmission characteristics over a band of frequencies and rapid attenuation outside. Their use enabled one sideband to be selected and the other sideband rejected with a carrier as high as 30 kc in the early installations (1916). As better magnetic materials became available in later years, higher carriers could be employed. With piezoelectric resonators used in the filters, carriers in the hundreds of kilocycles are possible. In all these systems, the carrier, if eliminated, is eliminated by a balanced modulating circuit, either of triodes or rectifiers.

Frequency Modulation

Radio telephone devices and tests so far discussed were all for AM, FM or "wavelength" modulation is occasionally mentioned in the early literature, but not favorably. The establishment of broadcasting and the spacing of channels 10 kc apart caused many engineers to give thought to how to narrow the band. FM was proposed. In 1922 J. R. Carson published his famous paper pointing out that FM would not provide a narrower band than AM, but might actually require a wider band. His analysis, however, led him to believe that FM inherently produced distortion in the signal and this he stated in his paper. For a time this subject seemed to have been disposed of. However, a few years later others who must not have noticed Carson's paper began working on the subject again under the impression "that the necessary frequency band for each transmitter may be reduced to any desired minimum." The

[3] J. H. Hammond, U. S. Patent No. 1,491,774; 1912.
[4] R. A. Heising, U. S. Patent No. 1,633,100; 1927.
[5] R. A. Heising, U. S. Patent No. 1,675,888; 1928.
[6] R. A. Heising, U. S. Patent No. 1,690,227; 1928.
[7] B. W. Kendall, U. S. Patent No. 1,734,132; 1929.

expected results did not materialize upon making tests, and published analyses by others supported most of what Carson had said. However, E. H. Armstrong who had been trying to circumvent static for many years discovered that if one widened the swing in FM much beyond the bandwidth used in AM, and used an effective limiter in the receiver, noise lower than the carrier could be substantially eliminated. This was an unexpected discovery, because it had been previously shown in connection with AM systems that once the noise got into the circuit nothing could be done to eliminate it without eliminating part or all of the signal. This property in a FM system, together with the "capture effect" and other advantages have resulted in the wide use of FM.

Any analysis of FM brings forth the subject of "phase modulation" as frequency shift and phase shift go hand in hand. Frequency and phase modulation transmitters may be used interchangeably by proper predistortion of the signal wave.

FM of a master oscillator may be produced in small amounts by variation in plate voltage, in larger amounts by a variable resistance (a triode with the signal on its grid) across the inductance or capacitance of the oscillator circuit, by variably introducing a 90° component, by core saturation of the inductance, as well as by numerous other arrangements. To insure that the carrier has suitable frequency stability, phase modulators excited by piezoelectric oscillators are often used. A phase-modulated wave of less than ±30° phase swing is the equivalent of an AM wave with the carrier shifted 90° from its amplitude position. By using a balanced modulator to produce the two sidebands without the carrier, and then adding the carrier shifted 90°, the phase-modulated wave can be originated using a crystal master oscillator. The correction circuit mentioned above must be used when FM is desired. Harmonic step-up increases the frequency and the frequency swing up to any desired amount.

Pulse Modulation

Modulation methods and devices are designed to go with the type of HF generator used and fit into the communication system involved. The systems discussed so far are for AM and FM and involve frequency selection to separate channels. Before the frequency selection systems were in commercial use in telephony, time-division systems were in use in wire telegraphy. Such systems arrange to synchronously switch a wire channel between sets of transmitting and receiving telegraph devices. These systems came into wide use with the printing telegraph. In a common system, a 5 dot code, called the Baudot code, formed the basis, where the various combinations of the 5 dots in their positions were transmitted or omitted, and the 32 combinations possible permitted covering the 26 letters of the alphabet, space between words, punctuations, shifting to figures, etc. In the synchronous printing telegraph, a full set of 5 "bits" or pulses were transmitted while one pair of instruments had the line, and the character was printed while the line was in use by others. Such a system had one advantage over frequency selection systems in that all pulses were of the same amplitude, and when repeated at a repeater station by relay-type devices, constant amplitude pulses were regenerated and sent out so no accumulation of noise is built up by successive repeatings. This, together with the difficulties of frequency selections at very high frequencies, led to the study of using pulse modulation and transmission systems in telephony. The principal systems are called:

pulse amplitude modulation	PAM
pulse duration modulation	PDM
pulse position modulation	PPM
pulse code modulation	PCM.

In the printing telegraph, the switching rate of the channel between users is slower than the rate of transmission of pulses. In telephony, that is not possible. In practice, the speech signal is sampled—say 6000 times a second if 3000 c is the highest frequency to be transmitted. One of the pulse-modulation systems is then utilized to convey to the receiving end information as to the magnitude of each and every pulse in succession. In pulse amplitude modulation, each pulse is of suitable short duration, and could be unidirectional in character in the apparatus at the ends, but is transmitted as a short burst of high frequency. The PAM pulses are transmitted, as generated, in time and amplitude. The length of each pulse will depend upon how many channels are planned, as pulses for the several channels will be interspersed.

In PAM the pulses may be produced in a variety of ways such as by a flip-flop circuit stable on one side, which remains on the other side for a suitably short time. The short pulses are then amplitude controlled in any of numerous ways used in AM that do not destroy their length. At the receiving location, after detection, a low-pass filter of 3000 c will average out the pulses into the original complex signal. Multiplexing will require synchronous switching of the channel between pulses to other users. This system is not free of noise accumulation problems on passing through a repeater as the pulses repeated must be proportioned to the amplitude of the incoming pulses.

In PDM the pulses can be produced in numerous ways. A simple manner is to apply the signal directly to the flip-flop circuit so as to control or modulate the duration of the "flip." This system, because of variation of pulse lengths, will not permit of as many channels upon multiplexing as PAM for the same average pulse length. However, all pulses are of the same amplitude, and suitable constant amplitude repeaters can avoid much noise accumulation. Reception, after detection, requires only a low-pass filter to smooth out the pulses.

Pulse position modulation depends upon the movement of pulses away from and toward a mean position, or back and forth across it to represent the amplitude being conveyed. It resembles in a way phase modulation. An early way of generating these pulses is to first generate pulses by PDM and then differentiate them, utilizing a uniconducting arrangement to pass pulses of one sign only.

Pulse code modulation utilizes a code to represent the amplitude of the sampling pulses. A common one is to use the code for the Baudot telegraph system in which a group of pulses 5 in number identify certain amplitudes instead of alphabetical characters. The code does not necessarily have to use 5 pulses, but can start with 2, and extend upward. The codes will not represent each and every gradation in a signal wave, but only a certain number of discrete amplitudes. A sine wave signal will be represented by a wave that possesses only horizontal and vertical components and will zig-zag back and forth across the sine wave trace. The wave will be approximated, or "quantized." The wave will approach the sine wave more and more as the number of pulses in a character is increased.

In PCM the pulses in a group represent different components of magnitude of the sample in geometrical relation 1, 2, 4, 8, etc., and in the receiver must each contribute its component to the resultant pulse. The circuits and devices for generating the code and reproducing the original sampling pulse are too numerous to describe in this article.[8]

Miscellaneous

With the invention of new devices for generating the extra high frequencies coming into general use, new methods and circuits for modulation have been devised. The reflex klystron, generating frequencies in the thousands of kilocycles, is used for pulse systems, and also for FM sustained wave telephony. The frequency varies with applied voltage and so it provides a convenient generator and modulator combined for FM. The two-cavity klystron is used largely as an amplifier, but can be used for generating oscillations. The traveling-wave tube, which operates at higher frequencies than the klystrons, is largely used as an amplifier. The magnetron can be made for operation in most parts of the radio spectrum and can deliver large amounts of power. It is used in pulse systems, especially radar systems. Electro-optical devices have been devised that will generate radiations in the infrared region and have been tested for telephony. The methods of modulating them have not been publicized.

[8] For details on modulation theory, see H. S. Black, "Modulation Theory," D. Van Nostrand Co., Inc., New York, N. Y.

Transmitters*

JAMES O. WELDON†, FELLOW, IRE

Summary—**The evolution of the type, form and scope of radio transmitters is traceable from the nascent technology of not yet a century ago to today's elaborately sophisticated systems, among which are included radar, multichannel microwave, aircraft and marine communications and navigation and commercial broadcast. Apparatus and techniques progressed from the early spark and arc transmitters to alternators of various designs and powers, through the era of the first vacuum-tube transmitters to the later highly efficient vacuum-tube types. More recently, transmitter design has utilized high-power tetrode and klystron amplifiers. Although code signaling was the first practical method of communicating intelligence, interest in voice transmission resulted in concurrent efforts to adapt the then existing transmitters to this purpose. The discovery of ionospheric phenomena permitted advances into areas of transmitter development previously considered impractical for other than short-distance communications. Regulations governing transmitter power, frequency and use were established by government agency as the need became apparent. The regulations have been adjusted to keep pace with the increase in power capabilities and the use of the frequency spectrum as the trend continues toward higher-power transmitters.**

* Received by the IRE, August 2, 1961.
† Continental Electronics Mfg. Co., Dallas, Tex.

IT is interesting to note the progress of the entire field of electromagnetic phenomena which has taken place since the inspired mathematical formulation of the electromagnetic theory by James C. Maxwell in 1864. Twenty-two years later, in 1886, Heinrich Hertz successfully demonstrated the truth of Maxwell's equations, and thirty-seven years later radio communications first became a reality with the transmission and reception of the first feeble code signals across the Atlantic by Marconi in 1901.

All the progress from Maxwell to the present has taken place during the lifetime of some older people, and almost all the progress has taken place during the last thirty years, within the life-span of our younger engineers.

The first emphasis was naturally upon communications for ships at sea, and the first interest in transmitter development was for ship-to-shore and ship-to-ship use. The spark transmitter, the only known system at that time, became the subject of intensive development. Considering the existing state of the art and the fundamental limitations of the spark transmitter itself, the development which followed was amazing.

From the open-spark direct-excitation method, with antenna selectivity only, elaborate circuitry evolved to reduce the bandwidth. The energy source, initially a battery and induction coil with a vibrating interrupter, developed into the alternator and step-up power transformer. The spark-gap underwent several stages of development from the simple 2-electrode gap to the non-synchronous and synchronous rotary gap, the quenched gap and other variations. The power of such transmitters was continuously increased until shore stations rated at 100 kw were common. Spark transmitters emitting damped waves for telegraph communications continued to be popular and were extensively used, on wavelengths between 600 and 800 m, as late as 1930.

In about 1912, the dominance of the spark transmitter was challenged by the arc transmitter. The Danish scientist, Valdemar Poulsen, invented the Poulsen arc in 1903, following original work by Elihu Thomson in 1892. Fig. 1 is a photograph of the first Poulsen arc brought to the United States. As early as 1912, 30-kw arc transmitters used by the Navy established communication between Arlington, Va., and Mare Island, Calif., and later from Arlington to Honolulu, Hawaii. Before vacuum-tube transmitters began replacing the older systems, arc transmitters rated as high as 1000 kw had been developed. The transmitter at the U. S. Navy LaFayette Station in Bordeaux, France, shown in Fig. 2, is an example. The arc transmitter represented the first introduction of continuous-wave telegraph systems.

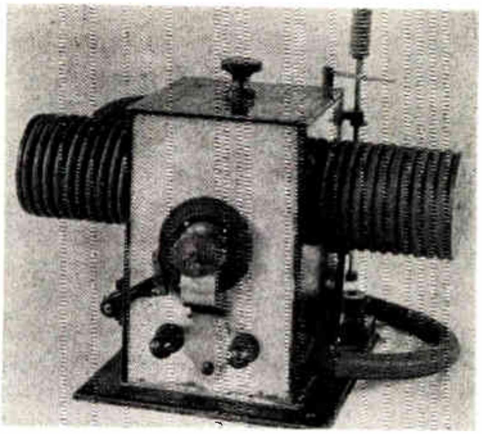


Fig. 1—First Poulsen arc brought to the U.S.A. from Denmark in 1909. Rated at 100 watts, it actually delivered about 50 w of RF power. With it, radio-telephone signals were sent from North Palo Alto to South Palo Alto, thereby demonstrating the efficacy of the Poulsen arc for radiotelephony. (Courtesy of *Electrical Communication*.)

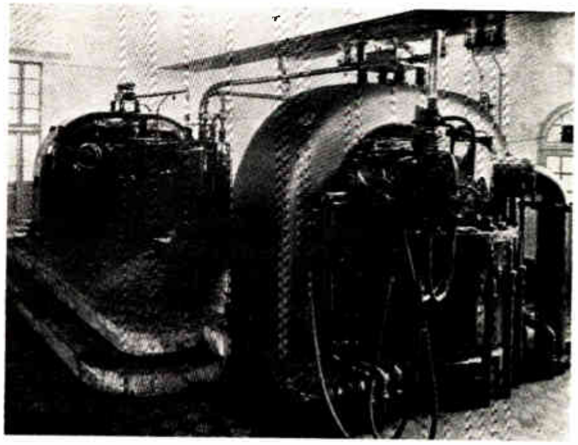


Fig. 2—Historic Lafayette radio transmitter station built by the Federal Telegraph Co. for the U. S. Navy near Bordeaux, France, during World War I. It was the largest transmitter ever built up to that time. (Courtesy of *Electrical Communication*.)

Another early system extensively used was the Alexanderson alternator, a high-speed, ac generator by which the high frequency was directly produced. Frequencies as high as 200 kc were generated in this manner with a machine having a capacity of about 1 kw; 2-kw machines generating up to 100 kc were constructed, and also higher-power machines at frequencies of 50 kc and lower. Such alternators rated as high as 200 kw have been constructed, and at least one is still operated by the U. S. Navy in Japan.

Another type, called the "Goldschmidt Alternator" after its developer, Dr. R. Goldschmidt, was based on a theory of conjugate rotating vectors, a principle different from the Alexanderson alternator. Alternators of this type, with the advantage of lower revolving speed relative to the output frequency, had a maximum output of 200 kw and a normal output of 100 kw. Such machines were used at Hanover, Germany, and at Tuckerton, N. J.

The Edison diode effect was controlled by the three-element vacuum tube invented by Dr. Lee De Forest in 1906. Fig. 3 is a three-stage De Forest audion amplifier of 1912. At that time it was considered a receiving device only. Vacuum-tube transmitter development followed at a much later date. Fig. 4 is a 400-w, LF transmitter built by the General Electric Company for the Signal Corps in 1923.

Extensive replacement of the older spark, arc, and alternator systems with communications transmitters using vacuum tubes for generating the radio frequency began in about 1927. With this major advancement in the state of the art, later transmitter development is related closely to several factors which shaped the requirements and directed the trend.

Fig. 3—Three-stage De Forest audion amplifier of the type first built by the Federal Telegraph Co. in 1912. This was the earliest known commercial cascade audio frequency amplifier, and was the type demonstrated to the U. S. Navy in September, 1912. (Courtesy of *Elec rical Communication.*)

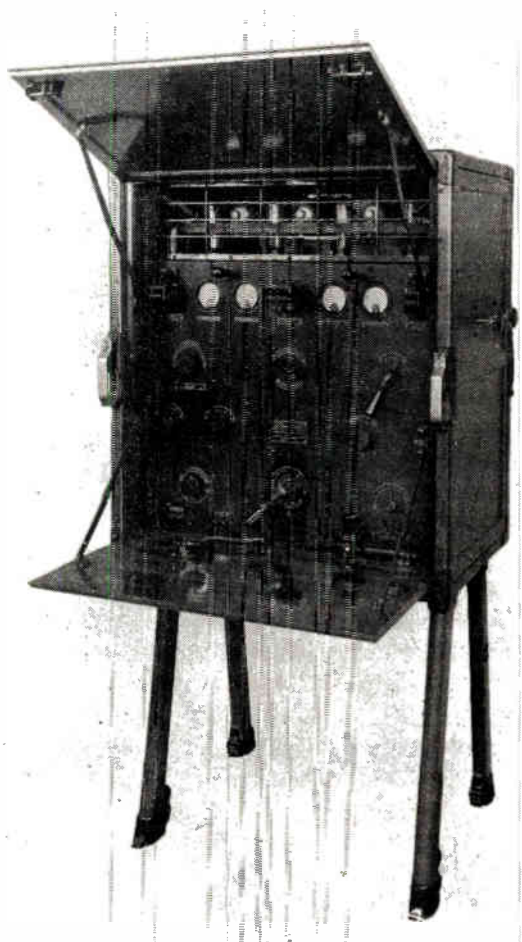


Fig. 4—Developed and designed, 1922–1923, by the General Electric Co., BC-127 400 watt, low frequency transmitter. It was transportable, and the first regular Signal Corps transmitter to use the MO-PA technique. Power-MG set. (Courtesy of *RCA.*)

Among the factors may be listed: 1) The rapidly expanding demand for such communications services by an increasing number of operating agencies. 2) The increasingly crowded condition of the frequency spectrum brought about by this demand, and by the use of the spectrum for types of services other than communications. 3) The continuous requirement for increasingly higher data rate to obtain greater use from a given frequency assignment. 4) Increasing knowledge regarding the modes of propagation of radio waves, permitting more intelligent assignment of frequencies to various types of services. 5) Collateral developments in circuitry, vacuum tubes and other components, which extended the state of the art and permitted transmitter improvement. Figs. 5 and 6 are representative of modern, high power transmitter installations.

International Morse Code was used for signaling by all the telegraph transmitters in use prior to the appearance of vacuum tube transmitters, and in all the early models of these. The International Morse Code characters were transmitted by interrupting the carrier (sometimes called "on-off" keying) to form the dots and dashes, with the interruptions of the carrier making the spaces between the elements of the characters. Variations of this type of transmission occurred in the case of the continuous-wave transmitters. In one method a "chopper" was keyed on and off to introduce interruptions into the continuous wave at an audio rate so that a tone could be heard. The carrier remained on continuously but, with the receiving systems used, it could not be heard except during the time it was interrupted at an audio rate by the chopper. Most of the receiving methods at that time required some modulation of the carrier for a signal to be heard. With the appearance of the vacuum tube receiving system, the regenerative detector and later the beat oscillator could be used to produce an audio note in the receiver so that reception of the continuous-wave signal with on-off keying became possible.

Keying was also accomplished by frequency shift on some of the earliest transmitters. Since the transmitter frequency was determined by the resonant antenna circuit and any associated resonant circuits, keying could be accomplished by short circuiting turns of an antenna inductor, for example, by means of a suitable contactor capable of carrying the current. The frequency of the carrier was shifted far enough by this method so the receiver tuned to it would not respond. The off-frequency radiation or "back side" of the signal was not used and, of course, represented a waste of spectrum space.

Because most of the receiving systems were designed to receive an interrupted carrier such as that generated by the spark transmitter rather than a continuous wave, many of the early vacuum-tube telegraph transmitters used ICW, interrupted continuous wave. The interruption was accomplished at an audible rate by a chopper similar to that previously described for use with the arc

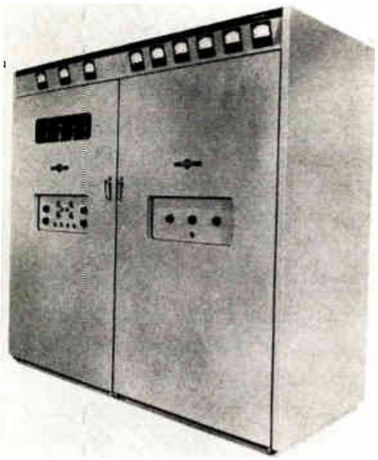


Fig. 5—Collins 205J-1 HF single-sideband power amplifier. This equipment is an automatically tuned 45-kw PEP linear power amplifier with an average power rating of 22.5 kw over the frequency range of 2 to 30 Mc. (Courtesy of *Collins Radio Co.*)

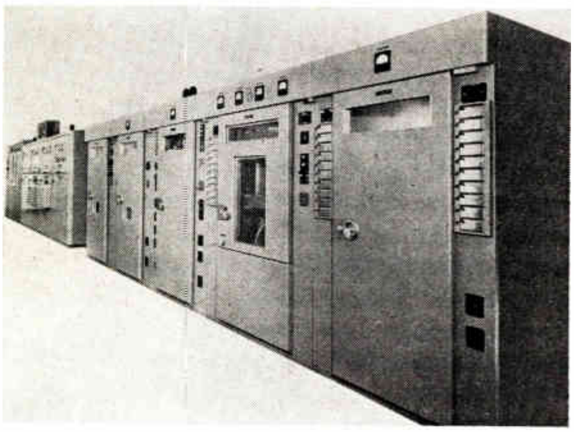


Fig. 6—Continental Electronics HF single-sideband 300-kw transmitter. Average power, 600-KwPEP.

or alternator type transmitter. Spark transmitters were modified in some cases by the substitution of tubes, the anodes of which were supplied from the 500-c alternator through a suitable transformer. Thus, the ac supply to the anodes of the tube provided a modulation which made the signal audible to such receivers without the necessity of a chopper. These transmitters were classified as self-rectifying types.

The method of signaling by interruption of a continuous carrier, on-off keying, is still used in many types of simple transmitters and also in high-power, VLF transmitters where the Q of the antenna system so limits the bandwidth that the application of other methods is difficult, and where even the continuous-wave-type of keying is limited to as low as 25 words per minute because of the narrow bandwidth of the antenna. More sophisticated methods of keying have been adopted for most high speed point-to-point systems and for many communications systems to ships and aircraft. A popular method is that of frequency-shift keying, by which the frequency of transmission is shifted above the assigned frequency for the "mark" signal and below the assigned frequency for the "space" signal. This is at present one of the most popular methods for single-circuit teletype operation using the 5-pulse-character teletype code.

Another early development in signaling methods was the use of telephone or voice transmission. Attempts were made to apply voice transmission to spark transmitters with very little success. Considerable interest developed with the appearance of the arc transmitter, and Dr. Poulsen and Dr. P. O. Pederson demonstrated the transmission of wireless telephone signals by a continuous-wave arc transmitter in Denmark as early as 1909.

Other attempts at the transmission of voice by radio were made using the alternator type of transmitter, with modulation introduced into the antenna system by the absorption method or by modulating the current in the field coil of the generator. However, the completely successful development of telephone transmission came only with the vacuum-tube transmitter.

The system using suppressed carrier and single- or double-sideband signaling has become relatively important because of advantages with regard to selective fading and for the conservation of spectrum space. In such a system, a single sideband without the carrier is capable of carrying many independent teletype channels at signaling rates of up to 100 words per minute; or a sideband having a width of 3000 c can carry a single telephone conversation while the other sideband is used either for another voice channel or teletype channels. Transmitters capable of handling this type of signaling system represent an important part of contemporary communications systems.

Early frequency allocations were assigned to various types of service. Prof. J. H. Morecroft wrote, in 1920, "The laws of the United States specify the following wave lengths:

High Power Stations, over 1600 meters.
Navy, 600 to 1600 meters.
Ship Stations, 300, 450, 600 meters.
Amateurs, below 200 meters."

For transoceanic communications, wavelengths as long as 20,000 m, corresponding to a frequency of 15 kc, were used with transmitters utilizing as much power as the state of the art would permit. About this time, the amateurs assigned to frequencies of 1.5 Mc and higher, working with transmitters of a few watts or a few hundred watts, began making long distance contacts on frequencies up to 20 Mc.

The phenomenon of ionospheric reflection was discovered, and frequencies above 1.5 Mc, previously considered suitable only for short distances, became important for long-distance and transoceanic point-to-point communication. This started the development of many types of transmitters covering the frequency range from 4 up to 20 or 30 Mc. Because of diurnal and seasonal variations in the maximum usable frequency for a given distance, and the necessity for selecting the proper frequency as related to the distance between stations, requirements were established for transmitters capable of operating over wide frequency ranges with reasonably quick change from one frequency to another. Early transmitters of this type required discontinuation of operation while manual adjustment was made on inductors and capacitors. In some cases, interchange of conductors by plug-in methods was made to select a new frequency. Intensive development work, particularly for the military, has resulted in continuously decreasing the time required for frequency change. Transmitters were introduced wherein the tuning was accomplished by motors associated with a servo system so that a number of preset frequencies could be selected by a single switch, and the entire transmitter would align itself on the new frequency in a period of one or two minutes.

A further extension of the fast frequency change system is the application of automatic tuning, by which a stable variable-frequency oscillator is used to excite the transmitter and the tuned circuits of the transmitter are controlled by servo systems directed by impedance and phase-angle sensing devices, causing the system to adjust itself automatically to the applied frequency.

The wide variety of communications transmitters developed for service in the frequency range from 2 to 30 Mc covers a wide range of applications such as airborne transmitters, transmitters for teletype, telegraph and telephone, ship-to-shore and shore-to-ship, ground-to-air, ground-to-ground, higher-power long-distance point-to-point transmitters and the transoceanic telephone services.

The power of the transmitters ranges from 100 watts and lower for airborne units, through 1 to 5 kw for shipboard equipment and higher power for land-based stations. A rating of 10 to 50 kw is common for continuous-wave frequency-shift keying transmissions and for the peak envelope power in single-sideband transmissions. A linear amplifier for single- or double-sideband suppressed carrier service, capable of continuously tuning the 4- to 30-Mc range, was developed by the U. S. Army Signal Corps in 1953. These transmitters, with a peak envelope power of 600 kw and an average power of 300 kw are now used by the Army, Navy and Air Force.

In recent years the phenomena of tropospheric and ionospheric scatter came under investigation as frequencies in the VHF and UHF ranges came into use. These phenomena presented the possibility of extremely reliable transmission over long periods of time at fixed frequencies or at frequencies that changed only seasonally.

Ionospheric scatter circuits required the development of higher-power VHF transmitters in the range from 30 to 65 Mc. Concurrent with the development of the continuously-tunable linear amplifier, the U. S. Army Signal Corps developed a 600-kw peak, 300-kw average power continuously-tunable linear amplifier to cover this range. This amplifier is now being used for research at Stanford University.

The tropospheric scatter, using UHF frequencies to cover shorter distances with greater reliability, also introduced requirements for a new type of communications transmitter. Requirements were developed for long-range communications circuits, established by a series of tropospheric and ionospheric scatter transmitters and receiving systems. An example is the communications system which connects the Distant Early Warning (Dew Line) radar system across Canada, where a UHF system interconnects all sites and carries the gathered information to a central location.

Typical transmitters in the UHF range for tropospheric scatter have been developed for various frequencies between 300 and 1000 Mc and from 10-kw to 100-kw peak envelope power.

Thus, from the early systems which operated as low as 15 kc, the frequency range of communications transmitters has been extended to the area of 1200 Mc, at which point microwave communications systems begin and use frequencies through several kilomegacycles.

In spite of the trend toward higher frequencies, emphasis is even now being placed on the very low frequencies. Only frequencies below 20 kc have a usable capability of penetrating salt water, so signals can be received by submerged submarines. With the U. S. Navy Polaris program for nuclear submarines, a new emphasis is placed on this type of communication, and in 1961 the Navy put into operation a new VLF station capable of covering the frequency range from 14 to 30 kc and operating with a continuous output power of 2,000,000 watts. Other VLF stations are in the planning stage, including one for NATO, in England.

The most important collateral development contributing to the modern communications transmitter has been in vacuum tubes. By 1925, practical water-cooled triode tubes permitted the design of transmitters having power outputs several times greater than those of previous models using radiation-cooled types. Water cooled types have since been improved so that anode dissipation in excess of 500 kw is possible. At the same time, the introduction of forced-air cooled tubes with a fin-type structure on the anode has permitted the simplification of transmitters having output powers up to 50 kw. The introduction of the high-gain, high-power tetrode, and its development into ever higher power designs, has permitted circuit simplification by eliminating the need for neutralization and by the reduction of the number of amplifier stages. Ingenious design has continued to raise the upper limit of operating frequency for higher-power tetrodes so that the useful frequency limit of these tubes extends into the UHF region. For UHF and higher frequencies the klystron amplifier, now available in extremely high-power models, has been an important factor in extending both the power and frequencies of communications transmitters.

Improvements in many types of components have resulted in increased reliability and smaller size. This includes the introduction of ceramic capacitors and the development of vacuum capacitors, both fixed and variable. Variable vacuum capacitors have been of considerable importance, particularly in the development of transmitters for wide frequency coverage in the HF range, where the extension to the lower frequencies required lumped-component circuits and where the HF end of the range approaches an area in which cavity-type circuitry might be more applicable.

Early continuous-wave vacuum-tube transmitters used a power oscillator as the generator of the RF energy. The requirement for better frequency stability soon caused a change to the master-oscillator power amplifier system. A major circuit advance for that period was the suppression of self-oscillation in the power amplifier by neutralizing the interelectrode capacities of the triode tubes which were then in wide use. Various methods of neutralization have developed since, and circuit improvements have been continuous. One variation in the use of the triode is the grounded-grid amplifier or grid-separation amplifier in which neutralization is unnecessary. The tetrode mentioned previously provides a screen between the control grid and the plate of the tube which, along with other advantages, eliminated the need for neutralization.

As the frequency spectrum became more crowded, the need for more stable control of operating frequency became increasingly urgent. In 1929, the Federal Radio Commission (now the Federal Communications Commission) began establishing requirements for piezoelectric crystal control of frequency, and these requirements were applied to more types of transmitters to maintain frequency stability. The development in crystal stability has greatly advanced since that date to the point where stability of one part in 10^9 is now being specified in extreme cases and one part in 10^8 has become common. At the same time, while still less stable than good crystal oscillators, the development of variable-frequency oscillators has advanced so that they provide satisfactory stability for many operations.

Each increase in the available communications facilities seems only to create a greater demand, and this ascending requirement for the transmission of information, whether it be voice, printed word or data transmission of various types, provides the background for the current work and possible future work in transmitter development.

The information theory discloses a theoretical maximum of information which can be transmitted within any given specified bandwidth in the frequency spectrum. Present signaling methods do not closely approach the theoretical ultimate. Therefore, laboratory workers are seeking to develop methods for increasing the speed of transmission, multiplexing of more channels in a given bandwidth and at the same time improving circuit reliability. These developments take various forms and will undoubtedly place certain additional requirements on the capability of communications transmitters. Already in single sideband systems the requirements, for reduction of intermodulation products which require improved linearity and frequency response in the amplifier, have been changed to permit maximum use. Specifications have been changed from 30 db below either of two tones in the standard two-tone test to present-day specifications which require a reduction of such distortion to 47 to 50 db.

Certain of these systems depend upon increasing frequency stability for proper operation, and still further improvement in frequency stability will result in the transmitters. In this connection, there is an increasing demand in some types of service for the capacity to operate on any one of a very large number of closely-spaced frequencies with the same extreme accuracy of frequency control. This results in the development of synthesizers of improved design which are controlled by a single, highly stable source and develop instantly-selectable frequencies at close intervals throughout the operating range.

This leads directly into the requirement of ever-faster frequency change for the entire transmitter. As this requirement develops, we may expect to see more work on the development of broad-band transmitter systems where the amplification will be independent of frequency, and the process of frequency change will be merely that of shifting the frequency at a source which may be a variable-frequency oscillator or a synthesizer. Such a broad-band system has no selective circuit for the suppression of harmonic radiation or other spurious radiation. A further critical requirement is placed on linearity of the amplifying system to avoid the introduction of distortion products, whether they are or are not harmonically related. Such broad-band amplifiers are already available in lower power. In the natural course of development these will probably be used as driver stages where high power is required, with fast tuning arrangements retained for the extremely high-power stages until further development permits broadbanding at such power levels.

Work continues toward increasing the power of transmitters in all categories. It might be said that the present state of the art in the lower frequencies makes any power technically possible that can be economically justified. However, as we pass through the medium-frequency range into the high-frequency range and above, there is, at present, a definite technical limitation on power vs frequency.

Notwithstanding the improvements in reliability as the result of new signaling systems and new knowledge of propagation, the demand for greater reliability has always resulted in some increase in transmitter power, usually with simultaneous improvements in antenna systems. Therefore, the trend toward higher-power transmitters will continue.

Microwave Communications*

JOSEPH H. VOGELMAN†, FELLOW, IRE

Summary—**An historical summary of the development of microwave communications is presented, together with an examination of the present state-of-the-art. The current research and development efforts are used to forecast the extension of the present applications and techniques into the future.**

Introduction

THIS paper is an attempt to summarize the field of microwave communications, past, present and future, within an imposed limit of 4000 words. With almost this number of references bearing on the history and technology of microwave communications and hundreds of important contributors to its development, the allowable space precludes tabulation of references and acknowledgement of the individual contributors. To those whose names are omitted and those whose papers are not referenced, the author offers his apologies. For the purpose of this paper, microwave communications will be limited to technology where wavelengths shorter than 40 cm are used. Within these limits four major communications media will be considered:

1) Microwave relay using line-of-sight propagation.
2) Beyond-the-horizon systems using tropospheric scatter propagation.
3) Beyond-the-horizon systems using man-made satellites as reflectors or repeaters.
4) Closed circuit systems using the TE01 mode of propagation in circular cross-section waveguides.

The first two have reached maturity in the form of numerous operating systems around the world. The latter two are by comparison in their early infancy, yet on the threshold of full-scale operations.

The Past

In 1864 James Clerk Maxwell laid the foundation for microwave communications with the electromagnetic wave theory he used to explain the phenomenon of light-wave propagation. In 1888 Heinrich Hertz generated and received electromagnetic waves with wavelengths of 60 cm using a spark transmitter and parabolic antennas. By 1898 other experimenters using similar methods had generated wavelengths as short as 0.4 cm.

In 1893 J. J. Thomson published the first analysis suggesting the possibility of waves in hollow pipes in his book "Recent Researches in Electricity and Magnetism." In 1897 Lord Rayleigh published an analysis of propagation through dielectrically-filled pipes, both of circular and rectangular cross section. In 1895 Victor von Lang propagated electric waves in pipes of appreciable length.

The apparent advantages of the longer wavelengths, as demonstrated by Marconi, and the advantages of vacuum tubes for longer wavelength operation temporarily curtailed progress in microwaves.

In 1920 the Barkhausen positive-grid oscillator provided a means for the efficient generation of 40-cm waves. This revived the interest in the centimeter waves. In 1929 Andre G. Clavier, then associated with Laboratoire Central de Telecommunications in Paris, started an experimental project to challenge the then accepted principle that wire or cable circuits should be used in preference to radio whenever physically possible.

In 1930 a link was started between two terminals in New Jersey using 10-ft parabolic antennas. Just as testing started, the project was transferred back to France. On March 31, 1931, Clavier and his associates demonstrated that microwave transmission provided a new order of economy, quality, dependability and flexibility in communications over a 40-km path between Calais and Dover. The circuit provided both telephone and teleprinter service using 17.6-cm waves transmitted in a 4° beam by means of a parabolic reflector 3 meters in diameter with a power output of a fraction of a watt. Andre Clavier went on to establish the first commercial microwave radio link in 1933 from Lympne, England, to St. Inglevert, France.

In 1933 Marconi published results showing 500-Mc propagation well beyond the horizon.

In 1936, from the Bell Telephone Laboratories, Carson, Mead and Schelkunoff published their mathematical theory on "Hyper-Frequency Waveguides" while G. C. Southworth published his experimental results. These papers provided the basis for the TE01 mode cylindrical waveguide. In that same year W. L. Barrow of M.I.T. published his work on the "Transmission of Electromagnetic Waves in Hollow Tubes of Metal." Before 1934 Southworth had transmitted telegraph and telephone signals at 15-cm wavelengths in a 5-in diameter hollow metal pipe 875 ft. long with relatively small attenuation.

In July, 1939, Clavier published the results of his experiments on millimeter wave transmission, pushing the hollow circular waveguide closer to the day when it would become a major communications medium. In the early part of this same year magnetrons producing small power at 30 Gc came into experimental use.

In 1941 the first "beyond-the-horizon" 3000-Mc propagation tests were undertaken by Andre Clavier around Toulon in the south of France from May to

* Received by the IRE, July 20, 1961.
† Capehart Corporation, Richmond Hill, N. Y.

December of that year under a contract with the French Department of Defense. The transmitter consisted of a 10-watt frequency-modulated klystron. The antennas were 25-degree horns with the transmitter on land and the receiver aboard ship. A sizable gain margin of received signal above the noise level was found at 103 miles from the transmitter and 43 miles beyond the horizon. These experiments furnished basic information for practical utilization of microwaves in beyond-the-horizon links.

In 1944 Andre G. Clavier and V. Altovsky published the results of experimental work which established the technique of frequency modulation as the optimum method for use in microwave line-of-sight links. In 1946 a 12-channel FM microwave link was incorporated in the French Postal Telegraph and Teletype System between Paris and Montmorency.

During the same period the Bell Telephone Laboratories under Harald T. Friis were engaged in extending the one-hop microwave radio link into a long-distance repeater network. Their microwave repeater research delayed by the five war years, culminated in a New York-Boston system; results were published by Friis in April, 1948, in the *Bell System Technical Journal*. The horn-reflector antenna and the shielded metal lens antenna joined the paraboloid as basic antenna types for microwave relay systems.

Since 1950 the National Bureau of Standards, the Lincoln Laboratory of the Massachusetts Institute of Technology, and the Bell Telephone Laboratories conducted many theoretical and experimental investigations of microwave tropospheric propagation beyond the horizon.

The forward scattering is explained by H. G. Booker, J. T. de Bettencourt and W. E. Gordon in terms of a granular turbulent troposphere with "blobs" of refractive indexes that differ from the average value. It is explained in terms of partial reflections on various refractive layers by T. J. Carroll and R. M. Ring, and H. T. Friis, A. B. Crawford and D. C. Hogg. K. Bullington, the Lincoln Laboratory research group, D. Davidson and A. J. Pote, and members of the National Bureau of Standards have contributed greatly to our knowledge of microwave tropospheric propagation. This work has formed the basis for the implementation of beyond-the-horizon microwave communication.

In 1955, The United States Air Force started the installation of an operational chain of tropospheric scatter stations in an inaccessible region of Labrador. Within the next two years the world's first commercial over-the-horizon link from Sardinia to Minorca was installed to provide direct telephone communications between Spain and Italy.

The first step towards a communications satellite was made on October 4, 1957, with the launching of Sputnik I by the U.S.S.R. While this device operated in the HF range, it was the first communication transmitter to orbit the earth.

On December 6, 1958, Pioneer III took the first microwave transmitter at 960.5 Mc into space for a 38.1-hour trip. This was quickly followed on December 18, 1958, when the United States placed the first satellite communication system in orbit as part of Project Score. While this system operated at VHF, it was the forerunner of microwave "store and forward" message transmission service.

The passive communication satellite repeater became a reality when the Echo Satellite was launched by the National Aeronautics and Space Agency on August 12, 1960. It provided an effective mechanism for microwave communications experiments by the Bell Telephone Laboratories and Jet Propulsion Laboratory of California Institute of Technology along a cross-country east-west path, by the Rome Air Development Center on a north-south path, and by Collins Radio between Iowa and Kansas.

At the time of this writing, the electronics-communications industry is negotiating for the launching of a repeater satellite by NASA to inaugurate commercial communications on an intercontinental basis. This will give microwave communications a global capability not heretofore possible.

The Present

The state-of-the-art in microwave communications has advanced markedly in the thirty years since the first demonstration of a practical microwave single-hop communications link.

The Line-of-Sight Relay

Today microwave line-of-sight repeater systems are found all over the world providing reliable broad-band channels between major terminals, carrying thousands of voice and data channels, as well as television broadcasts. The accepted modulation method is FM with diversity in frequency serving as the prime method for insuring propagational reliability. Line-of-sight microwave links provide a communications method for major military traffic between remote sites and communications relay centers. Power levels in the order of one watt are used with ratio-squared-combining in dual-frequency diversity or combined space-and-frequency, four-fold diversity systems. The antenna systems consist of parabolic antennas and plane reflectors, shielded metallic lens antennas, and horn-reflector antennas in various combinations and sizes. Frequencies as high as 12 Gc are in common use, and transmission bandwidths of 50Mc are not uncommon. Ferrite load isolators have been introduced to insure a high degree of decoupling of the antenna reflections from the klystron transmitter and to minimize intermodulation. These contribute to the achievement of the extreme linearity required in long-haul multichannel microwave systems for transmission of color television signals. The line-of-sight microwave relay has become a standard medium for public as well as private communication systems.

Tropospheric Scatter

The major obstacle to the use of the line-of-sight relay has been the long over-water hop. Tropospheric scatter provided the transmission medium to overcome this natural obstacle as well as adverse combinations of geography and climate. Since the original Sardinia-to-Minorca installation, tropospheric scatter has developed to a world-wide importance, predominantly as an essentially American long-range communication system. The major impetus behind this development has been provided by the Department of Defense agencies. The path length for a single hop of tropospheric scatter has been extended beyond 700 miles while the number of channels of 4-kc bandwidth has been increased to 240. Transmitter power levels of 50 kw have been attained with antennas as large as 120 ft in diameter. At the other extreme, transportable tropospheric scatter systems installed in vans or shelters are finding more and more applications in the various military services. Systems have been installed to operate in the 755–985-, 1700–2400-, 4400–5000-Mc bands and experimentally in the 7125–8500-Mc and higher-frequency bands. The feasibility of ground-to-air use of tropospheric scatter out to 1000 miles has been demonstrated, though never implemented, as an operating communication system. This communication medium has been integrated into a long-range repeater network system for commercial as well as military use. The largest system at this time is the combined U. S. Air Force White Alice Dew-Line system which ties together all Alaska and northernmost Canada. The White Alice system has become the backbone of the far-flung commercial communication system of the State of Alaska.

Data transmission tests of multiple-hop paths have proven the inherent reliability and over-all linearity of microwave communications medium. Since the basic propagation medium is subject to deep fading of very short duration, various diversity systems have been proposed, tested and applied to reach the required reliability. The system in most frequent use is space diversity where two or more antennas are provided at both ends of the link for receiving and transmitting. In combination with polarization diversity, dual-space diversity is extended to quadruple diversity over a duplex communications link. Time-diversity reception and multi-beam angle-diversity systems have also been demonstrated. An angle-diversity system, designed at the Rome Air Development Center, utilizes a very large (compared to the wavelength) antenna with multiple beams clustered about the great circle path. In combination with frequency diversity it provides reliable communications for long-range hops at frequencies of 8000 Mc and higher. The signals received over these uncorrelated paths are then combined for greater system reliability.

Various combining systems have been successfully used. These include selector, linear-adder, and ratio-square combiners. In the selector combiner, the largest signal is accepted while all others are rejected. This type of combining does not benefit from the power received from smaller signals. The linear adder combines all signals by simple addition, giving equal weight to all signals. In this type a deep fade on one path results in the addition of noise to the other signal, thereby degrading it. The ratio-square combiner squares the level of each signal before combining is performed. It thereby gives emphasis proportionate to the level of the signal above the noise and comes within a decibel of the idealized combiner. Combining has been used in the IF circuitry where the local oscillator is varied to phase the IF signals to produce two signals for optimum addition. Post-detection combining is also found in many systems. In several systems with quadruple diversity both IF and post-detection combining are used. Over-all reliabilities as high as 99.999 per cent have been demonstrated in transmission of digital data.

Satellite Communication Systems

Microwave communications by means of a satellite repeater became a reality during 1960. While the circuits are considered experimental, they have demonstrated operational capabilities.

The Echo passive satellite was placed in orbit on August 12, 1960, and signals were transmitted between the Bell Telephone Laboratory installation at Holmdel, N. J., and NASA's Jet Propulsion Laboratory at Goldstone, California. This passive repeater consisted of a 100-ft diameter metallized plastic balloon in orbit about 1000 miles above the surface of the earth. The very-low-noise receiving system uses a low-noise horn reflector antenna, a broad-band maser amplifier using a ruby crystal refrigerated by liquid helium and a receiver using a feedback technique for recovering an FM signal which would ordinarily be lost in the noise. The combination of antenna and maser provide a background noise power about 1/100 of that encountered in ordinary microwave radio systems.

Simultaneously the Rome Air Development Center has been communicating on a north-south path using the same passive repeater satellite. They were the first to receive a message reflected from the man-made satellite. The transmitter is located at the RADC Trinidad, British West Indies site, and the receiver site is at Floyd, New York. A radar tracker at Trinidad is provided with a separate feed for simultaneous transmission of radar and communications signals. The radar acquires and tracks the satellite using normal back-scatter signals. Direction-finding antennas at the Floyd receiving site detect the radar signal and orient the 33-ft diameter parabolic communications receiving antenna. The communications signals, teletype signals as well as data signals, at 2270 Mc consist of PB-50 word lists for articulation tests. Both of these paths provide microwave communication links capable of extension into operational military as well as commercial systems using passive satellite repeaters. Collins Radio has been operating between Iowa and Texas.

The active repeater satellite Courier was placed in orbit October 4, 1960. This satellite is a delayed-repeater communications system capable of voice, facsimile and high-speed digital transmission. The satellite is a 52-in-diameter sphere weighing 500 lbs. Eighty per cent of the surface is covered with 19,152 solar cells to provide power to the electronic components. The satellite receives at a frequency in the 1.7- to 1.8-Gc band and transmits between 1.8 and 1.9 Gc. The storage capability is 13,200,000 bits in each of four magnetic tape recorders. The transmission rate is 55,000 bits/second. The ground stations transmit at a power of 1000 watts while the satellite transmits 4 watts. The stored data is retransmitted to ground upon ground command. Four minutes are required to transmit the storage of each recorder. In addition to the delayed mode, the Courier satellite can be commanded to act as a real-time repeater and retransmit immediately. The orbit is at 750 miles above the earth. The satellite has been used to transmit over 6,000,000 words of teletype information daily between the ground station at the United States Army Signal Research and Development Laboratory, Fort Monmouth, N. J., and the Puerto Rico station located at Salines, during periods of visibility.

TE01 Circular Waveguide

The major interest in the circular electric-mode TE01 waveguide has been stimulated by the need for large bandwidth toll trunks capable of long-range usage. The essentially unlimited bandwidth available in the millimeter wavelength band and the ability to obtain any reasonable attenuation by increasing the diameter of the tube has promoted great interest among telephone companies. The Bell Telephone Laboratories in the United States have developed the necessary components for long-range application of this waveguide system. Necessary mode filters as well as techniques which will allow reasonable radii of curvature have become available. A helically wound copper wire structure of 2-inch diameter is used as the basic structure. The tightly-wound helix is covered with a mode absorbing compound and enclosed in a steel tube. In England Standard Telecommunications Laboratories of the Standard Telephone and Cables, Ltd., demonstrated a flexible cylindrical waveguide for TE01 mode operation. A guide of approximately 3 in. in diameter is formed with helically wound aluminum wire. This tightly wound helix is covered with a water-proofing and a strengthening coating. In this structure type the electromagnetic waves follow small bends in the tube without prohibitive loss or excessive mode conversions. This waveguide structure is expected to operate in the millimeter wavelength band with an attenuation of 2 to 3 db per mile. Repeaters will be required every 20 miles to maintain a reasonable signal-to-noise level at the output terminals. The first demonstration of an operating cylindrical waveguide was held in 1960 at Frogmore Hall, Herts, England, using a 3600-ft-long loop of $2\frac{3}{4}$-in-diameter waveguide. The loop combined an 1800-ft helical section with an 1800-ft optically straight tube. Good quality television pictures were transmitted. For long-range operation the differential phase delay and the residual unwanted mode energies contribute to the distortion of the signals transmitted. To permit undistorted regeneration of the original signals, pulse-code modulation is essential and is proposed as the applicable technique. Spectrum bandwidths of 10,000 Mc and greater are obtainable in this waveguide mode. A bandwidth of this size will be sufficient for at least 400 television stations or a million voice channels.

The Future

The growth of microwave communications in the last twenty years has been phenomenal, and it portends becoming the major communications medium of the future. The pressure of frequency unavailability has forced and will continue to force more and more services into the higher microwave frequencies. A realistic future must be an outgrowth of the research and development programs of today. On this basis the future is presented.

Line-of-Sight

The basic line-of-sight microwave system will continue to be the mainstay of the broadband network. The power levels used will increase. The operating frequencies will go higher. In addition to the relay function, there will be increased usage of connecting links between subscribers and a communications relay center for retransmission by a more complex long-range system. Space diversity will be used in addition to frequency diversity for added propagation reliability. (The greater power and higher antenna gains will permit greater spacing between stations or increased signal bandwidths.) Except for improvements in equipment and components, the major change will be in extending the microwave line-of-sight service to vehicular terminals, both air-breathing and orbital. Solid-state components are receiving major emphasis for these and other mobile uses. Suppressed-carrier single-sideband multiplex techniques will be used to compress more channels in the same bandwidth. Wide index FM will allow greater concentration of stations.

Tropospheric Scatter

There is an urgent need for tropospheric scatter operation at higher microwave frequencies to relieve the RF interference problems. In many parts of the world only frequencies above 5 Gc can be made available for use in tropospheric scatter. The multiple-angle diversity system, under experimental evaluation at the Rome Air Development Center, promises to overcome the higher losses and restricted transmitter powers at the shorter microwave wavelengths (less than 4 cm) by better matching between antenna aperture and medium. When the 200–250-mile hops are operating at the higher microwave frequencies, the lower end of the band can be designated for long hops of 750 to 1000 miles.

The number of multiplexed voice channels is gradually being pushed towards 600, while the transmitter power is approaching 100 kw. The use of parametric amplifiers in the receivers will increase the signal-path reliability with a resultant increase in the over-all reliability. This will insure long-range multiple-hop relay systems capable of handling high-speed digital data.

Forward scatter at 400 Mc has been received at great circle distances in excess of 2500 miles. The propagation mechanism involved is believed to be that of incoherent scattering induced by electron-density fluctuations in the common volume of the antenna beams. Enough is known to postulate a special-purpose communication system designed to use this region for forward scatter.

Satellite Relays

Global communications by satellite relay is currently being undertaken as a commercial venture by the communications industry of the United States. A wide variety of systems have been proposed for the future. Among the more promising are the following.

Passive Satellites: The passive satellite offers the advantage of reliability and minimizes the complexity of the space vehicle. It requires greater transmitter power and extreme receiver sensitivity on the ground, but where a great flexibility is needed it will probably be the logical solution. Four types of passive satellites have been proposed. The first is a large-diameter spherical reflector based on the successfully orbited Echo satellite. It offers advantages of simplicity, since it does not require any stabilization. The second is the diffuse scatterer, made up of a belt of chaff dipoles in orbit. While it simplifies the tracking problem, the launching is greatly complicated by the need to space the dipoles in their orbits and maintain them in precise relative positions. The dipoles would be subject to solar pressure. The third satellite is a partially stabilized Van Atta array whose distance from the earth is such that its beamwidth includes the whole earth. Since the array will return energy in the direction from which it is received, over an angle of 45 degrees, the requirement of stabilization is greatly reduced. Lastly, the Lambertion surface has been proposed as a diffuse scatterer. Any surface of sufficient irregularity and roughness will reflect according to Lambert's law. A Lambertion surface cannot be distorted enough to produce an undesirable effect, and therefore a spherical shape is not necessary to produce a uniform bistatic cross-section. Crumbled sheets of reflective foil expanded into a loosely fitted covering on a support structure would suffice if the wavelength is short in comparison to the distribution in elevation of the individual scatterers.

Active Satellites: The active satellite relay divides the equipment complexity between the repeater and the ground environment and results in a reduction in the power requirement. The two types of active satellite are the "store and forward" or Courier-type communications satellite and the instantaneous repeater which shifts the frequency while reproducing the modulation. The Courier type offers global coverage in a single satellite vehicle for those services where several hours delay are tolerable. The instantaneous repeaters required a multiplicity of satellites as well as ground relay points to produce a global system. An active Van Atta array using amplifiers between corresponding elements has also been proposed as a communications satellite.

Various orbits have been proposed for both active and passive satellites. The stationary or synchronous satellite (24-hour orbit at the equator) has very useful orbital properties, but its 22,400-mile height above the earth complicates the electronic components of the communications system and introduces a significant time delay. However, the reduced tracking requirement and the continuous availability compensate for these complications. Polar and 45 degree orbits at 1000- to 3000-mile heights above the earth are the most readily realizable in the immediate future. Their use requires extensive tracking antennas at both ground terminals as well as provisions for jumping from satellite to satellite if continuous operation of the communications system is to be achieved. Various other satellites have also been proposed with orbits from 90 minutes to 96 hours.

Circular Waveguide

The TE01 mode circular waveguide will ultimately replace the multiconductor cable and the coaxial line as the carrier of underground or overhead trunk traffic in urban areas and between adjacent urban areas.

Conclusions

In the nearly 100 years since Maxwell microwaves have indeed become the highway of our electronic communications, but this is only its infancy.

Section 9

COMPONENTS

Organized with the assistance of the IRE Professional Group on Component Parts

Historical Development of Component Parts Field*

J. T. BROTHERS†, FELLOW, IRE

Summary—**This paper reviews the history of the component parts field by equating the effects of the invention of active components to the resulting regenerative expans on of requirements for types and capabilities of components needed to make possible the development of new systems equipment to meet the demands of applications and, in turn, generate new applications for other system equipment.**

IN REVIEWING the history of the component parts field in the electronic equipment area, the most impressive factor that becomes evident almost immediately is the immensity of the magnitude of human endeavor in the entire field of natural philosophy, to use a historic term, that has resulted in our present day technological status. The contributions of Henry, Maxwell, Ampère, Volta, Kelvin, Ohm, Helmholtz, Herz, the Curies, Marconi, Loomis, Faraday, Edison, Fleming, to name a very few specifically, were known well enough to their contemporaries so that they have been signally honored; our recognition of their successes is perpetuated in some cases by the use of their names for units of measure, in other cases by the use of their names for certain effects. For each of these, however, there were thousands more who added their important contributions to the over-all sum of knowledge that formed the foundation of the wireless era at the beginning of the twentieth century. To all of these en masse we must pay homage and acknowledge our indebtedness.

The following sixty years may be arbitrarily divided into double decades and equally arbitrarily we can designate the first, to 1920, as the wireless generation; the second, to 1940, as the radio generation; and the third, to 1960, as the electronics generation. During the last fifty years, since 1912, which is at this point the historical part of the life of the IRE, there has been an explosive expansion of the capabilities in materials for, and concepts of, component parts for electronic equipment. This has been the effect of the continually increasing amount of research work done in the basic sciences field in both educational and industrial laboratories, the shift

* Received by the IRE, August 1, 1961.
† Philco Corporation, Philadelphia, Pa.

of emphasis in scientific work from the technique of experiment to the technique of analysis, and the aggressive exploitation of this work by further development work in the engineering field. These additional capabilities of the past 60 years are in turn due to the accomplishments of a myriad of workers in these fields and when we acknowledge the contributions of Alexanderson, Cady, de Forest, Dubilier, Eastham, Kusunose, Langmuir, Néel, Polydoroff, Rubin, Thordarson, we do it in the full realization that these are only a few of the important contributors who were responsible for the current state of the art and entitled to our eternal appreciation. It is beyond the scope of this paper to credit individuals adequately. It is also beyond its scope to cover the field completely or cite much more than a few examples.

In this short summary of the past which serves as an introduction to the other papers in this series it is necessary to bear in mind the cyclical nature of the end equipment requirements since the turn of the century. An example of this is the part of the frequency spectrum used. Although a great deal of work was done initially in the gigacycle per second region, by 1900 our transmitter frequency range for commercial purposes was in the 10- to 100-kc region; the emphasis shifted progressively higher in frequency and by 1920 the dominant range was 100 to 1500 kc (although there had been a great deal of emphasis on very low-frequency ground radiation telegraph systems during World War I for military communication), by 1930 the frequency range had shifted to between 200 kc and 20 Mc, by 1940 this had extended to 200 Mc, by 1950 important applications had been using the frequency range to the 10 Gc region, and by 1960 high-power stations were once again being built in the 15- to 50-kc range. At this point we have filled the gap between the 1900 spectrum usage, the one area of usage about a center frequency of 50 kc and the other area around a center frequency of 1 Gc, and extended the range to somewhat higher frequencies. Another significant point to keep in mind is the change in the method of generating radio-frequency power. Spark, arc, and mechanically rotated alternators have been partially superseded by vacuum tubes of all kinds, triodes, klystrons, magnetrons, and others, more recently the semiconductor devices have staked out a part of the field, each with its own applicable part of the power range and frequency spectrum, but it is interesting to note that the quenched gap spark system is still being used for low-frequency industrial and medical applications. It is equally significant that the magnetron remained a laboratory curiosity for 10 years until the need arose for a laboratory signal generator for UHF research and that another 5 years passed before the requirement for a significant amount of power in a specific frequency range accelerated the development of the tube to its present capabilities. Basic power supply systems have also gone through a full cycle. Fifty years ago transmitters were designed to operate from dc generators and/or storage batteries for the arc generators; the spark systems were operated from ac generators, (or mechanical interruptor which operated in turn from dc or low-frequency ac) that provided 250 cps, 500 cps, or 900 cps power, although some spark systems with nonsynchronous rotary gaps were operated from 50/60 cps lines. In the radio generation the dc required for vacuum tube transmitter plate supply was obtained successively from dc generators, mechanical rectifiers, pool type mercury arc, gas, and high-vacuum thermionic rectifiers, some code transmitters used raw ac of whatever frequency available in a self-rectifying circuit, and in 1930 there were three or four broadcast transmitters using storage batteries for plate supply. Airborne requirements were satisfied by batteries and dynamotors until the higher power requirements of the late 1930's and the consequent requirements for weight reduction revived the use of the higher frequency alternators, from 400 cps to 2400 cps with rectifier systems. This trend has continued in the electronic generation so that the larger airborne systems are now 3-phase 400-cps systems. The rectifying systems are also going through a transition phase from thermionic rectifiers to multicrystalline solid-state devices such as selenium, to single-crystal solid-state devices made from germanium or silicon, and now we have the silicon-controlled rectifier.

The historical background of the component parts field should be viewed from this perspective of ever-shifting requirements, purpose, and nature of the end equipment. In 1910, in order to build a receiver we would need a tuner, a detector, and a pair of headphones as shown in Fig. 1. The first element of the tuner, the loading coil *L*1, would be a coil in which the inductance could be varied by means of a tap switch; the second coil would be a pair *L*2 and *L*3, a loose coupler, with means for varying the coupling between them, preferably with different diameters and mounted coaxially so that the coupling could be adjusted by sliding one inside the other and the inductance of each variable by a tap switch similar to the loading coil. The detector could be one of several options:

1) A Marconi magnetic detector, the real basis of reliability for that era.
2) A Fleming tube—very unreliable because of short filament life and complex because it required a filament battery and rheostat.
3) A silicon carbide crystal detector—preferably entailed the use of a polarizing voltage, adjustable by means of a potentiometer to the most sensitive part of the rectification curve.
4) A de Forest triode—not very well known as yet, also very complex in that it required a filament battery and rheostat, a plate battery that was tapped or adjustable to a critical value by a potentiometer; further, it had a probable life as a detector of 20 to 30 hours and the filament could be expected to burn out in less than 50 hours. The

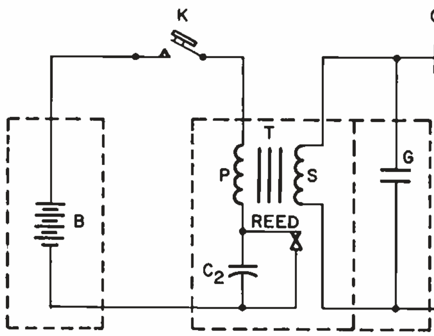


Fig. 1—Transmitter.

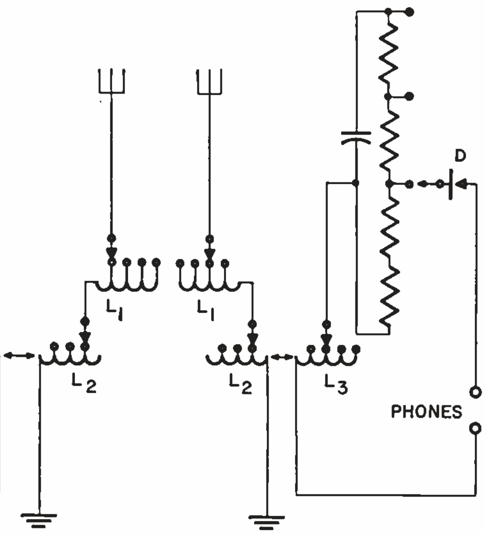


Fig. 2—Receiver.

headphones were pretty much as we know them today, preferably 3000 to 6000 ohms resistance for the pair in series, and a resulting impedance of 4000 to 8000 ohms at 1000 cps. A simple spark transmitter of the same era would be as shown in Fig. 2. *L*1, *L*2, *L*3 serve the same function as their counterparts in the receiver and if the power to be generated was within the dissipation capability of the receiver coils, they could be identical; usually the transmitter coils were wound of bare copper or brass strip, and were considerably larger physically because of the conductor size and spacing. The loose coupler could be cylindrical and coaxial just like the receiver coils, or it could be made as a pair of helical pancakes and the coupling adjusted by varying the spacing between the two coils on the axis, or shifting the axis of one coil. *C*1 was usually a Leyden jar, but it could have been made with brass electrodes inserted between flat glass plates. *G* is the spark gap which could be fixed or rotary. *T* is the interrupter coil, very similar in principle to an automobile ignition coil (as a matter of interest, somewhat later many amateur stations were built which used Ford coils), *P* being a few turns of heavy wire, *S* being a large number of turns of fine wire, both being wound on an open core composed of a bundle of iron wire, and the reed being physically located at the end of the core with the interrupter contact. *C*2 is the point capacitor and was made of sheets of mica interleaved between tinfoil or copper foil. *K* is the conventional key of today, *B* was the battery. As in the receiver, there were several alternatives. In the spark transmitters the battery shown in the *B* block could be replaced by a source of ac such as the 50/60 cps from a commercial power line, or the 115–120 cps generated on some ships, or the 250, 300, or 500 cps alternators specifically designed for this purpose, in which case the interruptor shown in the *T* block, from which frequencies from 250 to 1000 cps were obtained by design, was replaced by power transformers designed for this use. The simple spark gap in the *G* block could then be replaced by a nonsynchronous rotary gap for 50/60 cps, 115–120 and 250 cps supplies, or by a synchronous rotary gap or by a quenched gap for the higher frequency supplies. Similarly for an arc system the *B*, *T*, and *G* blocks were replaced by a dc generator and an arc converter. We should be reminded at this point that, despite the apparent crudity of the equipment of this period viewed with 1962 eyes, transoceanic wireless telegraphy was a flourishing business, ships at sea received news and directions, automatic code recorders were in use, voice transmissions had been sent over 350 miles and opera had been broadcast from both Chicago and New York City.

With equipment of this type, however, there were relatively few components as we know them today. Resistors were only used for adjusting the bias on silicon carbide detectors and these were usually air-supported turns of resistance wire connected to the contact points of an open rotary switch, miniature versions of the same type which were used to control field excitation on motors and generators. Because of the frequency ranges involved, receiving inductors were solenoids wound on large-diameter tubes (4 to 8 inches) adjustable by means of taps brought out to rotary switches, and the coupling varied by sliding one coil inside the other or rotating the axis of one coil with respect to the other. Transmitting inductances were helices of flat strap tapped by clamps that could be fastened directly to the strap, or a slider that could be rotated in continuous contact with the strap; the other popular style was flat strap wound edgewise into a cylindrical coil with similar clamps or a rolling or sliding contact that could be driven around the inner or outer periphery of the coil by a screw mechanism. Some of the lower power units were made with prefixed taps brought out to tap switches similar to those used for receiving purposes. Similarly, coupling between the coils was varied by sliding cylindrical coils inside each other, and the couplings between the pancake styles were adjusted by separating the axis edgewise or separating them on the axis. Supporting structures were wood, hard rubber or asbestos sheet depending on the size of the coil. Very few capacitors were used and most of these were Leyden jars arranged in series or parallel, depending on the voltages and capacitances required, in trays or holders. Where capacitance variation was required it was customary to remove jars or separate the trays. There were a few capacitors used which consisted of sheets of glass with electrodes on each side supported in racks immersed in oil. There were also a few variable capacitors used that were made much the same as present styles with semicircular parallel plates arranged so that the alternate plates could be rotated with respect to the fixed plates, but most of these were made so that they could be immersed in paraffin oil. Transformers were only required for power-supply purposes and were principally high-reactance types on shell cores, mostly open core and coil, although for very high power they were immersed in oil. Modulation was accomplished by means of carbon grain microphones cooled with oil, water, or forced air, inserted directly in the antenna circuit and usually arranged so that they could be rapped with a mallet when they stuck due to excessive current. The battery systems in vogue at that time were wet primary systems, such as bichromate, Bunsen, and Edison types, very few Le Clanché as we know present dry batteries, and the lead-acid secondary type.

It should also be remembered that the emphasis on receivers prior to this time had been visual indicators, such as galvanometers and coherers with their associated ink recorders, or perforators, on paper tapes rather than for audio indicators such as head phones; this is the reason why so much work was devoted to coherers, magnetic detectors, electrolytic detectors, etc. It was about 1910 that the emphasis changed to increasing receiver selectivity and sensitivity and the use of audio indicators (too many stations!).

The third quarter of this generation, 1910 to 1915, saw a marked change. One of the most important factors, from a long-range standpoint, was the recognition that the thermionic triode was capable of providing amplification and this almost simultaneously sparked cascade amplifiers and feedback; it was also recognized that the real function of the filament was to provide electrons and that high vacuum was not only desirable but necessary. This led to the replacement of the tantalum filament by tungsten, and accelerated means for obtaining higher vacuum. Simultaneously damped wave-spark transmission began to lose out to CW generated by arcs, alternators, and triodes. Special ballast resistors, iron wire in air and in hydrogen, were introduced to prolong triode filament life. The high input impedance of triodes provoked changes in coil design, the spider web coil was designed to decrease distributed capacity, facilitate coupling between coils, and increase Q; further emphasis on Q increased the usage of Litz wire for receivers. Improved magnetic systems were applied to headphones and the moving coil horn loud speaker with a separate electromagnetic field was produced.

The last quarter of the wireless generation, 1915 to 1920, had a considerably greater slope in the state of flux curve, accelerated beyond normal by the demands of World War I. The hard-vacuum triode was by now a reality, due to tungsten filaments, high-temperature bake-out of parts during evacuation, and diffusion vacuum pumps. Transcontinental wire telephony was made possible by the use of triode repeaters. Speech was transmitted to Honolulu and Paris from Arlington by means of a tube transmitter and it was demonstrated that the telephone land line could be coupled directly to the transmitter and used to talk to ships at sea. Simultaneously improvements in frequency and power made the alternator practical and the corresponding development of the magnetic amplifier for modulation combined to make practical another system for high-power radio telephony. During this period the lid of Pandora's box was opened wide and all of the problems of the past became small by comparison. Special resistors were required for grid bias supply and grid leaks were made by impregnating paper with carbonaceous inks or by carbonizing cotton string, power resistors for transmitter circuits became necessary and were made by winding bare resistor wire on ceramic and slate forms and protecting them with japan and vitreous enamel coatings. Grid coupling capacitors were required for receivers and made by stacking mica sheet and tinfoil together, the assembly rivetted together with protective insulation covers, and transmitter requirements for similar devices for tuning and coupling were solved by potting equivalent designs into ceramic and metal cases. Filter and bypass capacitors were required for high-voltage supply in transmitters and were made by winding rice paper and tinfoil into flat coils which were then impregnated with sulphur or castor oil and inclosed in metal containers which were sealed with bituminous waxes. Audio amplifiers created a demand for small signal transformers which were made with soft iron wire cores inserted through the core tube of the coil and then bent around the coil to form a closed core. It was not long before the leakage flux and the physical proximity of these transformers and the gain in the system made it necessary to enclose the transformers individually in cascaded iron and brass shield housings. Power transformers were required to step up the voltages for vacuum-tube rectifiers, and it was necessary to enclose the paper interlayer coils and cores in steel cans and seal them with bituminous pitch. Transmitting inductors started to shrink in size as the frequencies used increased and required better design and insulation. Honeycomb coils were invented for receiver tuner use and the wider frequency ranges required for tuners made the bank wound solenoid necessary. Broad-band iron core transformers were developed for untuned RF amplifiers. Spark transmitters were designed for aircraft to operate from batteries and separate wind driven generators. Small tube transmitters borrowed the old spark coil interruptor and developed it into a vibrator supply to transform 20-volt dc systems into high voltage at 350 to 500 cps for rectification by self-heated cathode gaseous rectifier tubes to supply 1000 volts dc at 50 ma for portable ground and aircraft use. The iron-wire magnetic recorder was invented.

The first quarter of the radio generation, 1920 to 1925, is perhaps best signified by the addition to the professional engineers, scientists, and experimenters previously engaged in the field, of an entirely new group of amateurs who became interested as a result of their exposure to wireless during the war, the publicity enjoyed by the "new science," or the advent of broadcasting. In any event, this was a period of digestion in the commercial field; silicon carbide, zinc chalcopyrites, and other semiconductor detectors were superseded by galena; more vacuum tube detectors came into use for shipboard use; arc and spark transmitters were the mainstay of the ships; arcs and alternators were the fixed location telegraph transmitters; and the vacuum tube reigned supreme in the new broadcast stations. The amateurs were banished to the nether regions below 200 meters where rotary gap spark transmitters created a veritable cacophony, and the serious amateur began to use vacuum tubes for both transmitting and receiving as soon as he could financially afford it. The principal changes in the component field were caused by the increase in production of radio receivers for the amateur band and the broadcast band. In both of these there were additional requirements for rheostats for filament controls which were made of wire wound on fibre, or asbestos covered steel forms, supported by brackets, or made of carbon, graphite and mica dust enclosed so that they could be compressed by a threaded screw plunger, or carbon disks similarly arranged. Filament voltage adjustment was the conventional means for controlling gain. Fixed and variable grid leaks made from carbonaceous inks on various substrates, with a slider for the variable ones, took the place of the previously used graphite pencil leads, and the ink or pencil marks across the grid condenser or piece of bakelite, fixed resistors for all purposes began to be marketed, one style of composition resistor being made of a mixture of binder resin, carbon, and silica which was compressed and cured into a rod to which wire leads were wrapped around the ends; another style was made of a silicon carbide mixture pressed into a rod and fired and finished with radial wire leads; another style was a carbon film on a glass rod which was enclosed in a glass tube, and connections, made to cast soft-metal terminations which sealed the assembly. High-resistance-value potentiometers and rheostats were made from pressed carbon and graphite half-toroids which could be bought as an element or complete with hardware as a completed unit. Variable-inductance tuning systems were found to be less suitable for shunt circuits and more attention was directed to the design of variable capacitors which began to appear in different forms; one style consisted of two coaxial, parallel, circular plates, separated by a small air space in which there was an insulating disk of mica, which were coupled by an insulated screw and nut so that the capacitance was varied by turning the screw and changing the air spacing; the book type consisted of two wooden blocks hinged at one edge arranged so that a cam could vary the angle between the two blocks—the metallic plates on the inner surfaces of the blocks were prevented from shorting, when closed on each other, by a sheet of mica dielectric. The mercury capacitor consisted of a closed cavity free to turn on a horizontal shaft in which a mercury pool in the lower half filled the space between two semicircular plates fastened to the inner surface of the cavity insulated from the mercury pool by mica sheets. Another multi-section capacitor for single-shaft operation had two circular die-cast zinc plates per section; each plate had a large surface area formed by concentric convolutions that were a mirror image of the other of the pair and the capacitance was varied by changing the spacing. Other styles of conventional rotary parallel-plate air capacitors were designed specifically so that they could be placed in tandem on a single shaft and varied simultaneously. Die-cast cradles and other mounting systems were introduced during this period. It was during the latter part of this period that variable capacitors began to be most widely used for varying resonance, and this led to the design of modifications of plate shape from the straight-line capacitance type to what were called straight-line frequency and straight-line wavelength shapes which had the effect of spreading the tuning characteristic more evenly over the dial. As rolled-paper capacitors came into more general use for lower voltage applications, beeswax, paraffin, and castor oil largely replaced sulphur as an impregnant. The introduction of the neutrodyne, and other forms of neutralizing tuned RF amplifiers, produced new varieties of coils, basket weave, spider web, and solenoids mounted at magic angles to each other to reduce feedback and interstage coupling, binocular style coils with limited fields, etc. The spherically shaped variometers were augmented by styles with **D**-shaped coils. Coils made by spraying molten copper on ceramic cylinders and cutting a thread into the ceramic (there were several variations of the process) to create a solenoid were used in low-power transmitters, anticipating the later efforts at "printed wiring." String drives and rack and pinion systems were added to loose couplers to provide more convenient means of coupling adjustment and facilitate back of panel installation. Belts and pulleys, and chain drives, were used to vary capacitance and inductance simultaneously, retrogressing to the Fleming cymometer wave meter of 1904 and the single-dial-control ideas patented in 1912 by Hogan. Sound-chamber adaptors to couple a pair of headphones to a horn "loud talker" were produced, and headphones with conical diaphragms and a variety of magnetic structures were devised for headphones and horns. The electrostatic speaker with a motor-driven drum appeared but won no friends. The friction sound amplifier and loudspeaker with motor power also flashed into the picture and disappeared. The double button carbon microphone made its appearance and contribution to higher quality and, in turn, the moving coil microphone made its contribution. Phonograph pickups, as we would call them now, of a magnetic type were placed on the market, and the enervation of the 1960's began to appear in the form of electric-motor drives to take the place of the spring-motor drive for phonograph turntables. Copper-oxide rectifiers and electrolytic rectifiers were competitors of the Tungar rectifiers and high-vacuum thermionic rectifiers as well as the gas-filled "S" tube. Le Clanché cells with the carbon serving as the container instead of the previously conventional zinc can appeared for a short time to disappear, until the 1960's. The dry-charged lead-acid storage battery was put on the market. The tape-code printer went through another phase and acquired the capability to print alphanumeric characters instead of the code characters and replaced the human being and typewriter link. Vacuum tubes acquired bases and the tube socket became a necessity.

The second quarter of the radio generation, 1925 to 1930, saw the broadcast receiver burgeon into reality principally due to the introduction of receivers capable of operation from electrical power lines. The first phase was the use of conventional filament tubes in series circuits with a dropping resistor directly from dc lines where they were available, and from a rectifier system using two Tungar rectifiers in full wave from ac lines; a simple double pole double throw switch served to convert the set from one service to the other. The second phase was the use of 60-ma battery-type tubes in series circuits from a full-wave rectifier system for all purposes except the final output stage; the output-tube filament could be operated from a filament winding, which provided raw ac, without too much hum. The third phase, of course, saw the application of indirectly heated cathode-type tubes and very low voltage and high current filaments that could be operated from the ac line by means of a stepdown transformer.

With the introduction of the tetrode and pentode, increased quantities of composition resistors were required, as well as greater quantities of wire-wound resistors, principally those made on ceramic forms and covered by japan enamels. Although the reflex systems made a strong bid, the great majority of receivers had tuned radio-frequency (TRF) amplifiers, and the inductors required were smaller in size so that they could be shielded either individually or in compartments, and inductance tolerance and matched sets of coils became a necessity for single dial tuning systems; simultaneously, Q-requirements increased greatly because of the multiplicity of stations. The mica capacitor molded in a bakelite housing was introduced, small paper capacitors were manufactured with paper wrappers and cylindrical cardboard housings, and chlorinated naphthalene became one of the most widely used impregnants for paper capacitors for low dc voltages. Wet electrolytic capacitors were introduced for power-supply filter applications and so-called dry electrolytics were introduced soon afterward. Variable air capacitors with three and four sections on a single shaft closely matched for capacitance vs rotation were required for the TRF sets with single-dial control, and trimmer sections were built into the gangs so that variations in circuit capacitances could be accommodated. Small compression mica capacitors were required to adjust circuit capacitances and for neutralization, although most of the neutralizing capacitors were made of tubing over part of the wiring. The cone loudspeaker became popular with a variety of magnetic drive systems and the flared exponential horn was designed for the phonograph industry. This was not adequate to stem the tide of radio, and the electrical phonograph went into production before the end of this period. Magnetic phono pickups were challenged by the capacity pickup, but the simplicity of the magnetic won out for that era. Improvements were made in the carbon-button microphones, but the condenser microphone became the preferred unit for high-quality work. The electrolytic rectifier and the storage battery supplies became history for applications where ac power was available. Another signpost was erected by the sale of the first battery-operated portable superheterodyne for entertainment use; rotating disk television was exhibited to the public and television signals were broadcast. The wideband coax cable was installed from New York to Philadelphia. One model radio receiver, TRF with 6 tubes, for battery operation, was made in which 90 per cent of the wiring, some of the capacitors, and the tube socket contacts, were automatically stamped from brass strips, positioned and self-rivetted to a bakelite sub panel, anticipating our modern printed wiring system.

The third quarter of the radio generation, 1930 to 1935, was extremely active and showed the direction that the industry was to follow. The super-heterodyne patents were released for license, and there was a frantic scramble to produce and merchandise the "supers." The combination of a radio receiver and a two-speed ($33\frac{1}{3}$ and 78 rpm) electrical phonograph was merchandised. $33\frac{1}{3}$-rpm records were released for general sale. Talking movies were produced in quantity, public-address systems came into general use, and radio receivers were installed in automobiles. Mechanically generated, but electronically displayed, television was demonstrated over the cable, and shortly thereafter all electronic television was demonstrated. All-wave radio receivers were introduced so that foreign broadcasts could be received, and transoceanic programs were broadcast throughout the United States on the newly established radio networks. The vacuum tube was becoming important for many industries, presaging the electronic generation and creating the electronic component industry. The use of composition resistors was increased by the additional requirements of the screen grid tubes, for decoupling and voltage dropping, the increased number of tubes in the super heterodynes, audio amplifiers, etc. and for use in the auxiliary systems, such as automatic volume control, automatic tone control, automatic frequency control, etc. In the capacitor field a similar expansion in demand took place and the bimetal temperature compensating capacitor became necessary to compensate for oscillator drift. Pre-tuned push-button station selectors amplified the use of mica trimmers. This also led to the development of the mechanical push-button system, which rotated the gang capacitor to predetermined positions; new mechanical systems for accomplishing this changed the appearance of gang capacitors considerably. In the inductance category, the super-heterodyne circuits permitted bigger coils and less shielding for the tuning system, but made greater demands on the shielding of IF transformers (of course, we soon found out that we had to do a better job of shield- the oscillator and antenna coils in order to reduce oscillator radiation). The iron-core permeability tuning system became a factor in design. Iron cores became useful in low-impedance antenna transformers for all wave receiving systems and desirable for other antenna coupling systems to improve signal-to-noise ratio. The wave switch became a major component in all wave receivers and made further demands on coil size reduction, and the use of Litz wire for broadcast frequencies became desirable. The push-button switches for the electrical station-selector systems assumed an important role. The electrodynamic cone loudspeaker system became the most universal reproducer, even to driving the horns of large theater and public address sound systems. Power-supply systems for the automobile radio receivers progressed from Le Clanché B-batteries to the dynamotor and then to the vibrator (here we are back to the spark interruptor again) transformer-rectifier system of providing plate power. Vibrator frequencies ranged from 75 to 250 cps, but the 100 and 110 cps vibrators were the most widely used. Electronic musical instruments (?) were introduced as well as electronic versions of the pipe organ.

The last quarter of the radio generation, 1935 to 1940, saw the extension of ac power lines to rural areas and the reduction of sales of battery-powered receivers in that field, but also saw the development of a whole new product line of battery-powered portable radio receivers due to the introduction of the low-drain battery-type tubes, and there was a real increase of sales in that area. Combination radio-receiver–phonographs with automatic record changers became more popular as the efficiency and performance of record changers increased and costs were reduced. The phonograph pickup emphasis shifted from magnetic devices to piezolectric pickups. During this period the performance of variable-composition resistors increased as the demands for lower noise level and greater stability increased. As the demand increased for more accurate tuning systems with less drift and push-button retrace error, particularly in automobile radio, the ceramic capacitors with fixed temperature coefficient of capacitance were introduced, semi-adjustable mica capacitors with bimetallic elements to provide temperature compensation became available, and the temperature coefficient of inductors was reduced or made more uniform. Temperature and time stability of IF transformers was increased by wider use of powdered-iron cores and silver-mica capacitors. The liquid-filled electrolytic capacitor gradually faded from the scene with the improvement of performance of the dry electrolytic, and the ability to make the larger capacitance values required in smaller cans as a result of etched anodes, and sprayed gauze and paper anodes.

This period was characterized by a further extension of radio applications and techniques to an area far beyond the transmission of information and entertainment. Scientific instruments of all sorts were improved or invented for a very wide range of uses in the laboratory and in industry for measurement and control of physical, chemical, and electrical properties, for use in medicine and surgery, for labor-saving applications in the home, etc. Television was almost at the point of emerging from the laboratory. FM radio was just beginning. Specialization in the various aspects of the field became more general, and as the number of radio-receiver manufacturers decreased there were more companies who specialized in more restricted aspects of the component parts field. The individual end equipment manufacturer did not, as a usual rule, have the volume nor the economic justification to pay for the research and development necessary for individual component parts and the mechanization necessary to compete with the component manufacturer who specialized in a limited product area, but who now had a considerably greater sales field than the receiver field alone.

The first quarter of the electronic generation, 1940 to 1945, started in a period that was to be another World War. All of the engineering and research personnel spared from other demands for manpower were engaged in developing new devices, such as radar, bomb sights, identification systems, navigation equipment, bomb and artillery fuses, gun-laying and fire-control systems, computers, and too many more to mention, for operation under a much wider range of environments than previously required or even thought practical. Many of these devices required parts which were beyond the state of the art in order to make the size and weight of the equipment practical.

The second quarter saw the expansion of research and development programs to meet these challenges and simultaneously the introduction of television broadcasting and increased activity in the frequency modulation entertainment field. A whole new class of components were required for the microwave spectrum. Variable inductance tuning systems reappeared and transmission line types of tuners were developed in which inductance and capacitance were varied simultaneously. Electrical and mechanical push-button-type tuners were devised for television and FM.

The third and fourth quarters of the electronic generation, 1950 to 1960, are too close to us to view with much of a sense of perspective. The most significant thing about this period is the introduction of the germanium class of transistors and diodes, the silicon devices that followed, and the other classes of semiconductor materials that are being developed. Cadmium-sulphide light-sensitive resistors, new materials, or the methods of synthetically producing old materials in more suitable shapes, sizes and purity, tantalum electrolytics, both solid and liquid electrolyte types, the magnetically soft and hard ferrite materials, and other new types of magnetic materials are just symptoms of this time. Infrared devices, photovoltaic cells, the reappearance of the magnetic-tape recorder, first for music, then for telemetry recording and now for television programs, thin-metal-film resistors, plastic dielectrics and very thin inorganic film dielectrics, and new ceramic dielectrics are only a few of the other things that should be mentioned, but are beyond the scope of this paper. For this purpose it is perhaps sufficient to say that once again a World War has changed the entire aspect of our scientific outlook and has wrought a revolution in the electronics industry. This acceleration and expansion of electronic techniques into all facets of industry probably has a much greater impact on our daily lives than most of us can realize because we are too close to it. Black and white television is now commonplace, color television after the false starts of 1928 and 1950 on a mechanically rotated disc-presentation system, has reverted to an electronic display system, the Hi-Fi that gave us problems in 1935 is now in 1960 commonplace, stereo phonograph records and reproducing systems are readily available in homes, we have been broadcasting stereo programs on one AM and one FM channel simultaneously and the rules have just been laid down so that both channels can be broadcast on one FM channel. The description of component parts and their requirements, for computers and business machines alone, would fill an entire issue of these PROCEEDINGS. The combination of components into subassemblies—such as the Module and Micro Module, PAC, the 2D ceramic substrate with lumped or distributed resistors and capacitors printed as units, the 2D systems of resistors and capacitors vaporized or deposited in successive layers on a substrate which in turn may be semiconducting material on which have been formed transistors or diodes, the selective doping of a single block of material to create a complete active circuit, printed wiring, automatically programmed component insertion machinery, etc.—have all had a great effect on design and manufacturing philosophies and promise even greater changes for the future.

The author would like to suggest that the 1960 to 1980 period may very well be the solid-state-physics generation, and to make the following gratuitous remarks. One of the most significant things to me in the history of our business is the recurrent manner in which parts, circuits, and philosophies are used and abandoned. There may be a moral here, only faintly concealed, that in our continual pressure for "Onward and Upward," we should not be too proud to stop occasionally and review our past lest we in ignorance not only retrace our steps, but what is worse, also repeat our previous mistakes in the application of parts and materials.

Although basic design concepts and even form factors of components may not seem to have changed very radically in the last fifty years, the design engineer of 1912 would be faced with an unbelievably kaleidoscopic vista if he were transported directly to the 1961 era. New materials, new methods of processing material, the ability to produce in quantities components which were once laboratory curiosities, the increasing complexity of end equipment and the resulting demands for more accurate control of component characteristics, the requirements for increased performance in more exacting environments, and the economic pressure to reduce cost have all combined to provide the present-day engineer with an array of parts with an extremely wide range of capabilities. The 1912 man might be almost frightened by this array, although in all due fairness he probably would not be anywhere nearly as nonplussed as the author was when, many years ago, he was ordered to repair and put in sufficiently reliable operation, so that the public could start and stop it with a push button, a tape sender and spark transmitter and a tuner, Wollaston detector and tape recorder of the vintage of 1910.

Resistors—A Survey of the Evolution of the Field*

JESSE MARSTEN†, FELLOW, IRE

Summary—**A survey is made of the evolutionary steps leading to the present state of the resistor field. The survey is confined to the resistance element which is the basic and most important factor in all resistors. The development is traced as nearly as possible from the origins of resistors, and proceeds from wire to composition, deposited carbon and films of metal and their compounds. Some future trends are indicated.**

I. Introduction

UNTIL the advent of radio broadcasting and particularly the use of the triode, industry needs were adequately served by wire resistors in a relatively small span of low to medium resistance values and from low to high power. The triode with its high-impedance input expanded the spectrum of values into the megohms region where resistances were required in ever increasing quantities by the burgeoning radio industry. However, size, cost and reactance limited the use of wire resistors.

Prior to radio, there was no organized research and development specifically directed to resistors. However, some very early work was so significant as to foreshadow the accomplishments today. There was considerable activity thereafter, largely disclosed in the patent literature. World War II, with its unusual environmental requirements, precipitated the major intensive research and development effort in resistors. Here special credit must be accorded the military for sparking, promoting and encouraging the major portion of this activity.

The factor common to all resistors, fixed and variable, which largely determines properties and performance, is the resistor element. It can be adapted to an infinite variety of types, shapes and uses of both fixed and variable resistors. It has therefore been the primary subject of significant investigations which fundamentally resolve into a search for new resistive materials. The evolution of the resistor field parallels this search.

II. Wire Resistors

Resistors utilizing alloys such as copper-nickel and nickel-chromium wound on insulating forms of varied geometry in a protective enclosure such as varnish, cement or vitreous enamel, are old and well-known. Advances have been the result of normal engineering improvements in structure and assembly, and need not be dwelled upon. The use of new resins for enclosures such as epoxy and silicones have made substantial improvements in humidity characteristics and have increased maximum operating temperatures.

Two developments of more technical merit warrant mention. Copper-nickel alloy has low temperature coefficient (20 ppm/°C) and medium resistivity (200 ohms/cmf). Nickel-chromium has higher resistivity (600 ohms/cmf) but high temperature coefficient (150 ppm/°C). In 1946, Lohr [1] disclosed a nickel-chromium-aluminum alloy, and in 1949, Allen [2] disclosed nickel-chromium-aluminum-copper alloy, both having a resistivity of 800 ohms/cmf with a temperature coefficient of 20 ppm/°C. The resulting gain was very substantial: greater temperature stability, increased maximum resistance available and space saving.

Differential expansion of the different elements in vitreous enamel resistors, causing shorted turns, loose terminals and crazing, results from the very high maturing temperature of the enamel. Leonard [3] in 1901 developed a glaze with low fusing point, 700°C instead of 1500°C, which helped but did not quite solve the problem. Ganci [4] in 1936 determined that the solution resided in the expansion coefficients of base, metal and enamel being properly related to each other. Formulations for the ceramic form and the enamel were developed which provide the proper coefficients relative to the metals employed. Thereafter it cannot be said that there were any major improvements. However, wire-wound resistors properly constructed are still the most stable and reliable in the limited values in which they are made, but in high values the size and cost are prohibitive. Future developments in other fields caused by these limitations are gradually narrowing the field of usefulness of wire resistors.

III. Composition Resistors

The urgent need for high resistance required by the triode was manifested in a rash of pencil and India ink line resistors. Future effort centered on "compositions of matter" as resistors. Such compositions have a long history. Thus in the field of solid compositions, Slattery [5] in 1885 described a high resistance comprising carbon and a refractory insulator molded into a cylinder with carbon terminations on each end. Voss [6] in 1896 shows a molding of carbon, metal oxides and binder fired to resistance value. In the field of composition film resistors, Gambrell [7] in England, 1897, disclosed a dispersion of carbon in a liquid medium containing a binder and volatile solvent which, applied to an insulating base, is heated to produce a resistive film.

Here we have the basic ingredients of both solid and

* Received by the IRE, May 2, 1961.
† International Resistance Company, Philadelphia, Pa.

film composition resistors of today. What was lacking to provide a satisfactory product capable of being produced uniformly in quantity, were the right materials and production techniques. These represent the technical contributions which resulted in today's product. As to material, it was the use of phenol-condensation products invented by Baekeland, used as molding powder and varnish binders which transformed the above early compositions into the present high quality product.

As for processing techniques which made possible prices of the order of one to four cents each, there were many workers in the field of solid composition resistors. The techniques were all based on hydraulic press molding of the materials, typical of which is that introduced in 1925 by Bradley [8] and Powers [9] disclosing specific formulations of materials and the simultaneous molding of the resistor body and end conductive terminations without wire leads.

Pender [10] invented an ingenious technique in 1925 for automatically and continuously depositing a composition film which, with engin ring refinements, is in use today. A solid or tubular glass fiber is continuously drawn to controlled dimensions through a liquid resistive solution and coated with a film of controlled thickness. The film is cured, measured for resistance value and cut to size, all in one operation.

The final stage in the development of the composition resistor, both solid and film, was the elimination of the so-called "uninsulated resistor," with thin varnish or lacquer coating. This was accomplished in 1933 by Pender [11] by molding an outer insulating casing around the resistor element, conductive terminations and leads. This is still the standard construction of not only the solid and film composition resistor, but precision film resistors as well.

The composition resistor is probably the most versatile of all resistors. It is miniaturized. It can be made in values from 1 to 10^9 ohms and higher. In the film form it is made for high frequency, high-voltage and moderately high power applications. As film, Brunetti and Khouri [12] introduced its use as a two-dimensional resistor for printed circuit applications. It is the lowest cost resistor, requiring minimum man-power and noncritical materials. For general purpose work requiring broad tolerances, it has been the resistor work horse of the electronics industry for four decades, still being used by the billions annually in the U. S. and throughout the world. Used within its limitations, it is now one of the most reliable components. However, it has two major deficiencies; stability and precision, properties so important to the war effort, which were directly responsible for the major effort in resistor research during and after the war.

IV. Deposited Carbon Resistors

The solution to the instability problem was the elimination of the organic binders in composition resistors. The use of deposited carbon resistors was the first important step in this direction. However, it is to be noted that this class of resistor was already developed and widely used in Europe in the twenties. Its use in the U. S. was precipitated by the need for precision resistors in great quantities in World War II and by the lack of facilities and critical materials for producing the necessary quantity of precision wire resistors.

The origin of this class of resistors appears to be German. The principle of cracking a hydrocarbon gas by pyrolysis to produce carbon may have had much earlier origin, but in 1925, Hartman [13] disclosed the basic method for depositing a hard carbon resistive film on an insulating substrate by pyrolysis of a hydrocarbon gas in vacuum or in a vessel containing an inert gas. Both methods are used today. The idea of increasing its resistance value by cutting a helical groove in the film was introduced in Germany much earlier in 1919 by Kruger [14]. Seibt [15], in Germany (1930), disclosed the process and equipment for producing the films by batch and continuous conveyor methods. Refinement of these basic disclosures resulted in the present semiprecision resistor having a stability in general terms three to five times that of composition resistors.

The resistance properties and performance of this film depends on its thickness. Resistance and temperature coefficient increase as film thickness decreases. Performance degrades with decreased film thickness. Maximum usable values are therefore limited. Temperature coefficient is variable and may be as much as three to four times greater at maximum usable resistance than at lowest resistance.

It is well-known that an alloy of metals shows increased resistivity and lower temperature coefficient than the individual components. This phenomenon also applies to films. In 1909, Weintraub [16] discovered that boron changed the temperature coefficient of carbon from negative to positive. Grisdale, Pfister and Teal [17] disclosed in 1950 a "boro-carbon" resistor film deposited pyrolytically from a mixture of a hydrocarbon and boron gas. Such resistors did show temperature coefficients very much lower than that of the conventional deposited carbon resistor, but only at low to medium values. Much higher values could also be made but deterioration of properties was so serious that its use was extremely limited.

Much activity continued along these general lines, and in 1952 Stein and Pugh [18] in research on high-stability resistors sponsored by the U. S. Signal Corps, described a deposited carbon film of carbon, silicon and oxygen which increased resistivity and improved humidity stability over that of carbon films. Such a resistor, originally called "pyrolytic carbon alloy" [19] in the trade, increased the usable maximum value fivefold, increased the hardness of the film so that resistance cannot be altered by abrasion or rubbing (a common procedure for adjusting the value of deposited carbon resistors), increased maximum operating temperature to 250°C for nonprecision work and enabled normal power ratings to be doubled for semiprecision applications. However, temperature coefficients were not improved. They were still too high and nonuniform for precision applications.

V. Metal Film Resistors

While carbon as resistive material was being fully explored, many parallel investigations were being conducted for other materials having greater stability and lower temperature coefficients, most being centered on films of metals and their compounds.

A. Noble Metal Films

The basic idea involved is the very old art of decorating chinaware with precious metals. In Germany in 1926, Loewe [20] developed a resistive film by atomizing a liquid solution of platinum resinate by forcing compressed air through it and applying the spray to an insulating base. Heating the film thus formed reduces it to the metal. Rosenthal Isolatieren Company [21] in 1937 included other metals in solution, *e.g.*, gold and silver, which lowered the temperature coefficient. Jira [22] (1942) gave minute detail of the entire procedure, which resulted in a practical product. While there is still some activity in this area, the improvements have not been noteworthy and the product has attained small use.

B. Tin-Oxide Films

The seed for this important contribution was provided by Littleton [23] (1931) who developed an iridized, conducting tin-oxide film coating for glass insulators. Its resistivity was sufficiently low to equalize potential across the insulator, thereby reducing corona effect, but too high for use in conventional resistors. Mochel [24] modified this film by the addition of antimony oxide which stabilized its electrical properties. By varying the tin-antimony proportions, negative or positive temperature coefficients are obtained.

The film is applied by atomizing a solution of tin and antimony oxides, applying the mist to a glass substrate heated to 500°C or more. Heating reduces the film to the metal-oxide state, integrally united with the substrate. This class of resistor has a very high degree of stability on all counts, can withstand temperatures of 225°C or more and has a wide field of use. It is, however, more limited in maximum value than either composition or deposited carbon resistors, and its temperature coefficient is high and of the order of deposited carbon.

C. Metal Films Produced in Vacuum

The resistor that most nearly approaches the properties and performance of wire resistors is one using a metal film produced in vacuum. There are two principal methods: 1) sputtering and 2) evaporation. While sputtering is not used today, the early developments have proved very valuable.

1) Sputtering: The principle is very old. A metal cathode is subjected to high voltage. Ionic bombardment of the cathode results in dislodgement of cathodic atoms which, falling on an insulating substrate, form a metal film. Richtmeyer [25] (1919) utilized this principle to produce resistors. The additional importance of his disclosure lies in the discussion of resistance instability caused by oxidation and in his use of heat treatment to stabilize the film. In the same year, Kruger [14] disclosed a high-resistance unit made by sputtering, in which adjustment of the film to desired value is accomplished by cutting a helical groove in the film, a technique which is almost universally used to-day.

2) Evaporation: The evaporated-metal film resistor is the latest useful step in the evolution of resistor development. While the general technique for producing evaporated metal films was known for a long time, very little was known about its properties. Extensive investigation by physicists [26] (the references cited contain extensive bibliography on the subject) disclosed that these films do not have the same properties as the bulk metal. For example, they show much higher resistivities, and temperature coefficients are lower and may be either positive or negative. Also, film properties are influenced by details of preparation and substrate history. Studies by Hoffman and Riseman [27] in research sponsored by the Air Force, indicated good possibilities for controlling these properties.

An early disclosure of metal evaporation was made in 1884 by Edison [28], who described the process of heating a metal to incandescence in vacuum, the vaporized metal depositing as a film on the material to be plated. This principle was utilized first by Loewe [29] in Germany, 1924, for making a resistor particularly for radio. An evacuated tube contains an insulator, with attached terminals coming out of the tube, surrounded by a metal cylinder heated by eddy currents induced in the metal. The evaporated metal deposits on the insulator, resulting in a hermetically sealed resistor.

In 1945 Weber and Johnson [30] described a procedure for evaporating nichrome on glass substrates and making precision resistors with temperature coefficients as low as 70 ppm/°C. The importance of evaporation parameters such as rate of evaporation, pressure, temperature, is emphasized. To protect against oxidation and aging, a protective insulating film is evaporated on the resistive film.

In 1954 Stein and Riseman [31] reported on the problems and conditions involved in the production of these resistors and disclosed the possibility of producing resistors with temperature coefficient controlled to any desired value and near zero. Not only can the temperature coefficient be controlled within limits, but unlike carbon and metal oxide films, it is relatively constant at all resistance values. Extensions of this investigation to precision resistors and power resistors, based on Signal Corps Contracts are described by Stein and Riseman [32] and Bohrer, Hauth, and Stein [33] and have led to successful production.

VI. Present Status and Future Trends

The absence of any important effort in wire resistor development is conspicuous and implies a declining use except for specific applications such as high power. Composition and film resistors have replaced wire in many applications and promise to continue doing so. The concentrated development effort on films has brought the deposited-carbon, tin-oxide and evaporated-metal film resistors to a stage of stability and reliability where they can meet most of the requirements in electronic and other applications. Even more so is the case with evaporated-metal film resistors since they can be made with temperature coefficients as low as wire and near zero and relatively constant over a wide range of values, thus affording greater temperature stability.

As to the future, while predictions are dangerous considering the rapid changes that are taking place, two trends are clearly perceptible. These trends are still dictated by the insistent demand for ever greater reliability and stability of components and by changes in electronic systems.

A. Inorganic Compositions

The route by which the present stage was reached proceeded from inorganic to organic compositions, to carbon films and finally to films of metals and their compounds. The trend now returns full circle to inorganic compositions with the expectation that much higher operating temperatures and stability will be obtained, and greater reliability by the elimination of catastrophic failures, occasionally experienced by film resistors.

As mentioned previously, very early work was not fruitful for lack of suitable materials. Later work by Hunt [34] in 1922 showed possibilities in fusing glass and carbon to form solid resistors. Some carbon-glass frit resistors were produced for a short time in Germany. This approach has been pursued by many investigators, and more recently by Hutter [35] in 1956 and Andrea [36] in 1958, who developed resistive and conductive glazes of borosilicate glass, fused with silver, copper and/or palladium. Other binding media for the conductor, such as ceramic, are also being investigated. Film resistors called "cermet resistors," have recently been produced by both methods. Many variations of this approach in film and solid form are being vigorously explored for use as composition, precision and power resistors. While it is too early to draw conclusions, the indications are that this promises to be a very important contribution with some potential in further limiting the uses for wirewound resistors.

B. Semiconductor Materials

1) The present trend toward micromodular systems and microcircuitry practically compels the use of film resistors made, because of space and geometry, with films of extremely high specific resistivity beyond that available now, if high values are needed. Research on thin semiconductor films of low conductivity is underway to meet this need. If successful, this could well eliminate the mechanical procedures of helixing and meandering currently used to increase resistance.

2) Film investigation in recent years has consumed most of the research and development man-power. It is not expected that this situation will change drastically since much remains to be done to make film production a science rather than an art. Also, other film materials will continue to be investigated. However, a departure is evident in increased activity in solid resistors as in Section VI-A, but also in bulk semiconductors utilizing doping methods to obtain necessary properties. Also, where discrete resistors are used, such materials, if successful, will be more amenable to subminiaturization.

3) Although the use of discrete resistors will still predominate, the semiconductor resistor effort will of necessity assume greater importance and result in greater activity because of the trend toward molecular electronics.

4) New areas in variable resistors will be explored. Noise and wear caused by sliding contactors are perennial problems. Programs [37] designed to investigate noncontact potentiometers utilizing semiconductors as the resistance medium are already underway. Among others, photo-conductive elements such as cadmium sulphide, and materials showing magneto-resistive properties [38] such as indium antimonide, are being studied. These promise to have value in specialized applications such as transducer functions. For general-purpose applications, there is nothing on the horizon to threaten the pre-eminence of the composition variable resistor [39].

References

[1] J. M. Lohr, U. S. Patent No. 22,460,590; 1946.
[2] V. O. Allen, U. S. Patent No. 2,638,425; 1949.
[3] W. H. Leonard, U. S. Patent No. 691,949; 1901.
[4] C. J. Ganci, U. S. Patent No. 2,179,212; 1936.
[5] M. Slattery, U. S. Patent No. 354,275; 1885.
[6] D. C. Voss, U. S. Patent No. 573,558; 1896.
[7] T. E. Gambrell, *et al.* British Patent, No. 25,412; 1897.
[8] L. Bradley, U. S. Patent Nos. 1,845,828 and 1,835,267; 1925.
[9] L. E. Powers, U. S. Patent No. 1,316,194; 1927.
[10] H. Pender, U. S. Patent No. 1,771,055; 1925.
[11] H. Pender, U. S. Patent No. 2,052,533; 1933.
[12] Brunetti and Khouri, "Printed Electronic Circuits," Electronics vol. 19, p. 104, April, 1946.
Brunetti and Curtis, "Printed Circuit Techniques," Nat'l. Bur. Stds., Washington, D. C. Publication 468, November 15, 1947.
[13] C. A. Hartman, German Patent 438,429; 1925.
[14] F. Kruger, German Patent 336.469; 1919 and British Patent 157,909; 1919.
[15] G. Seibt, German Patents 540,859 and 623,583; U. S. Patent Nos. 1,965,059 and 1,998,060; 1931.
[16] E. Weintraub, U. S. Patent Nos. 1,019,391, 1,019,393, and 1,019,394; 1909.
[17] R. O. Grisdale, *et al.* U. S. Patent No. 2,671,735; 1950.
R. O. Grisdale, A. C. Pfister, and G. K. Teal, *Bell Sys. Tech. J.*, vol. 30, pp. 271–314; April, 1951.
[18] S. J. Stein and A. L. Pugh, U. S. Patent 2,598,899; 1952. Research was sponsored by U. S. Signal Corps, Second Quart. Rept. Signal Corps Contract DA-36-039-SC-15332; October, 1951–January, 1952.
[19] C. Wellard and K. Gentner, "Pyrolytic Alloy High Temperature Resistor," *Proc. 1956 Electronic Components Symp.*, Washington, D. C., May 1–3, 1956; pp. 179–183. Also British Patent No. 839,118; 1956.

[20] S. Loewe, German Patent 591,735; 1926, and U. S. Patent 1,717,712; 1926.
[21] Rosenthal Isolatieren Co., British Patent 486,639; 1937.
[22] J. W. Jira, U. S. Patent No. 2,281,843; 1942. Also J. W. Jira, "New Advances in Printed Circuits," Natl. Bur. Stand. Washington, D. C., NBS Publication No. 192, p. 25; 1948.
[23] J. T. Littleton, U. S. Patent No. 2,118,795; 1931.
[24] J. M. Mochel, U. S. Patent No. 2,564,707; 1947. Reissue 25,556.
[25] F. K. Richtmeyer, U. S. Patent No. 1,388,373; 1919.
[26] L. Holland, "Vacuum Deposition on Thin Films," John Wiley and Sons, Inc., New York, N. Y., p. 260; 1956.
M. Mostovetch and B. Vodar, "Semi-Conducting Materials," Academic press, New York, N. Y., 1952.
[27] D. M. Hoffman and J. Riseman, "Theoretical and Experimental Investigations on Thin Film Resistance Elements," Wright Air Dev. Ctr., Dayton, Ohio, WADC Tech. Rept. 57-215, July, 1957.
[28] T. A. Edison, U. S. Patent No. 526,147; 1884.
[29] S. Loewe, German Patent 500,794 and U. S. Patent No. 1,751,-587; 1924.
[30] E. Weber and S. A. Johnson, "Precision Metallized Glass Resistor Units," Polytechnic Institute of Brooklyn, Brooklyn, N. Y., Rept. R-116; 1945. U. S. Patent No. 2,586,752; 1946.
[31] S. J. Stein and J. Riseman, "Evaporated Metal Film Resistors," *Proc. 1954 Electronic Components Symp*, Washington, D. C., 1954; pp. 100–106.
[32] S. J. Stein and J. Riseman, "Precision and high temperature metal film resistors," *Proc. 1956 Electronic Components Symp.*, Washington, D. C., May 1–3, 1956; pp. 171–175. (Based on Signal Corps Contract DA-36-039 SC-72726 Repts.)
[33] J. J. Bohrer, *et al.*, "Metal film power resistors," *Proc. 1956 Electronic Components Symp.*, Washington, D. C., May 1–3, 1956; pp. 166–170. (Based on Signal Corps Contract DA-36-039 SC-75973 Repts.)
[34] M. H. Hunt, U. S. Patent No. 1,764,311; 1922.
[35] D. E. Hutter, U. S. Patent No. 2,837,487.
[36] J. B. d'Andrea, U. S. Patent No. 2,924,540.
[37] U. S. Navy Contract NobsR 77548 Repts. Quarterly, January–September, 1960; October 1960–July 1961.
[38] H. Weiss and H. Walker, U. S. Patent No. 2,894,234; 1954. *Magnetic-Resistor Bulletin* (MS-41), Ohio Semi-Conductor Company, Columbus, Ohio; 1961.
[39] For a general comprehensive bibliography on the subject of resistors see P. R. Coursey, "Fixed resistors for use in communication equipment," PROC. IEE, vol. 96, pp. 169–182; May, 1949.
Note: Patent dates are filing dates.

Capacitors*

LEON PODOLSKY†, FELLOW, IRE

***Summary*—A review of the history and bibliography of the capacitor art of the 18th, 19th, and 20th centuries, prior to the founding of the IRE in 1912, compared to the explosive developments of the fifty years to 1962, shows the evolutionary nature of the advances in capacitor engineering. The dependence and parallelism of capacitor engineering with the mainstream of development in the basic sciences are clear and are further indicated for the future.**

IN THIS Fiftieth Anniversary Issue, the compulsion to review historically any sectional field of the electronics industry is inescapably strong. In the field of capacitors it is not only strong, but especially valuable to do so as a guide for our expectation of the future. For surely in these fifty years, as viewed from this time in the second half of the "Electronics Century," revolution and dramatic new invention in the capacitor field must have taken place. Certainly in the half century that saw atomic and electronic theory twice basically revised, and nuclear technology entirely developed, dramatic and fundamental changes must have taken place in our *concepts* of capacitors. Or so it would seem to the present generation of electrical and electronics engineers. But is it so?

Omitting most of the voluminous literature on dielectrics in general, there were more than *four hundred primary* references in the world's literature on the subject of "condensers" and their general properties and measurement at the time of the Institute's founding in 1912. In reviewing this extensive bibliography it is refreshing, sometimes shocking, to note the continuous flow of capacitor *ideas* and technology from the early 18th century, through the 19th century, and into the 20th century, well before the IRE was conceived. In addition to the extensive theoretical and practical publications in connection with the Leyden jar and its associated experiments beginning in 1745, there is a continuous literature of capacitor *ideas* dating at least from S. Grey's article "Electrical Experiments" in the *Philosophical Transactions of the Royal Society*, vol, 21, No. 366, page 104, of 1720. It is startling to find in this extensive literature before 1912 well-developed references, reports, and experiments on ideas, techniques, and structures which are today extant in the newest capacitors of our electronics technology.

To provide a frame of reference, Table I lists the most important properties and parameters of capacitors as we conceive them today in the light of our current needs.

Surely high performance in these parameters, or most of them, must evolve from the needs of our modern electronic apparatus and could hardly have been a matter of concern more than half a century ago! Let us see.

Without reference to any specific type of capacitor, what does the literature show us historically with regard to each of the major parameters of modern interest? The following fractional excerpt from the extensive bibliography will serve to show some of the more significant references to each of these important capacitor properties which existed *before 1912*.

* Received by the IRE, April 3, 1961.
† Sprague Electric Company, North Adams, Mass.

TABLE I

Capacitance (per unit volume, or volume efficiency)
Capacitance stability with time
Capacitance stability with temperature
Capacitance stability with applied voltage
High-frequency properties and effects
Voltage rating and dielectric strength
Insulation resistance
Insulation resistance variations with temperature
Dissipation factor (or dielectric loss)
Effects of humidity
Dielectric hysteresis
Pulse operation and surge rating
Effects of radiation
Effects of vibration and shock

Capacitance (Per Unit Volume, or Volume Efficiency)

1782—A. Volta, "Method of rendering very sensible the weakest natural or artificial electricity," *Phil. Trans. Roy. Soc. London*, vol. 72, p. 237; 1782.

1810—J. Wingfield, "A new method of increasing the charging capacity of coated electrical jars," *Phil. Mag.*, vol. 36, p. 259; 1810.

1855—W. Thompson, "On the electro-statical capacity of a Leyden phial, etc.," *Phil. Mag.*, pt. 4, vol. 9, pp. 531–535; 1855.

1862—J. M. Gaugain, "On the limit of the charge of condensers," *Phil. Mag.*, pt. 4, vol. 24, pp. 495–496; 1862.

1873—M. Neyreneuf, "On electrical condensation," *Telephone J.*, vol. 1, p. 257; 1873.

1879—C. V. Boys, "A condenser of variable capacity," *Phil. Mag.*, pt. 5, vol. 7, p. 108; 1879.

1884—R. T. Glazebrook, "Method of measuring electrical capacity of a condenser," *Phil. Mag.*, pt. 5, vol. 18, pp. 98–105; 1884.

1884—L. R. Wilberforce, "On some experiments on the measurements of the capacity of a condenser," *Proc. Cambridge Phil. Soc.*, vol. 5, pp. 175–182; 1884.

1900—L. Lombardi, "An Improved Process for Manufacturing Thin Homogeneous Plates, More Particularly Applicable for Use in Electrical Condensers," Brit. Patent No. 9133; filed May 17, 1900.

1904—G. F. Searle, "On the calculation of capacities in terms of coefficients of electrostatic induction," *Proc. Cambridge Phil. Soc.*, vol. 12, pp. 378–387; 1904.

1906—A. Zeleny, "Capacity of mica condensers," *Phys. Rev.*, vol. 22, p. 651; 1906.

1908—A. Zeleny and Andrewes, "The capacity of paper condensers," *Phys. Rev.*, vol. 27, pp. 65–75; 1908.

1909—E. H. Weiss, "Conditions for minimum volume of a condenser," *Lumière élec.*, vol. 8, pp. 326–29; December 11, 1909.

Capacitance Stability with Time

1874—Schiller, "Variation of dielectric constant with time of charge," *Poggendorfs Annalen*, vol. 152, p. 535; 1874.

1897—J. Hopkinson and E. Wilson, "On the capacity and residual charge of dielectrics as affected by temperature and time," *Phil. Trans. Roy. Soc. London*, vol. 189A, pp. 109–137; 1897.

Capacitance Stability with Temperature

1897—J. Hopkinson and E. Wilson, "On the capacity and residual charge of dielectrics as affected by temperature," *Elec. Rev.*, vol. 40, pp. 181 and 830; 1897.

1905—E. Terry, "On the variation of a capacity with temperature," *Phys. Rev.*, vol. 21, p. 193; 1905.

Capacitance Stability with Applied Voltage

1887—W. E. Sumpner, "On the variation of the coefficient of induction," *Phil. Mag.*, pt. 5, vol. 25, pp. 453–475; 1887.

High-Frequency Properties and Effects

1889—J. J. Thomson, "Specific inductive capacity of dielectrics at very high frequencies," *Proc. Roy. Soc.*, vol. 46, p. 292; 1889.

1891—R. Blondlot, "On the determination of the dielectric constant of glass by the action of very rapid electrical oscillations," *Phil. Mag.*, pt. 5, vol. 32, pp. 230–232; 1891.

1891—E. Lecher, "On the measurement of dielectric constants by means of Hertz oscillations," *Phil. Mag.*, pt. 5, vol. 31, pp. 172–182; 1891.

1897—E. S. Ferry, "The specific inductive capacity of dielectrics in slowly and rapidly changing electric fields," *Phil. Mag.*, pt. 5, vol. 44, pp. 404–414; 1897.

1898—J. Hanauer, "Influence of frequency on capacity," *Ann. Phys. u. Chem.*, vol. 65, p. 789; 1898.

1907—J. G. Coffin, "The effect of frequency upon the capacity of absolute condensers," *Phys. Rev.*, vol. 25, pp. 123–135; 1907.

1911—J. A. Fleming and G. B. Dyke, "The energy losses in condensers in high-frequency circuits," *Proc. Phys. Soc. London*, vol. 23, pp. 117–146; February, 1911.

Voltage Rating and Dielectric Strength

1881—E. Villari, "On the internal discharges of electrical condensers," *Phil. Mag.*, pt. 5, vol. 11, p. 541; 1881.

1892—W. H. Preece, "On the dielectric strength of a condenser," *Brit. Assoc. Paper*, Sec. A; August, 1892.

1893—A. MacFarland and G. W. Pierce, "On the electric strength of solid, liquid and gaseous dielectrics," *Phys. Rev.*, vol. 1, pp. 161–165; 1893.

Insulation Resistance

1897—P. A. Bates and W. C. Barnes, "Dielectric resistance," *Proc. AIEE*, vol. 14, p. 253; 1897.

1902—E. R. von Schweidler, "Conductivity and polarization in dielectrics," *Sitzungsberichte der Kaiserlichen Akad.*, pt. 111, p. 579; 1902.

1911—W. W. Stifler, "Resistance tests of insulators at high temperatures," *Phys. Rev.*, vol. 32, pp. 429–432; April, 1911.

Insulation Resistance Variations with Temperature

1889—H. Kouler, "On the electrical resistance of insulators at high temperatures," *Phil. Mag.*, pt. 5, vol. 23, p. 147; 1889.

1889—W. H. Schultze, "On the electrolytic behavior of mica at high temperatures," *Phil. Mag.*, pt. 5, vol. 27, p. 368; 1889.

1897—R. Appleyard, "Temperature corrections for dielectric resistance," *Elec. Rev.*, vol. 40, p. 737; 1897.

Dissipation Factor (or Dielectric Loss)

1891—P. Cardew, "Energy loss in condensers," *J. Inst. Elec. Engrs.*, vol. 20, p. 447; 1891.

1897—L. Houllevigue, "Electrical heating of condensers," *J. Phys.*, vol. 6, p. 120; 1897.

1899—E. B. Rosa and A. W. Smith, "Condenser losses," *Phys. Rev.*, vol. 8, p. 1; 1899.

1900—Pellant and Beaulard, "Energy losses in condensers with sinusoidal ac potentials," *Comptes Rendus*, vol. 130, pp. 1457–1460; 1900.

1901—C. P. Steinmetz, "Dielectric losses in condensers," *Elec. World and Engr.*; June 22, 1901.

1904—R. F. Earhart, "Absorption of energy in condensers," *Phys. Rev.*, vol. 19, p. 8; 1904.

1911—L. W. Austin, "Condenser Losses at High Frequencies," *J. Washington Acad. Sci.*, vol. 1, pp. 143–145; 1911.

1912—K. W. Wagner and A. Wertheimer, "Measurement of the dielectric angle of loss and the capacity with the ac bridge," *Physik. Z.*, vol. 13, pp. 368–375; May, 1912.

Effects of Humidity

1892—A. Johnstone, "Action of water on mica," *J. Chem. Soc.*, vol. 62, p. 573; 1892.

1898—E. Lethoule, "Effect of heat and moisture on dielectrics," *l'Eclairage Electrique*, vol. 17, p. 439; 1898.

Dielectric Hysteresis

1864—J. M. Gaugain, "On the residual charge of electrical condensers," *Phil. Mag.*, pt. 4, vol. 28, pp. 77–78; 1864.

1893—F. Bedell, N. Ballantyne and R. Williamson, "Alternate current condensers and dielectric hysteresis," *Phys. Rev.*, vol. 1, p. 81; 1893.

1895—A. W. Porter and D. K. Morris, "The question of dielectric hysteresis," *Electrician*, vol. 34, pp. 735–736; April 12, 1895.

1899—W. Schaufelberger, "Dielectric polarization and hysteresis," *Ann. Phys. u. Chem.*, vol. 67, p. 307; 1899.

1907—L. Malcles, "After effects in solid dielectrics," *Comptes Rendus*, vol. 144, p. 264; 1907.

1912—L. Decombes, "Electron theory of residual dielectric phenomena," *J. Phys.*, pp. 181–196; March, 1912.

Pulse Operation and Surge Rating

1864—W. Siemens, "On the heating of the glass plate of the Leyden jar by the discharge," *Berliner Berichte*, p. 612; 1864.
1889—J. Borgmann, "The heating of glass of condensers by intermittent electrification," *Phil. Mag.*, pt. 5, vol. 23, p. 472; 1889.
1908—B. V. Hill, "The variation of apparent capacity of a condenser with the time of discharge," *Phys. Rev.*, vol. 26, pp. 400–405; 1908.

Effects of Radiation

1896—J. J. Thomson and J. A. McClelland, "On the leakage of electricity through dielectrics traversed by roentgen rays," *Proc. Cambridge Phil. Soc.*, vol. 9, pp. 126–141; 1896.
1902—J. Curie, "Dielectrics under radium," *Comptes Rendus*, vol. 134, p. 420; 1902.
1904—Becker, "The influence of cathode rays on good dielectrics," *Ann. Phys.*, vol. 12, pp. 394–421; 1904.
1908—G. Jaffe, "On the ionization of fluid dielectrics by radium," *Ann. Phys.*, vol. 25, p. 257; 1908.
1909—B. Hodgson, "The conductivity of dielectrics under the action of radium rays," *Phil. Mag.*, pt. 6, vol. 38, pp. 252–258; 1909.

Effects of Vibration and Shock

1909—P. L. Mercanton, "Effect of mechanical shock on the residual charge," *Comptes Rendus*, vol. 149, p. 591; 1909.
1913—P. L. Mercanton, "Vibration and residual charge," *Arch. Sci.*, vol. 35, pp. 48–56; 1913.

Surprising as it may seem, the characteristics which are deemed essential today were under study in the old art, even to effects of supposedly modern environments such as acceleration and radiation. Surely, however, we might expect to find radical departures in types of materials used in capacitors. Let us look at the new materials in comparison with the old.

The most significant types of "fixed" capacitors in general use at this time, classed according to dielectric or structure, are those shown in Table II.

TABLE II

Glass
Mica
Vitreous enamel
Ceramic
Impregnated paper film and foil
Plastic film and foil
Composite dielectrics (paper and plastic films)
Metallized paper and plastic films
Vacuum or pressurized gas
Electrolytic—aluminum or tantalum oxide

Did the knowledge and use of these dielectrics and structures spring out of our mid-20th century electronics requirements? Let us again look at history before 1912 to gain perspective.

Glass

1788—E. W. Gray, "Observations on the manner in which glass is charged with the electric fluid," *Phil. Trans. Roy. Soc. London*, vol. 77, pp. 407–409; 1788.
1877—J. Hopkinson, "Residual charge of the Leyden jar—dielectric properties of various glasses," *Phil. Mag.*, pt. 5, vol. 4, pp. 141–142; 1877.
1891—R. Blondlot, "On the determination of the dielectric constant of glass by the action of very rapid electrical oscillations," *Phil. Mag.*, pt. 5, vol. 32, pp. 230–232; 1891.
1904—I. Mosicki, "Improvements in Electric Condensers," Brit. Patent No. 1307; filed January 18, 1904.

Mica

1874—M. Bauer, "Physical properties of mica," *Z. deut. geol. Ge.*, vol. 26, p. 137; 1874.
1888—C. Doelter, "On the artificial manufacture of mica," *Comptes Rendus*, vol. 107, p. 42; 1888.
1890—E. Bouty, "Mica condensers," *Comptes Rendus*, vol. 110, p. 486; 1890.
1893—P. Janet, "The Hysteresis and Dielectric Viscosity of Mica for Rapid Oscillations," Paris Academy of Sciences, Paris, France; February 20, 1893.
1906—A. Zeleny, "Capacity of mica condensers," *Phys. Rev.*, vol. 22, p. 651; 1906.
1910—H. L. Curtis, "Mica condensers as standards of capacity," *Bull. Bur. Standards*, vol. 6, pp. 431–488; November, 1910.

Vitreous Enamel

1912—W. Torikata and E. Yokoyama, "Enamelled condensers for radio telegraphy," *Electrician*, vol. 69, pp. 1072–1073; October, 1912.

Ceramic

1900—L. Lombardi, "An Improved Process for Manufacturing Thin Homogeneous Plates, More Particularly Applicable for Use in Electrical Condensers," Brit. Patent No. 9133; filed May 17, 1900.
1908—H. W. Brady, "Electrical porcelain," *Elect. Rev.*, vol. 62, p. 927; 1908.
1908—H. F. Haworth, "Electric properties of porcelain," *Proc. Roy. Soc. London*, vol. 81, p. 221; 1908.
1910—Rebora, "Experimental researches on insulators of glass and porcelain," *Atti Assoc. Electrotech., Italiano*, vol. 14, p. 665; 1910.

Impregnated Paper Film and Foil

1898—X. Boucherot, "Paper condensers," *l'Eclairage Electrique;* February 12, 1898.
1906—A. Campbell, "Electric inductive capacities of dry paper and solid cellulose," *Proc. Roy. Soc. London*, vol. 78A, p. 196; 1906.
1908—G. F. Mansbridge, "The manufacture of electrical condensers," *J. Inst. Elec. Engrs.*, vol. 41, p. 535; May, 1908.
1908—A. Zeleny and Andrewes, "The capacity of paper condensers," *Phys. Rev.*, vol. 27, pp. 65–75; 1908.
1911—F. W. Grover, "Experiments on paraffined paper condensers," *Bull. Bur. Standards*, vol. 7, p. 495; December, 1911.
1913—G. E. Bairsto, "Experiments on tinfoil contact with dielectrics," *Proc. Phys. Soc. London*, vol. 26, pp. 301–323; August, 1913.

Plastic Film and Foil

1910—"Measurements on a celluloid condenser," *Atti Assoc. Elettrotech. Italiano*, vol. 14, pp. 227–235; 1910.
1910—A. Hakansson, "Plastic materials for insulation," *Elektrotech. Z.*, vol. 31, pp. 953–958; September 22, 1910.
1911—G. L. Addenbrooke, "Electrical properties of celluloid," *Electrician*, vol. 66, pp. 629–630; January, 1911.

Composite Dielectrics

1892—A. Hess, "Heterogeneous insulators," *J. Phys.*, vol. 1, p. 459; 1892.
1910—W. C. M. Lewis, "Notes on the energy of a 'double layer' condenser," *Phil. Mag.*, pt. 6, vol. 19, pp. 573–575; 1910.

Metallized Paper and Plastic Films

1705—F. Hauksbee, "Experiments on the attrition of bodies in vacuo," *Phil. Trans. Roy. Soc. London*, vol. 24, 304, p. 2156; 1705. (First article on vacuum metallization.)

Vacuum or Pressurized Gas

1850—M. Faraday, "On the magnetic and diamagnetic condition of gases," *Phil. Mag.*, pt. 6, vol. 37, pp. 88–101; 1850.
1892—F. J. Smith, "Compressed air condensers," *Proc. Phys. Soc.*, vol. 11, p. 355; 1892.
1904—Lord Kelvin, "Electric insulation in vacuum," *Phil. Mag.*, vol. 8, p. 534; 1904.
1905—R. A. Fessenden, "High pressure electric condensers," *Electrician*, vol. 36, p. 112; 1905.

ELECTROLYTIC

1873—G. Little, "Condensers for the little automatic telegraph," *Telegraph J.*, vol. 1, p. 207; 1873.
1891—E. B. Rosa, "Specific inductive capacity of electrolytes," *Phil. Mag.*, pt. 5, vol. 31, pp. 188–207; 1891.
1894—Sheldon, Lietche and Shaw, "The capacity of electrolytic condensers," *Phys. Rev.*, vol. 2, pp. 401–411; 1894.
1897—A. G. Davis, "Electrolytic condensers," *Elec. Rev.*, vol. 40, p. 33; 1897.
1898—C. Pollak, "Improvements in electrolytical condensers," Brit. Patent No. 933; filed January 4, 1898.
1901—Siemens Brothers, "An improvement in electrolytic polarization cells having an aluminum electrode," Brit. Patent No. 17480; filed August 31, 1901.
1912—"An electrolytic condenser employed in connection with a sparkless contact," *Elektrotech. Z.*, vol. 33, pp. 889–891; 1912.

It is evident that not only the critical properties but also the basic types are the same today as a half century ago. Are we then to conclude that there is nothing new in the capacitor art, or that no progress has been made in these fifty years? Most certainly not! For although the bibliography clearly shows a long history of theoretical and practical ideas in capacitor art and engineering predating the formation of the IRE in 1912, modern capacitors are vastly different in every category of performance and structure from those of fifty years ago. The important properties are the same, and even the generic names of the principal dielectrics are the same, but a revolution in science has come between.

Advances in capacitors have actually been very real and very large, and have ensued from intensive and continuous development effort involving progressive aplication of the newest findings in physics and chemistry.

In an article of this length it is impossible to explore the full ranges of structure and performance which are available today. However, Table III lists the improvement factors for each of the major parameters originally

TABLE III
COMPARISON OF PERFORMANCE PARAMETERS

	1912	1962*
Capacitance (per unit volume)	2 mf/cu foot	>3000 mf/cu inch
Capacitance stability with time	±1 per cent per year	±0.01 per cent per year
Capacitance stability with temperature	±500 ppm/°C	±10 ppm/°C
Capacitance stability with applied voltage	Uncontrolled	Effect can be made negligible
High-frequency effects	Uncontrolled	±1 per cent, 1000 cycles to 1000 Mc
Voltage rating	100 volts per section 500 volts per capacitor	10,000 volts per section Unlimited per capacitor
Insulation resistance	5 to 10 megohms×mf	>1,000,000 megohms×mf
Dissipation factor	0.01 to 0.1 at 1 kc	3×10^{-4} at 1 kc
Effects of humidity	Somewhat uncontrolled	Eliminated by hermetic sealing
Dielectric hysteresis	No reference standard	Can be made negligible
Pulse operation	Highly destructive	Can be made negligible
Effects of radiation	Uncontrolled	Can be minimized or eliminated
Effects of vibration and shock	Highly destructive	Can be minimized or eliminated

* This is *not* to imply that *all* these properties can be obtained with a single dielectric material, structure, or rating, but only to indicate the performance which can be obtained in general commercial capacitors, at least for single parameters of importance.

TABLE IV
DIELECTRIC MATERIALS AND STRUCTURES IN COMMON USE

	1912	1962
Glass	Plate glass—0.1 in thickness.	1 mil thickness ribbon. Borosilicate or high K glasses.
Mica	3 to 5 mil thickness—generally unselected.	1 mil or less—high grade selected—silvered. Mica paper.
Vitreous enamel	Glass frit—low K.	High temperature vitreous enamel; K=20 to 120.
Ceramic	Porcelain—K approximately 20.	Titania, barium titanate, stannate, zirconate bodies. High K with controllable TC K=10 to 10,000.
Impregnated paper	1 to 3 mil sulfite paper. Beeswax, paraffin, or oil impregnation.	0.2 mil high purity Kraft—Polyisobutylene, chlorinated diphenyl, chlorinated napthalene impregnation.
Plastic film	Celluloid, 3 to 5 mils.	Polyethylene, Mylar, Teflon, polystyrene. 0.25 mil. 0.10 acetate.
Composite Dielectrics	Paper and shellac, etc.	Paper and plastic films.
Metallized paper and plastic films	Unused.	Highly developed. Vacuum metallized paper and plastic films of all types.
Vacuum or Pressured gas	Air or low vacuum.	High vacuum, pressured nitrogen, sulfur hexafluoride.
Electrolytic	Wet cell aluminum.	Dry electrolyte, aluminum foil, tantalum foil, tantalum porous pellet, solid electrolyte tantalum pellet.

shown in Table I which have resulted from the advances in the art in these fifty years.

What of the dielectric materials in use for capacitor structures of today? Table IV shows the outstanding materials in use now, for the categories of capacitors originally shown in Table II, as compared to those of 1912.

There are many other capacitor materials and techniques in more exploratory stages of application at this time which are worthy of mention. Among these are:

High-temperature mica (500°C).
Rolled-glass film.
Solid electrolyte—aluminum oxide—pellet.
Lacquer film (0.1 mil or less).
Mica paper.
Silicon dioxide.
Titanium, columbium oxide, etc.
Metallized spun film.
Polycarbonate film.
Cyanoethylcellulose film.

The most fundamental improvements which have been made in capacitors in the mid decades of this century have been in the following categories:

1) Improvements in reliability (at least 1000:1).
2) Reduction in size (at least 10:1).
3) Increased operating temperature range (−70°C to +150°C for general types).
4) Improved stability (at least 5:1).
5) Development of capacitors suitable for use in nuclear environments (progress).

What shall we say of the future that will have any significance to posterity on the 100th Anniversary of the IRE? We already have in operation electronic circuitry in which all the basic circuit parameters of capacitance, resistance, and inductance are produced within solid-state semiconductor devices. Can we predict the ultimate demise of integral capacitor units as circuit elements?

We must learn the answer from the long continuum of capacitor history. Examination of the extensive bibliography and history of the art shows clearly that improvements in capacitor materials and design have followed closely the increased knowledge in the fundamental sciences, and most particularly in chemistry and physics. All of the major advances of this half century have resulted from basic advances and the improvement of knowledge of the structure and performance of matter and the improvement of methods of measurement.

The advances in capacitor art and engineering have been largely evolutionary, seldom revolutionary, and have been one with the long continuum of advance in basic science. Undoubtedly this has also been true of all other electronic component fields.

There were no companies of record in 1912 which specialized solely in capacitor manufacture. Capacitors were made by the apparatus manufacturers for use in their own equipment. It is therefore impossible to obtain comparative figures on the numbers of capacitors produced or their dollar value in 1912. However, today the capacitor industry in the United States alone consists of more than one hundred fifty manufacturers. Production in 1960 of the major capacitor types shown in Table II is estimated at approximately 1,200,000,000 units with a total sales value of $274,000,000. This represents approximately 23 per cent of the total value of all electronic component sales, excluding tubes and semiconductors, in that year and 2.68 per cent of the value of all electronics manufacture in the United States.

No man can predict what the next half century will bring, just as no prediction made in 1912 could possibly have foreseen the directions of growth of the electronics industry or the development of nuclear science. We can only be sure that developments in the capacitor field will advance constantly with the mainstream of science and that the art will progress to meet the needs of the electronics industries for every application whether they be terrestrial or in outer space.

Acknowledgment

The author is indebted to colleagues at the Sprague Electric Company and to D. A. McLean of the Bell Telephone Laboratories, for reading of the manuscript and constructive suggestions.

Piezoelectric Effect and Applications in Electrical Communication*

VIRGIL E. BOTTOM†, SENIOR MEMBER, IRE

Summary—**The phenomenon of piezoelectricity, discovered in 1880, is the basis of a number of devices which are essential to the art and practice of electrical communication. The discovery of the phenomenon, the development of the quartz resonator, and its application as a frequency-stabilizing element in oscillator and filter circuits are described. Some of the technical problems encountered in the design and fabrication of quartz-crystal units are discussed. Using quartz crystal units for short-term stability and atomic resonators for long-term stability, frequency standards which are reliable and reproducible to 1×10^{-11} are now available.**

The Piezoelectric Effect

MECHANICAL strains in certain crystals are accompanied by electrical polarization, the magnitude of which is directly proportional to the strain. This is the *direct piezoelectric effect.* The reciprocal effect in which the application of an electric field causes a mechanical stress in the crystal is called the *converse piezoelectric effect.* Both effects were demonstrated experimentally by Pierre and Jacques Curie [3] in the years 1880–1881. For the next forty years piezoelectricity remained a laboratory curiosity. During World War I Langevin revived interest in the phenomnon by using plates of quartz to generate and detect high-frequency sound waves in water.

While investigating the use of Rochelle salt as a transducer, Cady, in 1918, observed a peculiar reaction between the piezoid and the driving circuit when the driving frequency was near that of a mechanical resonant frequency of the piezoid. Cady described his discovery of a paper presented to the IRE on November 21, 1921, and in a subsequent paper, published in the PROCEEDINGS the next year. In this paper Cady [4] described the use of the piezoelectric resonator to stabilize the frequency of a vacuum tube oscillator circuit. The pioneer work of Cady was extended by Pierce [5] and others who devised a number of circuits employing piezoelectric resonators as frequency-stabilizing elements.

The Equivalent Circuit

The first investigation of the theory of equivalent circuits appears to have been made by Butterworth [6] in connection with his studies of the vibration galvanometer in 1915. In his paper Butterworth showed quite generally that any one-dimensional mechanical system excited by the interaction between electrical charges and an electric field may be represented by a circuit consisting of an inductance, a capacitance and a resistance in series. During the years 1925–1928, Van Dyke [7] and Dye [8] developed the theory of the equivalent circuit of the piezoelectric resonator and gave the equations relating the values of R, L, C and C_0 of the equivalent circuit (Fig. 1) to the piezoelectric, dielectric and elastic coefficients, and the density and dimensions of the resonator. Using the concept of the equivalent circuit, Espenchied [9] showed how to combine piezoelectric elements with conventional circuit components to provide improved band-pass filters. His work led to the development of the high-frequency carrier current telephone system and laid the foundation for coaxial and microwave transmission systems in which several hundred signals are transmitted simultaneously over the same communication channel.

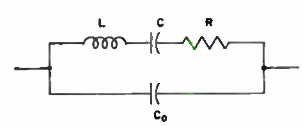


Fig. 1—Equivalent circuit of piezoelectric resonator.

Development of the Quartz Resonator

Although the piezoelectric effect is very much greater in several other crystals, quartz has always been a favored material for many applications because of its superb mechanical qualities. The earliest quartz piezoids were X-cut plates, so called because the plates are cut from the crystal in such a way that the XD-crystallographic axis is perpendicular to the faces of the plate. The application of an electric field in the X direction causes the plate to change its thickness, thus resulting in the propagation of a longitudinal wave in the thickness direction. The resonant frequency of an X-cut quartz plate decreases about 20 ppm/°C.

In 1933 a patent was issued to Tillyer [10] disclosing the Y-cut plate. The temperature coefficient of the Y-cut is much worse than that of the X-cut, being about 100×10^{-6} per degree C. Nevertheless the Y-cut soon came to be favored over the X-cut because the mode of vibration, which results from a transverse wave traveling in the thickness direction, permits the plate to be clamped, thus resulting in a more rugged device. In 1934 workers at the Bell Telephone Laboratories [11] and in Germany and Japan discovered almost simultaneously that by inclining the plane of the Y-cut plate at certain angles with respect to the optic or Z axis the resonant frequency could be made nearly independ-

* Received by the IRE, June 19, 1961.
† McMurry College, Abilene, Texas.

ent of temperature over fairly wide temperature ranges. These cuts were named the *AT*- and *BT*-cuts.

The development of the *AT*- and *BT*-cuts was quickly followed by the development of a number of others designed to have low temperature coefficients of frequency and to operate over different ranges of frequency. One of these is the *GT*-cut, discovered by Mason [12] in 1940. By ingenious use of coupling between modes of vibration having different temperature coefficients, the *GT*-cut provides a standard of frequency which is substantially independent of temperature over a range of 100° of temperature. Oscillators controlled by *GT*-cut plates have been used as standards of time and frequency almost universally since that date and are only now being replaced by more stable devices.

Use of Quartz Crystal Units in Communication

The radio-frequency oscillators in early radio transmitters were subject to severe frequency variations because of varying environmental conditions and loading. It was early recognized that precise control of frequencies is essential to reliable communication. In 1926 WEAF became the first commercial broadcast station to be crystal controlled. Between then and 1940 crystal control became universal in broadcasting and many amateurs used crystal units to control the frequencies of their transmitters. By 1939 the growing use of radio in mobile communication had created a small market for quartz crystal units, most of which were made by hand by a few skilled craftsmen using simple equipment.

Slight use was made of crystal control in military communication equipment until immediately prior to the outbreak of World War II when the advantages of reliability, convenience and reduced weight were demonstrated in field tests. The decision to convert military communication equipment to crystal control practically coincided with the entrance of the United States into the war. A crash program was organized to provide the needed crystal units with such success that an industry which produced fewer than 30,000 units in 1939 produced over 100 million units during the years 1943–1945. This vast increase in production was made possible by the development of a number of ingenious machines including diamond saws, lapping machinery and X-ray diffraction equipment for use in orienting the plates precisely with respect to the atomic planes in the crystal of quartz. This was the first commercial application of X-ray diffraction to a production process.

Improvements in Design

The modern quartz-crystal unit bears little resemblance to its predecessor of twenty years ago. Fig. 2 shows a modern crystal unit beside its counterpart of 1940. The earlier unit was enclosed in a plastic holder having a volume of approximately 3 cubic inches, whereas the current design employes a metal holder having a volume of about $\frac{1}{8}$ cubic inch. The 1940 unit operated with a frequency tolerance of ± 0.02 per cent within the temperature range $-40°$ to 70°C, whereas the modern unit meets a frequency tolerance of ± 0.005 per cent from $-55°$ to 90°C. The earlier unit contains a quartz plate $\frac{3}{4}$ inch square clamped between two stainless steel plates which serve as electrodes and provide mechanical support. The modern unit contains a circular plate having a diameter of less than half an inch with metal electrodes deposited on the quartz.

When the highest stability is required, units are often encapsulated in glass holders using the techniques developed for the production of electron tubes.

The maximum frequencies available with quartz crystals prior to 1945 was about 10 Mc. Today crystal units operating at frequencies above 100 Mc are readily available. Still higher frequencies are desired.

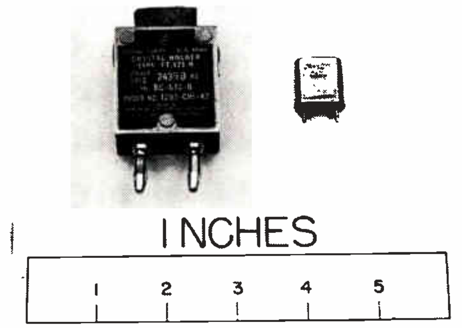


Fig. 2—Typical quartz crystal units of 1940 (left) and 1960 (right).

Technical Difficulties

The extreme precision required of the quartz crystal unit naturally imposes severe design and production problems. The thickness of a 10-Mc *AT*-cut plate is about 0.0065 in. To adjust the frequency of such a plate to within 100 cps of its assigned frequency, which is often necessary in production, the thickness of the plate must be adjusted with an accuracy of about 6×10^{-8} inch, which is equivalent to only two or three layers of atoms or about three thousandths of the wavelength of visible light.

With such critical dependence upon thickness, it is not surprising that the resonant frequency of a quartz plate should be highly sensitive to surface effects. Prior to 1943 the frequencies of quartz plates were commonly adjusted by removing quartz mechanically to obtain the desired frequency. Units prepared in this way are subject to "aging" which results from weathering of the surface of the quartz. Millions of units fabricated in this way were found, in 1943, to have become useless after a short period of storage. To prevent aging of this type, it is necessary to adjust the frequency of the quartz plate by chemical rather than physical means.

The deposition of as little as a monatomic layer of a foreign material on the surface of a quartz plate is often sufficient to render it useless for its intended purpose. It is therefore necessary to establish the most exacting standards of cleanliness and quality for the materials used in the fabrication of crystal units.

Other forms of aging, having less serious effects upon the device, are under investigation. These include changes in the physical structure of the metallic electrodes, rearrangement of the crystalline defects within the quartz and relaxation of stresses within the quartz and its supporting structure. Continued research and development has been necessary to keep abreast of the demand for tighter frequency control.

The variation of frequency with temperature is critically dependent upon the orientation of the piezoid with respect to the atomic planes of the crystal. In order to meet the requirements of modern specifications it is necessary to control the orientation with an accuracy of one quarter minute of arc. This accuracy is obtained using X-ray diffraction apparatus employing double-crystal techniques developed for the purpose.

Synthetic Quartz

Despite many investigations, no substitute has been found for quartz as a material for fabricating precision piezoids. However, quartz suitable for fabrication of quartz crystal units is relatively scarce and expensive and no significant domestic source has been discovered. To provide a reliable source of raw material, a program to develop techniques for producing synthetic quartz was instituted following World War I. This program was successful in developing techniques for producing synthetic quartz at a cost which is competitive with that of the natural product. Slight differences between the natural and the synthetic product make the present synthetic product inferior to natural quartz for some applications but superior for others. With further work it may be possible to produce synthetic quartz with special properties for specific applications.

Atomic Frequency Standards

Astronomical events such as the mean solar day have long been used as standards of time for reasons of permanence, availability and reliability. For the past thirty years quartz crystals have been used to interpolate between these relatively large time units and to provide standards of frequency based upon them. Quartz crystal units developed for the purpose and widely used as national standards of frequency are capable of accuracies of the order of 1×10^{-10} per day. If the clock controlled by such a crystal unit continued to operate with such precision it would deviate by only about 0.3 sec per century. But, however reliable a crystal clock is made, it cannot serve as a primary standard of time or frequency. The short-term accuracy is enough, however, to reveal serious discrepancies in the regularity of astronomical events.

Within the past ten years, devices utilizing atomic events have been devised which provide standards of time and frequency which are reproducible and reliable over periods of a few months to about 1.5×10^{-11}. Such devices will soon provide new standards of time and frequency far more accurate than those based upon astronomical observations. Most of these devices incorporate a high-quality quartz-crystal unit to furnish short-term stability, while the atomic standard provides reproducibility and long-term stability by correcting the quartz-crystal unit for aging and other effects.

W. P. Mason has remarked that the quartz crystal "... can be regarded as one of the cornerstones of the whole art and practice of electrical communication." It seems likely that it will continue to provide the foundation for a long time to come.

Bibliography

Very complete bibliographies on the subject of piezoelectricity and associated topics are found in [1] and [2].

[1] W. G. Cady, "Piezoelectricity," McGraw-Hill Book Co., Inc., New York, N. Y.; 1946.

[2] J. P. Buchanan, "Handbook of Piezoelectric Crystals for Radio Equipment Designers," Wright Air Dev. Ctr., Wright-Patterson AFB, Dayton, Ohio, WADC Tech. Rept. 56-156, ASTIA Document No. AD 110448; October, 1956.

[3] J. Curie and P. Curie, "Development by pressure of polar electricity in hemihedral crystals with inclined faces," *Bull. Soc. Min. de France*, vol. 3, p. 903; 1880.

[4] W. G. Cady, "The piezo-electric resonator," Proc. IRE, vol. 10, pp. 83–114; April, 1922.

[5] G. W. Pierce, "Piezoelectric crystal resonators and crystal oscillators applied to the precision calibration of wavemeters," *Proc. AAAS*, vol. 59, pp. 81–106; 1923.

[6] S. Butterworth, "On electrically maintained vibrations," *Proc. Phys. Soc.*, vol. 27, pp. 410–424; 1915.

[7] K. S. Van Dyke, "The piezoelectric resonator and its equivalent network," Proc. IRE, vol. 16, pp. 742–764; June, 1928.

[8] D. W. Dye, "Piezoelectric quartz resonator and equivalent electrical circuit," *Proc. Phys. Soc.*, vol. 38, pp. 399–457; 1926.

[9] L. Espenschied, "Quartz crystals used in ladder form circuits to obtain band pass filters," U. S. Patent No. 1,795,204; filed January 3, 1927, issued August 8, 1933.

[10] E. D. Tillyer, "Y-cut crystal vibrating in a shear mode," U. S. Patent No. 1,907,613; issued May 9, 1933.

[11] F. R. Lack, G. W. Willard, and I. E. Fair, "Some improvements in quartz crystal circuit elements," *Bell Syst. Tech. J.*, vol. 13, pp. 453–463; 1934.

[12] W. P. Mason, "A new quartz crystal plate, designated the GT, which produces a very constant frequency over a wide temperature range," Proc. IRE, vol. 28, pp. 20–223; May, 1940.

Relays and Switches*

A. C. KELLER†, FELLOW, IRE

Summary—**Relays and switches are among the most important components in the general area of electromechanical devices. Both of these have been in use for many years and continue, in modern form, to be essential elements in many modern switching applications including electronic systems.**

Early forms of relays were simple and controlled a single contact. Telephony created the need for more complex forms of relays and switches early in its history, and has continued ever since to require advances.

Switching functions such as counting, registration, translation, etc., are accomplished by relays and switches. The most difficult and complex switching functions arise from centralized or common control types of automatic control systems including telephone central offices and computers. These are often referred to as the "brains" of the system.

Modern trends in relays and switches are in miniaturization, hermetic sealing, improved reliability in the presence of shock and vibration and in higher operating speeds. Glass-enclosed contacts of the "dry reed" and mercury types have had an important part in achieving better performance. Another important trend is for them to become competitive, compatible and combinational with solid-state switching devices and systems as illustrated by the ferreed.

Research is continuing actively with miniaturized relays, piezoelectric relays, glass-enclosed contact relays, higher-speed relays, transistor-operated relays and combinational devices such as the ferreed. One form of ferreed will operate from a 5 μsec pulse and remain operated until released by a similar pulse.

RELAYS and switches are among the most important components in the general area of electromechanical devices. Both of these have been in use for many years and continue, in modern form, to be essential elements in many modern switching applications including electronic systems.

For the purpose of this paper *relays* are, essentially, remotely controlled switches which can open or close electrical contacts when suitable electrical conditions are met and are, therefore, two-state devices, *i.e.*, on or off.

Switches on the other hand are often two-state devices, but are also multi-state or multi-position devices. The names of switches often indicate the intended use or function such as, Limit Switch, Sensitive Switch, Coaxial Switch, Pushbutton Switch, Momentary Contact Switch, Pressure Switch, Thermal Switch, Lever Operated Switch, etc. Multi-state switches include Rotary Switches, Stepping Switches, Unidirectional Switches, Two-Motion Step-by-Step Switches, Crossbar Switches, Commutating Switches, Scanning Switches, etc.

Many switches are of the manual type, but equally important are those electrically operated in which case they might be considered to be relays.

* Received by the IRE, June 16, 1961.
† Bell Telephone Laboratories, Inc., New York, N. Y.

Switches and relays have many problems in common but none of more importance or more involved than that of the contact materials and the contact supporting members. Contact materials vary from expensive noble metals used for the best reliability with low-contact forces to the less expensive base metals often used with large contact forces. Modern relays and switches are often hermetically sealed to prevent contamination from external dust, dirt, and other environmental conditions. However, in some cases of sealed relays and switches, the entire structure is encased including the actuating coil, insulators, etc., which usually contain organic materials. These conditions are likely to result in polymer formation and the resulting high-contact resistance when certain precious metals are used for the contact surfaces. This is particularly true for metals in the platinum family including paltinum, palladium, rhodium, etc., when used in "dry" circuit applications, *i.e.*, those in which no contact arcs are involved and the circuit current is neither made nor broken. Important improvements have resulted in such cases by a separate sealed chamber in which all organic materials are eliminated.

Contact arcing is another type of problem which may cause excessive erosion of the contacts in certain types of circuits. Contact protection in the form of electrical components or networks is used, at times, to extend the life of contacts. To be successful in reducing contact erosion the protection must control the voltage to keep it below the breakdown voltage and also control the current to keep it below the minimum arcing current for the particular materials involved, and for the surface conditions of the materials, *i.e.*, clean or contaminated, smooth or rough, etc.

Early forms of relays were simple and rather large, controlling one contact. These were actually modifications of telegraph instruments and, in fact, the name relay came from telegraph practice, *i.e.*, the Morse Relay, etc.

Telephone created the need for more complex forms of relays and switches early in its history and has continued ever since to require advances. The advent of automatic or dial telephone switching systems resulted in rapid and important new forms of relays and switches (see Fig. 1). These were required to have more contacts, higher speeds, etc., and to be smaller in size. Other new features were switch selectivity involving 100 or more circuits and relays that responded to dial pulses or, on the other hand, were of the slow-release type to be held operated during a series of pulses.

As automatic dial-switching systems developed to become more completely mechanized, as automatic-con-

Fig. 1—Some steps along the road to modern relays. *Right*—the wire spring relay used in modern telephone switching systems.

trol systems for many other uses developed and as computers came into use, switching functions became more complex. Switching functions and switching units were given names, sometimes new and sometimes duplicate names; such as counting, registering, translation, selectint, decoding, directing, sender, marker, memory, etc. Many of the most difficult and complex switching functions arise from the common control types of automatic control systems. These are often referred to as the "brains" of the system. In a modern crossbar telephone system which uses essentially only repetitive units of crossbar switches and various kinds of relays, one marker may have 1500 relays. This type of circuit unit performs many logical operations. These are chiefly based on the two-state relay. Circuit requirements in terms of these can set up logical interrelationships of circuit paths and apparatus to meet the needed conditions. Twenty such two-state relays can be used to give more than 1,000,000 combinations. Logic such as "and," "or," "but," etc., can be easily obtained with simple relay circuits as well as "add," "substract," "divide," etc. These form the basic building blocks for the brains or common control switching functions.

Electromechanical relays and switches are now made in more forms by more manufacturers in larger quantities than ever before. This is of interest in view of the recent rapid development of solid-state switching devices and systems. In general, solid-state devices operate in microseconds compared to milliseconds or longer for electromechanical devices. The reasons for the continued expansion of the uses of electromechanical switching devices are several: 1) they often result in equipment designs which are simple and inexpensive, and yet operate fast enough to make unimportant any increase in operating speed; 2) they can be used singly and in small numbers without the associated common equipment often required to take full advantage of the sensational speeds of solid-state devices; 3) the rapid expansion of switching of all kinds (which requires more of everything including electromechanical and solid-state devices); and 4) relays and solid-state devices are developing a compatibility, and in fact, combinations of both have been developed.

The most widespread present trend in switches and relays in all types of switching systems is in miniaturization and in hermetically sealed types. (See Fig. 3). Relays have been manufactured and used in size steps down to "postage stamp" or "crystal can" sizes and smaller. These are not, in general, fully functionally equivalent to the larger sizes but they are being used in increasing quantities in many applications.

Several other important trends in relays have been: 1) code card operation by which notches in the actuating card determine the contact sequence; 2) magnetic latching which provides relay operation or release under the influence of pulses; and 3) frequency selective relays.

Reliability has also become increasingly important and low failure rates are required under severe conditions of shock, vibration, temperature and humidity as in missile and other military switching systems. The miniature relays are often better under shock and vibration than larger types of the same form because of the smaller inertia of the moving parts.

Another important trend for switching systems in general is for relays and switches to be competitive, compatible and combinational with solid-state switching devices and systems.

Glass-enclosed contacts represent another trend in improving the reliability of relay and switch applications. The glass-enclosed contacts are either the "dry reed" type or the mercury type. (See Fig. 2.) In either case they are free from external influences such as dust, corrosive atmospheres, ambient pressures, etc. In addition, relays using glass-enclosed contacts, particularly of the miniature type, are fast in operation and release, often less than 1 msec for each of these. Glass enclosed contacts of the mercury type have the least chatter (often none) and have the longest operating life (greater than 1,000,000,000 operations) of any relays yet designed.

Research is being actively continued in the areas of miniaturized relays and switches, printed-circuit relays, thermal relays, piezoelectric relays, glass-sealed contacts using dry and mercury contacts, higher-speed relays, frequency-sensitive relays (selectors), relay arrays, transistor-operated relays, and combinational devices such as the ferreed.

One form of the ferreed (see Fig. 3) is a combination

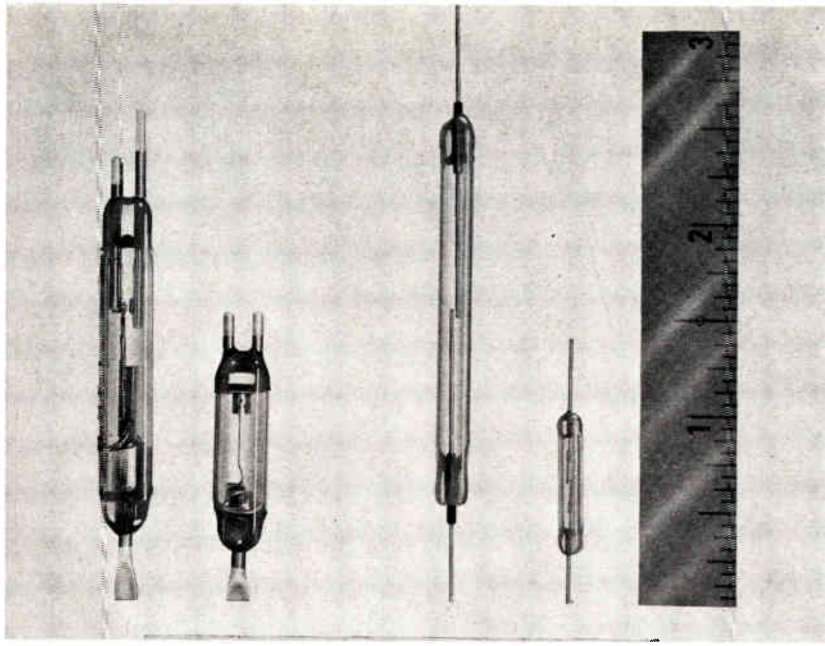


Fig. 2—Glass enclosed contacts for relays. *Left to right*—mercury-wetted type, small mercury-wetted type, dry reed type and miniature dry reed type.

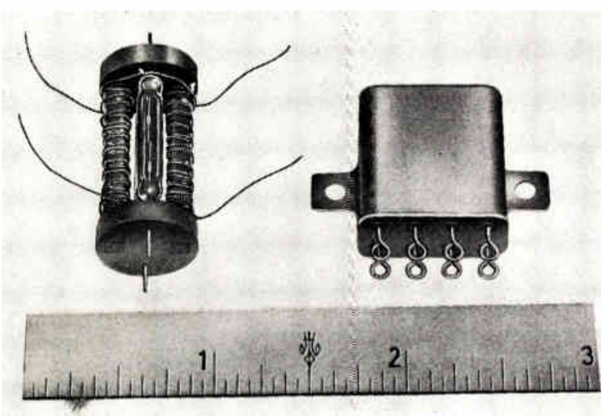


Fig. 3—Modern trends in relays. *Left*—ferreed with two contacts; *right*—DPDT "crystal can" relay.

of one or more miniature glass-sealed contacts and one or more pieces of ferrite or other suitable remanent material. Operating pulses as short as 5 msec can be used to operate the contacts of a ferreed. This is the time required to change the magnetic state of the ferrite to a magnetized condition. Later, perhaps $\frac{1}{2}$ msec later, the glass enclosed contact will operate. Another way to describe a simple ferreed is to say that it is basically a "magnetically latched" relay that can be operated or released by pulses. Ferreeds do not need holding power to hold them in the operated condition. The basic properties of the ferreed, *i.e.*, sealed contacts, control times in the μsec range, coincident selection, memory without holding power, and small size, make them compatible with solid-state circuit applications.

Transformers, Inductors and Filters*

HAROLD W. LORD†

Summary—**Transformers, inductors and filters have continually played important roles throughout the field of electronics. Relatively crude electromagnetic devices were employed in early telephone and radio communication systems and today their vastly improved descendants are still found in most electronic systems.**

The sizes of electronic power transformers have been reduced through the development of new core materials and suitable high-temperature insulating materials. Design methods for audio and video transformers have been evolved which, in combination with improved core materials, made it possible to provide broad-band transformers having low distortion effects. Pulse transformers were developed for radar systems and made practical the hydrogen thyratron type pulse modulator.

The performance characteristics of intermediate-frequency and radio-frequency transformers have been improved and their sizes reduced through the development of pressed-powdered iron and ferrite core materials.

The linearity and stability of iron-cored inductors for filters have been improved through core materials development. Nonlinear inductors were made more nonlinear through the development of rectangular hysteresis loop magnetic alloys and ferrites, providing many new types of electromagnetic devices for electronic circuitry.

Advances in electronics have made obsolete some applications of transformers and inductors but also have opened up other opportunities for inductive devices of novel or new characteristics.

INTRODUCTION

THE transformer, the inductor and the filter have continually played roles throughout the field of electronics. As relatively crude electromagnetic devices, they were employed in early telephone and radio communication systems. Today, vastly improved, their descendants are still found in most electronic systems.

TRANSFORMERS

General

A transformer performs one or more of the functions of 1) electrical isolation between circuits, 2) impedance matching, 3) voltage or current step-up or step-down, and 4) phase inversion. The voltage appearing between the terminals of a transformer coil must be an alternating voltage with a cyclic average value of zero, except for the effects of resistance voltage drops produced by a direct current through a coil. This characteristic is an advantage in many applications but prohibits their use in other circuits where their isolation and impedance matching functions might be advantageously employed.

Transformers for electronic applications are often divided into the general categories designated by the terms *Power*, *Audio*, *Video*, *Pulse*, *Intermediate frequency* (*IF*) and *Radio frequency* (*RF*).

* Received by the IRE, June 29, 1961.
† Res. Lab., General Electric Co., Schenectady, N. Y.

Power Transformers

Electrical isolation and voltage-changing are the chief functions of power transformers. For electonic use power transformers cover a range of ratings from a few watts to a few megawatts, the very large sizes being employed in rectifier circuits.

The most common power transformers are those supplying filament voltages and/or rectifier input voltages. The supply frequency is usually fixed and is from 50 cycles to a few kilocycles. A nominal frequency of 400 cycles is employed in many military applications.

Since 1945, considerable effort has been expended on reducing the size and weight of electronic power transformers with the most dramatic size reduction attained in 400 cycle operation. The development of reduced core loss materials at 400 cycles with high values of peak flux densities, coupled with the development of insulating materials suitable for use at high temperatures has led to this reduction. Recent developments in high-temperature transformers has culminated in units which operate for over 1000 hours at hot-spot temperatures of 600°C.

Audio and Video Transformers

Audio transformers have been employed since the infancy of electronics. Their chief functions are impedance matching, but all of the other basic functions of transformers are often employed. Prior to the development of high voltage-gain electron tubes, interstage transformers were employed to provide a part of the voltage gain between the relatively low output impedance of a tube plate circuit, and the high input impedance grid circuit of the following tube. Impedance-matching transformers made it practical to develop the wide-frequency range dynamic speaker with its low-impedance voice-coil and to drive it by efficient and low-distortion class AB push-pull amplifiers. Audio transformers range in size from those used in transistor audio amplifiers to the high-level modulation transformers in high-power AM transmitters, which include ratings up to several hundred kilowatts.

For many years, the designs of audio transformers involved some cut-and-try procedures. Even as late as 1948, the designs of the wide-range audio transformers for FM studio equipment often required a development effort consisting of several successive trial designs and tests in the specific electronic circuit in which they were to be used. Today, the design of high-quality low distortion, audio transformers capable of operating over a frequency range of 20 cycles to 20 kc and higher is standard engineering. This has been accomplished through core materials development and research studies of the interrelations of transformer parameters and electronic circuit parameters.

When the frequency range of a wide-band transformer includes frequencies much above the audio range, it is often referred to as a video transformer even though it does not cover the range of frequencies required by television video amplifiers. The development of thin high permeability alloys has made possible low-power tube-to-line transformers with practically linear response from about 50 cycles to above 1 mc in suitable circuits.

Pulse Transformers

Pulse transformers have played a key role in the development of radar. Most microwave transmitters include a pulse modulator circuit for supplying the input pulses to the microwave transmitter tube. In the first magnetron radar transmitters, the negative-going plate of the hard-tube pulse modulator was coupled through a capacitor to the cathode of the magnetron. The impedance level of the coupling between pulse modulator output and transmitter input was from 1000 to 1500 ohms. The pulse power level was about 100 kw, and the voltage level was 10 to 12 kv. Cables for these levels were not available, so the transmitter tube was required to be located close to the pulse modulator.

Recognition of the advantages of efficient transformers, capable of handling rectangular pulses of one or two microseconds duration and peak powers of 100 kw, led first to the development of transformers capable of matching the pulse modulator to a 50-ohm cable and the 50-ohm cable to the magnetron input. Low-power pulse transformers were developed along with the high-power types. These led to the development of pulse-forming blocking oscillators capable of operating at high repetition rates. The blocking oscillator replaced the less reliable low-level thyratron pulse-forming circuit employed in early hard-tube pulse modulators. In addition, a third winding on the blocking-oscillator transformer inverted the output pulse and eliminated the awkward "boot-strap" circuit employed in early pulse modulators.

The development of cable-to-magnetron pulse transformers made practical the hydrogen thyratron modulator. This consists of a resonant charged pulse-forming-network which the hydrogen thyratron discharges periodically into the cable supplying the pulse power to the cable-to-magnetron pulse transformer.

Present pulse transformers range in size from the tiny torroid-type employed in computer logic circuitry to large types weighing several thousand pounds for handling peak pulse powers of more than 50 Mw.

Intermediate Frequency (IF) and Radio Frequency (RF)

Intermediate-frequency transformers are in reality band-pass filters. They were an important part of the invention of the superheterodyne receiving circuit as they are the chief means for accomplishing high selectivity with reasonably uniform response within the pass-band.

The first superheterodyne receiver for entertainment use employed an intermediate frequency of 40 kc. The small IF transformers had cores consisting of thin laminations of iron. When, in later models, the intermediate frequency was raised, much larger air-core coils in shield cans were used as the Q of laminated iron core and coils was too low.

Subsequently, pressed powdered cores and more recently ferrite cores, have permitted efficient 456 kc IF transformers to have a small fraction of the volume of an equivalent air-core IF transformer. The size of ferrite cored IF transformers are commensurate with the sizes of transistors.

Radio-frequency transformers were originally air-core type coils. Impedance matching was effected by tapping or by a separate untuned secondary coil coupled to a tuned primary or vice versa.

Coefficients of coupling of air-core coils are too low for efficient operation as untuned broad-band transformers. Ferrite core materials have made possible the design of close-coupled untuned transformers for impedance matching purposes which operate satisfactorily over frequency ranges of three or four to one. Ferrite has also provided means for reducing the sizes of RF transformers where the operating frequencies are less than 20 Mc.

Inductors and Filters

General

Inductors employed in electronic circuitry vary from small loops of wire with a fraction of a microhenry inductance, to large filter chokes in dc power supplies. Those for high-frequency operation are either air-core or have cores of pressed powdered iron or ferrite. Those for operation at frequencies below about 5 kc may have laminated iron cores.

When part or all of the magnetic circuit linking a coil is ferromagnetic, the inductance of the coil is not a linear function of the current. For a given magnetic material, the nonlinearity of inductance increases with increasing ratios of core material length to the total length of the magnetic circuit.

Some magnetic materials are much more nonlinear in permeability than others. Temperature effects are quite different for various magnetic materials. If a variation of inductance is deleterious to the application of an inductor, care must be exercised in choosing the core material to be used and in choosing the percentage length of air-gap.

Linear Inductors

Linear inductors are required for most tuned circuit and filter applications. Magnetic materials have been specifically developed to meet requirements for linear inductors and have virtually eliminated the use of air-core inductors in filters and tuned circuits for frequencies below 20 Mc at low power. At high-power levels and/or at high frequencies, air-core coils are still used almost exclusively.

Nonlinear Inductors

There are a number of electronic circuit devices which depend for their operation upon the nonlinearity of the magnetization characteristic of magnetic materials. The magnetic amplifier is such a device. It is interesting to note that Dr. E. F. W. Alexanderson invented the magnetic amplifier to provide a means for voice-modulating the current which was supplied to the antenna by his high-frequency alternator. Thus the magnetic amplifier originally performed a function which was later performed by high-power electron tubes.

The utilization of nonlinear inductors was not very extensive until about 1940. Since then, new rectangular hysteresis loop magnetic alloys and ferrites have been developed. Nonlinear inductor devices employing these new magnetic materials are now used in combination with electronic devices (especially semiconductor devices) to perform functions which include: 1) amplification and control of voltage and/or current, 2) forming pulses for operating radar transmitters and firing thyratrons, ignitrons and silicon controlled rectifiers, 3) static switching, 4) frequency multiplication and division, 5) computer logic and memory, and 6) static permeability tuning.

Trends in the Areas of Transformers, Inductors and Filters

The trend toward miniaturization of low-power electronic circuitry may reduce the role which inductive devices play in low-power electronics. When one or more inductors (L) are combined with one or more capacitors, (C) to form a resonant circuit or a filter, the frequency selective functions can be performed by combinations of electro-mechanical transducers and mechanically resonant structures of metal or ceramic. The upper frequency limit of such electro-mechanical filters is presently about 5 Mc. Their size can be made less than LC type resonant circuits of comparable selectivity. This fact, plus microminiaturization, where the equivalent of LC functions are combined with the functions of other basic circuit elements in a semiconductor material, are factors which may cause inductive devices to disappear from low-power electronic circuitry.

The trend toward electronic devices of increased power capacity will call for rectifier transformers and possibly pulse transformers of higher voltages and power ratings. While the development of new electronic devices sometimes displaces certain inductive devices, they often create new uses for other inductive devices. The designers and manufacturers of inductive devices for the rapidly changing electronics industry may find it difficult to forecast the demands for a specific device, but if past performance can indicate future trends, the introduction of a new electronic device is likely to need an inductive device having novel characteristics.

Printed Circuits and Microelectronics*

S. F. DANKO†, MEMBER, IRE

***Summary*—In the two decades since the concept of ceramic based "printed circuits" was first suggested for an Army Ordnance application, substantial changes have taken place in the techniques of electronic circuit construction in the United States. Today, such ceramic circuits, in simple and complex networks, are complemented by printed wiring assemblies in several variations. Both technologies, now well established and in mass use on production lines, represent the current plateaus in miniature circuit construction for general commercial and military usage.**

Other construction philosophies and technologies are now shaping in the country's industrial and military laboratories, all aimed at new orders of size reduction of electronic equipments. In the several microelectronic techniques under development, the elemental electronic part appears destined to lose its logistic identity completely, and yield its classical position as a building block to black boxes called "circuit functions."

The paper provides a summary review of the evolution of current ceramic printed circuits and printed wiring practices and, in the light of today's microelectronic activities, frames the trend in equipment design in the years ahead.

* Received by the IRE, July 10, 1961.
† Electronic Components Research Dept., U. S. Army Research and Development Lab., Fort Monmouth, N. J.

DOZENS OF REFERENCES (some dating back 100 years) have been compiled to trace the subtle origin of some of the concepts and technologies upon which today's varieties of printed circuit systems are based. But the earliest coherent history of the printed circuit arts in the United States began in the Ordance Development Division of the National Bureau of Standards (NBS) in 1942 when a decision to explore two-dimensional printing of electronic parts for the Army's VT fuze was made by Ellett[1] and Diamond.[2] The successful implementation of this decision by Khouri,[3] and by a NBS research team under Brunetti[4] led to the first mass production of screened ceramic based printed circuit fuzes in 1945, and triggered the sequence of developments in the United States that, in the next two decades, profoundly affected the size, weight and physical appearance of commercial and military electronics.

The post-war electronic industry found this concept of screening of resistor inks and silver pastes useful for the manufacturer of miniature hearing aids, and radio and TV networks (Fig. 1). Aside from size reduction, the cost advantages inherent in mass-produced pre-assembled networks, inspection reduction, maintenance simplification, etc., were some of the attributes that attracted volume usage by commerical equipment manufacturers, and introduced the "module"[5] as a useful entity in electronic assembly. By 1950 a substantial market had developed for such ceramic based circuits and many innovations, and process refinements were introduced by the original contractor (Centralab) and other firms. By May, 1962, over 140 million screen deposited circuits have been manufactured by the Centralab Division alone in hundreds of varieties of simple and complex network configurations.

Continued military development of the related technologies by Franklin[6] led to usage of other techniques of part deposition and attachment, use of recessed substrates and improved protective coatings; today, this military derivative[7] of the screened ceramic circuit concept is specially oriented to weapons applications.

This early screening technology, however, could not satisfy the full spectrum of electronic circuit usage. The necessity for attaching "nonprintable" components (tubes, transistors, quartz crystals, transformers, etc.) and limitations in power-handling capacity did not permit realization of some of the optimistic projections made in 1946–1948 regarding the ultimate universal applicability of these screen deposition techniques. However, with this war-born success as a stimulant, Brunetti, in a 1947 Symposium,[8] sparked national interest in other technologies applicable to miniature circuit fabrication. Although the 1947 Symposium was (in retrospect) truly a brainstorming conference, it served as a springboard for dozens of independent and fruitful pursuits of electronic miniaturization and mechanical assembly (Fig. 2).

The U. S. Army Signal Corps, hard pressed for reduction of volume and weight of its mushrooming electronics systems, had established a miniaturization group in 1946 at the U. S. Army Signal Research and Development Laboratory (USASRDL) at Fort Monmouth. Following an intensive study and experimentation period, Danko and Abramson in 1949 announced the Laboratory's development of a rapid one-shot solder-dip system of component assembly to copper circuit (printed wiring) boards.[9] Etching of copper coils in runs of millions was then current at RCA under the technical guidance of Mackey and Black;[10] experimental copper-forming techniques were being used by J. Beck;[11] etched circuits on flexible plastic with surface attached parts were being fabricated by Henderson-Spalding,[12] and the Franklin Airloop Corporation[13] was embossing copper coils on plastic; the copper metal patterns being produced by all four were found compatible with the suggested Signal Corps solder-dip assembly system.

In the period 1949–1951, collaboration between Fort Monmouth and industrial laboratories led to considerable experimentation with solder-dipped etched circuits for commercial usage. The comparatively simple process steps involved in the fabrication of etched circuits from commerically available copper-clad laminates led many of the large manufacturers to establish in-house facilities for manufacturing printed wiring boards. Swiggett,[14] in organizing one of the first vendor sources exclusively for production of printed wiring and assemblies in 1951, provided a substantial stimulus to usage by a cautious industry. Figs. 3 and 4 reflect the growth of etched printed wiring usage in the subsequent years.

As with the screen stenciled ceramic circuits, innovations and process improvements in the printed wire concept came quickly as industry acceptance developed (Fig. 5). Single-sided printed wiring was followed by double-sided circuit boards, then plated circuits and plated-through holes, and multi-layer etched circuits; the manual dip-soldering technique of the early 50's yielded to mechanized soldering stations, wave soldering and selective soldering. In addition to the original copper-clad phenolic laminates for etched circuits,

[1] Dr. A. Ellett, Chief, Division A, National Defense Research Committee; now with Zenith Corp.

[2] H. Diamond, Chief, Ordnance Development Division, NBS (deceased).

[3] A. S. Khouri, Centralab Division of Globe Union, Inc. (contractor to Ordnance Corp.).

[4] Dr. C. Brunetti, Chief Engr., Ordnance Development Division, NBS; now with the Grand Central Rocket Co.

[5] In its current popular usage, "module" refers to a building block of electronic circuitry; this may be a passive network or an active circuit function with included tubes or semiconductor devices. A more disciplined definition has yet to be promulgated and accepted.

[6] P. J. Franklin, Ordnance Development Division, NBS; this research organization is now called the Diamond Ordnance Fuze Laboratories (DOFL).

[7] Identified as the *2D technique* by DOFL.

[8] Symposium on Printed Circuits, sponsored by NBS, October, 1947, Washington, D. C. (see Bibliography).

[9] "Auto-Sembly," U. S. Patent No. 2,756,485 assigned to U. S. Army; July 31, 1956.

[10] D. Mackey and O. D. Black, RCA Victor Division, Camden, N. J.; printed coils were being fabricated for use in the RCA TV tuner. Over 10 million etched coils were produced in this run.

[11] Beck's Inc., St. Paul, Minn. (U. S. Patent No. 2,683,839).

[12] Henderson-Spalding, London, England; the work of this firm was based on an etched foil patent by Dr. Paul Eisler (U. S. Patent No. 2,441,960).

[13] Franklin Airloop Corp., Long Island City, N. Y. (U. S. Patent No. 2,431,393).

[14] R. L. Swiggett, Photocircuits Corporation, Glen Cove, Long Island, N. Y.

melamine, teflon, and epoxy laminates became available. By 1955, mechanized lines for production of etched and plated circuits were in operation and component insertion machines for sequential assembly of parts on boards were commercially available.[15] Several mechanized production lines for electronic radio and TV sets came into being and "automation" became a new theme for symposia discussions.

Design innovations were introduced to increase the density of parts that could be placed on a printed wiring board. Three-dimensional stacking of commercially available miniature parts in cordwood fashion represented one of the more fruitful efforts to achieve greater packaging densities in military applications while retaining the inherent simplicity of the printed wiring concept.

These process, assembly, and design improvements, however, though individually contributing to the increasing quality and producibility of printed wiring circuitry, did not provide the "seven-league" steps toward further equipment size reduction that the future seemed to demand, particularly for the military's mushrooming electronic complexes. In 1957, this growing pressure for a radically new dimension for electronics veered many of the country's laboratories into: 1) modifying, extending or reducing to practice some of the miniaturization techniques and concepts suggested and tried years before, or 2) probing the realm of solid-state physics for the purpose of designing processed single crystal semiconductors to perform electronic functions conventionally performed by aggregates of individual parts. This revitalized pursuit of size reduction of electronics immediately picked up the term *microminiaturization* as a general identification; in the next few years dozens of proprietary names were attached to processes and technologies used to fabricate ultra-small experimental electronics (Fig. 6).

This microminiaturization activity can be broadly classified into three major categories: 1) *high-density packaging*[16] in which individually fabricated parts in extremely small sizes or in special shapes from established vendor sources are compactly packaged and interconnected by handwiring, printed wiring, or welding, 2) *thin film techniques*[17] in which two-dimensional parts are formed in place directly on a common substrate by any appropriate technique such as screening, vacuum evaporation, sputtering, etc., 3) solid-state circuits, one form of which, called *integrated circuits*, consists of single crystal semiconductor chips which are processed by etching, diffusing, alloying, and evaporation to duplicate, on an equivalent part-by-part basis, the conventional circuit schematic; within this general solid-state classification is another more general concept called *molecular electronics* which does not recognize this part-for-part equivalence, but attempts to apply appropriate materials phenomena to perform the same circuit function.[18]

Whether the circuit function is fabricated using high-density packaging, thin films, or solid-state phenomena, it appears that the next decade will see substantial growth in the acceptance of the electronic functions in the form of "little black boxes" with specific performance characteristics guaranteed (Fig. 6). It appears very likely that each of the systems being explored today will find individual "black box" applications in which it will show distinct performance, reliability, or cost advantages. The history of electronics, as a matter of fact, reflects continuous hybridization of technologies. Our electronic systems today, for example, are hybrid constructions using handwiring, printed wiring, ceramic printed circuits, module construction, and thin film techniques applied to the manufacture of some of the parts; even examples of the molectronics concept (perhaps not in the microminiature dimension) can be pointed out in today's systems. It is a certainty that industry and the military will accept and use any technological advance in electronic parts, in assembly techniques, or in the physical form of the elemental building blocks if the economics, reliability, logistics, and producibility of the final product (the equipment) are enhanced or, at the very minimum, stay the same. Accordingly, an open and inviting field for accomplishment and acceptability waits for the products of all three of the current approaches on a broad competitive basis which includes more than just microminiaturization. It is highly unlikely that tomorrow's designer, once convinced that the use of packaged electronic functions is to his advantage, will arbitrarily reject one technology for the black boxes he procures, any more than he would question the mixture used in the composition resistors he uses today. Performance, reliability and price will be the principal criteria for acceptance.

The current substantial research and development effort by industry and government in the microelectronics area is evidence enough of the importance of the microminiaturization objectives for the military purpose. However, it is expected that, as in the case of the ceramic based printed circuits and printed wiring, the by-product benefits of this military motivation will accrue to the entire body of commercial and industrial consumers.

[15] United Shoe Machinery Corp., Boston, Mass.

[16] The Micro-Module (RCA-Signal Corps), using standard-shaped (0.31 in×0.31 in) notched wafer parts stacked and interconnected by riser wires is in this category. So are the several varieties of cordwood assembly and the several variations of using miniature parts with one disciplined dimension (thickness or width) recessed within the thickness of a board which is then surface printed to provide interconnections.

[17] Thin film circuits today are still hybrids of deposited and attached parts; this approach is rooted in the screened ceramic circuits of World War II (Fig. 1) which first developed the concept of two-dimensional parts processed *in situ*. The ultimate goal of thin film technologists is the substrate deposition of most (if not all) parts, including the semiconductor devices.

[18] Molecular Electronics (also called Molectronics) is the principle behind an Air Force-Westinghouse microminiaturization program.

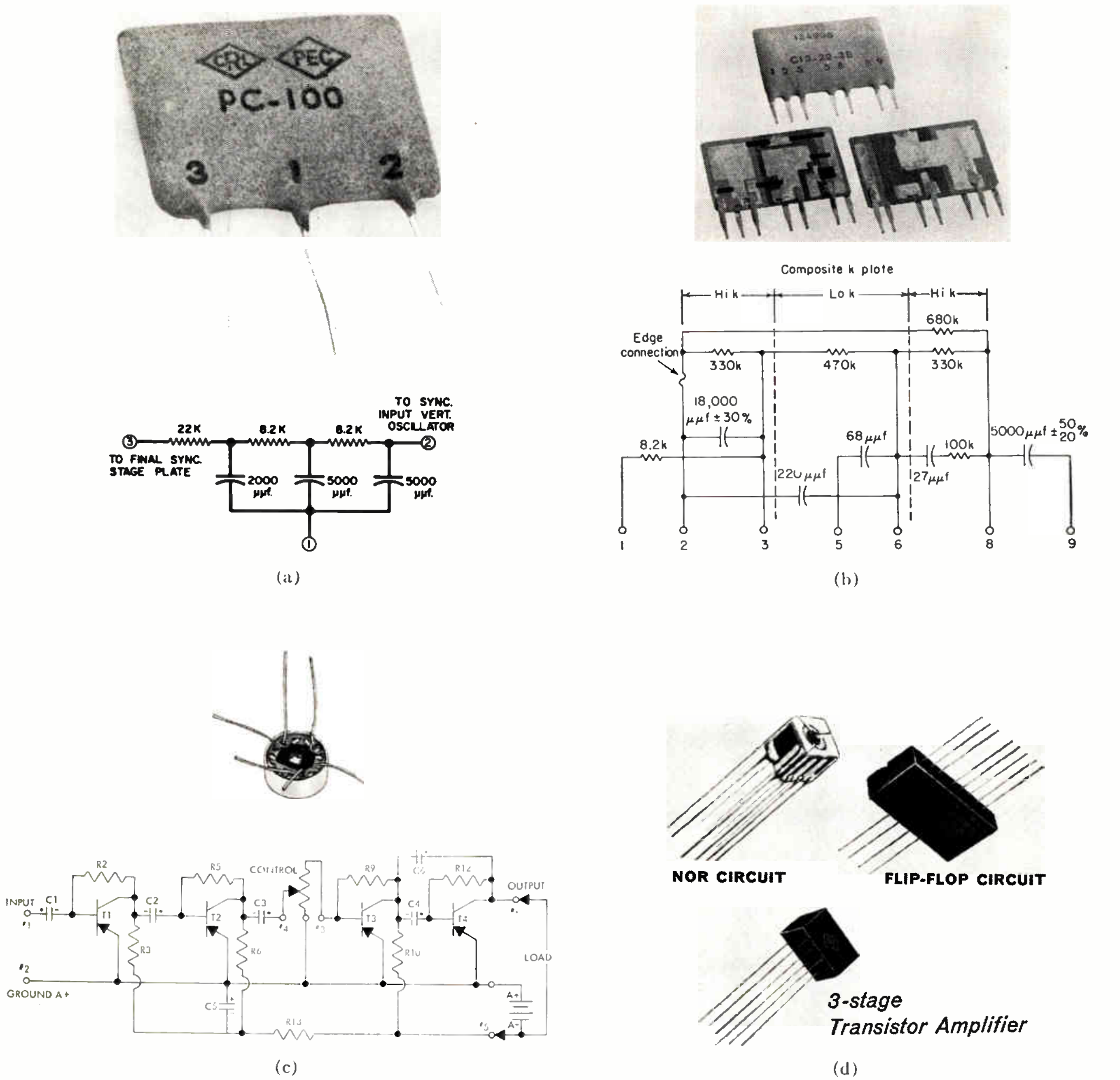


Fig. 1[19]—*Ceramic Based Printed Circuits:*

Screen deposition of resistor inks and silver pastes today is a matured, versatile technology applicable to two-dimensional RC networks for interconnecting tubes or semiconductor devices, and to fabrication of three-dimensional complete circuit functions with enclosed conventional semiconductor elements. Well over 140 million of such ceramic circuits have been manufactured by Centralab since World War II.

(a) A relatively simple vertical integrator for TV applications uses a high K plate as a substrate; the $1\frac{3}{32}$ in $\times\frac{13}{16}$ in $\times\frac{1}{8}$ in unit has 450 v dc working voltage capacitors and $\frac{1}{5}$ watt ± 20 per cent resistors. Over 16 million units of this integrator have been manufactured, indicative of the mass producibility, and the performance and economic acceptability of this form of packaged RC circuitry. Steatite is also used as a substrate in other designs.

(b) A more complex network showing use of a composite substrate having different dielectric constant areas. Ratio of capacity values up to 1000 to 1 and resistance values up to 10 to 1 can be achieved in one screening operation. Maximum capacity values approximate 0.03 μf. Resistor inks and geometric patterns are selected to suit resistance values desired (5 to 100 megohms); resistor tolerances down to ± 5 per cent can be controlled in production if desired.

(c) A three-dimensional audio amplifier module made up of stacked ceramic circuits and transistor elements. Unit is hermetically sealed, and typifies the capability for higher density packaging using stacking techniques. Unit size is 0.531-in D$\times$0.228 in thick.

(d) Other three-dimensional module configurations based on assembly of planar ceramic circuits: a nor circuit (top left) in a $\frac{3}{8}$ in $\times\frac{3}{8}$ in $\times\frac{13}{32}$ in configuration containing 8 resistors and 1 transistor; a flip-flop circuit (top right) $\frac{1}{4}$ in $\times\frac{1}{2}$ in $\times 1$ in (approximately) containing 8 resistors, 6 capacitors, 6 diodes and 2 transistors; a three-stage transistor amplifier (bottom) $\frac{1}{4}$ in $\times\frac{1}{4}$ in $\times\frac{1}{2}$ in. Module parts densities up to about 1 million parts/cubic ft can be realized in practical, mass producible designs.

[19] All photographs courtesy Division of Centralab of Globe Union.

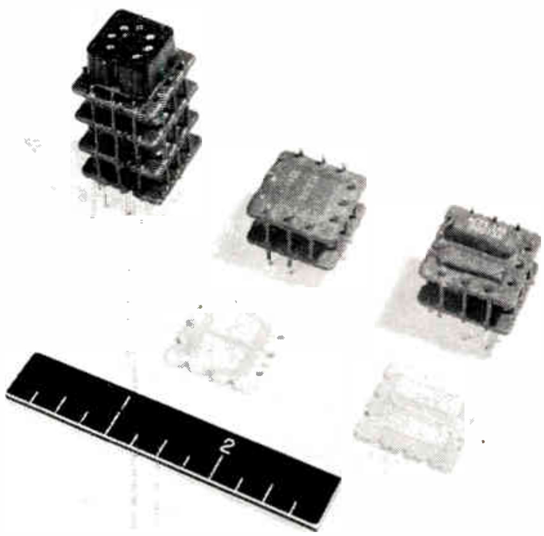


Fig. 2—*Mechanize Production of Ceramic Wafer Assemblies*

In the early 1950's, Robert Henry of NBS, working under Navy sponsorship, completed development of an experimental mechanized line for producing finished electronic assemblies using a novel stacked ceramic wafer approach with riser wire interconnections. The system (called "Tinkertoy" at that time) considered component manufacture from processed materials as part of the mechanized production complex. Assemblies such as these with considerably re-designed machinery are being produced to-day by Paktron, Division of Illinois Tool Works; about 5 million of these stacked wafer modules have been manufactured as of May, 1962.

In an independent 1957 study of feasible approaches to automatic production of microminiature circuits, RCA also concluded that a physical system such as Henry's, based on uniformly shaped miniature parts with riser wire interconnections, offered the best features for automatic assembly (see Fig. 6).

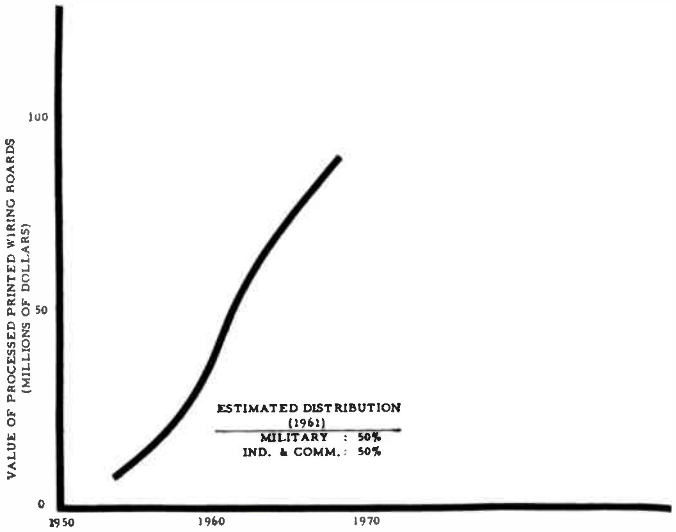


Fig. 3—Estimated and projected growth of U. S. printed wiring board industry.

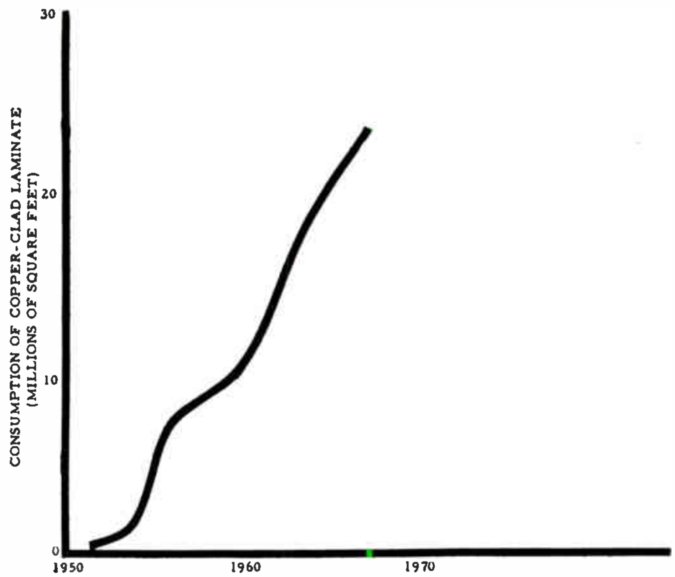
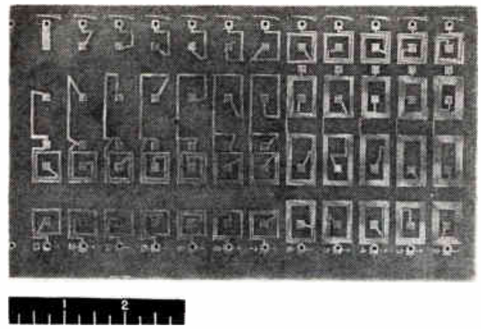
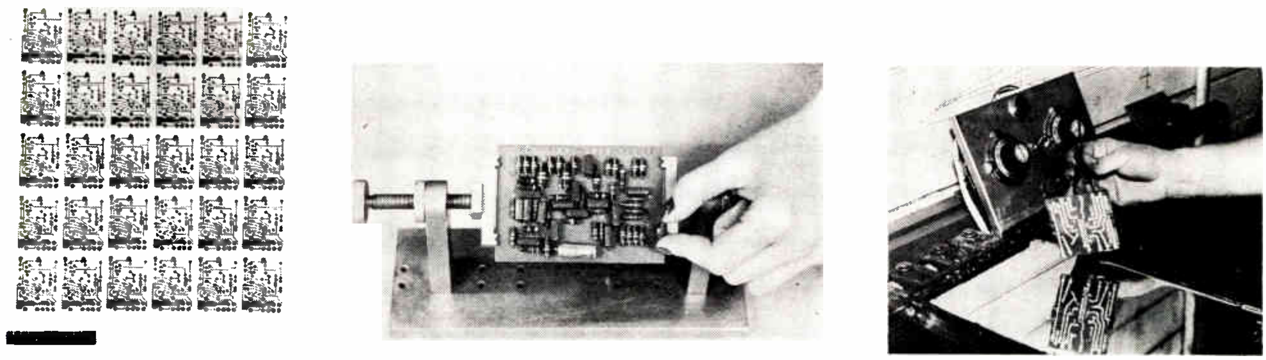


Fig. 4—Estimated and projected growth in U. S. usage of printed wiring boards.

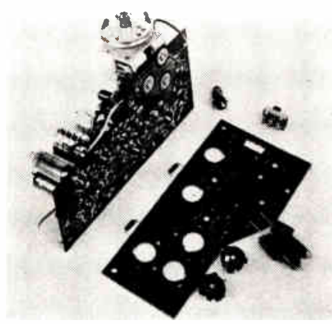


(a)



(b)

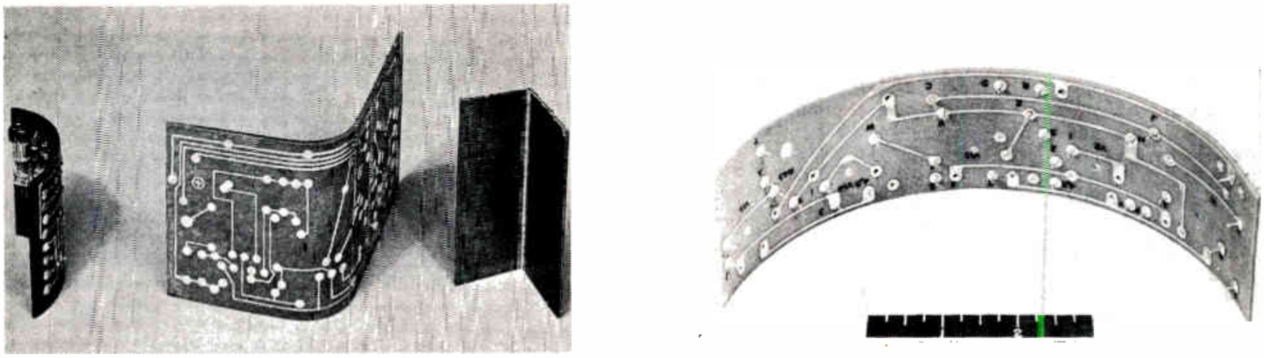
(c)



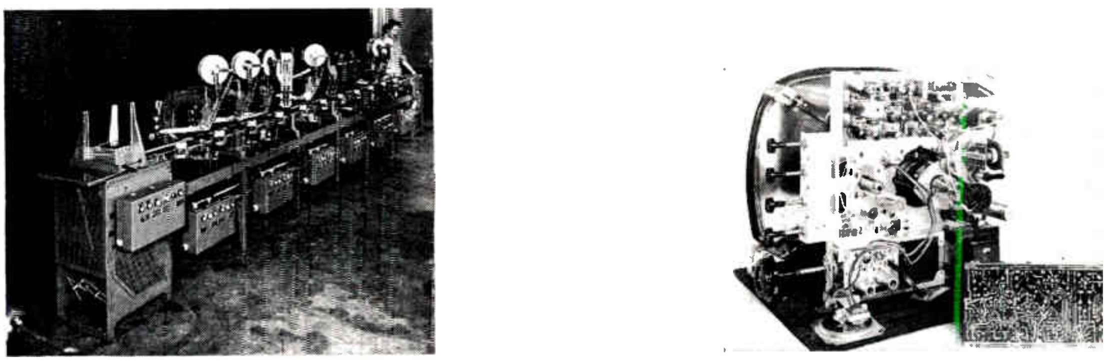
(d)

Fig. 5—*Printed Wiring:*
(a) 1948–49—TV tuner coils mass produced by photo-resist etching of foil clad laminates (RCA).
(b) 1949—Rapid multiple-pattern fabrication via photo-resist etching, insertion of conventional electronic parts into pre-punched holes, and one-shot solder dip assembly provided size and production advantages which attracted industrial and military equipment designers to the printed wiring technique (Signal Corps).
(c) 1952–1953—The first mass produced printed wiring (PW) home radio sets appeared in the 1952–1953 period, with manually inserted parts and solder dipping being done manually or by semiautomatic machines. Motorola (left) introduced its plated circuit (Placir) process as a mass production technique; laminate etching process was used by Admiral (center) and Hallicrafters (right).
(d) 1953—Selective solder-dipping (solder resistricted to joints only) was used by Philco in its mass production of this home radio receiver.

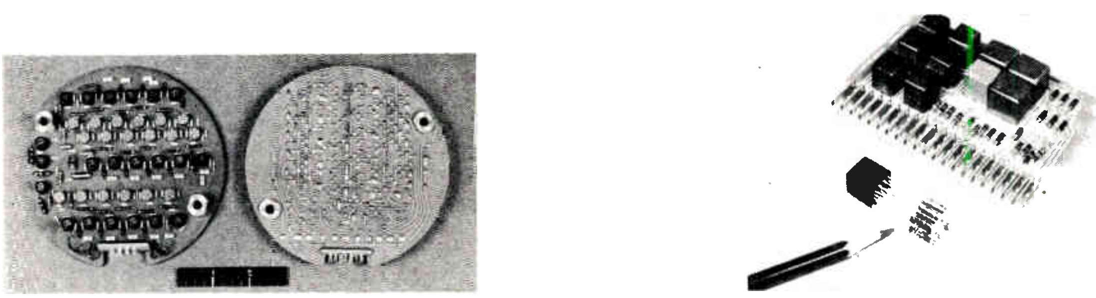
(*Fig. 5 continued opposite page*)



(e)



(f)



(g)

Fig. 5—(*continued*)

(e) Curved laminates suggested by Signal Corps (left) found one of its first military applications in an Army Weapons System (right).

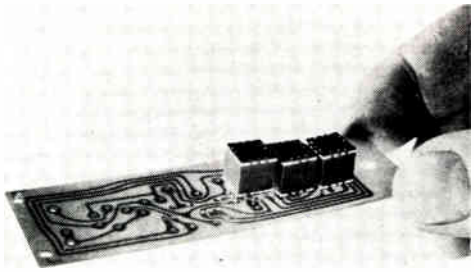
(f) 1953–1954—Mechanization of the electronic circuit assembly operation, a thwarted ambition of Sargrove,[20] became a reality in the United States in the early '50's. The United Shoe Machinery Company (left) developed individual component insertion machines which could be coupled in multiple to a conveyer system and a solder dipping apparatus to assemble most of the PW electronic chassis automatically; hand insertion and attachment of special components was required to complete operation. In this period too, Motorola was operating its highly mechanized Placir process for pattern fabrication. In 1954, the Admiral Corporation (right) operated an impressive automatic assembly line for its 21-in TV receiver, fabricating a substantial part of the circuitry (shown beside set) on a mechanical line having sequential insertion stations fed by reeled components and an automatic dip-soldering position. Other mechanized assembly lines were being independently developed at Sylvania, IBM, RCA, etc.

(g) 1956–1960—Military usage of printed wiring gathered momentum in this period; one out of every 3 new equipments developed at the Fort Monmouth Laboratories of the Signal Corps used printed wiring in whole or part. Weapons systems, computers, satellites, portable and mobile communication receivers, airborne electronics swelled the number of applications of printed wire so that in 1961, authoritative estimates placed the military consumption of processed printed wiring boards at about 50 per cent of the dollar value. Photo shows a high density deck of the Explorer VII Satellite, fabricated with PW solder-dip technqiues by the Army Ballistic Missile Agency; cumulative experience and improved techniques permitted greater and greater use of such high-density designs in ground and airborne equipments. Four military specifications and standards[21] were in effect in 1960 and helped stabilize the quality and performance of assemblies. A technique of packaging parts between two printed wiring end-plates ("Cordwood"), suggested by Roman[22] a number of years previously, matured into commercial and military applications; the figure on the right shows a hybrid of such molded "Cordwood" modules and conventional PW constructions, which can yield up to two times better space efficiency (Republic Aviation Corp.).

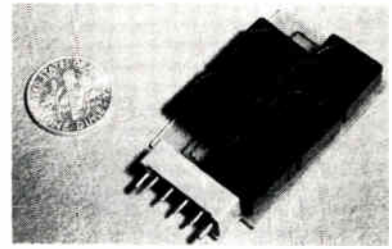
[20] In 1943, J. A. Sargrove, England (U. S. Patent No. 2,474,988) had suggested a production line system of circuit fabrication involving spraying metal into a grooved phenolic base for form conductors, capacitors, coils and switches and spraying graphite suspensions to form resistors. The system was abandoned after considerable effort was put into physical construction (see Bibliography).

[21] MIL-P-13949B—Plastic sheet, laminated, copper clad (for PW); March 15, 1960.
MIL-P-55110—Printed wiring boards; September 6, 1960.
MIL-STD-275A—Printed wiring for electronic equipment; September 7, 1960.
MIL-STD-429—Printed circuit terms and definitions; December, 1957.

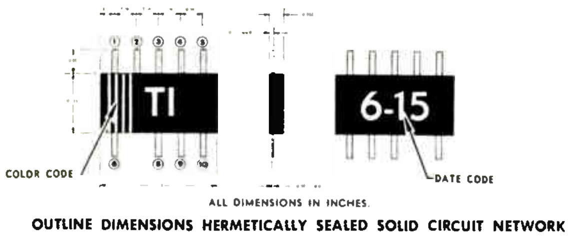
[22] Robert J. Roman, Eastman Kodak Company. (See Bibliography)



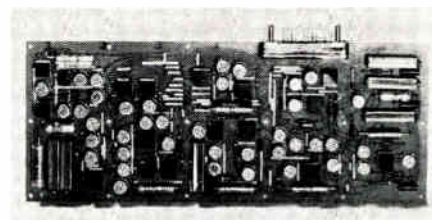
(a)



(b)



(c)



(d)

Fig. 6—*Functional Circuits* (*Illustrative*):

(a) *Micro-modules* (*Signal Corps-RCA*): A high-density packaging approach consisting of standard shaped 0.31 in×0.31 in ceramic based parts (micro-elements) vertically stacked and interconnected by[12] riser wires soldered in notches. From 5 to 10 times size reduction of modules has been achieved over the best conventional printed wiring constructions. Final product can be treated as a functional entity for machine insertion and solder dipping on printed wiring boards.

(b) *Thin film circuits* (*International Resistance Co.*): Thin film circuits on substrates of glass are plastic encapsulated and appear as functional nuggets with stiff wire projections for use with PW boards. Illustration shows a half-adder consisting of five such circuits mounted on a multi-layer PW board. Thin film circuits consist primarily of conductors, R's, and C's at this stage of the art. Plastic encapsulation for climatic protection and ruggedization very substantially increases the effective size of the basic thin film circuit and this problem is one of the prime targets of researchers in the field. Size reductions of 20 to 1 over conventional PW circuits using miniature R's and C's appear attainable.

(c) *Solid-state circuits* (*Texas Instruments*): Processing of single crystals of silicon to duplicate circuit schematics yields this tiny hermetically sealed flip-flop circuit. Applications up to now have been limited almost exclusively to digital circuits. Practical size reductions of modules up to about 350 to 1 over conventional PW units are expected in the full application to computer circuitry.

(d) Illustration of hybrid use of conventional parts, printed wiring and functional circuit modules (micro-modules in this instance). It is expected that hybrid use of conventional thin film, solid-state, and high-density packaged functional circuits will become increasingly common in the next decade (RCA).

Acknowledgment

The author appreciates the generous help and constructive observations of a number of individuals and organizations in the printed circuit field, in particular: Robert Henry, Paktron Division; Robert Swiggett, Photocircuits Corporation; Institute of Printed Circuits; Robert Gerhold, USASRDL; J. G. Reid, Fort Huachuca; Robert Curran, Electralab; W. S. Parsons, Centralab; Harold Stral, Stral Company; Vincent Kublin, USASRDL; C. Rayburn, Paktron; Donald Mackey, RCA, and W. J. Bornemann, Hamilton Standard.

Bibliography

Ceramic Based Circuits

[1] C. Brunetti and R. W. Curtis, "Printed Circuit Techniques," National Bureau of Standards, Washington, D. C., Circular 468; November 15, 1947. (This complete review of the then known printed circuit field was also published under the same title in Proc. IRE, vol. 36, pp. 121–161; January, 1948.

[2] T. A. Prugh, J. R. Nall, and N. J. Doctor, "The DOFL microelectronics program," Proc. IRE, vol. 47, pp. 882–894; May, 1959.

[3] M. Kahn, "Ceramic-based micro-circuits—a heterogeneous approach to miniaturization," *Proc. 1960 Electronic Components Conf.*, Washington, D. C., May 10–12, 1960; pp. 13–19.

[4] G. J. Selvin, "The Sylvania microminiature program," *Proc. 1960 Electronic Components Conf.*, Washington, D. C., May 10–12, 1960, pp. 46–51.

[5] J. Kirscher, "Packaged electronic circuits," *Electronics World*, vol. 63, pp. 35–38; May, 1960.

[6] C. C. Rayburn, "Design trends in low-*K* substrate modules," *Electronic Ind.*, vol. 20, pp. 92–95; March, 1961.

Printed Wiring and High Density Packaging

[1] D. Mackey and E. Suss, "Utilization of PC components in television tuner," *RCA Rev.*, vol. 12, pp. 293–302; September, 1951.

[2] S. F. Danko, "New developments in the Auto-Sembly technique of circuit fabrication," *Proc. National Electronics Conf.*, Chicago, Ill., October 22–24, 1951; pp. 542–550.

[3] R. L. Swiggett, "Printed circuits on foil clad plastics," *Mod. Plast.*, vol. 28, pp. 99–100, 105–107, 110–111; August, 1951.

[4] R. A. Gerhold, "Techniques of producing printed wiring boards," *Proc. Symp. on Printed Circuits*, Philadelphia, Pa., January 20–21, 1955, RETMA; (a compendium of papers reflecting the contemporary status of printed circuits).

[5] *Ibid.*, R. L. Swiggett, "Printed circuit switches and commutators."

[6] S. G. Bassler and M. Hinebaugh, "Principles of electronic circuit packaging," *Elec. Mfg.*, vol. 56, pp. 77–85; August, 1955.

[7] N. R. Cass, "Double dip soldering," *Electronics*, vol. 28, pp. 226–230; December, 1955.

[8] R. M. Barrett, "Microwave printed circuits—a historical survey," IRE Trans. on Microwave Theory and Techniques, vol. MTT-3, pp. 1–10; March, 1955.

[9] R. J. Roman, "A subminiature printed circuit system," *Proc. Symp. on Printed Circuits*, Philadelphia, Pa., January 20–21, 1955.

[10] R. Strauss and A. Barnes, "The flowsolder method of soldering printed circuits," *Electronic Eng.*, vol. 28, pp. 494–496; November, 1956.
[11] R. L. Swiggett, "Introduction to Printed Circuits," John F. Rider, Inc., New York, N. Y.; 1957.
[12] C. Rhine and S. Lewis, "A survey of materials, methods and equipment for printed circuits," *Plastics World*; November, 1957.
[13] P. Eisler, "The Technology of Printed Circuits," Academic Press Inc., New York, N. Y.; 1959. (Includes comprehensive bibliography of printed circuits.)
[14] S. F. Danko, W. Doxey, and J. P. McNaul, "The micro-module: a logical approach to microminiaturization," Proc. IRE, vol. 47, pp. 894–904; May, 1959.
[15] D. Mackey, "The micro-module program," *Proc. 1960 Electronic Components Conf.*, Washington, D. C., May 10–12; pp. 51–59.
[16] N. Schuster and W. Reimann, "Multilayer etched laminates," IRE Trans. on Product Engineering and Production, vol. PEP-5, pp. 114–128; June, 1961.
[17] S. F. Danko and S. J. Lanzalotti, "Auto-Sembly of subminiature military equipment," *Electronics*, vol. 27, pp. 94–98; July, 1951.

Thin Film Circuits

[1] G. N. Queen and J. J. Bohrer, "Preparation and evaluation of thin film circuit functions," *Proc. 1960 Electronic Components Conf.*, Washington, D. C., May 10–12, 1960; pp. 39–45.
[2] "Thin Film Circuit Functions," Final Rept., Signal Corps Contract DA36-039 SC-84547 (IBM); August, 1959–July, 1960.

Integrated and Molectronic Circuits

[1] "Molecular Electronics—Dendritic Approach," WADD Tech Notes 60-97 and 60-295 (Reports by Westinghouse Electric Corp., DYD-45106 and DYD-45158, dated February and July, 1960, prepared for WADD, USAF).

General State of the Art and Automation

[1] "New Advances in Printed Circuits," National Bureau of Standards, Washington, D. C., publication 192; November 22, 1948.
[2] J. A. Sargrove, "Machine produces complete radios automatically—three a minute," *Elec. Mfg.*, vol. 41, pp. 116–120, 170; June, 1948.
[3] E. Waverling, "Mechanization of wiring—key to automatic assembly," *Electronics*, vol. 25, pp. 140–142; January, 1953.
[4] R. L. Henry and C. C. Rayburn, "Mechanized production of electronic equipment," *Electronics*, vol. 26, pp. 160–165; December, 1953.
[5] *Proc. Symp. of Automatic Production of Electronic Equipment*, Philadelphia, Pa., April 19–20, 1954.
[6] "Fabricating television circuits," *Automation*, vol. 2, pp. 31–34; February, 1955. (Admiral Corporation's mechanized production line.)
[7] "Machine automates assembly of printed electron circuits," *Automation*, vol. 2, pp. 65–68; April, 1955. ("Autofab" concept developed by General Mills.)
[8] *Proc. RETMA Symp. on Automation*, Philadelphia, Pa., September, 26–27, 1955.
[9] "Microelectronics and Systems," Issue IRE Trans. on Military Electronics, vol. MIL-5; July, 1961.
[10] A. R. Gray, "Automation for electronics—a 1956 status report," IRE Trans. on Production Techniques, vol. PT-2, pp. 29–45; April, 1957.
[11] V. O'Gorman, "An automatic dip soldering machine," 1957 IRE National Convention Record, pt. 6, pp. 43–48.
[12] C. Brunetti, "A new venture into microminiaturization," 1957 IRE National Convention Record, pt. 6, pp. 3–11.
[13] J. Markus, "Mechanized production of TV wiring boards," *Electronics*, vol. 30, pp. 138–143; April, 1957.
[14] *Proc. 1960 Electronic Components Conf.*, Washington, D. C., May 10–12, 1960.
[15] J. R. Black, "Microelectronics and the systems engineer," *Ibid.*, pp. 36–39.
[16] H. A. Stone, "Components of 1970," presented at the 1960 Electronic Components Conf. (Paper available from Bell Telephone Labs., Inc., Murray Hill, N. J.)
[17] A. E. Javitz, "Key problems in microelectronics," *Electro-Tech.*, vol. 67, pp. 61–66; June, 1961.
[18] V. J. Kublin, S. F. Danko, and R. A. Gerhold, "Miniaturization" (a 4-part series), *Electronic News*, February 29, March 14, 21, 1960.

Electronic Materials, 1912-1962*

PRESTON ROBINSON†, MEMBER, IRE

***Summary*—"Electronic Materials" is taken here to include dielectric materials with plastics and gases; magnetic materials, including permanent magnet materials, soft magnetic materials, magnetic powders, magnetic ceramics; resistive materials; ferroelectric materials; wire insulation; insulators; semiconductor materials; and conductors. In a general statement, as given here, fine distinctions have been avoided, and there is necessarily some overlap in these categories.**

Introduction

All the original electronic materials are still with us. The first condenser recognized as such was the Leyden jar, employing glass as a dielectric, and glass is still extremely popular today. The oil-paper dielectric of the condenser Galvani used to accumulate the charge from his galvanic cells is still with us. The galena, as used with the first cat's whisker, is still used in many photoelectric devices. Mica is still a popular dielectric. Lodestone, until very recently, was still used in some cores for permeability tuning. Further, many of these old materials are among those most stable to high temperatures and nuclear radiation, so that a good question is: what have we been doing the last 50 years?

* Received by the IRE, August 28, 1961.
† Sprague Electric Co., North Adams, Mass.

Dielectric Materials

As stated, glass, mica and oil-paper are still with us. The glass has a higher dielectric constant now and is thinner. Higher temperature glasses also are available. There is now synthetic mica which employs fluorine instead of hydroxyl ion in the molecule constituting the mica, and this permits a mica to be melted and retain many of the other good characteristics of mica. Films consisting of compacted mica particles and also of bentonite particles are in use. The oil-insulated paper is thinner, and while originally vegetable fibers were used, today wood pulp treated by the Kraft process has perhaps wholly replaced the other because of the superior temperature stability and lower power factor. And the vegetable oils originally used for impregnating purposes have, with the possible exception of castor oil, been replaced by minerals oils together with synthetics such as chlorinated diphenyl and liquid polymers of isobutylene. For solid impregnants, chlorinated naphthalene and high melting olefins have replaced the natural waxes. Silicone oils have found use as impregnants to a considerable extent, and liquid perfluoro compounds to a much lesser extent.

Plastics

As a sheet dielectric for use in capacitors as well as a spacer for cables, there are a number of new materials. Polystyrene, with its vanishingly small power factor and low temperature coefficient of capacity, is widely used in low temperature applications for condensers which previously used mica. Mylar is increasingly used as a dielectric in condensers where an extended temperature range and less severe temperature requirements make it possible. Newer materials are the polycarbonates, which have excellent temperature characteristics, the cyanoethyl cellulose films, which, while they have not found their way into production, are of interest because of the high dielectric constant, 15 to 20. At higher temperatures, fluorinated materials are available of which polytetrafluoroethylene (Teflon) has the highest temperature, although other fluorinated materials for somewhat lower (but still high) temperatures are also available.

Plastics are widely used as containers for electric components, and these usually consist of the phenolics, the polyesters, epoxies, diallyl phthalate and triallyl cyanurate.

Monomers which polymerize to form plastics have two functional groups which can form links with other monomers, and which, for example, can be called head and tail. Ordinary polymers, now called atactic, have large numbers of monomers joined up randomly. Syndiotactic polymers have monomers joined head-to-head and tail-to-tail. And, finally, isotactic polymers where the monomers are uniformly joined head-to-tail, have much better thermal and mechanical properties than the earlier versions. Thus, isotactic (high density) polyethylene and polypropylene are having spectacular success. Isotactic polystyrene is still in the laboratory but has promise for the future.

Linear polymers which have been cross-linked to form fairly rigid three-dimensional structures by the effect of radiation have found success, particularly with cross-linked polyethylene. Radiation is also one method used to form so-called "block and graft" polymers, where polymers of one species are grafted to the chain comprising a polymer of another species. These, too, are largely in the laboratory stage.

Gaseous Dielectrics

Where once we used air, nitrogen and carbon dioxide, there are now other gases available to permit higher breakdown field strengths, of which the most notable are, sulfur hexafluoride, selenium hexafluoride and the family of fluorocarbons, and of the fluorochloro carbons known as freons. These are all "electronegative" gases, and, while not as stable as nitrogen under all conditions, can be used with proper precautions.

Finally we should not forget "nothing" as a dielectric. While "vacuum" condensers have been known for a long time, new outgassing techniques and new "ion-pumps" have made the attainment of much higher vacua possible, and this permits increased ratings of these condensers.

Magnetic Materials

Permanent Magnet Materials

In the early days of radio, available magnet steels containing tungsten, chromium and cobalt, or all three, had too low an energy product to be seriously considered for speaker applications. With the discovery of the aluminum-nickel-cobalt magnetic compounds in Japan, commonly called alnico in this country, together with the similar compounds containing titanium, called ticonal in the Netherlands, their high energy product has made possible the use of permanent magnetic speakers and has largely done away with the solenoid magnets. More recently, ceramic magnets consisting of the ceramic chemical $BaFe_{12}O_{19}$ have existed in two forms: a randomly oriented ceramic and an oriented, one where the crystals are lined up with the hard direction of magnetization of the individual crystals aligned. These materials, while having a lower energy product than the alnicos, have higher coercive forces and are thus permissible to use in magnets thinner than alnico without demagnetization setting in. Another material, manganese bismuthide, of even higher coercive force, was discovered in France and has been the object of a great deal of study but has not found its way into many commercial applications as yet.

There are also available now permanent magnet metal foils, which, while having a relatively low energy product, are ductile and can be manufactured into various shapes quite readily, whereas the alnicos and the ceramics are somewhat refractory to deal with. These alloys are called Cunife and Vicalloy.

With the discovery in the 1930's that fine iron particles of single domain size can be agglomerated to make a permanent magnet, there has been an extensive development program in many countries. The highest energy products are now obtained with particles of ellipsoidal shape, of the order of 5×10^6. The theoretical maximum is much higher and work is still progressing.

Soft Magnetic Metals

Metallic iron at the start of this period had a maximum permeability in the neighborhood of 5000. Over the years it has continued to be made with ever higher permeability both as a result of removing impurities and in removing strains from the metal. Maximum permeability as high as 350,000 has been reported. However, iron of this quality is a little too delicate for use in most applications and, therefore, iron alloys have largely taken over the field.

The addition of a few per cent of silicon to iron increases the resistivity, thus cutting down the eddy current losses, and this metal is universally used as a core material for power transformers and low-frequency transformers. Aluminum can be substituted in whole or in part for silicon in these steels, but these materials, of which alfenol is an example, have not yet achieved large production. A major breakthrough in the development of silicon steel was achieved by Goss in the 1930's. He developed a process for the grain orientation of silicon steel whereby the easy direction of magnetization is in the direction of rolling. (In America, where it was discovered, it is called grain-oriented steel; in Europe it is called "Goss-iron)".

Iron nickel alloys—the permalloy family—started during this period with the work of Elmen. The 78 per cent nickel, 22 per cent iron permalloy has nearly zero magnetostriction and an extremely high permeability. Over the years this permeability has been increased by greater skill in refining the impurities out of the material and in superior heat treatments. The resistivity of permalloys has been increased by additions of other metals, notably chromium and molybdenum; molypermalloy containing 4 per cent molybdenum is possibly the most popular permalloy at present, and possibly represents the greatest advance to date in the control of manufacture and in meeting more rigid specifications.

Supermalloy contains about 70 per cent nickel, 5 per cent molybdenum, 15 per cent iron and 0.5 per cent manganese, all of extreme purity. Maximum permeability in the neighborhood of 1,000,000 is possible with this material if properly heat-treated.

The 45 per cent nickel permalloy is also used and the 50 per cent nickel have also found wide use particularly as cores for pulse transformers and for magnetic amplifiers, because it can be made to have a rectangular hysteresis loop. Iron-cobalt-nickel alloys called Perminvars are characterized as having permeabilities independent of field strength over a wide range.

Permalloy-type materials with 5 per cent copper added are known as mu metal and are useful for magnetic shielding as well as core material for electromagnets.

Iron-cobalt alloys have the highest saturation induction (and probably the highest Curie temperature). These have been given the name "Permendur" and also were invented by Elmen.

Magnetic Powders

Magnetic metals reduced to fine-particle size, coated with insulating material (usually of polymerizable resin), and molded have extensive use particularly in permeability tuning and in high-frequency application. The materials used in addition to magnetite are iron and permalloy, all of which have found wide use, and also include sendust, a Japanese alloy consisting of 14 per cent silicon, 14 per cent aluminum and the rest iron. This is a brittle material, easily reduced to powder and fabricated. While it has found extensive use in Japan and Germany during World War II, it does not seem to be able to compete successfully with permalloy dust cores.

Magnetic Ceramics

In addition to the original lodestone (or magnetite), a large family of synthetic magnetic ceramic compounds called ferrites was originally developed by Snoek in the Netherlands. These materials have lower permeabilities and lower saturation magnetization than the magnetic metals but, because of their high resistivity have more greatly reduced eddy currents than the metals, permit coils to have much higher Q's, and can be used at much higher frequencies than the metals, up to hundreds of megacycles. The principal use commercially is made of manganese zinc ferrites which have the higher permeabilities but also moderately low resistivity; that is to say they have high resistivity compared to metals but low as compared to other ferrites. They are useful up to the broadcast band. At higher frequencies, nickel zinc ferrites with much higher resistivity are used, and the still higher frequency ferrites containing cobalt are used. Ferrites containing cobalt may also be annealed in magnetic fields to produce B-H curves quite similar to the Perminvar metals mentioned above.

Many of these ferrites are also used in the microwave region, where use is made of the fact that their application of external magnetic fields causes the plane of polarization of the microwave radiation within a cavity containing the ferrite to be rotated, thus permitting for the first time the existence of non-reciprocal electrical networks.

More recently, a family of magnetic materials called garnets, of the general formula

$$(R_2O_3)_3\ (Fe_2O_3)_5,$$

where R stands for a rare earth metal, has been developed in France. These materials, in contrast to the ferrites, do not contain any ferrous ions and therefore can be made to have even higher resistivities than the ferrites and thus are even more efficient when used at microwave frequencies. The width of the resonant absorption band is also much narrower.

Another new series of magnetic ceramics discovered in the Netherlands has been called Ferrox-Plana. The crystals making up these ceramics are hexagonal and have asymmetric properties in that the C axis is often a hard direction of magnetization whereas the plane of the crystal has two easy directions of magnetization. These materials can be oriented in the manufacturing process similar to the oriented barium ferrite mentioned above. While unusual uses could be found for these materials at present they have been chiefly used to extend the frequency range in which ferrites can be used.

Families of ferrites containing copper or magnesium have been made to have rectangular hysteresis loops and these have found extensive use in saturable pulse transformers and as information storage elements in digital computer memories.

Resistive Materials

While the nickel-chromium alloy, Nichrome, is still with us and still to be improved on, various forms of carbon are perhaps in more general use. These go all the way from various inks consisting of mixtures of various carbon blacks with synthetic resins, through the form of carbon known as deposited or cracked carbon, which is the form of carbon obtained by decomposing an aliphatic hydrocarbon gas on a hot ceramic surface at a temperature of several hundred degrees C, up to some graphites. Graphite is unusual in that it is made up of hexagonal crystals with low resistance along the plane and high resistivity in the C axis. Oriented graphites are now available, and possibly use will be made of this asymmetric conductivity. Nickel-chromium alloys with aluminum added are also in wide use. Also, recently a number of organic salts sometimes called π complexes have been found with low value of resistivity together with some asymmetric properties, and these may possibly be considered as electronic materials.

Alloys of the noble metals (platinum, rhodium, palladium, etc.) formed from resinates and sputtered tantalum films are also being used for resistors.

Temperature sensitive resistors called thermistors use Ge, B, U_3O_8 and compounds comprising the transition metal oxides in two valence states, so that electrons can migrate from a metal atom in one state to another metal atom in the other state, for example,

$$Li_xNi_{(1-x)}O$$

where some of the nickel atoms are divalent and the others trivalent. These all have strongly negative temperature coefficients of resistivity.

Voltage sensitive resistors called varistors are largely limited to silicon carbide with or without some added carbon. Bismuth and more recently the semiconductor indium antimonide have been used as magnetic-field-sensitive resistors, while doped germanium and silicon have been used as pressure sensitive resistors.

Resistors of positive temperature coefficient are made from highly doped silicon, and from barium-strontium titanates containing a fraction of a per cent of lanthanum or indium (these types sometimes called "controlled valency semiconductors").

Ferroelectric Materials

For a long time Rochelle salt was the only known ferroelectric, but even here, the phenomenon was discovered as recently as 1921. This material had a comparatively low Curie temperature of 30°. This was increased to some extent by substitution of deuterium for hydrogen in the salt. While a number of other materials were found to have ferroelectric properties at temperatures below room temperature, the big impetus to this field was given early in the forties by Wainer who discovered that barium titanate was ferroelectric with a Curie temperature of 120°C.

Since that time a large number of other ferroelectrics have been discovered. These consist not only of compounds which contain barium titanate but other titanates as well, for example, lead titanate. Niobates and tantalates have been found to be ferroelectric, for example sodium niobate and lead niobate. (A number of antiferroelectric compounds have been formed, for example lead zirconate, as well as solid solutions of lead zirconate and lead titanate up to a certain percentage of titanate, when they become ferroelectric.) Of these, compositions can be made to go through from the antiferroelectric to the ferroelectric state by application of a high enough field. A new family of bismuth titanates (in which partial substitution of other metals is possible) gives high Curie temperatures which can to some extent be preselected. In addition, a number of organic solids have also been found to be ferroelectric of which possibly the most important one at the moment is guanidine aluminum sulfate hexahydrate, commonly called GASH.

Previously polarized ferroelectrics have found wide use as transducers (phonograph pickups, ultrasonic generators and receivers) and show some promise for microwave devices and as substitutes for IF transformers.

In piezoelectrics proper, as contrasted to the ferroelectrics above, there is substantially nothing new since the Curie brothers discovered the phenomenon in 1880 in quartz, tourmaline and Rochelle salt. However, LeCraw and co-workers at Bell Laboratories have recently discovered that oscillators using magnetostrictive yttrium iron garnet (also, under some circumstances, cobalt ferrite) have Q's several times the Q of oscillators based on quartz oscillators.

Wire Insulation

Wherein in an earlier day varnishes, cotton and silk were used, these have been to a considerable extent replaced respectively by polyvinyl formals (sometimes with the addition of phenolics), cellulose acetate or cellulose acetate butyrate, and the vinyl chlorides. At higher temperatures silicones and Teflon insulations are available. At still higher temperatures ceramic insulation with or without some of these plastics are also used.

Insulators

While originally steatite and porcelain were used as insulators, a number of additional ceramics, forsterite, spinel, magnesia, alumina and zircon are used, particularly at high frequencies. (While porous alumina has been with us for a long time, high density alumina as used for radomes and the like have only reached production in the last 20 years.)

Single crystals of alumina doped with a suitable transition metal element such as chromium are the functional elements in masers and lasers. Mycalexes comprising mica and a glass, have been extended to higher temperature ratings with the advent of synthetic fluorine mica. There are several new materials proposed for high temperature use including boron nitride, aluminum nitride and silicon nitride. Considerable effort is also being expended to prepare these materials in sheet form as dielectric spacers.

Semiconductors

The original materials used in cat's whiskers, rectifiers, and photocells were selected from lead sulphide (galena), lead telluride, selenium, cuprous oxide and cadmium sulfide. While these are all still used for one purpose or another, the present semiconductor art is mainly concerned with devices based largely on germanium and silicon, often in the form of extremely pure single crystals which have been doped with selected impurities to a precise level. Intermetallic compounds of elements lying in the IIIB and the VB columns of the periodic table, for example indium antimonide and gallium arsenide, have been synthesized and permit the manufacture of improved Hall generators, magnetic field detectors, Esaki diodes and so on. Some of these materials also have found use in thermoelectric generators and thermoelectric coolers, but presently solid solutions of lead selenide and lead telluride are more widely used. There are several new families of semiconductors containing as many as four or five elements which show great promise for use in this field. In an earlier day when the thermocouple materials, chromel, constantan, alumel, only were always available, such applications were given little consideration.

Conductors

While copper always was and always will be the principal conductor used, in recent years it has taken new form in the form of so-called "wiring boards", where sheet copper is laminated to plastic, either paper based Bakelite or polyester-impregnated Fiberglas. For high-frequency applications, copper conductors are now often plated with silver. For high temperature application they are plated with nickel. For applications where tensile strength is of great importance, iron wire with a heavy coating of copper is in common use. Conducting "paints" of silver and gold are in common use, as well as "conducting epoxies" consisting of metal particles dispersed in a polymerizable resin.

More recently superconductors lead, tin, niobium and tantalum have been used in switching circuits near helium temperatures. Niobium nitride at slightly higher temperatures has been used as a sensitive infra-red detector. In addition to superconducting alloys of superconducting metals, a new family of superconductors has been found at Bell Telephone Laboratories consisting of alloys of metals.

The first superconductors, now called "soft" superconductors, lose their superconductivity when the (self-generated) magnetic field exceeds a critical amount. The superconducting region is confined to a thin surface layer. "Hard" superconductors is the name for a family in which the maximum current density is uniform over the cross section of the material. Thus the "hard" superconductors will permit the establishment of high magnetic fields. In the order of increasing possible field strength these new materials are: Mo_3Re, Nb_3Zr, V_3S_1, Nb_3Sn, V_3Ga.

Hybrid Materials

We recently have had a glimpse into the future of electronic materials. Thus, the Bureau of Standards has synthesized materials which show *both* ferroelectric and ferrimagnetic properties. And Bell Telephone Laboratories have synthesized alloys which show both superconducting and ferromagnetic properties. In addition to the asymmetrical resistance materials, we can expect to utilize materials with asymmetric dielectric properties. Thus, to sum up, we can look forward to a complicated and exciting future.

Conclusion

While the over-all review of 50 years shows considerable accumulated result, progress in electronic materials has been slow and painful. The skills necessary to produce new or better materials require so many different scientific disciplines, that each new development must pass through many hands before maturing.

Future of the Component Parts Field*

P. S. DARNELL†, FELLOW, IRE

Summary—**Increasingly complex electronic systems are creating the need for greater numbers of components, with consequent serious problems of systems reliability and maintenance. Rockets, satellites and space vehicles demand small, lightweight, highly reliable electronic equipment. It is noted that miniaturized assemblies, integrated circuits, and functional solid-state devices provide solutions to these critical demands. Opinions are offered as to the time scale of availability of these new approaches. However, a continuing need for conventional parts in certain applications is noted, and possible trends in further development of several kinds are discussed. To derive maximum benefits inherent in the application of truly functional devices, the necessity of changes in points of view and design philosophies of both systems and device engineers is indicated. Certain factors other than those of a technical nature are considered which may have an important influence on the rate of transition from classical components to functional entities. Finally it is opined that the device engineer of the future should have a breadth and depth of intellectual disciplines which not only allows him to appreciate systems objectives, but also enables him to implement these objectives by functional devices through the media of basic properties of substances.**

Introduction

THE GREAT importance of electronics in performing communication, detection and guidance functions in World War II, and the emergence of solid-state technology into increasing prominence at the mid-point of this century, gave rise to conditions profoundly influencing the future of electronic component parts. As if those factors were not enough, the forces shaping the future for these parts are becoming even more intensified and diverse as the demands of the Space Age necessitate consideration of the effects of extraordinary terrestrial and interplanetary environments. At the bewildering pace with which these forces now are driving the field of electronics ahead, it is impossible to prophesy the long-range future with any assurance. Moreover, trends in technological development will be sensitive to influences created by the political, social and economic pressures associated with the international contests for power and position in the world today. Recognizing the vagaries of long-range prediction let us turn our attention to trends which it seems reasonable to expect during the next decade or so.

Systems Complexity

Electronic systems developers are continuing to design systems of greater and greater complexity, requiring more and more parts. For example, Baldwin, Crowley and Rosenthal [1] in discussing the application of computer techniques to the design of digital systems remark that " . . . the complexity of future digital systems leads one to the conclusion that such techniques are a virtual necessity. There are many systems now contemplated as utilizing hundreds of thousands, or even millions, of active elements." Another frequently quoted example is that more than 90,000 electronic component parts are required for a B-58 bomber. Of course, coupled with the unlimited growth in numbers of parts are the increasingly serious problems of interconnection and reliability. How many years more can the elaboration of electronic systems be continued on the basis of the classical concepts and approaches of the past? It seems that the time is not far off when this issue must be resolved. Although conventional components as we now know them will be with us for a long time to come, other ways and means must be pressed into service to alleviate the "tyranny of numbers."

* Received by the IRE, August 8, 1961.
† Bell Telephone Laboratories, Inc., Whippany, N. J.

Functional Entities

Types

As is well known, other ways and means in terms of physical embodiment under very active development today are packaged assemblies, integrated circuits, functional devices, etc. These vary in method of accomplishing in a physically unitary form the same function as a collection of discrete components. Packaged assemblies are essentially modular arrangements of regular circuit elements, so constructed and miniaturized as to result in high-density packaging of components. Integrated circuits may be either two-dimensional thin-film type on an insulating substrate, or a three-dimensional variety on a base of semiconductor material. In such forms of integrated circuits, the electrical components of which they are comprised are electrically distinguishable, but mechanically inseparable. In functional devices, or functional electronic blocks as they are sometimes called, the realization of the desired circuit function is accomplished by direct application of basic material properties. In this instance, circuit elements are no longer either electrically or mechanically identifiable as classical entities. Additional examples of functional devices are the maser, the Esaki diode, and the four-region semiconductor switch.

Miniaturization

A primary objective of packaged assemblies, integrated circuits, functional devices and the like is reduction in size and weight to yield miniature and so-called microminiature electrical building blocks. The importance of such miniaturization is emphasized by the demands for minimum size and weight of the electronic equipment for space vehicles. Noble [2] points out that in high-performance Air Force systems and space craft, " . . . the ratio of airframe and fuel weights to payload weights may run as high as one thousand to one." Hence the payoff for low weight load may be very high, and so justify the use of very costly electronic parts and equipment in the interest of minimizing weight.

Schedule

As regards the time when the types of functional assemblies above discussed will find practical usage, there are many divergent opinions. It seems reasonable to expect that initial applications will be for military purposes. Industrial and commercial usage undoubtedly will follow, but the extent to which functional units are employed by industry will be strongly influenced by economic consideration. The following time schedule is suggested.

1) The state of development of packaged assemblies and the less complex forms of thin-film type integrated circuits is now at a point where operational systems or equipment embodying these approaches are to be expected within a year or less. Also, some of the simpler structures of semiconductor integrated circuits will become commercially available during this period.
2) The period 1965–1970 should see the commercial use of circuits based upon semiconductor substrates.
3) Possibly toward the latter part of this decade some types of functional electronic blocks may be in pilot line production. However, the current state of the art is such that large quantity manufacture of functional electronic blocks before 1970 seems unlikely. It is believed that considerably more research and development effort is needed in this area to bring devices of this kind to the point where they can be produced on a commercial scale.

Conventional Component Parts

While packaged assemblies and functional electronic blocks provide approaches on which systems of the future may be founded to a large degree, there will continue to be a need for discrete components. For example, there is as yet no generally satisfactory solution to the problem of replacing the inductive functions of coils and transformers in two-dimensional arrays and in blocks of semiconductor material. Moreover, although functional entities are ideally suited to miniaturization and microelectronics, they are not compatible with high-power applications where large voltages and currents must be accommodated. The ability to put a new circuit together readily in the early design phases with discrete parts provides a flexibility to circuit engineers which may not be attainable with functional units. Until a new philosophy and technology of systems and circuit function design are evolved which provide other means of quickly demonstrating circuit feasibility, circuit designers will be loath to relinquish present procedures employing conventional components.

During the past several years, improvements in the classical forms of discrete components have been incremental in nature [3]. This situation will continue during the 1960–1970 decade with particular emphasis on performance in satellite environments and reduction in failure levels in all environments. Advantage will be taken of new materials as they become available. Generally speaking, no breakthroughs are foreseen which will radically affect the form, structure or function of our present line of classical components. Specifically, the following opinions relating to the future of these components are offered.

Resistors

Marsten [4] observes that there will be further development in inorganic resistance elements, such as carbon-ceramic structures. He also foresees continuing activity in the development of film-type resistors and increasing emphasis on semiconductor resistive devices. New areas in variable resistors will be explored to get away from the perennial problems of noise and wear caused by sliding contacts. In this connection it should be noted that the solid-state field-effect tetrode described by Stone and Warner [5] may be used as a voltage-controlled resistor exhibiting either positive or negative resistance depending upon the circuit configuration in which the tetrode is connected.

Capacitors

In reviewing the developments by which capacitors have evolved, Podolsky [6] notes that the " . . . history of the art shows clearly that improvements in capacitor materials and design followed closely the increased knowledge in the fundamental sciences, and most particularly in chemistry and physics." He recognizes that there is now " . . . electronic circuitry in which all the basic circuit parameters of capacitance, resistance and inductance are produced within solid state semiconductor devices." Of course, one well-known example of a semiconductor capacitor is the variable voltage diode [7], [8]. In conclusion he offers the opinion " . . . that developments in the capacitor field will advance constantly with the mainstream of science and that the art will progress to meet the needs of the electronic industries for every application whether they be terrestrial or in outer space."

Relays and Switches

With respect to relays and switches, Keller [9] states that the most widespread present trend in these devices in all types of switching systems is in miniaturization, and in hermetically sealed types to increase reliability. He notes the development of new high-speed relays such as the ferreed, a magnetically latched relay that can be operated or released by pulses as short as 5 μsec. He observes that, despite the recent rapid development of solid-state switching devices and systems, electromechanical relays and switches are being manufactured in more forms and larger quantities than ever before. In his opinion an "... important trend for switching systems in general is for relays and switches to be competitive, compatible and combinational with solid state switching devices and systems."

Coils and Filters

In considering trends in the field of transformers, inductors and filters, Lord [10] holds the view that miniaturization of low power electronic circuitry may very significantly reduce the role which discrete inductive devices play in low-power electronics. Particularly with respect to size reduction of filters, he cites the trend to combinations of electromechanical transducers and mechanically resonant structures of metal or ceramic to accomplish frequency selective functions. Others [11] have also noted efforts to circumvent transformers and inductors as circuit elements because of miniaturization difficulties and because of the difficulties of duplicating their functions in solid-state circuits. In the latter connection, however, Stone and Warner [5] have shown that the solid-state field-effect tetrode may be made to exhibit the behavior of a low-Q inductor by using its properties as a gyrator to invert capacitive admittance.

Selection Criteria

In his review of printed circuits and microelectronics, Danko [12] foresees high-density packaging of components, integrated circuits, and functional electronic blocks becoming competitive in equipment applications. Among these various approaches the criteria of choice will be the economics, performance, reliability, producibility and logistics of the end equipment rather than the particular technology associated with a functional assembly.

Frequency Control

In considering frequency-control devices, Bottom [13] notes that quartz crystal units used for national standards of frequency are capable of accuracies in the order of one part in ten billion (1×10^{-10}) per day. He further states, "Within the past ten years devices utilizing atomic events have been devised which provide standards of time and frequency which are reproducible and reliable over periods of a few months to about 1.5×10^{-11}." In this connection, recent work with the hydrogen maser at Harvard shows that frequency stabilities in the order of 1×10^{-13} part per day are feasible. Looking ahead, a frequency stability of 1×10^{-15} may be attained within the decade by means of the maser approach.

Materials

As Robinson [14] indicates, the development of new materials and improvements in the purity and processing of present materials holds the key to significant advances in the field of component parts. Not only will such activities permit improvements in conventional components, but they are essential to the advance of new concepts in component development and design. For example, the great impact of the ferrites during the past decade in the development of improved types of conventional components, and in opening up entirely new areas for microwave devices is well known. In the decade ahead, thin-film forms of conductors, dielectrics and semiconductors will continue to be of extensive interest in relation to further miniaturization of components and integrated circuits. Single-crystal purification and growth will also continue to receive extensive support. New semiconductor materials with new properties will be investigated with the objective of enabling the accomplishment of new functions in both discrete devices and functional electronic blocks. The effects of radiation on materials will continue to be studied both as to the properties of materials in components, and as a tool for the polymeric combinations of materials not normally susceptible of polymerization.

Reliability

Whatever are the roles played in the years ahead by the continuation of classical approaches and by the application of new integrated and solid-state devices as above described, the matter of reliability will continue to be of paramount importance. For example, an assumption often used is that the maintenance of military electronic equipment throughout its lifetime may cost as much as ten times the purchase price. Hence the initial cost becomes relatively unimportant. If extremely complex electronic terrestrial systems are to have economically reasonable mean-time periods between repairs, and if extraterrestrial systems are to give acceptable performance, failure rates must be known and controlled so that systems operation can be predicted and evaluated. Furthermore, it appears that continual improvement of classical components to the extent consistent with their continued application will be required. Similarly, the reliability aspect of integrated and functional devices will have to be kept up to the demands of over-all systems performance. Although some information is now available on packaged assemblies which show good reliability capabilities, there are not yet sufficient data at hand to permit assessment of functional solid-state circuits from the reliability viewpoint. It may be a matter of years before enough of such data are available to justify valid statements concerning the reliability of integrated and solid-state circuit devices [15]. It is generally believed that since functional electronic blocks will be derived from the same general materials and processing techniques as are now employed for transistors, they should have a high potential level of reliability. Furthermore, virtual freedom from the element interconnection problem should enhance their life expectancy.

Design Philosophies

Thus far, component parts and functional electronic entities have been reviewed principally from the standpoint of their technical aspects. Perhaps factors which have even greater influences than purely technological considerations on trends in the field of electronics are the development and design philosophies of people working in this field. As Morton [16 has stated, "Our heritage in components is essentially the physical embodiment of the mathematical concepts of circuit theory. With this viewpoint, the system designer has translated his over-all system requirements to those for components, thinking only in terms of classical inductance, capacitance, resistance, tubes and transistors. The component designer, adopting this viewpoint, therefore has been limited in his permissible solutions only to finding new techniques and materials for the classical elements. . . . After all, the aim of electronics is not simply to reproduce physically the elegance of classical circuit mathematics—rather, it is to perform desired electronic system functions as directly and as simply as possible from the basic structure of matter."

It is quite clear, therefore, that realization of the maximum benefits of truly functional approaches will necessitate major reorientation in the design philosophies of both system and device engineers. The system designer in following the principles of systems engineering [16] must start at a new level of synthesis and specify his needs only in terms of basic system functions and not in terms of the means by which the functions are to be accomplished. In turn, the device engineer must seek the attainment of these desired functions by going directly to the basic physics and chemistry of matter without becoming preoccupied with electrical circuit element concepts.

In achieving the higher levels of technical sophistication necessary to meet the challenges of the future new combinations of intellectual disciplines will surely come to pass. As Noble [2] observes, the future device engineer will have to be a "versatile specialist." Not only will he need to understand fully the purposes of systems functions in order to formulate device requirements, but he also must know how to cause electrons, atoms, phonons and photons to do what he wishes to accomplish the desired functions. The demarcation between systems, circuits and devices which has generally prevailed during the past will certainly become increasingly vague and may disappear entirely as more and more sophisticated devices provide complete electronic systems functions.

The Future

In summation, present trends and influences in the component parts field are of such a nature that the future of this field may be characterized as follows:

1) There will be a continuing need for discrete or single function components for many years ahead. *Relative* to the total requirements for electronic devices to support manufacture of systems, the production of conventional component parts will gradually diminish as new functional devices play an increasingly important role. However, if the current unprecedented rate of expansion in scientific and technological fields continues [17], the *absolute* level of demand for conventional components may show little change during the 1960–1970 decade.
2) The use of more sophisticated devices to perform complex circuit functions will gradually increase. Physically unitary forms of functional circuits will replace arrays of discrete conventional parts. Growth in systems complexity will inevitably force this replacement. Transition from classical to advanced approaches will be most rapid in applications where major benefits result. For example, the Aerospace Industries Association [18] in forecasting usage trends in aerospace electronic equipment predicts that in general, in 1970, conventional parts will constitute about 80 per cent of total usage. The remaining 20 per cent will be provided by microminiaturized discrete parts and by functional electronic devices, there being roughly equal division between these two approaches. However, in specialized applications, such as space craft, missiles and certain aircraft, miniaturized assemblies and functional devices may account for as much as 50 per cent of the total electronic componentry.
3) Improvements in the operation of conventional components will continue to be made. Reliability will become a quantative performance parameter of paramount importance with respect to both classical components and the new embodiments of functional circuits.
4) Continuing research into the basic nature and properties of matter will pave the way for the development of new components which individually or in combinations will perform functions that were either not possible or practical of accomplishment within the past state of the art. For example, the Peltier effect has been known since 1834 but it was not until the advent of modern semiconductor technology that the practical implementation of this effect in thermoelectric cooling could be considered [19], [20]. Another illustration is superconductivity. Discovered by Onnes in 1911, it is only recently that this phenomenon has received attention for technological purposes [21].
5) The rate of transition from conventional techniques to new approaches in the device field will depend in part upon many factors other than those of purely technological nature. There is a momentum tending to perpetuate the status quo, generated by the great amount of engineering effort in the past decade on improvement and refinement of classical

components. Development procedures and design practices founded upon years of application of classical methods have become strongly ingrained and the implementation of advanced concepts in system and device engineering will force a break with this past. The consequent reorientations in attitudes and thinking will not occur rapidly. The education and training of a new breed of "versatile specialists" to equip them with the breadth and depth of scientific knowledge necessary to cope with the problems inherent in advanced concepts of functional devices will take many years. In Morton's judgement [22], "The most profound understanding of solid state physics and chemistry will be needed to bring forth functional devices; indeed, it is likely that functional device invention will produce new science as did the transistor."

In conclusion, the wish is expressed that future efforts in the field discussed here could be devoted mainly to the instrumentalities of peaceful purposes. The political climate abroad in the world today makes this a forlorn hope and dictates major attention to keeping military facilities on the forefront of technological advance. Whatever the path ahead in these regards, the future of electronic componentry in its various forms will provide both stimulating challenges and exceptional opportunities for all those pursuing this exciting field of scientific and engineering endeavor.

References

[1] G. L. Baldwin, T. H. Crowley, and C. W. Rosenthal, "Design Automation—A Look at the Future," AIEE Conf. Paper CP-61-261; February, 1961.

[2] H. V. Noble, "The molectronics engineer—a new species," *Ind. Res.*, vol. 3, pp. 24–27; June–July, 1961.

[3] P. S. Darnell, "Present and future developments of electronic components," IRE Trans. on Component Parts, vol. CP-5, pp. 124–129; September, 1958.

[4] J. Marsten, "Resistors—a survey of the evolution of the field," this issue, p. 920.

[5] H. A. Stone, Jr., and R. M. Warner, Jr., "The field-effect tetrode," Proc. IRE, vol. 49, pp. 1170–1183; July, 1961.

[6] L. Podolsky, "Capacitors," this issue, p. 924.

[7] L. J. Giacoletto and J. O'Connell, "A variable capacitor germanium diode for UHF," *RCA Rev.*, vol. 17, pp. 68–85; March, 1956.

[8] H. A. Stone, Jr., "Component development for microminiaturization," 1957 IRE National Convention Record, pt. 6, pp. 13–20.

[9] A. C. Keller, "Relays and switches," this issue. p. 932.

[10] H. W. Lord, "Transformers, inductors and filters," this issue, p. 935.

[11] See for example, J. Holahan and B. Kovit, "Electronic components lag behind systems demands," *Space/Aeronautics*, vol. 34, pp. 93–100; December, 1960.

[12] S. F. Danko, "Printed circuits and microelectronics," this issue, p. 937.

[13] V. E. Bottom, "Piezoelectric effect and applications in electrical communications," this issue, p. 929.

[14] P. Robinson, "Electronic materials, 1912–1962," this issue, p. 945.

[15] A. E. Javitz, "Key problems in microelectronics," *Electro-Tech.*, vol. 67, pp. 61–66; June, 1961.

[16] J. A. Morton, "New Directions in Component Development," keynote address presented at the 1959 Electronic Components Conf., Philadelphia, Pa., May 6, 1959. (Copy available from Bell Telephone Labs., Inc., Murray Hill, N. J.)

[17] See for example, A. M. Weinberg, "Big Science, Marvel or Menace?" The New York Times, Section 6, pp. 15 *et seq.*; July 23, 1961.

[18] "1960 Aerospace Forecast of Technical Requirements," Aerospace Industries Assoc., Los Angeles, Calif., p. 30.

[19] See for example, H. A. Stone, Jr., "Development Trends in Electronics," presented at a meeting of the Merrimack Valley Subsection, IRE; February 6, 1961. (Copy available from Bell Telephone Labs., Inc., Murray Hill, N. J.)

[20] W. H. Clingman, "New concepts in thermoelectric device design," Proc. IRE, vol. 49, pp. 1155–1160; July, 1961.

[21] See for example, T. A. Buchhold, "Applications of superconductivity," *Sci. Am.*, vol. 202, pp. 74–82; March, 1960.

[22] J. A. Morton, "Functional components and integrated circuits," "Communications and Electronics—2012 A.D." Section.

Section 10

EDUCATION

Organized with the assistance of the IRE Professional Group on Education

Electrical Engineering Education Today*

FREDERICK EMMONS TERMAN†, FELLOW, IRE

Summary—**Education in electrical engineering has changed greatly since World War II and continues to change. There is greater orientation towards science and much more emphasis on graduate work, while student interest in electrical power as a field of specialization has virtually disappeared. Student participation in research conducted at a highly sophisticated level has become a major element in the training at the Ph.D. level, and the Ph.D. degree has come to symbolize the electrical engineer with superior technical training.**

Faculty members in electrical engineering departments have available rewarding and stimulating opportunities such as never before existed. Generous research support is available, there are numerous opportunities for consulting, and the services of professors are in demand for government advisory committees, membership on boards of directors, etc.

Educational institutions with strong graduate programs in electrical engineering are becoming centers for the development of growth industries. They are therefore important economic influences in the industrial world.

In summary, our universities today are turning out young electrical engineers far better equipped to meet the new challenges that are ahead of them than they did in the period before World War II. Concurrently the importance in today's world of the electrical engineering departments and faculties of our better schools is high, and is growing steadily.

EDUCATION in electrical engineering today is in a state of flux. Great changes have occurred during the fifteen years since the end of World War II, and changes still continue. Noteworthy is the greater emphasis being placed on fundamentals at the expense of attention to engineering practice and applications, the increasing orientation toward science, and the virtual disappearance of student interest in electrical power as an area of specialization. The latter has made electrical engineering and electronics virtually synonymous. Of particular importance is the enormous expansion of graduate work as compared with the prewar period, and the fact that graduate work carried on in electrical engineering departments now deals with many topics that were once the exclusively property of the physicist, such as electromagnetic theory, electron dynamics and much of solid-state electronics.

In order to understand the significance of the changes that have taken place in recent years in electrical engineering education, it is necessary to review some history. Before World War II, electrical engineering curricula were strongly influenced by an electrical power tradition, so that even those in the "communications" option were required to have a thorough background in rotating machinery and electric-power systems. Also, at this time only a very small fraction of the students graduating with a B.S. in electrical engineering engaged in graduate work, and very very few carried their studies as far as the doctorate. With the coming of World War II, electrical engineers with this four-year background augmented by practical experience were found to be wanting in the exciting new areas of electrical engineering involving microwaves, pulse technology, computers, diodes, etc. Instead, Ph.D. physicists who knew very little conventional electronics and no engineering, but whose education included three or four years of graduate work and a sound background in classical physics and mathematics, took over almost to the exclusion of the electrical engineer. At the end of the war, many of these men went back to being nuclear physicists, but they had shown the electrical engineers what was needed to handle the newly developing areas. Since then electrical

* Received by the IRE, June 5, 1961.
† Stanford University, Stanford, Calif.

engineering education has never been the same.

Graduate training, at least to the master's level, has become almost a necessity for the man with good ability who wishes to make full use of this ability in the technical aspects of such challenging new fields as solid-state electronics, computers, communication theory, controls, microwave electronics, etc. The nature of these fields is such that practical "on the job" experience is not a substitute for formal training. A man does not learn the quantum mechanics involved in semiconductor behavior from the experience gained on a beginning job in industry—such knowledge requires formal study, best carried on in an organized class under the guidance of an instructor. Today nearly 20 per cent of those graduating in electrical engineering go on to the master's degree, while two or three per cent of the B.S. graduates complete work for a doctorate. The doctor's degree is now beginning to be almost a prerequisite for a faculty position in the better electrical engineering departments, and is also given a high value by industry.

Research has become a major factor in electrical engineering education. Before World War II, only a few educational institutions carried on really significant research in electrical engineering, and even at these places the volume of such activity was not large because funds were so very limited. In contrast, the availability from government sources of large sums for basic research in electronics, and even larger sums for applied work in this field, has made it possible for many institutions to develop strong research programs. This has had its effect on the character of electrical engineering education. Undergraduate teachers either do research or are closely coupled to research activities, and so bring to the classroom a vitality and a spirit of adventure based on an interest in new ideas. This is in contrast with the prewar era, when the wisdom of the faculty member tended to be in a knowledge of the existing practices of the industrial world. In balance, this change has been for the good.

At the graduate level research tends to be an integral part of the educational program in electrical engineering. Through participation in research activities the graduate student not only obtains financial support, but also receives training in the methods of research. As a result, a large fraction of the products of our electrical engineering departments who carry their training to the doctorate level are true engineer-scientists. Moreover, they have the background necessary to enter new fields that open up, just as the physicists were equipped to go into microwave radar at the beginning of World War II. Never again will electrical engineers be caught short, as they were at that time.

The faculty man in electrical engineering never had more interesting and rewarding opportunities than are available to him today. If he has demonstrated creative ability, he can obtain research funds in sufficient amount to exploit his talents effectively, and he will have the help of bright, eager, and well-grounded graduate students to aid him in working out his ideas. As a result, he will be able to accomplish several times as much during his active career than his counterpart of an earlier generation, and will probably train five or ten times as many students at advanced levels. During the summer, he can stay on the campus and concentrate on his research activities, while receiving summer pay from his research project; this increases both his research productivity and his income. In addition, the electrical engineering faculty member will commonly spend up to one day a week as a consultant for one or more companies in connection with advanced development work which relates to his own special knowledge. For this he is well paid, and sometimes has opportunities to purchase stock at the employee's price. He may sit on the board of directors of a company for which he consults, or which was founded by one of his friends or former students. If he likes to get around, today's professor finds plenty of opportunity to travel in connection with his government contracts and consulting activities, and in connection with the affairs of professional groups. Today's electrical engineering department is no ivory tower, and its faculty members tend to be men of the world.

Finally, educational institutions with strong graduate programs in electronics attract industry. This is because industry is beginning to discover that its creative activities are best carried on near a center of brains, *i.e.*, near a good educational institution, and that for creative work such a location is far more important than a location near markets, raw materials, transportation, or factory labor. Substantial concentrations of research-oriented electronic activities are already associated with many of our best educational institutions, and this trend is rapidly increasing. The result is that a strong graduate program in electronics is an important economic influence in the industrial world, and not merely a place where young people can get a good technical education.

In summary, electrical engineering education has responded to the new age of electronics, and is now capable of meeting new challenges of this field as they arise.

Future Trends in Electrical Engineering Education*

J. D. RYDER†, FELLOW, IRE

Summary—**Maxwell started the electrical field on the mathematical analysis path, and the field will continue to successfully exploit the mathematical method. It can be expected that engineering will more completely take over the teaching of its own basic principles, as engineering becomes more accurately described as applied science. Future progress in electrical education will be built about mathematics, and especially employ systems analysis and materials science. Without further advance in his science base, the engineer may be in danger of becoming an order-taking technician. Such progress in fundamental will require the elimination of many present-day barriers between areas of specialization.**

SOME may say it started with Thales and his amber, or Galvani and his frogs, or Volta and his batteries. But the path of modern electrical education was really established by Maxwell, with his 1864 statement of the basic laws of electrical science in mathematical form—precise, rigorous, and so far, all inclusive. From this mathematical research also came the statement that light was electromagnetic in character, and the additional prediction of electromagnetic radiation of energy at other and nonvisible wavelengths. Thus was the future pattern of electrical research and education set—that knowledge of electrical science should come from logical reasoning and be expressed in mathematical language.

By its nature invisible, electricity made its presence known only by indirect phenomena, variously described by the early experimenters in terms of imperfect concepts of flow, flux, or force, borrowed from the field of mechanics. With this quasi-qualitative background, no matter how right the quantitative results, electricity could have remained much longer an area of metaphysics and legerdemain, had not Maxwell's basic statements channeled research into the mathematical realm.

To be honest, one must also admit that electrical phenomena, in their usual forms of exhibition, obeyed linear differential equations to which mathematical solutions were already available. Thus answers to many problems in physical situations were provided to the early investigators. In this, electrical science differed from some other early engineering fields which were advanced by the experimentalist or the practioner who made things work, without much use of the mathematical or analytical process. On their side, it must be recorded for fairness that many of the early problems involved turbulence, boundary layers, microscopic material phenomena, or nonlinear processes not even yet fully amenable to mathematical statement and solution.

Following the lead of Maxwell, similar mathematically-trained investigators unearthed and applied the methods of Heaviside, LaPlace, Fourier, Cauchy and others to develop further electrical theory. The mathematical method has continued in use, and has provided the profession with a medium of communication and a method of analysis, not only useful in electrical discovery but also suited to the solution of problems in related fields. Thus it should have been no surprise for the electrical engineer to aid and abet the development of the electronic computer in logic as well as in circuitry, or for him to apply mathematical logic to the switching problem, or quantitative methods to information transfer or communication in general.

There is no reason to believe that this continued usage of and profiting from the methods of mathematical analysis will cease. Therefore it seems possible to predict with reasonable surety that future progress in electrical science will be coincident with progress in application of mathematics to that science. In fact this writer has previously proposed[1] as a theorem: "That a field of engineering can advance only as fast as the mathematical abilities of its members advance." Electrical science, a term here used as inclusive of all *applied science* concerned with electrical phenomena, seems an area providing adequate proof of this theorem. The recent rapid advances of other fields of applied science and engineering, as they have begun to utilize higher mathematical methods, seems to provide further confirmation.

If this prediction of the mechanism by which we will further advance in electrical science and engineering is correct, then the converse of the theorem is important. If we are to advance, better and additional mathematical education must be provided to future members of our profession. As far as can be foreseen, these programs will be well suited to students in all of the engineering or science-engineering disciplines, since advance in all fields is to be hoped for, if our science-based civilization is to continue forward.

Engineering college mathematics will continue to begin with the calculus, giving to the high school or preparatory school the responsibility for laying the foundation in prior forms of mathematics. This will call for further changes in an already-changing high school pattern to provide more work in trigonometric identities,

* Received by the IRE, July 14, 1961.
† Dean of Engineering, Michigan State University, East Lansing, Mich.

[1] Poles and Zeros, "D-day in engineering education," PROC. IRE, vol. 46, p. 1571; September, 1958.

more algebraic manipulation, and a thorough understanding of analytic geometry. The more advanced high schools will provide introductory work in the calculus itself, but such opportunities in the average high school may be beyond the time scale of this paper, unless greater encouragement is made available to attract more capable people into high-school teaching, and more rigorous mathematics courses are required for their training.

College departments of mathematics may have to be reoriented and in some cases, reorganized to insure interested teaching of the kinds of mathematics desired and needed by engineers and other workers in applied science. Assuming such reorientation, then the college teachers of calculus must recognize that the engineer of tomorrow (and tomorrow is nearly here) will work almost always with a computer, of at least reasonable capability, figuratively in the next room. The treatment of a Table of Integrals as something approaching the level of a clandestine crib must cease, and the Table recognized as a place of first resort upon meeting an integral sign. The classroom development of the solutions to the integrals of a great many functions can be eliminated, and the time formerly so utilized given to insuring that the student has a thorough understanding of the significance and limitations of the process of integration. This work should include a much expanded discussion of change of variable, limits, integrability, existence, and other areas of mathematical rigor now treated much too lightly. Given a thorough understanding of the process and limitations of integration, the end point of any integration will then most usually be the Table of Integrals, or an application of Simpson's rule. The latter will lead directly to a piece of tape or a program in the computer library, and a short time on the local computer.

It will then be apparent to the teachers that the treatment of differential equations in a separate and distinct course is undesirable. Just as the teaching of methods of solving a wide variety of integral forms is unnecessary, so also is it unneccessary to teach methods of solution to a variety of differential equation forms. The computer will provide the answers; student abilities in setup of the original equations, and manipulation into computer-suited forms can become the emphasized topics. Solution of a differential equation can be presented as a problem in simultaneous integration, and at this point the methods and analysis employed in setting up an analog computer will justify the presence of such a computer in the mathematics classroom.

It is realized that appropriate textbooks and course syllabi do not yet exist in most schools. However, if new material is appropriate to a course or curriculum it should be introduced and taught with whatever aids are available. The textbook will surely follow.

So much is prologue—a part of the preparation for the professional education in electrical science to follow. In view of the diversity and breadth of the field served by today's electrical engineers—from costs and studies of new communications systems to development of erotic methods of space-ship propulsion, from abstract logic to gallium arsenide diodes, from transfer functions to information channel capacity, from sales to research—it is not easy to determine the trends. An extrapolation of present movement allows the prediction that the electrical engineer will become an electrical scientist-engineer, a man interested in the discovery of new knowledge which he hopes to apply in useful new devices and systems.

To become such he will require a thorough knowledge of electrical and allied science as his background, a rigorous coverage of mathematics as his language, and methods of analysis and synthesis in engineering application as his tools. He will not lack for scientific fundamental nor for ability in engineering analysis and synthesis; his interest will be in the application of scientific phenomena for human benefit. Where scientific fundamental may still be lacking, he will have an engineer's ability to design around the lack.

In tomorrow's starkly competitive scientific world such an education will be required of all who will originate, innovate, face the challenging problems, make the broad decisions, and reap the large rewards; any education less than this will fit the recipient only to take and carry out orders. Which do we select for the engineer of tomorrow? Can the engineer afford to be second best?

To provide the training for leaders in electrical and applied science a college must attract the best students, must excite and challenge their abilities from the moment of entrance, and this cannot be done leisurely. Included must be basic work in physical phenomena on a broad front; at least the areas of solid and fluid mechanics, heat, and atomic and nuclear structure, taught at a level which recognizes the parallel advancement in the mathematical ability of the student. Engineering schools in this near time scale will have overcome the traditionalism which today keeps the teaching of classical chemistry and physics in those science departments.[2] Subject matter in the classical physical science areas is now the basis on which engineering is built, and, as engineering becomes more a field of applied science, it will no longer be logical that engineering should turn over the basic preparation of their students to others. A basic mechanics or heat course, taught by a theoretically-inclined engineer will not differ from a course taught by a physicist, and will be much better than one taught by a graduate assistant. Engineers today teach such classical material at the graduate level; can we long tolerate the duplication of effort at lower levels? While we have here named chemistry and physics, many of the same points need to be discussed on the teaching of mathematics, as well.

[2] J. G. Brainerd, "Cooperation among electrical engineers, physicists and mathematicians in engineering education," *Proc. Syracuse University Sagamore Conference*, p. 78; 1960.

The teaching of the fundamental science material in courses supervised and taught by engineers of science bent will provide definite gains by coordination of treatment with later engineering courses. As an example, a new approach to mechanics will be required. It should not be necessary for the electrical engineer to analyze trusses to learn about forces, and he should care little whether the forces are coplanar or noncoplanar. Work, energy, power will be introduced with greater breadth and more mathematical and physical rigor than has been past practice. It will be found that learning in this field can be expedited by resort to some of the fundamentals common to all fields,[3] but usually associated with network theory, namely: forces to a point sum to zero, displacements around a path sum to zero. There are other similar broad analogies covering many fields. Such unification of treatment cuts across traditional departmental lines, and barriers between departments must be removed to allow such free interchange of ideas, and free movement of teaching staff.

Electrical studies may begin with field or continuum theory as one stem, network theory or lumped-element treatment as the other. Field work will later expand into fluids and other field processes, and into transmission and radiation of energy, all areas where the partial differential equation rules. Such unity in the study of continua would also cut across the departmental barriers of today—perhaps we are already hearing the trumpets of Joshua before these walls of Jericho?

Network theory will be developed not only for its answers to response of networks, or for design to stated conditions, but also to develop skill in the application of rigorous and systematic methods of solution for systems of equations. The voltages, currents, frequency response, or transient performance may still be needed in engineering application, but often these figures will be determined by the computer after the student has set up and manipulated the basic system equations by matric techniques. Electronic study will be more systematized than it is today—being recognized as the study of the active network. Generality will be the byword, and devices when discussed will be largely solid-state in nature.

In the electrical area, as it was in mathematics, we find that what has been already said is largely prologue. Major emphasis at higher levels will go to two areas—systems analysis and synthesis, and solid-state phenomena or materials theory. A third, but less important, area will employ statistics or statistical mechanics in the study of noise, random phenomena, and the information capacity of communication systems.

The study of a system, as contrasted with the "isolated machine" concept of an earlier day, probably began as a formal discipline with the interconnection of components or machines for control purposes three decades ago. With such interconnection the transient performance was at the heart of the study; the steady state so avidly studied in an earlier day is now seen as only an occasional end-point of a transient. The popularization of the LaPlace transform and the fundamental work on feedback by Black, Bode, Nyquist, and others lead to understanding. Thus were poles and zeros, complex plane analysis, Nyquist diagrams, and stability brought into electrical parlance.

Understanding, however, is not enough for the electrical engineer. It remains to analyze the system in terms of its interconnections and the characteristics of its component parts, and here again mathematics supplies at least a partial answer in topology and linear graph theory. Through such mathematical methods the analysis and synthesis of systems, regardless of operating media or type of component, will become a major responsibility of the electrical engineer. This will require that he once more break down the barriers between today's disciplines; he will require basic knowledge of pneumatics, hydraulics, heat transfer, thermodynamics, even chemical reaction. Given the proper breadth in his basic education, he will discover that his mathematical training plus his network analytic methods can readily carry his systems analysis and design into nonelectrical areas.

Materials science will receive great emphasis in the future educational process of the electrical engineer. Needing knowledge of the magnetic, thermal, and conductive processes in metals and alloys, in the past he all too often received instruction in the mechanical properties of steel and concrete, and so he chose to study none. Tomorrow he will study materials in the atomic sense, since his devices will employ lattice arrangements to secure their operative properties, and his design of circuit elements may depend on crystal growth. Courses in this area are today far from usual; in fact, syllabi and texts must still be written. However, when a complete amplifier or subsystem can be designed into a wafer of germanium or silicon, or, when new operating properties appear by reason of quantum tunneling of thin films, there can be little reason to doubt the importance of the area.

Interest in broadening of the engineer in nontechnical fields will continue. His path in these directions must also be prepared by questioning of tradition, and coordination and integration of subject matter. Perhaps a few walls of Jericho may fall in the areas of the social sciences as a result. We must keep constantly in mind that, if the engineer of the future is to make the proper decisions for the major problems he will face, he must be able to live, to work, to communicate, and to lead in a world of nontechnical but well-educated people.

While all engineering will tend toward further emphasis on research, the electrical engineer will look upon this as his major function—graduate preparation for this and other areas of work will become commonplace for all capable of carrying on with their education. Engi-

[3] H. E. Koenig and W. A. Blackwell, "Electromechanical System Theory," McGraw-Hill Book Co., Inc., New York, N. Y., 1961.

neering will have become more fully recognized as a field of applied science, and the electrical engineer will have become a scientist-engineer, interested in the search for new knowledge and in the development of new devices and services employing that knowledge. Engineering education in general will have blown down the walls of Jericho which isolated its many fields, and will have achieved a closer understanding in a technical sense, through broadening and deepening of our common science base. Engineering will also have achieved a better understanding of the relationships between, and the differing objectives of, scientists, scientist-engineers, engineering operators, and engineering associates.

Thus will real unity be achieved through teams of diverse function and training, all working to employ nature's phenomena for the benefit of man.

Graduate Study in Electrical Engineering*

ERNST WEBER†, FELLOW, IRE

Summary—**Graduate study in electrical engineering developed significantly after World War I when radio engineering came of age. The many new scientific discoveries preceding World War II and exploited technologically during and following that war lent tremendous impetus to the further growth of graduate study, particularly towards the doctor's degree. The keen competition for the outstanding students between industrial research laboratories and the faculties of the leading institutions has retarded the growth of the genuine full-time student body and has enormously accelerated the part-time enrollment for graduate degrees with the result that the numbers of degrees awarded have increased at a lesser rate. This has accentuated evaluative studies of graduate education and drawn the attention to national needs; namely, large-scale support of a sufficient number of outstanding graduate study and research centers with adequate support for genuine full-time graduate students in order to supply faculty needs as well as the needs of industry in this highly technological society.**

The Evolution of Graduate Study

GRADUATE EDUCATION in the United States could be said to have become significant about 1876, when The Johns Hopkins University in Baltimore, Md., established a graduate school organized pretty much along the German pattern of the time, which strongly emphasized research. In his rather exhaustive study, B. Berelson,[1] in fact, distinguishes the following periods concerning graduate study in the United States: Prehistory (before 1876), University Revolution (1876 to 1900), Consolidation and Standardization (1900 to World War I), Growth and Diversification (World War I to World War II), and Revival and Reappraisal (since World War II).

Engineering generally, and electrical engineering particularly, had very little graduate study interest and, therefore, facilities before 1900 were few. As a matter of fact, the total graduate enrollment in engineering schools in 1921 was 368 students in 40 institutions, whereas 124 engineering schools had an enrollment of over 64,000 undergraduate students! By 1933, however, interest in graduate study had become significant enough so that the U. S. Office of Education in cooperation with the Society for the Promotion of Engineering Education initiated a survey[2] which was published in 1936 and which could point to the fact that the 126 doctoral degrees in engineering awarded in 1934 gave it second place in the science group, chemistry being first with 590 degrees, and physics third with 121 degrees. Engineering had thus leaped over more than 50 years of evolution and, of course, had created quite a host of problems concerning standards, objectives, and administrative functions. The impetus to this first growth came clearly from the general growth of population with the increasing emphasis upon college education, from the general prosperity of the early twenties which induced tremendous industrial developments, and from the technological innovations that had transformed the transportation, communication, and manufacturing systems of the country, and had created the modern home with all its appurtenances.

Alarmed by the seemingly boundless expansion, the Committee on Graduate Study of the Society for the Promotion of Engineering Education prepared a Manual of Graduate Study in Engineering[3] in 1945, which set down the philosophy and objectives and established guide lines for the organization, conduct, and administration of graduate study in engineering, both towards the Master's and the Doctor's degrees. Though in principle still valid, a revised edition was issued in 1952

* Received by the IRE, December 6, 1961.
† Polytechnic Institute of Brooklyn, Brooklyn, N. Y.
[1] B. Berelson, "Graduate Education in the United States," McGraw-Hill Book Co., Inc., New York, N. Y.; 1960.

[2] W. C. John and H. P. Hammond, "Graduate Work in Engineering in Universities and Colleges in the United States," Office of Educ., Dept. of the Interior, Washington, D. C., Bull. No. 8; 1936. See also "Graduate study, report of progress," *J. Engrg. Educ.*, vol. 2b, pp. 313–355; December, 1935.

[3] "Manual of graduate study in engineering," *J. Engrg. Educ.*, vol. 36, pp. 615–652, June, 1945; also issued as reprint, revised ed., 1952.

to take care of special problems connected with the changing patterns of graduate study. Finally, a graduate study commission[4] concentrated on these same problems of motivation, financing, patterns of study and development of faculty in greater detail.

It is of interest to note that all of these studies and surveys encompass all disciplines of engineering and therefore are of a compromise type between the strongly professionally oriented representatives, who would like to see in engineering primarily design and development, and between those representatives who see engineering as applied science with strong emphasis upon research. Depending upon one's views, experience, and personal ambition, the evaluation of trends and programs and the statement of problems will differ.

The Growth of Graduate Study in Electrical Engineering

Graduate education in electrical engineering has exhibited a growth pattern that runs rather parallel to the development of IRE, and in fact, has confronted us in recent years with the same kind of explosive growth which has caused serious reappraisal of objectives, standards, and modes of operation.[5] Peculiarly enough —or perhaps for very logical reasons—the two world wars have provided the greatest stimuli for growth.[6] World War I initiated the development of the world-wide communication systems with the need for systematic analytical as well as experimental treatment of signal shaping, propagation, and reception. The foundations of network theory,[7] operational analysis,[8] and early applications of electromagnetic theory[9] were laid in that period and introduced into the graduate programs of progressive schools of engineering. World War II brought the fruition of the feedback concept,[10] and its extensions into automatic control systems,[11] opened the whole new world of microwave systems of radar, navigation, air traffic control, and communication links,[12] and fathered the penetration of modern quantum physics into material technology.[13] Today we find ourselves exposed to a dynamic development at an exasperating pace with which even the professional group system of the IRE has difficulty to keep step, notwithstanding its unrivalled flexibility and practically unlimited scopes of interests.

Statistical data are not reliable before the first survey of graduate study in the so-called Hammond report,[2] but we find it illuminating to compare enrollment figures and degrees awarded from 1933 on. Table I gives the data published by the ASEE for five selected years avoiding the war years and the immediate post-war period. Even so, the number of Doctor's degrees in 1948 is low as a post-war effect because of the residence requirement for doctoral studies, and the number of day students in 1952 is low because of the effects of the Korean war. Over-all, these facts stand out:

1) Total enrollments for Master's degrees have increased by a factor of 14.5 between 1933 and 1960, but at a much greater rate in the part-time category than in the full-time group.

2) The number of Master's degrees awarded per year has increased in the same period only by a factor of 6.4, suggesting a high and apparently increasing attrition rate which one might not expect in graduate study.

3) Total enrollments for the Doctor's degree were not significant before 1939, and have grown since 1948 at a greater rate than enrollments for Master's degrees.

4) The number of Doctor's degrees awarded per year has not kept step with the enrollments in spite of the fact that the doctoral enrollments are mostly classified as day students.

TABLE I

Graduate Students in Electrical Engineering

	1933	1948	1952	1956	1960
Master's programs					
Enrollment:					
Day (Full and part time)	469	2267	1686	3035	5710
Evening and Special	214	1133	3575	3822	4168
Total	683	3400	5261	6857	9878
Degrees awarded	301	866	990	1128	1934
Doctoral programs					
Enrollment:					
Day (Full and part time)	105	361	567	747	1548
Evening and Special	7	106	269	208	230
Total	112	467	836	955	1778
Degrees awarded	21	36	121	136	203

[4] "Facilities and opportunities for graduate study in engineering," *J. Engrg. Educ.*, vol. 48, pp. 938–974; June, 1958.

[5] "Sagamore Conference on Graduate Electrical Engineering Education," IRE Trans. on Education, vol. E-2, pp. 30–63; April, 1959.

[6] E. Weber, "The challenges in the development of graduate programs," *ibid.*, pp. 39–43.

[7] G. A. Campbell, "Theory of the electric wave filter," *Bell Sys. Tech. J.*, vol. 1, pp. 1–32; November, 1922.
T. E. Shea: "Transmission Networks and Wave Filters," D. Van Nostrand, Co., Inc., New York, N. Y.; 1929.

[8] H. Jeffreys, "Operational Methods in Mathematical Physics," Cambridge University Press, England; 1924.
J. R. Carson, "Electric Circuit Theory and Operational Calculus," McGraw-Hill Book Co., Inc., New York, N. Y.; 1926.

[9] G. W. Pierce, "Electric Oscillations and Electric Waves," McGraw-Hill Book Co., Inc., New York, N. Y.; 1920.

[10] H. W. Bode, "Network Analysis and Feedback Amplifier Design," D. Van Nostrand Co., Inc., New York, N. Y.; 1945.

[11] G. S. Brown and D. P. Campbell, "Principles of Servomechanisms," John Wiley and Sons, Inc., New York, N. Y.; 1948.
J. G. Truxal, "Automatic Feedback Control System Synthesis," McGraw-Hill Book Co., Inc., New York, N. Y.; 1955.

[12] J. C. Slater: "Microwave Transmission," McGraw-Hill Book Co., Inc., New York, N. Y.; 1942.
S. Ramo and J. Whinnery, "Fields and Waves in Modern Radio," John Wiley and Sons, Inc., New York, N. Y.; 1944.
"M.I.T. Radiation Lab. Series," 28 vols., McGraw-Hill Book Co., Inc., New York, N. Y.; 1947–1951.

[13] W. Shockley, "Electrons and Holes in Semiconductors," D. Van Nostrand Co., Inc., New York, N. Y.; 1950.

As a matter of fact, the number of Doctor's degrees awarded in electrical engineering has remained at about two per cent of first degrees, which is the same ratio as for all graduate education in the United States, whereas in physics and chemistry the corresponding ratio is more like 14 per cent. This leads necessarily to a discussion of the

Objectives of Graduate Study

As a profession, engineering has been defined by the Engineers Council for Professional Development (ECPD)[14] in terms of application of " . . . a knowledge of the mathematical and natural sciences gained by study, experience, and practice, . . . with judgement to develop ways to utilize economically, the materials and forces of nature for the progressive well-being of mankind." In its earliest phases, engineering was the art of providing roads and river crossings, of supplying fresh water and disposing of waste, of building homes and implements for defense, briefly what we call civil engineering. This is clearly illustrated by the very first school of engineering, founded by Napoleon as the Ecole Polytechnique in 1794, to provide properly trained men for his génie civile corresponding to our present Corps of Engineers. Significantly, the earliest engineers in civilian practice in the United States were graduates of the U. S. Military Academy, organized in 1802, underlining the tremendous value that the military have always placed upon the systematic utilization of technological achievements. The very large expenditures of the military establishments all over the world for the development of technologically more perfect weapons of defense, and also of offense, are clear evidence of this fact today.

Perhaps we should keep in mind that the faculty at the Ecole Polytechnique included the mathematicians Lagrange and Laplace who evolved the systematic formulation of mechanical science and of potential theory, as well as the chemist Berthollet and the crystallographer Haüy![15] Unfortunately, this example was not emulated generally, and the early schools of engineering in this country stressed the art of application of mathematical and natural sciences and thus could rest satisfied with a four-year undergraduate curriculum until the turn of the last century as outlined earlier.

In electrical engineering, it was the advent of wireless or radio communications coupled with the invention of the vacuum tube and the selective filter circuits that introduced higher mathematics and electromagnetic theory into the technical literature. True, wire communications and electric power transmission along long lines led in their theoretical treatment to the complex hyperbolic functions. However, it was the partnership with physics and applied mathematics, which the young profession of radio engineering cultivated, that eventually reacted catalytically upon electrical engineering education and generated the graduate programs for Master's degrees in the middle twenties. We had the strange anomaly that, for practically two decades, the undergraduate curricula—with few exceptions—were entirely power oriented whereas many of the Master's programs offered a majority of communications related subjects with emphasis upon advanced mathematics and modern physics.

The objective of the early graduate programs was thus primarily to bring the student up-to-date again with few significant exceptions.

It is quite appropriate to quote here from the so-called Hammond Report the "general characteristics" of graduate study[16] in 1935:

> If the situation as to graduate work in engineering can be characterized in a sentence, it may be said that work for the Master's degree has become the regular order of things in more than half of our institutions and for an important fraction of our students, while work for the Doctorate is beginning to assume a place of significant importance. The committee ventures to express the opinion that it would be just as well if work for the doctorate were not to expand too quickly or spread too widely. Expansion of work for the master's degree has been extremely rapid. . . . It may be well for engineering education to consolidate this gain and to strengthen its position in graduate work before making any wholesale expansion in to the still more advanced field of work for the doctorate.

Little could the committee and, indeed, the Office of Education, anticipate the fantastic change in the rate of scientific discoveries and technological applications which occurred in the span of a few years and which confronts us today, at least in electrical engineering, with the indomitable demand to re-evaluate our objectives and to critically examine our curricula.

It has now become clear that all electrical engineering education must have a sound and solid foundation in the basic sciences followed by extensive exercise in the engineering sciences and culminating in the senior year in project work, in integrative efforts, and in systems concepts. Above all, the student must be made to realize that he is at the start of knowledge and that he needs to cultivate the ability to learn and to study. As yet, the four-year program will be satisfactory in many cases where truly professional contributions are not expected. However, to be able to contribute significantly in advanced areas of development, it has become necessary to add another year, usually leading to the Master's degree. This additional year must include advanced physics and mathematics as well as professional subjects which lead to basic understanding and mature exercise of the design function, the primary objective of engineering education in accordance with the definition quoted earlier. It has been suggested to recognize this five-year program which includes one year of graduate study as the first truly

[14] "26th Annual Report of the Engineers Council for Professional Development for 1957–1958," Rept. of the Committee on Recognition, September, 1958.

[15] Sir Eric Ashby, "Technology and the Academics," Macmillan & Co., Ltd., London, England, p. 19; 1959.

[16] *Op. cit.*, ref. 2, *J. Engrg. Educ.*, p. 321.

professional degree and even make it a general requirement for the engineering profession, relegating the four-year program to a nonaccredited status.[17] However desirable it might be to decree such requirements, the present patterns of graduate study and industrial employment practices do mitigate against it.

The objective of graduate study towards the Doctor's degree has traditionally been guidance towards independent investigational work. It is true that this objective is adopted from the graduate schools of arts and sciences. On the other hand, the vast expansion of the horizons in science and engineering has opened a large sector that can be classified as engineering research, independent investigation that sets itself certain more or less clearly defined goals. The difference between science and engineering is, then, one of attitude rather than of method or depth of training. Pure scientific research is entirely aimed at increase of knowledge, at exploration of the universe in systematic and methodical manner with a philosophical view towards better comprehension of the observed phenomena. Engineering research is related to the basic engineering function of design, even though it uses frequently the same methods and devices as scientific research. We need only to scan the titles and scopes of the Professional Groups of the IRE to realize the close partnership of science and engineering on the one hand, and the vast areas open to engineering research, say directed research, or "applied science"-research, on the other hand.

Because the objectives of the Master's and the Doctor's programs are different, though not mutually exclusive, it is most desirable to guide the senior student into the appropriate graduate program best fitting his abilities. Particularly gifted students could, if they are so inclined, be exposed to guidance towards independent research even within their undergraduate studies, be it in formal honor's programs or on the initiative of an individual faculty member.

A major objective of graduate study in most areas other than engineering, has traditionally been the training of teachers. Engineering on the other hand, has primarily furnished professional talent for industries, relying upon the normal feedback into faculties because of innate interest of a sufficient number in teaching as a career. Within the last decade, however, industrial operations have become more and more complex so that their requirements for manpower with advanced degrees could not be fully met. This has now placed faculty recruitment into a state of urgency, so that we must add as a significant special objective of graduate study, preferably to the Doctor's degree, the education and training of teachers.

Patterns of Graduate Study

The traditional pattern of graduate study as prevalent at the time of the Hammond report[2] was also taken over from the schools of arts and sciences whose faculty in many of the larger universities had full cognizance over all graduate work. Actually, as we see from Table I, the majority were day students, most of them studying full time for their advanced degrees, only about $\frac{1}{3}$ of the master's and $\frac{1}{4}$ of the doctoral students studying part-time, having some assistantships or junior instructorships to enable them to study at all. Master's theses were required. Doctoral dissertations were carried on under tutelage of a faculty member who had his own research laboratory and his well-known field of interest. The student's choice was in a sense free, in a sense dictated by available counsel and guidance.

Just as World War II changed the whole structure of science and technology as it created research and development efforts on gigantic scales, as it organized the whole scientific potential of the universities, so also has it left behind a complete social reorientation and with it a change in concept of support of our higher educational system. The impact upon graduate study in science and engineering has been enormous.[18] Whereas in the earlier days the student made a very weighty decision when he embarked upon graduate study and frequently anticipated years of struggle for a living and hard learning under rigorous and unrelenting tutelage, he is now apt to expect complete support and definitive requirements which, if met, practically guarantee the advanced degree. He has practically been promised that a considerably higher salary level will ensue so that graduate study is evaluated in terms of an investment. Of course, we still have the innately creative type of student whose urge to produce spurs him to effort, but more often than not, he will command an industrial position and not return to graduate school except for specific information.

To appreciate the radical change in patterns of graduate study and the effects upon the output of advanced degrees, it is necessary again to refer to statistics such as are given in Table II and prepared for ASEE[19] in order to clarify the almost meaningless categories of day and evening students, often confused with full time and part time. Unfortunately, the figures are available only for all of engineering but they allow readily the same conclusions for electrical engineering:

1) Genuine full-time enrollment has increased between 1933 and 1960 about fourfold; most of these students have full fellowships, available from various sources. Practically all foreign students are included because they cannot earn money, they subsist on fellowships and must study full time.

[17] M. P. O'Brien, "Professional graduate study in engineering," *J. Engrg. Educ.*, vol. 51, pp. 571–580; March 1961.

[18] E. Weber, "Types of graduate subsidy and their relation to educational values," *J. Engrg. Educ.*, vol. 44, pp. 188–191; November, 1953.

[19] C. Wandmacher, "Follow-up survey of enrollments and staff in various patterns of graduate study," *J. Engrg. Educ.*, vol. 51, pp. 410–418; February, 1961.

2) The number of students employed by the university has increased over tenfold, the large majority being engaged in research which under today's conditions is almost all government sponsored research.
3) Part-time students, employed outside the university, *i.e.*, by industry or government, constitute today the largest sector of graduate student enrollment, having had comparatively little significance in 1933. To the largest part, these are evening students.

TABLE II
THE CHANGING PATTERN OF GRADUATE STUDY IN ENGINEERING

Graduate students, enrolled (Fall)	1933	1948	1960
Full time*	1695	4100	7064
Part time, employed at University			
In Teaching	585	1234	2274
In Research		1053	3931
		2287	6205
Part time, employed outside University			
Studying on main campus	360	4607	12,996
Studying at University Center	—	546	3088
Studying off-campus	—	110	1608
		5263	17,692
In cooperative programs	109	42	102
Total enrollments	2749	11,692	30,063
Master's degrees awarded (June)	1071	4303	7128
Doctor's degrees awarded (June)	126	252	786
Total advanced degrees	1197	4555	7914
Ratio Enrollment / Advanced degrees	2.3	2.6	3.8

* Taking a minimum of 12 credit hours per week; employment, if any, is only incidental.

Little more can be said about the full-time student enrollment than that it is far too low and that every effort should be made, nay *must* be made, on a national scale to entice the really bright seniors to study on in graduate schools for the good of the nation. The total number of doctoral degrees given in any one year in electrical engineering excluding foreign students, who must return to their respective countries, could just about suffice to furnish the needed faculty members for the engineering schools, or it could cover about half of the requirement of industry and government, but it is hopelessly too small for the total demands.

University employment of graduate students has been a device of self defense in order to retain the qualified students who could not gain sufficient support with a fellowship even with the more liberal stipends that are now often available. The tremendous expansion of government sponsored research at universities has created a new kind of partnership which, with proper regard for academic freedom, has made possible the present scope of graduate work. Again this phenomenon is related to the social reorientation of the nation, which has changed the tax structure so radically that university income has been impaired. In turn, this meant that universities no longer can afford the costly and complex modern research equipment out of their own funds. The growth of sponsored research has also raised questions of interference with the freedom of university faculties,[20] inasmuch as availability of funds might dictate the orientation of research programs. Enlightened university administrations will, however, value the full unfoldment of faculty talents and will strive for support where it is of greatest benefit to the educational objectives.

The evening programs of graduate study, initiated by the Carnegie Institute of Technology in 1912, were expanded substantially by the institutions in the New York metropolitan area[21] in the 1920's. The real growth took place during and at the end of World War II. Many programs are centered in heavy industrial areas and have, in fact, had the character of cooperative programs in the broader sense.[22] The benefit to students employed in industry, and therefore to industry itself, is very great. This has been recognized by the tuition refund plans of many industrial concerns. Perhaps the most attractive mode of cooperation is release of employees for substantial fractions of day time to permit attendance of normal campus classes.[23]

The growth of part-time, and in particular, of evening programs throughout the United States has caused recurrent discussions of the desirability of accreditation at least of Master's programs in the same manner as the Engineers Council for Professional Development (ECPD) accredits the undergraduate curricula in engineering. It is agreed, however, that the caliber of graduate study rests squarely upon the caliber of the faculty so that accreditation would essentially mean a critical appraisal of faculties and this would raise the pertinent question of how to select the judges. It might also be argued that graduate students, being more mature and having professional experience, can gauge fairly well by themselves the caliber of graduate courses and thus appraise the value of an advanced degree in certain programs. The need for protective action might thus be less urgent than in undergraduate programs where the choice is often made or supported by the parents.

[20] N. Kaplan, "Research overhead and the universities," *Science*, vol. 132, pp. 400–404; August 12, 1960.
L. A. DuBridge, "Research overhead," *Science*, vol. 132, pp. 1746–1747; December 9, 1960.
P. R. Trumpler, H. K. Schilling, and R. J. Martin, "Impact of engineering college research on graduate programs and faculty development," *J. Engrg. Educ.*, vol. 51, pp. 618–640; April, 1961.
[21] E. Weber, "Evening graduate programs in electrical engineering," *J. Engrg. Educ.*, vol. 38, pp. 256–261; December, 1947.
A. B. Giordano, "Part-time graduate study in New York Metropolitan area," *J. Engrg. Educ.*, vol. 50, pp. 574–581; March, 1960.
[22] "Cooperative Education at the Graduate Level, Report on Conference at Troy, 1949," published by Division of Graduate Studies, ASEE, University of Illinois, Urbana, Ill.; January, 1950.
[23] J. M. Petit, "Part-time graduate electrical engineering programs in the San Francisco Bay area," *J. Engrg. Educ.*, vol. 50, no. 7, pp. 582–576; March, 1960.

Centers of Graduate Study

The fact that the growth of graduate enrollment cannot be accepted as a measure of available manpower with advanced degrees has given concern.[24] A recently published study by Dean Hollister of Cornell University[25] tries to provide an over-all picture of the performance of engineering schools at various academic levels. Separating the schools which offer strong graduate programs and taking as criterion the award of at least 25 Master's degrees and at least 5 Doctor's degrees in any one of the three years from 1952 to 1955, and using the figures published by the U. S. Office of Education for 1957–1958, the statistics in Table III result, leading to these conclusions:

1) Of the 220 institutions giving degrees in engineering, only 37 (20 privately supported, 17 state institutions) did offer the minimum number or more of advanced degrees specified above.

2) The total of these 37 institutions actually awarded 40 per cent of first degrees, 67.5 per cent of the Master's degrees and 91 per cent of the Doctor's degrees, thus establishing them essentially as the centers of graduate study in the country.

3) These same institutions also awarded about half of the Master's degrees in physics and mathematics, and more than two thirds of the Doctor's degrees in these fields, thus emphasizing the modern concept that strong graduate work in engineering must have as a corollary strong graduate programs in the basic sciences. Chemistry is also listed in Table III though its closeness is primarily important for chemical engineering.

Graduate study and research are inseparable components of advanced work so that one can also expect that the same institutions carry a large part of the sponsored research programs.

A recent study by the Engineering College Research Council of the ASEE bears out this correlation[26] as shown in Table IV. Taking the data as supplied by the institutions responding to the questionnaire on research expenditures for the fiscal year 1958–1959 and combining with the ASEE data on degrees conferred, we find 24 institutions that gave more than 10 Ph.D. degrees in engineering or 83 per cent of all Ph.D. degrees in engineering and whose research expenditures averaged $2,360,000 per institution. Converting then the Ph.D. degrees into equivalent Master's degrees by a factor of 3, we find an average annual expenditure per equivalent Master's degree of $10,900, which is remarkably close to the same averages computed for the other two groups B and C, demonstrating perhaps a uniform educational efficiency, though, of course, the larger programs afford larger and often much more expensive research equipment.

TABLE III

Summary of Degrees Granted in 1957–1958

Field and Degree	Total (220 Institutions)	Group A	Group B	Both Groups	Per Cent of Total
Engineering:					
First degrees	35,332	5325	8733	14,058	40.0
Master's degrees	5788	2275	1633	3908	67.5
Doctor's degrees	647	326	260	586	91.0
Mathematics:					
Master's degrees	1097	202	293	495	45.2
Doctor's degrees	210	80	65	145	69.2
Physics:					
Master's degrees	795	197	215	412	51.8
Doctor's degrees	464	186	133	319	69.0
Chemistry:					
Master's degrees	1125	178	259	437	39.0
Doctor's degrees	939	252	350	602	64.6

Notes: Group A comprises the 20 privately supported institutions that granted at least 25 Master's and 5 Doctor's degrees in engineering at least once during the period 1952–1955. Group B comprises the corresponding 17 publicly (state) supported institutions.

TABLE IV

Relationship Between Research Expenditures and Engineering Graduate Degrees Conferred 1958–1959

Group	Number of Colleges	M.S. Degrees	Ph.D. Degrees	Equivalent M.S. Degrees	Total Research Expenditures 1958–1959, Dollars	Annual Expenditure per Equiva-M.S., Dollars
A	24	3386	594	5168	56,703,000	10,900
B	32	1715	115	2068	20,296,000	10,000
C	46	948	—	948	9,436,000	10,000
Totals	102	6049	709	8184	86,435,000	
Grand totals	125	6562	714			

Notes: A—Institutions who awarded 10 or more Ph.D.'s in engineering in 1958–1959.
B—Institutions who awarded less than 10 Ph.D.'s in engineering in 1958–1959.
C—Institutions who only awarded Master's degrees in engineering in 1958–1959.

It is very interesting to note the inverse relationship of numbers of institutions in the three groups A, B, and C and their total research expenditures. In fact, one could argue that established strength appears to provide attraction for graduate students so that regenerative forces will build strength to greater strength and make it nearly impossible for new centers of graduate study to build up. On the other hand, should graduate enrollments continue their present trend, the problem of

[24] E. Weber, "Trends and Problems in Graduate Electrical Engineering Education," paper presented at the IRE International Convention, New York, N. Y.; March 22, 1961.

[25] S. C. Hollister, "Analysis of education potential in engineering colleges," *J. Engrg. Educ.*, vol. 48, pp. 586–593; April, 1958.

[26] R. J. Martin, "Impact of engineering college research on graduate programs and faculty development, view point of a research director," *J. Engrg. Educ.*, vol. 51, pp. 632–640; April, 1961.

handling large numbers of candidates in any one institution will arise, attention to the individual is bound to suffer and standards of performance might—unwittingly—be sacrificed. This has happened to a fair degree in the Master's programs where in many places the thesis has become optional or has been completely abolished. Certainly, this must not happen with doctor's programs or the necessity of formalized post-doctoral programs will appear!

Concentrating our attention upon electrical engineering, we find that 114 institutions gave Master's degrees during 1959–1960, and of these, 39 also awarded Doctor's degrees.[27] It is again interesting to note as shown in Table V, that 11 institutions of these 114, or about 10 per cent, conferred 25 or more Master's and 62 per cent of all Doctor's degrees! One must expect the question as to whether there exists a critical scope of graduate offerings supported by strong research and outstanding faculty members that leads to a natural grouping on a national scale. Certainly, the comparative data emphasize that graduate students, if given the opportunity, make choices of institutions on the basis of more pertinent information than undergraduate students.

TABLE V

Graduate Degrees in Electrical Engineering 1959–1960

		Master's		Doctor's	
		Number	Per cent	Number	Per cent
11	Institutions giving 25 or more Master's and 5 or more Doctor's degrees.	736	37	125	62
4	Institutions giving 5 or more Doctor's degrees but less than 25 Master's degrees.	53	3	26	13
24	Institutes giving less than 5 Doctor's degrees.	346	17	52	25
39	Institutions awarding 1 or more Doctor's degrees.	1135	57	203	100
75	Institutions only awarding Master's degrees.	858	43		
114	Institutions awarding graduate degrees.	1993	100		

[27] W. E. Tolliver and H. H. Armsby, "Engineering Enrollments and Degrees," Office of Educ., Dept. of Health, Educ., and Welfare, Washington, D. C., Circular No. 638; 1961.

Problems Confronting Us

Graduate education in electrical engineering has experienced a very dynamic expansion, a radical reorientation, and a spectacular growth of research. It is obvious that many problems have been created for which solutions are not at hand. Some of these problems, actively discussed in many educational circles as well as by the general public, are:

1) *Faculty:* Securing and holding faculty members, particularly those who contribute creatively in basic research and who stimulate graduate students to associate with them, has become a more severe problem[4] in the face of insufficient output of men with doctoral degrees in spite of rising graduate enrollments.

2) *Students:* Genuine full-time graduate students have become a relatively small fraction of the total graduate enrollment so that the number of advanced degrees awarded each year is lagging more and more severely behind the number of students admitted to graduate study.[24] A systematic effort is required to make full-time graduate study more attractive to students and to industry who employ the students.

3) *Support:* The costs of graduate study facilities and of student support are increasing beyond any feasible increase of university incomes. The federal government laid the original foundation for this country's growth of technological education by the Morrill Act in 1862. It will require a similar farsighted and broad act by the federal government to assure the adequate development of a sufficient number of first-class graduate study and research centers[28] to satisfy the national need.

4) *Curricula:* A very definite trend towards interdisciplinary graduate study and research has been evident. This poses questions of maintaining, altering, or abolishing traditional departmental patterns of operation, immediately in the graduate schools, and in the longer range, probably in undergraduate schools as well.

5) *Accreditation:* The proliferation of programs and relative increase in part-time study impress more urgently the need for setting appropriate standards and practices of graduate study, though it is generally agreed that the stature of graduate study depends primarily upon the faculty associated with it.

[28] L. V. Berkner, "Renaissance in the southwest," *Saturday Rev.*, pp. 42–47; June 3, 1961.

Section 11

ELECTRON DEVICES

Organized with the assistance of the IRE Professional Group on Electron Devices

Recent Developments in Space-Charge Control Tubes*

KARL R. SPANGENBERG†, FELLOW, IRE

***Summary*—In spite of being the oldest in the family of electron tubes, the space-charge control tube or conventional tube continues to experience continual development and improvement. Recent developments include extensions of power-frequency limits, improvements in emission, and extensive improvements of structure. Of particular novelty are a high g_m tube using secondary emission multiplication and a distributed-amplifier type tube.**

I. Introduction

IN A TIME when many stirring developments and new devices are coming into being, it is possible for the steady progress in certain well-established electronic devices to escape almost unnoticed. Such is the case with the conventional vacuum tube types which employ a control electrode to modify a current flow through space-charge modulation. Although the basic forms of such space-charge control tubes have been known for many years, there have been steady advances in almost all features of their operation. This paper will be concerned with developments within the last ten years of the operation of such tubes up to and including the ultra-high frequencies which is the portion of the frequency spectrum where they are most effective. These tubes have undergone significant recent improvements in form and operating capabilities, in addition to which the theory of their operation has been significantly advanced.

In a paper of this length it is not possible to refer to all of the developments in detail. Accordingly, an effort will be made to mention at least the significant developments and to give the pertinent references in which more detailed information can be found.

II. Emitters and Emission Studies

The theory and operation of thermionic electron emitters continues to be the subject of considerable study. Developments of oxide emitters are covered in several excellent review papers [1]–[3]. L-type emitters find increased application [4], likewise thoriated tungsten emitters have found increased application for heavy duty requirements [5]. Cold emission from magnesium oxide coatings has not developed to the point where extensive practical application seems imminent [6].

* Received by the IRE, July 5, 1961; revised manuscript received, November 16, 1961.
† Engrg. Consultant, 2100 Page Mill Rd., Palo Alto, Calif.

The development of film cast oxide cathode, known as Sarong Cathodes, has excited considerable interest. These cathodes may be formed with a thickness tolerance of 0.05 thousandths of an inch, which is an order of magnitude better than can be achieved with spray cathodes. Resulting improvement in lower filament power and more careful control of electrode spacings and current are significant [7]. Attention to design has made possible the reduction of cathode ray tube filament power by factors of one-third [8]. There have also been developed cathodes with a 5 second warmup time, 0.5 seconds for filamentary cathodes [9]. RCA has developed a "black" heater insulation which reduces heater wire temperature and so increases reliability and life.

III. Low Voltage Operation

The hybrid circuit design concept in automotive radios calling for 12-volt operation has led to the development of a series of space-charge grid tubes, especially for this purpose. Such tubes include the types 12K5, 12DS7, 12DV8, and 12AL8. These tubes contain tetrode or pentode sections, which are operated with a low positive voltage on the first grid and a slightly negative voltage on the second grid, which is the so-called space-charge grid and control grid for this type of operation [10]. This same principle has been used in certain cathode-ray tube guns [11].

IV. Developments in Form

A. Grid Structures

Many of the major tube companies have gone over to what is commonly called a frame grid construction. Notable among these are the structures of Philips and Sylvania [12], [13]. These structures make use of a rectangular grid frame which is designed to have a greater stiffness in all directions as well as increased ease of mounting. Techniques for winding extremely fine grids have been improved. It is now possible to form grids from 500 wires per inch with a positional accuracy of $\pm 5\mu (\pm 0.2$ mils) [14].

B. Metal Envelopes

Small and medium power electron tubes were originally made with envelopes which were mostly of glass. In the 1930's a number of metal envelope receiving type tubes were developed, which later were abandoned in favor of glass. Once more, low power metal envelope tubes have begun to appear. These include the RCA Nuvistor. This tube was designed to operate up into ultra-high frequencies and also have as low a filament power as possible. The first objective was achieved by using what is essentially the medium powered transmitting tube electrode configuration which supports all electrodes from a common base with a minimum of insulating material between electrodes. The principal features of these tubes are small size, low filament power, and low noise [15]. The use of the metal envelope combined with a ceramic base, giving a tube with no glass, makes possible operation at higher temperatures and, hence, a smaller tube for a given power level.

C. Metal-Ceramic Tubes

Metal-ceramic tubes were first manufactured in Germany during World War II. Later, there was considerable development of such techniques in the United States, notably by the General Electric Company for very small tubes and by Eitel-McCullough, Inc., for larger sized tubes. Some of the most recent additions to the metal-ceramic family include the RS1032 of Seimens and Halske (10 kw at 790 Mc, TV service), the TSP347 of Nippon Electric (100 watts CW at 2 kMc), and numerous tubes by American manufacturers including particularly RCA and the General Electric Company.

At still somewhat higher frequency, RCA has for several years been manufacturing a beam power tube, type 6952, which gives two Mw peak power output at 425 Mc. Another recent development is the RCA type 2054 super-power triode which gives 5 Mw peak power output in long pulse service at 440 Mc. More recently, RCA has described an integral circuit high-power amplifier triode for 420 Mc operation known as the coaxitron [16]. This tube is capable of 5 Mw peak power output over a frequency band extending from 385 to 465 Mc. Most of the tubes mentioned above are air cooled. Some related developments which are able to use glass because of careful water cooling include the TH470 of Compagnie Francaise Thomson-Houston which is capable of 2 Mw pulse operation at 200 Mc [17]. This tube uses the vapotron principle of cooling which makes use of the heat of vaporization of water to effect a maximum heat transfer by boiling water in a chamber about the anode. Improvements in smaller sized earlier type vapotrons take the form of improved corrugations in the anode for better heat transfer [18].

V. Super-High Frequency Operation

There have recently been extensive analyses of super-high frequency operation of conventional tubes both for small signal operation [19] and for large signal operation as part of the extensive studies of Professor Gundlach's group at the Technical University of Berlin [20], [21]. These studies have been supported by experimental observations [22], [23].

A remarkable series of super-high frequency lighthouse type triodes have been developed by Philip engineers. These tubes cover the range from 4000 to 6000 Mc with output powers as high as 1 watt at 6000 Mc for an air cooled tube and 10 watts at 6000 Mc for a water cooled tube [24]–[27]. Most of these tubes are, however, in a developmental form with only two types, namely the EC157 (1.5 watts at 4000 Mc amplifier operation) and the 22EC (5 watts at 4000 Mc amplifier operation) available as production items. These tubes use L-type cathodes and feature cathode-to-grid spacing as low as 25 microns (1/1000 of an inch). Such spacings have also been achieved in certain Western Electric and General Electric types.

VI. Ruggedized Tubes

Many modern tube applications require tubes which have been ruggedized to withstand rather severe shock and vibration. Outstanding among these are the RCA metal-ceramic pencil triodes types 7553 and 7554. These tubes are capable of being mounted in an integral pencil tube-cavity unit capable of withstanding 10,000 G's of longitudinal acceleration and 850 G's of transverse shock [28].

VII. Distributed Amplifier Tube

Another interesting recent development is a distributed amplifier tube, a new type of wide-band amplifier tube, in a single vacuum envelope. The anode and grid of this tube are helical lines across which an electron beam passes. An early model of this tube has given a gain of 17 ± 5 db over a 150 Mc video band. Gain drops from the low to the high frequencies but could be equalized to give an output constant within 3 or 4 db. The grid line has an impedance of 200 ohms and the anode line has an impedance of 450 ohms [29].

A later version of the tube gave a voltage gain of 9.0 ± 1 db (power gain of 7.0 db) from a 150-ohm grid line to a 200-ohm plate line over a 320-Mc bandwidth [30].

VIII. High-Current Grid-Controlled Electron Multiplier Tube

A grid-controlled secondary-emission electron multiplier tube has been designed and built which can deliver an output pulse of 5 amperes into a load impedance of 100 ohms, with a rise time of less than 10 nsec and a transit time of less than 20 nsec. The measured transconductance of the tube is 0.6 amperes per volt, and it can provide a positive output pulse with a positive grid input [31].

IX. Conclusion

A study of power-frequency limits shows an inverse square law barrier for space-charge control tubes. This situation continues with the barrier having moved up by a factor of ten in the last decade or two. Conceivably another factor of ten could be achieved in the next decade or two. More likely the most significant new devices will be unconventional structures performing more complex operations in a single vacuum envelope.

References

[1] L. S. Nergaard, "Thermionic emitters," *RCA Rev.*, vol. 20, pp. 191–205; June, 1959.

[2] L. S. Nergaard, "The physics of the oxide cathode," *RCA Rev.*, vol. 18, pp. 486–511; December, 1957.

[3] J. Ruf, "Ueber die Emissionsgleichung der Oxydkathode," *Die Telefunken Röhre*, vol. 37, pp. 79–132; October, 1960.

[4] G. Diemer, K. Rodenhuis, and J. G. Wijngaarden, "The EC57, disc-seal microwave triode with L cathode," *Philips Tech. Rev.*, vol. 18, No. 11, pp. 317–324; 1956/57.

[5] M. V. Hoover, "Advances in the techniques and applications of very-high-power grid-controlled tubes," *Proc. IEE*, vol. 105, Part B, Suppl. No. 10; 1958.

[6] A. M. Skellett, B. G. Firth, and D. W. Mayer, "The magnesium oxide cold cathode and its application in vacuum tubes," Proc. IRE, vol. 47, pp. 1704–1712; October, 1959.

[7] W. J. Kohl, "Materials and Techniques for Electron Tubes," Rheinhold Publishing Co., New York, N. Y.; 1960.

[8] H. E. Smithgall, "A low-heater-power cathode-ray tube," *Sylvania Engrg. Information Service*, vol. 7; July, 1960.

[9] P. J. Erdle, Sylvania Electric Products, Emporium, Pa., private communication, January 26, 1961.

[10] Tech. Information Sheets, Sylvania Types 12K5, 12DS7, 12DV8, and 12AL8.

[11] P. H. Gleichauf, "Space-charge grid high-transconductance guns," Proc. IRE, vol. 46, p. 1542; August, 1958.

[12] Tech. Information Bulletins, Philips Types E810F, PC86, PCC88, PC95, and EF183; 1961.

[13] Sylvania Engrg. Data Service, Types 6GK5, 7803, 6939, 6ER5, etc.

[14] A. E. Widdowson, D. C. Gore, and C. H. Butcher, "A new method of making accurate fine-wire grids for use in radio valves," *Proc. IEE*, vol. 105, pt. B, Suppl. 10, pp. 567–576; 1958.

[15] G. M. Rose, "Nuvistor—New Look in Tube Design," Electron Tube Div., RCA, Harrison, N. J., Publication ST-1654; May, 1959.

[16] W. N. Parker, "The coaxitron tube," *The RCA Electronics Pioneer*, pp. 10–12; February, 1961.

[17] J. Péhé, "Étude du Tube TH470," *Revue Technique*, Compagnie Francaise Thomson-Houston, pp. 13–20; December, 1956.

[18] C. Beutheret, "Technique des Vapotrons," *Revue Technique Compagnie Francaise Thomson-Houston*, pp. 53–85; December, 1956.

[19] M. T. Vlaardingerbroek, "Small signal performance and noise properties of microwave triodes," *Philips Res. Repts.*, vol. 15, pp. 124–221; April, 1960.

[20] H. Paucksch, "Die Elektronenlaufzeit in einer ebenen Raumladungsdiode," *Arch. Elek. Übertragung*, pp. 171–176; 1955.

[21] G. Heinzmann, "Das Verhalten der Triode im Laufzeitgebiet beihohen, Aussteuerungsgraden," *Arch. Elek. Übertragung*, pp. 13–20; 1959.

[22] K. Hennings, "Leitwerts-Messungen an planparallelen Diode im Laufzeitgebiet," *Nachrtech. Z.*, pp. 459–464; 1959.

[23] F. W. Gundlach and H. Lonsderfer, "Leitwerts-und Steilheitsmessungen an Dioden und Trioden im Laufzeitgebiet," *Fernmeldetech. Z.*, vol. 8, pp. 1–5; June, 1955.

[24] V. V. Schwab and J. G. van Wijngaarden, "The EC59, a transmitting triode with 10 watts output at 4000 mc," *Philips Tech. Rev.*, vol. 20, No. 8, pp. 225–233; 1958–1959.

[25] E. Mentzel and H. Stietzel, "A metal-ceramic disc-seal triode for frequencies up to 6000 mc," *Philips Tech. Rev.*, vol. 21, No. 3, pp. 104–108; 1959–1960.

[26] M. T. Vlaardingerbroek, "An experimental disc-seal triode for 6000 mc," *Philips Tech. Rev.*, vol. 21, No. 6, pp. 167–172; 1959–1960.

[27] H. Groendijk, "Microwave triodes," *Proc. IEE*, vol. 105, pt. B, Suppl. No. 10, pp. 577–582; 1958.

[28] RCA pencil tubes, Radio Corp. of America; 1960.

[29] T. Kojima, "Distributed amplifier tube," *Rev. of the Elec. Commun. Lab.* (Japan), vol. 8, pp. 343–368; July–August, 1960.

[30] T. Kojima, "Feedback type distributed amplifier tube," *Rev. of the Elec. Commun. Lab.*, vol. 9, pp. 150–164; March–April, 1961.

[31] G. L. Stambach, W. J. Graham, and T. E. Hanley, "High-current grid-controlled electron multiplier tube," IRE Trans. on Electron Devices, vol. ED-7, pp. 143–146; July, 1960.

The Development of Gas Discharge Tubes*

J. D. COBINE†, FELLOW, IRE

Summary—**The paper is an historical review of the development of gas discharge tubes. The material is divided into five classifications based on discharge types, namely:**

1) **Ionization tubes,**
2) **Cold-cathode discharge tubes,**
3) **Hot-cathode arc discharge tubes,**
4) **Liquid metal arc tubes, and**
5) **Plasma tubes.**

INTRODUCTION

ELECTRICAL DISCHARGES in gases have had a long and fascinating history. If the natural phenomena such as lightning, St. Elmo's fire, etc. are excluded, one of the earliest manifestations of electrical discharges in gases was observed in 1683 when Otto von Guerick [1], of Magdeburg produced spark and corona discharges by a static-electricity machine. Low pressure glow discharges in "exhausted" glass spheres were observed about 1700 by Newton and Hawksbee also using static-electricity generators. The first steady electric arc was produced by Davy in 1808–1810. He first observed transient arcs in 1801 when opening a short circuit on a voltaic pile. Aside from flames, these examples were the first time man had produced the strange state of matter named *Plasma* by Langmuir (c. 1923) that has been receiving great attention recently in both science and engineering.

In 1856, Heinrich Geissler [1]–[3], an artist and skillful glass blower of Bonn, Germany, originated the low pressure discharge tubes that were to bear his name. The Geissler tubes were long, small-bore glass tubes, usually shortened by the use of many coils and bends, which were filled with various gases at low pressures and originally excited by high-voltage alternating current. Many beautiful effects could be produced by Geissler tubes filled with different gases and they were often used for decorations. As, for example, a display used to commemorate Queen Victoria's Diamond Jubilee [4]. However, sputtering of the electrodes together with gas clean-up resulted in a short life for the tube. The principal use was for spectral analysis and lecture demonstrations.

The first public demonstration of a mercury arc lamp was by Prof. J. T. Way on the Hungerford Suspension Bridge in London on September 3, 1860. Two British patents were issued to him in 1857. The electric arc was first used commercially for illumination in Paris in 1863. Much later, low-pressure arc "tubes" were used for illumination. In 1879 John Rapieff described mercury arc lamps in British Patent No. 211 but there appears to be no evidence that they were built. Peter Cooper Hewitt showed in public his mercury-arc lamp on April 12, 1901. Georges Claude, a French inventor, demonstrated the first Neon sign [4], an improvement of the Geissler tube, at the Grand Palais in Paris in 1910. Developments in luminous tube discharges were made by Moore [5], in 1920. Since these tubes did not have a high light output, they were largely confined to sign applications.

Low-level ionization was present in some of the early radio tubes, and had considerable effect on their characteristics. The focusing of early Braun cathode-ray tubes and the Roentgen "gas" x-ray tubes depended on the focusing action of positive ions produced in the residual gas by the electrons. These ions, of course, resulted in slow response and short cathode life. The necessity of converting ac to dc gave rise to the mercury-pool arc rectifier. The first observation of the rectifying properties of an arc between a mercury and a carbon electrode was made by a French scientist, Jamin [6], in 1882. Probably the earliest scientific study of the low-voltage high-current arc was by Arons at Berlin [7]. In 1903, Cooper Hewitt announced [8] his mercury-pool rectifier. As early as 1905 Cooper Hewitt made metal-tank mercury-pool power rectifiers. In those days when electric automobiles and trucks enjoyed considerable favor, many two-anode mercury-pool, glass envelope rectifiers were in use for battery charging. After the dominance of the gasoline engine automobile, these pool tubes continued to be made for automobile-starting battery charging, being finally displaced by the *Tungar* [9] arc tube. Tungar was a trade name (G. E., Westinghouse called it *Rectigon*) resulting from the contraction of tungsten and argon. This arc tube had a tungsten cathode, which was operated at higher than normal temperature to attain several amperes of emission, and argon gas at a pressure of several Torr to prevent excessive evaporation. The Tungar has since, in many applications, been replaced by various members of the growing family of solid-state devices. The family of gas discharge tubes of interest to the radio and electronic engineer has been growing at a slow but steady rate ever since. There have been many contributions to this development. Most, if not all, were empirical solutions to special problems. Usually, theoretical explanations were developed after the fact, if at all. The discharges are extremely complex phenomena and have occupied the attention of many of the world's best scientists.

* Received by the IRE, August 23, 1961; revised manuscript received, November 15, 1961.
† General Electric Research Laboratory, Schenectady, N. Y.

The gas tubes with which we will be concerned may be grouped for convenience in five divisions:

1) Ionization tubes, *i.e.*, non-self-sustained discharges,
2) Cold-cathode glow discharge tubes,
3) Hot-cathode arc discharge tubes,
4) Liquid metal (or pool) arc tubes,
5) Plasma tubes.

These types of gas tubes will be reviewed in this order in the following sections. In order to provide proper perspective for the following discussion, the relations of the various discharges are shown by the composite volt-ampere characteristics of Fig. 1 and will be discussed briefly.

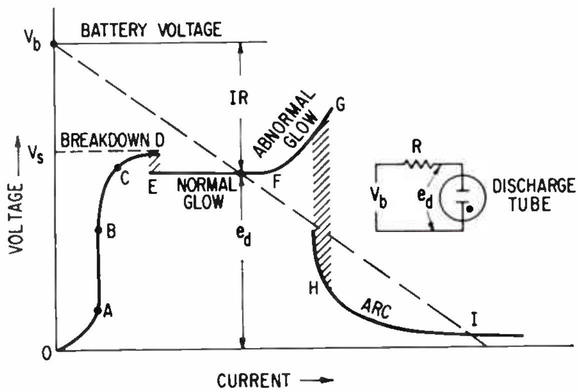


Fig. 1—Gas discharge volt-ampere characteristics.

The region *O-D* is roughly representative of the ionization type of tubes, group 1. In *O-A* partial collection of background ionization is represented. From *A-B* the tube is "saturated" and all ionization produced by external means is collected. From *B* to *C* is represented by the simple gas amplification of the Townsend [10], [11] type where the number of electrons increases exponentially with distance, by collision ionization. From *C* to *D* additional ionization processes develop, usually involving additional emission from the cathode as a consequence of earlier emission and ionization increases greatly to result in breakdown at *D*. Following *D*, all the discharges are self-sustained, *i.e.*, independent of the initiating ionization event. The region *E-F* is a normal glow, characterized by a large and substantially constant voltage drop of the cathode. This is the low pressure discharge of the familiar voltage regulator and voltage reference tubes of group 2, the indicator glow lamps are also cold-cathode tubes operating in the region *E-G*. Region *F-G* is the abnormal glow, in which the current density of the cathode surface increases with current, and the cathode glow becomes more intense. The glow-arc transition occurs at about *G*. Hot-cathode (group 3) and pool arcs (group 4) operate in the region *H-I*. These are essentially low-voltage high-current discharges. Both the cold-cathode glow and the hot-cathode arcs have regions of glowing gas outside the cathode glow region that are known as positive columns. This zone has a low voltage gradient, as it is a state of near electrical neutrality usually called a plasma. This part of the discharge may be used in the plasma tubes (group 5) for light, or for microwave purposes. The point *D* represents a breakdown of the gas, and in the case of the spark gaps, a very rapid transition takes place to some point such as *I* on the arc characteristic.

It is clearly impossible to cover properly the scientific aspects of the phenomena involved in the above discharges. The interested reader is referred to texts and treatises on the subject, some of which are included in the bibliography [12]–[15].

Ionization Tubes

The study of "simple" ionization in gases has contributed enormously to man's knowledge of the universe, ranging from the constitution of the stars and the transformation of matter into radiant energy, to the density of space. In 1785, Coulomb [16] noted the loss of electricity to the air from a charged insulated body. Both Geitel [17] and Wilson [18] found that the rate of leakage from an insulated body did not increase in proportion to the applied potentials, but soon saturated for a closed volume. With the determination of the charge in the electron in 1909 by Millikan, the number of ionizing events per second per cubic centimeter could be determined accurately. In 1912, McLennan [19] found this rate to be 9 over land and 6 over the sea.

About this time, experiments were being conducted with the collection of current by a positive wire and negative cylinder arrangement that were to have a profound effect on science. These experiments were reported in a paper by Rutherford and Geiger [20] which showed that the number of charges of an ionizing event could be multiplied several thousand times by the ionizing action of electrons in the high field region near the wire. This was the start of what for a time were called *Geiger-Müller counter* tubes and now simply *Geiger counter* tubes. The technique of the proportional type of counter was established [21] in 1928, and the following year, schemes for determining the coincidence of ionizing events were presented [22]. Thus direction, scattering, absorption, etc. types of experiments were possible and the modern era of cosmic-ray and nuclear research developed.

The mobility of electrons is very high compared to that of the positive ions and an electron avalanche in the high field region develops with great speed, so that, as far as the electrons are concerned, the count of an initial ionizing event can be made in less than 10^{-9} second. The ions may take of the order of 100 μsec to cross the gap. Spurious counts and self-sustained discharges can be produced by the emission of electrons at the cathode by ion-bombardment. Organic vapors were found [23] to quench the discharges and spurious counts were suppressed by attachment of the ion-emitted electrons. This speeded up the counter. It was found necessary to avoid pure inert gases because of the long life of the metastables formed by the action of initial avalanche. The counters for a time represented a considerable body of art, but with advancing knowledge these have become reliable and trustworthy devices for a wide range of applications.

The simple triode ionization gauge [24] is one of the oldest devices for the measurement of low gas pressures. With various modern refinements, it is the most widely used device for vacuum measurements. The device is based on the simple ionization of residual gas by electrons traversing the space between cathode and anode. As long as the electron energy is considerably greater than the ionization potential of the gas and the pressure less than that for which the electron mean-free-path is of the order of the interelectrode distance, the positive ion current will be proportional to the gas pressure. The anode is an open wire spiral, and the ion collector in the early tubes was a solid cylinder surrounding the other elements. Because of the open construction of the anode, many electrons oscillate on approaching it so that the effective distance traversed by the electron is considerably greater than the anode-cathode spacing. Because of the large number of electrons emitted by an incandescent filament, a measurable positive-ion current is present at pressures as low as 10^{-8} Torr. Below this pressure, the lowest that the tube could measure before 1948, the ion current for a gauge having 150 volts on the anode and a sensitivity of 10/Torr was just equal to the photoelectron current from the collector due to the soft x-rays produced when the 150-volt electron struck the anode. The effect was pointed out by W. B. Nottingham in 1947. In 1950 a modified structure was devised [25] which greatly reduced the area of the collector that could intercept x-rays. This was done by making the ion collector a centrally located wire, which reduced the lower pressure limit to 10^{-10} Torr. The lower limit was recently further reduced to 10^{-14} by the development of a hot-cathode magnetron ionization gauge [26]. This tube has the added advantage of being of ceramic construction which permits a high temperature bakeout of the tube and greatly reduces the helium diffusion into the tube. The use of the magnetic field results in a great increase in the distance an electron travels before reaching the anode so that its probability of having an ionizing collision is proportionately increased.

Several other vacuum gauges depending on ionizing the gas can only be mentioned. These are the Alphatron [27], which depends on collecting the ion current produced by alpha particles from a radioactive source, and the Penning type ionization gauge (PIG) [28]. The type of discharge found in the Penning gauge is often loosely referred to as a "PIG" discharge. The latter is a cold cathode tube in which electrons emitted by the collector are forced by a magnetic field to make many oscillations before reaching the anode, thus greatly reducing the minimum pressure that can be measured. The mass spectrograph tubes [29] are much more discriminating ionization gauges, *i.e.*, partial pressure gauges, which are becoming very important in fusion and space studies.

Cold Cathode Discharge Tubes

Experiments with low-pressure glow discharges started very early in the electrical art. Most of them were in small diameter glow tubes, often in the form called Geissler tubes. The first recognition of the fact that such a glow tube having its two electrodes of different size was capable of rectifying the oscillating current from an induction coil appears to have been that of Gaugain [30] in 1855. However, for many years, the only use for these discharges was as light sources. The early glow lamps all required high voltage excitation. With the availability of the rare gases neon and argon and by means of low work-function cathodes, glow tubes were developed for low voltage applications [5].

The first glow discharge rectifier tube appears to have been what was called an "S-tube." This tube is described in a paper by Bush and Smith [31] and consisted of closely spaced electrodes, the cathode being hollow. The authors attribute its rectifying properties to the characteristics of the low pressure branch of the Paschen curve [11], where short path breakdown voltage is high, but a much lower voltage will break down the gas if a long path is available. The hollow cathode permitted long path breakdown when that electrode was negative, but the close spacing of the two electrodes prevented breakdown on the inverse voltage. The tube discussed was rated for a current of 50 ma and an inverse voltage of 1500 volts. Magnetic control of the glow discharge was also demonstrated in the above paper.

In 1927 Bareiss [32] published a paper on a small glow discharge rectifier and showed the rectification to be due to the different volt-ampere characteristic obtained with the two sizes of cathode. Referring to Fig. 1, onset of the abnormal glow F, with its rapidly rising volt-ampere characteristic, occurs at a much smaller current for the small cathode than for the large cathode. Thus, although current (ac) will flow during both half-cycles of an impressed voltage, more flows when the large electrode is cathode and a net dc current results [33]. The low-voltage glow lamps, better called negative or cathode glow lamps, have found many applications [34] as long-life, dependable, voltage indicators and relaxation oscillators. The relatively constant potential characteristic of the normal glow discharge has permitted their use as voltage-regulator tubes and as voltage-reference tubes. In the former application, the tubes are made with quite large cathodes to permit as wide a range in current fluctuation as possible. For reference purposes as in a regulated power supply [35], the tube can be made smaller as it works at a small and substantially constant current. An important advance in construction of glow tubes was made by Lafferty [36] with the ceramic voltage regulator and voltage reference tubes. The high temperature (~1000°C) to which these tubes are heated during processing ensures that all parts are thoroughly outgassed. The gettering action by the titanium parts as they cool down is very great so that any chemically active gases that may be present in the fill gas are completely removed. Thus a truly clean tube is possible. This eliminates the discontinuous steps taken by most voltage regulator tubes as current variation causes the negative glow to spread over cathode areas whose emission properties are nonuniform due to a random distribution of impurities. The individual ceramic glow tubes can be made in disk form so that they can be stacked to form a high voltage regulator. Regulator tubes for voltages of 2,000 volts per inch of tube length are possible. These new ceramic tubes are more rugged than conventional tubes and can be operated at temperatures up to 500°C. Fig. 2 shows a variety of forms of developmental ceramic glow discharge tubes [36].

Voltage regulator tubes maintain a constant dc voltage where the input voltage changes slowly. However, the glow discharge represents an extremely complex circuit element whose impedance may be quite frequency sensitive [37]–[40]. For this reason they should not be expected to remove the higher harmonics of a ripple voltage.

Controlled glow discharge tubes were developed rather early [41], [42]. Until rather recently these have been confined largely to standby applications where a minimum of nonoperating power is required. With the increasing use of counting devices, computers, etc. the glow tube, either in the form of a multicathode device [43]–[46] or as tubes with a third control electrode [47], [48], has found many applications. The speed with which the glow tubes can react to recurrent events is limited by their de-ionization time. The upper frequency limits have been for 40–100 Kc, although in one case [47], megacycle frequencies are cited.

Early attempts [49] were made by Nienhold in the period 1916–1919 to utilize the plasma of a cold cathode glow discharge as the source of electrons for a multi-element "vacuum" tube. Noise and fluctuations in the primary discharge prevented this idea from being practical. Interest has recently revived in the possibilities of this discharge application [50].

The new ceramic techniques have been applied to spark gaps with the development of sealed gaps of great reliability and capable of operating at temperatures as high as 400°C. As simple two-electrode devices, these tubes are used to protect capacitors from over-voltages. With the addition of a third electrode, to be used as a trigger, they are used to control the stored energy in certain systems. The type Z-5369 is an example of this triggered spark gap, the construction of which is shown in Fig. 3. Gaps have been designed for application in the range of 2 to 10 Kv with a very rapid glow-arc transition [41], [36].

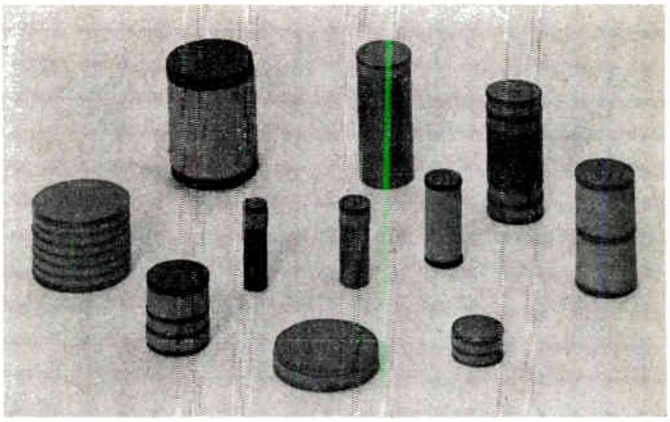


Fig. 2—Developmental high-temperature, glow-discharge tubes.

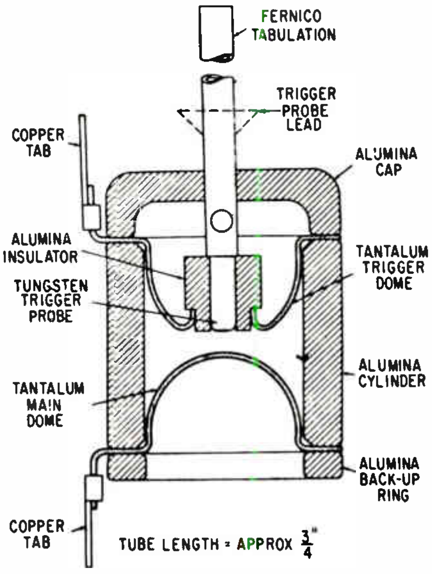


Fig. 3—Ceramic triggered spark gap Z-5369.

Hot-Cathode Arcs

In common with most gas discharge devices, the applications of hot-cathode grid-controlled arcs developed experimentally long before a reasonable understanding of the phenomena involved was attained. In fact, even today it is not difficult to find arc phenomena that are not understood. It was only natural that attempts should be made to control the flow of current in a gas discharge by means of a grid such as was used in the vacuum triode. Apparently, both Pierce [52] at Harvard University and Langmuir of the General Electric Research Laboratory were experimenting along these lines in the period 1913–1914. In 1914 Langmuir applied for a patent [53] on a method of controlling the current in a mercury arc. In this application he was particularly interested in controlling the flow of high-frequency power from an alternator, presumably an "Alexanderson" alternator, for antenna radiation. The Langmuir grid controlled arc tube is shown in Fig. 4 where it is seen to be an adaptation of the old two-anode pool-type mercury-arc rectifier mentioned in the introduction. At very low currents it was observed that the grids in these gas tubes functioned like the grid of a vacuum tube in controlling current. However, once the gas was ionized the grid lost its control, *i.e.*, it could prevent or start the flow of current, which once started was independent of the grid potential. This was a baffling behavior and one that was not explained until Langmuir wrote his classic papers on plasma [54]–[56]. It is interesting to note that this scientific study that so greatly advanced knowledge of the plasma state was undertaken as a consequence of the stimulation provided by practical problems, namely mercury-arc rectifiers. Langmuir's study demonstrated that the ability of the grid to control the flow of current through an ionized gas was due to the screening of the grid potential by a positive ion sheath. Thus, a negative potential placed on the grid resulted in the collection of a positive-ion current from the plasma and a positive-ion space charge (obeying the Langmuir-Child relation [14]) that completely shielded the grid from the balance of the discharge. Originally the name thyratron (Greek θυρα, a door) was given to a grid-controlled arc, both pool- and hot-cathode. Usage has confined the term to the hot-cathode, grid-controlled arcs. Following the basic papers by Hull and Langmuir [57]–[59] on hot-cathode thyratrons, these tubes have found many applications.

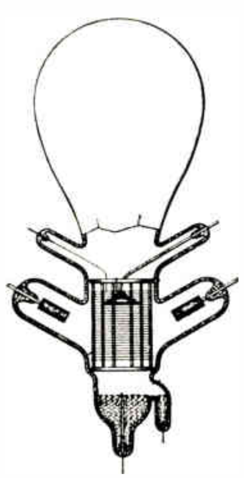


Fig. 4—Langmuir's thyratron.

Prior to World War II, a considerable number of thyratron types had been developed using either mercury vapor or one of the inert gases. Most of these tubes had rated currents falling in the range from about 50 ma to 5 amperes. The practical design of the various tubes are empirical and the result of engineering compromises to meet specific requirements. Many of the factors involved have been authoritatively summarized by Knight [60]. These tubes were ideal for many relay and control operations, motor speed regulators, etc. The requirements of radar modulators provided a special stimulus for the development of a thyratron with a short de-ionization. In 1949 papers were published [60]–[62] which reported the development of hydrogen thyratron devices which exhibited two advantages over the conventional thyratron. First, the de-ionization time was about one-tenth that of conventional thyratrons; and second, the oxide cathode was not destroyed by cathode-fall voltages below 600 volts—the maximum cathode fall permissible for mercury tubes where the positive ions are very much heavier was 30 volts. Hydrogen thyratrons for pulse power ratings from 10 to 2500 kw were developed and in production during World War II. Loss of hydrogen (clean-up) has been eliminated by the use of hydrogen reservoirs, usually consisting of temperature-controlled filaments of titanium. In recent years the increasing speed of airplanes and the development of high-speed missiles has greatly stimulated developmental activity in high-power hydrogen thyratrons for very long distance radar application. Ceramic construction applied to the hydrogen thyratron has made possible single tubes of the very high power rating [63] of 50 megawatts and greater for modern radar modulators. The hydrogen thyratron shown in Fig. 5 has a

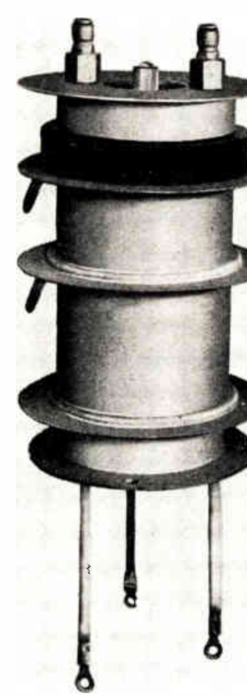


Fig. 5—Ceramic hydrogen thyratron GL-7890. Peak power=48 megawatts at 44 kw; peak current=2400 amperes; body length $=13\frac{1}{4}$ inches; ceramic diameter $=4\frac{1}{2}$ inches.

rating of 48 megawatts peak power and makes an interesting comparison with the original thyratron of Fig. 4. The new ceramic construction [36] has been adapted to both inert gas thyratrons and rectifiers [64] which can be operated at ambient temperatures up to 500°C and in the presence of considerable atomic radiation and shock.

The arc rectifier tubes, often called phanotrons (Greek φαινω appear or will be visible), developed along with the thyratrons [65]–[67]. They overcame one of the important limitations of the Tungar tube, namely its low inverse voltage rating. Both higher currents and very much higher voltage ratings were possible. The tubes had wide applications, but are slowly being replaced by solid-state devices. However, ability of the gas tubes to stand high surge currents and very high inverse voltages has not been equaled yet in the new devices. The relatively high arc drop of the gas tubes (8–12 volts) has been greatly reduced by the use of cesium vapor. The cesium-vapor rectifier [68] tubes have arc drops of from 2–4 volts for currents from 600–2400 amperes.

In common with all arc discharges and most gas discharges, thyratrons generate noise and oscillations over a wide frequency range. The variational voltages were studied [69], [70] as early as 1923 and analyzed in some phases by Tonks and Langmuir [71]. The need for noise sources for airborne radar-jamming transmitters during World War II led to the development [72]–[74] of high level gas discharge noise generators. The inherent noise output, which is of the order of 16 volts peak-to-peak for an argon tube, can cause considerable disturbance in adjacent circuits. It was found [75] by an important and thorough study [76], [77] of hot cathode arcs that, by operating a tube in the anode glow mode, the noise could be largely eliminated. A tube designed to operate in the anode mode was called a Tacitron [75]. This same series of arc studies led to the development of a new type of arc tube called the Plasmatron [78], [79] (Bendix RXB-103005). This is a rediscovery of a phenomena that was once known as the "Found Effect" [80], [81] and which had been forgotten. In the plasmatron tube, a plasma is established by an electron beam from an auxiliary cathode and its "constricting" grid. The main anode-cathode potential may then be as low as a few volts and the current can be modulated by varying the beam current. The tube of LeVan and Weeks [82] made use of a screened plasma as a source of electrons which could be modulated much as had been done earlier with glow discharges.

The gas-, or vapor-, type thermionic converter is a new type of hot-cathode gas discharge tube of considerable potential value for the conversion of heat to electricity. It was anticipated by the study of the vacuum thermionic converter by Schlicter [83]. The basic phenomena of the gas phase type was first reported on by Medicus and Wehner [84]. Later studies [85], [86] developed the principles in considerable detail and produced relatively efficient devices. In this tube a discharge is maintained by the difference in temperatures and work functions of the cathode and anode. Electrons emitted thermionically from a cathode have *thermal* energy (small) corresponding to the cathode temperature and *potential* energy corresponding to the work function of the emitter. If these electrons are collected by an anode whose work function is lower than that of the cathode, current may flow in an external load connected between cathode and anode due to the difference in work-function potentials. Since the range of work functions is rather limited, the output voltage is relatively low at best and it is extremely important that internal losses be minimized. Cesium, whose ionization potential is 3.87, is rapidly ionized by contact with hot tungsten whose work function is 4.52. Thus, a plasma may be produced and also the electron space charge at the cathode may be neutralized by the resulting positive ions. Cesium vapor condensed on the colder anode will result in the desired lower work function. Since cesium vapor must be maintained at a pressure of the order of 1-Torr, the temperature of the coldest point within the envelope (where cesium starts to condense) must be of the order of 275°C. Thus, a lower limit for the anode temperature is set. Efficiencies of 17 per cent have been realized and of 20 to 25 per cent may be reasonably expected of present developments. These converters have potential application to space vehicles for the conversion of nuclear or solar heat to electricity as well as many military applications requiring light weight, silent conversion of heat (nuclear or fossil) to electricity.

Liquid Metal Arc Tubes

As indicated in the introduction, one of the earliest applications of gas discharge tubes was as rectifiers. Initially, these were two-anode, glass-envelope mercury-pool tubes. By 1905 metal tanks were tried and developed into the large multianode, metal-tank mercury-pool power rectifiers such as described by Hull and Brown [87]. An important advance in the rectifier art was made by Slepian [88], with the discovery of the resistance, or immersion, ignitor. This is a semiconducting rod of boron carbide or silicon carbide immersed in the mercury pool. The discharge of a pulse of current from the ignitor to the mercury initiates a cathode spot. This permitted single-anode-pool tubes to be made and reliably started at any point in a positive half-cycle of anode potential. Thus, large current, controlled arcs could be used for straight rectification with voltage control, or the tubes could be applied to half-cycle welding, large motor control, etc. Although in some rectifying applications the pool tubes are being replaced by solid-state devices, the pool tubes possess the unique advantage of being able to recover from extremely high surge currents and withstand high inverse voltages without damage. As examples of the size of ignitrons now being used, two welding ignitrons (7151) in inverse parallel can switch 4,800 kva at 250 to 600 volts controlled precisely from 1 to 30 cycles if desired, while the locomotive ignitron (6878) shown in Fig. 6, has an average current rate of 675 amperes, a maximum peak anode current of 2500 amperes and a maximum peak inverse anode voltage of 4000 volts.

Ceramic construction [89], the use of metal anodes and the use of the alkali metals [90] added to mercury have greatly extended the useful possibilities of ignitron tubes. These metal anode ignitrons have been used in

Fig. 6—Locomotive rectifier. Body length $23\frac{1}{4}$ inches; peak anode current = 2500 amperes; peak fault current = 40,000 amperes; maximum peak anode voltage = 4,000 volts.

nuclear studies to control 20 kv, 1,000,000 ampere (GL-7703) capacitor discharges. The use of the metal anode avoids the release of gas which took place from the old carbon anodes and consequently produced a rapid deterioration of voltage capabilities. The alkali metal additives have reduced the arc drop from the order of 12 volts to about 4 volts. Although this is somewhat higher than the drop of the solid state devices, the large overload capacity and high inverse voltage ratings of the pool tubes with alkali additives should prove attractive for many applications.

The resistance ignitor requires rather high values of peak power to start an arc cathode spot satisfactorily. Attempts have been made since Cooper-Hewitt's band or capacitor starter to develop a low-current ignitor. A capacitor-type arc ignitor would require quite low power and would greatly simplify the control apparatus. Ignitor tubes of this type [91] have been made to operate for some time. However, none have proved commercially reliable over extended periods of time. This problem may eventually be solved using new materials and better vacuum techniques.

Plasma Tubes

The tubes to be considered here are those that function because of the special properties of a discharge plasma. They may utilize the plasmas of the negative glow or the positive column of a glow discharge or the plasma of a low-pressure arc discharge. All applications are as microwave devices. Early observations of high frequency discharges were made by Lodge [92], Thompson [93], Tesla [94]. A considerable number of experimental studies of high frequency discharges, mostly below 300 Mc, were performed over the years without yielding a very clear understanding of the phenomena. No important applications were made until 1941 when the requirements of radar opened up a new field of gas discharge applications and studies [95], [96]. In order to use a single antenna for both transmission and reception, a reliable device was needed to protect the receiver from the initial power pulse. This requires a device with an accurately reproducible breakdown voltage and uniformly low impedance to the power pulse and rapid de-ionization of the discharge at the end of the pulse. At first a simple spark gap was used. The next stage of development was to alter a local oscillator vacuum-tube by the addition of gas at a moderately low pressure. This "new" gas tube was called by the British a TR box, or TR tube, for the function transmit-receive. If the device is so placed in the wave guide array to disconnect the transmitter during the receiving period it is called an ATR tube for anti-transmit-receive. The TR and ATR tubes, as components of a waveguide system, developed empirically, and quickly became a vital part of radar systems. These TR tubes and similar devices are now grouped under the general term "duplexer" [97]. Because of the stimulus provided by the war, many workers both in the United States—particularly at Bell Laboratories and at the MIT Radiation Laboratory—and in England studied and modified the TR tube. The author has been unable to determine who made the first one. The long de-ionization time of pure gas discharges was reduced to the order of microseconds by the simple expedient of adding water vapor to the low pressure argon, thus permitting ranging over relatively short distances. The availability since the war of microwave apparatus of considerable power and versatility has made possible a wide range of scientific studies of gas discharges and has helped provide solution for many problems [98].

The present requirements for very long range radar necessitates very high power transmitters. This has required the development of new techniques to extend the power capabilities of the TR tubes. It has been found [97], [99] that tubes filled with quartz crystals, hard glass fragments or hollow ceramic beads, would have about ten times the power capacity of the water vapor tubes. A further factor of 10 is obtained by a "folded" tube [97], [99], in which the active zone is in a closely-spaced concentric cylinder portion of a tube with a dc "keep-alive" discharge in the larger cylindrical portions at one end.

Gas discharge tubes can be adapted to the controlled attenuation of microwaves [100]. Varying the dc current alters the electron density and temperature so that variation in attenuation is effected. These tubes may introduce noise of two types. The first is the low frequency variational effects [72], [73] due to space charge fluctuations, etc., and the second is a pure noise similar to resistance noise which depends on the electron temperature [101]–[103]. The electron temperature may be typical in the range 5000° to 30,000°K. Positive column discharge tubes, as, for example, fluorescent lamps, can be used as stable, broad-band, noise generators for use in making measurements of the noise figure of radio receivers in the microwave region [101].

Rotation of microwaves by discharge tubes has been accomplished by what are known as polarization duplexers. These consist of a series of small discharge tubes placed diametrically across a circular wave guide, each tube being rotated relative to the preceding tube. The gyro magnetic properties of a plasma in a constant magnetic field can also be used to produce rotation of an electromagnetic wave [104]. This is analogous to the Faraday effect for light.

Switch Tubes

Discharge devices such as circuit breakers, switches, etc. are clearly beyond the scope of the present review. However, mention should be made of the vacuum switch. Small vacuum switches [105], [106] have been used for many years as low capacitance devices for antenna switching. The vacuum switch will hold off very high voltages for a small separation of the contacts. When they open ac under load, current continues to flow for one-half cycle, the discharge being an arc supported by metal vapor from the electrodes. After the first current zero, the metal vapor disperses with great speed and the full strength of the gap is attained in a few microseconds. Recently, interest in vacuum switches has been revived with new developments that greatly extend the range of applicability [107].

Conclusion

The study of the physics of gas discharges is very old, and many famous scientific names are to be found in its voluminous literature. The author, in the limited space permitted for this review, has been able to touch on only the more important practical developments in gas discharge tubes. In each development, many interesting scientific facts have been revealed, many of which have been the subject of extensive study. These interesting studies obviously could not be included. However, it may be said safely that in spite of years of intensive study there remain many baffling and unanswered scientific problems in the field of gaseous electronics. As these problems are solved, the field of application of discharge phenomena will continue to expand.

Acknowledgment

The author wishes to acknowledge helpful suggestions made by Dr. J. M. Anderson and H. C. Steiner. The unpublished historical compilations of W. C. White were very useful source material. The triggered spark gap of Fig. 4 was the joint development of H. S. Dunkerley and J. Illingsworth. The author is indebted to H. S. Dunkerley for the figure.

Bibliography

[1] A. A. Bright, Jr., "The Electric-Lamp Industry," The Macmillan Co., New York, N. Y., p. 218 *et seq.*; 1949.
[2] J. P. Gassiot, *Proc. Roy. Soc.* (*London*), vol. 10, p. 36; 1859.
[3] W. DeLaRue, H. W. Müller and W. Spottesiwoode, *Proc. Roy. Soc.* (*London*), vol. 23, p. 356; 1875.
[4] H. A. Miller, "Luminous Tube Lighting," Chemical Publishing Co., Brooklyn, N. Y.; 1946.
[5] D. McFarlan Moore, *Trans. AIEE*, vol. 39, p. 2021; 1920.
[6] M. M. Jamin, *Compt. rend. Acad. Sci., Paris*, vol. 94, p. 1615; June 19, 1882.
[7] L. Arons, *Ann. der Phys.*, vol. 47, p. 767; 1892; vol. 58, p. 73; 1896.
[8] *Electric World and Engineer*, p. 121; January 17, 1903.
[9] A. W. Hull, *Trans. AIEE*, vol. 47, p. 753; 1928.
[10] J. S. Townsend, "Electricity and Gases," Oxford University Press, 1914.
[11] J. D. Cobine, "Gaseous Conductors," Dover Publications, Inc., New York, N. Y., p. 143 *et seq.*; 1958.
[12] G. Francis, "Ionization Phenomena in Gases," Academic Press, Inc., New York, N. Y.; 1960.
[13] S. Flugge, Ed., "Handbuch der Physik," Springer-Verlag, Berlin, vols. 21 and 22 (in English); 1956.
[14] J. D. Cobine, "Gaseous Conductors," Dover Publications, Inc., New York, N. Y.; 1958.
[15] A. von Engel, "Ionized Gases," The Clarendon Press, Oxford, England; 1955.
[16] C. A. Coulomb, "Memoires de L'Academie Royale des Sciences," p. 612; 1785.
[17] H. Geitel, *Phy. Zeit.*, vol. 2, p. 116; 1900.
[18] C. T. R. Wilson, *Proc. Camb. Phil. Soc.*, vol. 11, p. 32; 1900.
[19] J. C. McLennan, *Proc. Roy. Soc.*, vol. 68, p. 151; 1901.
[20] E. Rutherford and H. Geiger, *Proc. Roy. Soc.* (*London*), *A*, vol. 81, p. 141; 1908.
[21] H. Geiger and O. Klemperer, *Z. Phys.*, vol. 49, p. 753; 1928.
[22] W. Bothe and H. Kolhoorster, *Z. Phys.*, vol. 56, p. 571; 1929.
[23] A. Trost, *Z. Phys.*, vol. 105, p. 399; 1937.
[24] O. E. Buckley, *Proc. Natl. Acad. Sci.*, vol. 2, p. 683; 1916.
[25] R. T. Bayard and D. Alpert, *Rev. Sci. Instr.*, vol. 21, p. 571; 1950.
[26] J. M. Lafferty, *J. Appl. Phys.*, vol. 32, p. 424; 1961.
[27] J. R. Downing and G. Mellen, *Rev. Sci. Instr.*, vol. 17, p. 218; 1946.
[28] F. M. Penning and K. Nienhauis, *Philips Tech. Rev.*, vol. 11, p. 116; 1949.
[29] J. M. Lafferty and T. A. Vanderslice, "The interplay of electronics and vacuum technology," Proc. IRE, vol. 49, pp. 1136–1154; July, 1961.
[30] J. M. Gaugain, *Compt. rend. Acad. Sci., Paris*, vol. 40, p. 640; 1855. See also O. Lehmann, *Ann. der Phys.*, vol. 56, p. 304; 1895.
[31] V. Bush and C. G. Smith, *Trans. AIEE*, vol. 41, p. 402; 1922.
[32] M. Bareiss, *Zeit. Tech. Phys.*, vol. 8, p. 449; 1927.
[33] A. E. Shaw, "Cold cathode rectification," Proc. IRE, vol. 17, pp. 849–863, May, 1929; D. D. Knowles, *Elec. J.*, vol. 27, pp. 116, 232, 1930.
[34] H. M. Ferree, *Trans. AIEE*, vol 60, p. 8; 1940.
[35] F. V. Hunt and R. W. Hickman, *Rev. Sci. Ins r.*, vol. 10, p. 6; 1939.
[36] J. M. Lafferty, "A process for making clean gas discharge tubes," IRE Trans. on Electron Devices, vol. ED-5, pp. 143–147; July, 1958.
[37] H. J. Reich and W. A. Depp, *J. Appl. Phys.*, vol. 9, p. 421; 1938.
[38] C. Van Geel, *Physics*, vol. 6, p. 806; 1939.
[39] G. M. Kirkpatrick, "Characteristics of certain voltage-regulator tubes," Proc. IRE, vol. 35, pp. 485–489; May, 1947.
[40] F. A. Benson and L. J. Bental, *Wireless Engr.*, vol. 32, p. 330; 1955.
[41] D. D. Knowles and S. P. Sashoff, *Electronics*, p. 183; July, 1930.
[42] S. B. Ingram, *Trans. AIEE*, vol. 58, p. 342, 1939.
[43] J. J. Lamb and J. A. Brustman, *Electronics*, p. 92; November, 1949.
[44] R. C. Bacon and J. R. Pollard, *Elec. Engrg.*, p. 173; May, 1950.
[45] M. A. Townsend, *Elec. Engrg.*, vol. 69, p. 810; 1950.
[46] T. P. J. Botden, *Philips Tech. Rev.*, vol. 21, p. 267; 1959–1960.
[47] C. H. Tosswill, *Philips Tech. Rev.*, vol. 18, p. 128; 1956–1957.
[48] A. H. Beck and T. M. Jackson, *Elec. Commun.*, vol. 35, p. 251; 1959.
[49] J. Nienhold, German Patent Nos. 319,803, December 7, 1916; 331,029, April 25, 1919; 345,921, January 18, 1920; and 345,276, November 30, 1919. Also see A. Hund, *Electronics*, vol. 6, p. 6; 1933.
[50] J. M. Anderson and L. A. Harris, *J. Appl. Phys.*, vol. 31, p. 1463; 1960.
[51] H. C. Steiner, *Electronics*, p. 65; April 29, 1960.
[52] G. W. Pierce, U. S. Patent No. 1,127,371; February 2, 1915.
[53] I. Langmuir, U. S. Patent No. 1,289,823; December 21, 1918.

[54] I. Langmuir, *Science*, vol. 58, p. 290; 1923.
[55] I. Langmuir, *GE Rev.*, vol. 26, p. 731; 1923.
[56] I. Langmuir and H. Mott Smith, *GE Rev.*, vol. 27, pp. 449, 538, 616, 762, 810; 1924.
[57] A. W. Hull, *Trans. AIEE*, vol. 47, p. 753; 1928.
[58] A. W. Hull and I. Langmuir, *Proc. Natl. Acad. Sciences, U. S.*, vol. 51, p. 218; 1929.
[59] A. W. Hull, *GE Rev.*, vol. 32, pp. 213, 390; 1929.
[60] H. de B. Knight, *Proc. IEE*, vol. 96, p. 361; 1949.
[61] K. J. Germeshausen, "Pulse Generators," G. N. Glasoe and J. V. Lebacoz, Eds., McGraw-Hill Book Co., Inc., New York, N. Y., p. 335; 1948.
[62] H. H. Wittenberg, *RCA Rev.*, vol. 10, p. 116; 1949.
[63] H. N. Price and A. W. Coolidge, *Trans. AIEE*, vol. 78, p. 76; 1959.
[64] W. J. Kearns, Conference Paper, AIEE General Meeting; January 31–February 5, 1960.
[65] A. W. Hull, *Trans. AIEE*, vol. 50, p. 744; 1931.
[66] A. W. Hull, *Trans. AIME*, vol. 102, p. 1; 1932.
[67] A. W. Hull, *Physics*, vol. 4, p. 66; 1933.
[68] A. W. Hull, E. E. Burger, and R. E. Turrentine, *GE Rev.*, vol. 54, p. 16; 1951.
[69] E. V. Appleton and A. G. D. West, *Phil. Mag.*, vol. 45, p. 879; 1923.
[70] C. Eckart and K. T. Compton, *Phys. Rev.*, vol. 24, p. 94; 1924.
[71] L. Tonks and I. Langmuir, *Phys. Rev.*, vol. 33, p. 195; 1929.
[72] J. D. Cobine and C. J. Gallagher, *J. Franklin Inst.*, vol. 243, p. 41; 1947.
[73] J. D. Cobine and C. J. Gallagher, *J. Appl. Phys.*, vol. 18, p. 110; 1947.
[74] J. D. Cobine and J. B. Curry, "Electrical noise generators," Proc. IRE, vol. 35, pp. 875–879; September, 1947.
[75] J. A. Olmstead, W. M. Webster and E. O. Johnson, "The tacitron, a low noise thyratron capable of current interruption by grid action," Proc. IRE, vol. 42, pp. 1350–1362; September, 1954.
[76] L. Malter and E. O. Johnson, *RCA Rev.*, vol. 11, p. 165; 1950.
[77] L. Malter, E. O. Johnson and W. M. Webster, *RCA Rev.*, vol. 12, p. 415; 1951.
[78] E. O. Johnson, *Electronics*, p. 107; May, 1951.
[79] E. O. Johnson and W. M. Webster, "The plasmatron, a continuously controllable gas-discharge developmental tube," Proc. IRE, vol. 40, pp. 645–659; June, 1952.
[80] C. G. Found, *Phys. Rev.*, vol. 45, p. 519; 1934.
[81] C. G. Found, U. S. Patent No. 2,034,571; March 17, 1936.
[82] J. D. LeVan and P. T. Weeks, "A new type of gas-filled amplifier tube," Proc. IRE, vol. 24, pp. 180–189; February, 1936. See also H. G. Boumeester and M. J. Druyvesteyn, *Philips Tech. Rev.*, vol. 1, p. 367; 1936.
[83] W. Schlicter, Ph.D. dissertation, University of Göttingen, Germany, pp. 85–98; 1915.
[84] G. Medicus and G. Wehner, *J. Appl. Phys.*, vol. 22, p. 1389; 1951.
[85] K. G. Hernquist, M. Kanefsky and F. H. Norman, *RCA Rev.*, vol. 19, p. 244; 1958.
[86] V. C. Wilson, *J. Appl. Phys.*, vol. 30, p. 475; 1959. See also J. M. Houston, AIEE Elect. Engrg. Division-Work Shop, University of Kentucky, Lexington, Ky.; June 29–30, 1961.
[87] A. W. Hull and H. D. Brown, *Trans. AIEE*, vol. 50, p. 744; 1931.
[88] J. Slepian and L. R. Ludwig, *Trans. AIEE*, vol. 52, p. 643; 1933.
[89] G. H. Reiling, *Trans. AIEE*, Paper No. 61-731.
[90] G. H. Reiling, *Bull. Am. Phys. Soc.*, Series II, vol. 5, p. 371; June 15, 1960.
[91] K. J. Germishausen, *Phys. Rev.*, vol. 55, p. 228; 1939.
[92] O. Lodge, *Nature*, vol. 41, p. 462; 1890.
[93] J. J. Thompson, *Phil. Mag.*, vol. 32, pp. 32, 445; 1891.
[94] N. Tesla, *Elec. Engrg.*, vol. 7, p. 549; 1891.
[95] L. D. Smullin and C. G. Montgomery, Eds., "Microwave Duplexers," McGraw-Hill Book Co., Inc., New York, N. Y., p. 139; 1948.
[96] A. L. Samuel, J. W. Clark and W. W. Mumford, "Radar Systems and Components," Bell Laboratory's Staff, D. Van Nostrand Co., Inc., New York, N. Y., p. 310; 1949.
[97] For a general review of duplexing tubes and systems see A. F. Harvey, "Duplexing systems at microwave frequencies," IRE Trans. on Microwave Theory and Techniques, vol. MTT-8, pp. 415–431; July, 1960.
[98] S. C. Brown, "Handbuch der Physik," XXII, pp. 531–575, Springer-Verlag, Berlin, 1956.
[99] H. C. Alexander and A. G. Barrett, *Microwave J.*, vol. 4, p. 85; 1961.
[100] E. M. Bradley and D. H. Pringle, Jr., *Brit. Inst. Radio Engrs.*, vol. 15, p. 11; 1955.
[101] W. W. Mumford, *Bell Sys. Tech. J.*, vol. 28, p. 608; 1949.
[102] M. A. Easley and W. W. Mumford, *J. Appl. Phys.*, vol. 22, p. 846; 1951.
[103] L. Parzen and L. Goldstein, *Phys. Rev.*, vol. 79, p. 190; 1950.
[104] L. Goldstein, M. Lampert and J. Heney, *Phys. Rev.*, vol. 82, p. 956; 1951; vol. 83, p. 1255; 1951.
[105] A. Kling, *GE Rev.*, vol. 38, p. 525; 1935.
[106] W. Teare, *GE Rev.*, vol. 41, p. 280; 1938.
[107] J. D. Cobine, *Elec. Engrg.*, vol. 81; January, 1962.

History of the Microwave-Tube Art*

J. R. PIERCE†, FELLOW, IRE

Summary—**Many ideas important to microwave tubes, including the interaction of electrons with waves, occurred very early. An orderly understanding came much later. The RF diode equation (1933–1944) gave quantitative results concerning microwave tubes rather than insight. The klystron and the concept of velocity modulation and phase focusing (1938–1939) led to a clear picture of important phenomena as well as to powerful microwave devices. Insight and performance were extended by the traveling-wave tube and the wave analysis of its behavior (1946–1947). This led to explorations in terms of the behavior of coupled circuits (1951, 1954) and negative-energy waves (1951). Backward-wave oscillators provided further understanding and a new tool (1952–1953). This background of understanding which grew up concerning microwave tubes led to analysis of and improvement in noise behavior (1950–1959). Finally, the invention of the parametric amplifier led to very-low-noise microwave tubes (1959–1961). Crossed-field tubes have always been highly effective but difficult to analyze and design. Nonetheless, useful high-power amplifiers and oscillators have been built.**

* Received by the IRE, October 6, 1961; revised manuscript received, December 8, 1961.
† Bell Telephone Laboratories, Inc., Murray Hill, N. Y.

IT IS VERY HARD to give any satisfactory account of discovery, invention and development in the field of microwave tubes during the last half century. The period began close to the first effective use of triodes, and it closes with a current art which is almost beyond the power of description.

It is particularly difficult to give a proper account of early work on Barkhausen-Kurz tubes and on magnetrons. By 1931 the Barkhausen-Kurz tube [1] had enabled Clavier to demonstrate microwave telephone transmission across the English Channel [2]. In 1936 Kilgore had produced a CW power of 100 w at a frequency of 600 Mc by using a magnetron [3], and Cleeton and Williams had used magnetrons to produce small CW powers at wavelengths as short as 6.4 mm [4]. Thus, microwave tubes had a considerable record of accomplishment at a very early date.

Too, the variety of early microwave devices is astonishing. Llewellyn's negative-resistance critical-transit-time diode goes back to 1937 [5]. Wave circuits in connection with electron streams were used by Haeff [6] prior to 1936, and were proposed by Lindenblad [7]. Other uses of wave circuits are shown in a patent granted to Llewellyn late in 1940 [8]. In 1935 A. A. Heil and O. Heil described a tube which we would now consider a special form of klystron [9].

If we were to give a historical account of all of this progress in terms appropriate to its day, the account would be very hard to follow. Early workers sometimes did not fully understand the operation of devices they used. Sometimes they had valid but special theories of their operation. Today, we can look back and understand early devices easily, but in doing this we are necessarily led into a misinterpretation of the past. Still, in an article of reasonable length the sort of intellectual organization which has now come into being affords the only reasonable framework for discussion, and we will have to follow it.

In retrospect, then, we see that there have been several extremely important advances in the understanding of microwave electronic phenomena. Each can be associated with certain inventions, but the influence of these advances in understanding has gone beyond the inventions involved, illuminating past work, and making further work possible. Thus, certain areas of understanding afford the best means for describing and assessing progress in the field of microwave tubes during the last half century.

Diode Equations

The exact and useful formulation of the high-frequency electron flow between parallel grids was a vital step toward our understanding of microwave tubes. These equations were worked out in various forms by Müller [10], North [11], Benham [12], Llewellyn [13] and Llewellyn and Peterson [14]. Their original purpose was to provide an understanding of the high-frequency operation of triodes, tetrodes and pentodes, but their consequences were much greater. They led Llewellyn to the negative-resistance diode. They formed and still form part of the stock in trade of men who work with all sorts of microwave tubes.

The weakness of these equations, which give the RF properties of electron flow between parallel grids, is not so much their specialization as their failure to furnish physical insight. They describe phenomena which we would now speak of in terms of velocity modulation and wave phenomena; indeed, Gray derived the equations in a waveform [15]. But early workers did not recognize these features in the equations. To early workers, the equations provided more a mathematical tool than a deep understanding.

Production and Focusing of Electron Beams

Klystrons, traveling-wave tubes and some other microwave devices make use of long electron beams of considerably higher current than is usual in electron optical systems. The successful production and focusing of such electron beams has been extremely important to the success of microwave tubes. Some important advances were a rational method of gun design [16], "Brillouin" flow [17], [18] and periodic focusing [19].

Klystrons and Velocity Modulation

Perhaps the first great step in understanding the phenomena in microwave tubes came with the invention of the klystron. Brüche and Recknagel discussed "phase focusing" in 1938 [20] and the work of the Varians [21], Webster's theoretical treatment of the klystron [22], and the work of Hahn and Metcalf [23] were published in 1939.

With the klystron came a well-thought-out theory of its operation, the concept of velocity modulation, and a full appreciation of the value of microwave resonators.

With the klystron too, there came from contemporary workers, and notably from Hahn [24] and Ramo [25], the idea of space-charge waves and their use in explaining the operation of velocity-modulated tubes. But the chief impact of the wave phenomena of electron flow was to be in another field—that of traveling-wave devices.

The idea of velocity modulation is powerful because it gives a direct physical insight into the operation of microwave devices. If the electrons of a stream are alternately speeded up and slowed down prior to their injection in a drift tube, the faster electrons will catch up with the slower electrons, and the beam will become bunched, so that it can produce power in an output resonator through which it passes. Similarly, when a velocity-modulated beam is injected into a retarding field in the reflex klystron, the faster electrons go further and return later, and so the beam is bunched through a somewhat different mechanism.

Although klystrons produced powers as great as 50 w by 1939, and EMI klystrons [26] produced pulsed powers of around 30 kw during World War II, as power sources they were overshadowed by magnetrons during that war. They were, however, widely used as local oscillators in radar receivers. Particularly, it was found to be possible to build reflex klystrons for 300-v operation, using glass with copper disk seals and an external resonator at wavelengths around 10 cm, and an all metal construction around 3 cm [27].

Such low-voltage tubes had high-perveance low-impedance beams, and changing the reflector voltage changed the frequency of oscillation by tens of megacycles. This electronic tuning has proved to be an extremely valuable feature of reflex klystrons. They are now widely used, not only as local oscillators, but as sources of frequency-modulated waves in microwave transmitters, and, of course, as convenient sources in test equipment. Today, reflex klystrons are available in the power range from milliwatts to tens of watts, and for wavelengths ranging from a few millimeters to the long-wavelength end of the microwave band, wherever that may be.

While magnetrons beat the klystron out as high-power sources during World War II, klystrons have subsequently provided the highest microwave powers to be generated. In the klystron there is a clean separation of functions: gun, buncher, drift tube, catcher and collector. Rational and optimum design of each part is possible. The greatest early progress was through work at Stanford by Hansen, Ginzton, Chodorow and others, in connection with the construction of an electron linear accelerator. By the end of March, 1949, a pulsed power of 14 Mw with an average power of 1 kw had been produced at a wavelength of 10 cm.

The subsequent development and manufacture of high-power klystron amplifiers have provided powerful tools for all sorts of microwave applications including accelerators, electronic heating, radar and space research and communication. The fact that the klystron can act as a stable high-power *amplifier* has made feasible such sophisticated radar techniques as chirp.

Table I gives a rough idea of the sorts of gains and powers that are commercially available in klystrons today.

interact powerfully with a beam of electrons if the wave velocity and the electron velocity were nearly the same. He built a tube and found a gain of around 10 db for a wave traveling in the same direction as the electrons. A mathematical analysis agreed with this performance.

The traveling-wave tube turned out to be a device which amplifies over unprecedently broad bands, and which can function over an astonishingly wide range of frequencies and powers. A traveling-wave tube with a helix circuit can amplify over a frequency range of more than an octave. More typically, bandwidth is limited to 500–1000 Mc by input and output couplers.

Understanding and analysis of the phenomena of traveling-wave tubes were made easier when Pierce expressed their low-level behavior in terms of waves on a medium consisting of the electromagnetic circuit and the electron beam [29]. For example, this made it easy to investigate the effect of circuit loss.

The first traveling-wave tubes had considerable gain in the forward direction and negligible loss in the backward direction; reflections at source and load led to oscillations above a moderate gain. By means of a wave analysis, it was easy to show that the forward gain would be reduced by only about a third of the amount of loss, measured in db, which was introduced into the circuit. Thus, it was practical to make a truly stable tube with a backward loss greater than the forward gain. Lumped loss extending for a fraction of the length of the tube has led to even superior results.

While the wave theory of the traveling-wave tube made calculations easy, a simple intuitive understanding of traveling-wave devices required further insight. This has been provided in two ways. In 1951 Chu pointed out [30] that the slower of the two space-charge or plasma waves of an electron stream has a negative energy [31], in the sense that energy is taken from the electron stream in setting up the slow wave. Hence, the increasing wave in a traveling-wave tube comes about through a continual increase in excitation and hence in energy of the electromagnetic part of the wave, and a continual increase in excitation and hence decreases in energy of the space-charge or plasma part of the wave.

TABLE I

KLYSTRONS

Frequency kMc	Peak Power Mw	Average Power kw	Efficiency Per cent	Band-width Mc	Gain db
0.4	8	30	40	50	30
1.3	25	80	36	100	40
3	22	22	45	15	60
8	CW	20	32	29	44

Traveling-Wave Tubes and Wave Phenomena

We have noted that progressive waves and electron flow occurred in some of the early thinking about microwave tubes. It remained, however, for Kompfner to take the decisive step of reasoned approach and effective experiment which gave us the traveling-wave tube and led to a host of related devices [28].

Kompfner reasoned that an electromagnetic wave on a slow-wave structure, and especially a helix, should

This concept led naturally to a consideration of the coupling between electromagnetic waves and plasma waves [32], a subject which Mathews [33] had investigated at an early date. We now understand the behavior of many electronic devices in terms of the coupling between slow negative-energy space-charge waves and various environments. Near a lossy medium such a wave grows by losing its energy to the environment (resistive-wall amplifier) [34]. Periodic variations along a beam can give gain through coupling a fast positive-energy space-charge wave to a slow negative-energy space-charge wave (velocity-jump [35] and scalloped-stream [36] amplifications). Or, gain can be obtained by coupling the fast space-charge wave of one electron stream to the slow space-charge wave of a nearly or interpenetrating stream (double-stream amplification) [37]–[39].

While wave phenomena are important in explaining the low-level operation of traveling-wave tubes, their large-signal behavior can be understood chiefly through digital computation [40]–[42] and through experiment [43], although Paschke has made progress toward an analytical theory [44]. Actually to make effective high-power traveling-wave tubes required a great deal of work on circuitry. Particularly, satisfactory iterated resonator-like structures had to be used instead of the simple helix, and problems of unwanted modes and spatial frequencies and of matching and terminations had to be solved. Typically, the bandwidths of tubes with such circuits is around 10 per cent. In addition, many advances had to be made in techniques of construction.

In the early 1950's traveling-wave tubes giving a few watts power were in use in British and Japanese microwave radio relay systems. Today they are used in microwave radio relay in the United States and in Continental Europe.

Today, high-power traveling-wave tubes are available for a number of commercial and military purposes. Table II gives some ideal of what is commercially available.

TABLE II
TRAVELING-WAVE TUBES

Frequency kMc	Peak Power Mw	Average Power kw	Efficiency Per cent	Gain db
2.8	1.5	6	23	30
5.7	3	5	25	35
9	0.1	500	20	54
50–60	CW	1 w	5	40

Backward-Wave Oscillators (O-Type Carcinotrons)

In 1951 Millman [45] reported oscillations in a traveling-wave-type tube which were due to an interaction between the electrons and a spatial harmonic of the circuit wave with a phase velocity opposite to the direction of energy flow in the circuit. Work by Guenard, Doehler, Epszstein and Warnecke [46] and Kompfner [47] deliberately exploited such interactions to give oscillators of very wide voltage-controlled tuning range, sometimes of several fold. Backward-wave oscillators are particularly useful in test equipment, and they provide good millimeter-wave sources. Table III indicates the sort of performance that is commercially available.

TABLE III
BACKWARD-WAVE OSCILLATORS

Frequency kMc	Power mw
2.0– 4.0	50
8.2–12.4	50
12.4–18	65
23.5–37.5	250
27 –40	5

Low-Noise Tubes

All good low-noise microwave tubes, except parametric amplifiers, are of the traveling-wave type. Theoretically, low-noise klystrons could be made, but practically, traveling-wave tubes are what we have. This is partly because of the convenience and stability of an untuned broad-band structure. Partly, it is because of the strong interactions which can be attained without lossy, high-impedance circuits. Partly, it may be historical accident.

In inventing the traveling-wave tube, Kompfner was seeking a low-noise amplifier. The attainment of useful low-noise traveling-wave tubes, however, came about only with increased theoretical understanding. The understanding came in several steps.

The first was a wave analysis of the behavior of the traveling-wave tube [29].

The second step was an experimental demonstration by Cutler and Quate of the fact that the noise in an electron stream propagates as waves, just as a signal impressed on the stream would [48].

The third step was the reduction of the effect of noise due to velocity fluctuations at the cathode by using the velocity jump technique to reduce the amplitude of the noise waves on the electron beam. By this means Watkins [49] and Peter [50] attained noise figures around 10 db.

The fourth step was a realization, foreshadowed by Watkins, that there are two independent sources of noise, associated with current and velocity fluctuations at the cathode. This led to theorems indicating a minimum noisiness of an electron beam from a cathode of a given temperature [51], and a corresponding minimum noise figure for *any* beam-type amplifier [52].

The fifth step was the experimental and theoretical demonstration by Siegman, Watkins and Hsieh that in space-charge-limited flow the noisiness may be less than the supposed minimum [53]. Later, Currie and Forster showed how a very-low-noise beam can be attained over a broad band of frequencies [54].

If a backward-wave oscillator structure, and especially a backward-wave oscillator structure with a two-section circuit, is operated at a current lower than required for oscillation, and if a signal is applied at the downstream end of the circuit, the device acts as a narrow-band voltage-tuned backward-wave amplifier. Low-noise backward-wave amplifiers are commercially available.

Table IV shows the sort of noise performance attained in commercially available traveling-wave tubes and backward-wave amplifiers.

TABLE IV
LOW-NOISE TUBES

Type	Frequency kMc	Noise Figure db	Gain
TWT	2.6– 3.5	4	26
TWT	5.25– 5.75	4.5	22
TWT	10.7–11.7	6	35
BWA	2.4– 4.0	5	20
BWA	8.5– 9.6	5.5	20

With the coming of the parametric amplifier, Adler devised a transverse-field electron-beam parametric amplifier and attained a noise figure of 1 db at a frequency between 400 and 800 Mc [55], [56]. Further work elsewhere has led to a noise figure of 0.8 db at a frequency of 4140 Mc [57]. The Cuccia coupler [58], which at once removes noise from the electron beam and transfers signal to it, has been an important contribution to such low-noise parametric amplifiers.

CROSS-FIELD TUBES

The magnetron has a venerable history, dating back to ideas of Hull as early as 1921 [59]. There was much early work of considerable merit [60]–[62]. The idea of traveling-wave operation, a condition for it, and the considerable efficiencies which could be attained were discussed by Posthumus [63] in 1935. Structures of a multicavity type were described in 1936 [64] and 1940 [65]. Yet, the practical development and use of the magnetron occurred suddenly during World War II, following the production of the first high-power high-efficiency pulsed multicavity magnetrons by Randall and Boot [65].

The ability of the cavity magnetron to produce the short intense pulses required for radar depended on a hitherto unsuspected ability of oxide-coated cathodes to emit currents of many amperes per square centimeter in pulses of short duration.

The crossed-field configuration has a tremendous practical advantage. Electrons can give up energy to the RF circuit and yet stay in step with electromagnetic wave; as they lose energy to the RF field they gain it from the dc field.

The performance of microwave tubes making use of electron beams, with or without longitudinal magnetic focusing fields, can be analyzed in some detail and with some accuracy, so that performance accords well with design calculations. Thus, progress in the klystron and traveling-wave-tube art has exemplified a strong interaction between experiment and rather detailed theory. Crossed-field tubes are effective but very complicated in their operation. This is particularly true of oscillators. Thus, in the heyday of the magnetron workers were guided by a few theoretical considerations, such as the Hartree line, which relates operating voltage to magnetic field and mode number by appropriate means for expressing the operation, as the Rieke diagram, in which power and frequency are plotted against load impedance by cold circuit measurements and by endless building and measurement in the tradition of experimental physics. Happily, there are two fine records of this wartime work on magnetrons [67], [68].

Wartime work saw a rapid rise in the power and efficiency of magnetrons, so that by the end of the war powers of over a megawatt had been attained, efficiencies were well above 50 per cent and operation with tens of kilowatts of power had been attained at a wavelength of 1.25 cm. Further, tunable tubes had been produced, chiefly through pins which moved into and out of resonators.

The chief problem encountered with the multicavity magnetron was moding. Once this was understood to be a natural phenomenon in a system of coupled resonators, two remedies were found. One was strapping, invented by Sayers [69] and Spencer [70], that is, tying various anode elements together with short leads. Strapping separated the modes in frequency. Through strap breaks, degeneracy was overcome, so that all modes were coupled to the output. Besides this, strapping distorted the field pattern of undesired modes.

In the rising sun structure mode separation was attained by using alternate resonator slots of different depth or frequency. Some Japanese wartime magnetrons had rising sun structures.

Postwar advances have led to more highly developed magnetron oscillators of improved performance. One notable new advance in tuning and stability has been the CEM (circular electric mode) magnetron [71], in which a tunable high-Q cavity with a circular electric excitation is ingeniously coupled to the interaction slots. This gives a tube with a high degree of frequency stability.

Another tube in which high stability has been achieved is the stabilotron, in which the circuit is interrupted and the gap bridged by coupling through an external cavity [72].

Attempts to make wave-type magnetron amplifiers date back almost as far as the traveling-wave tube [73]. Early experimenters had many rude shocks awaiting them. It is hard to control electron flow in crossed electric and magnetic fields. Such flow can be inherently unstable [74]. Early attempts to make crossed-field amplifiers led to oscillation and noise but not to gain.

It has been possible to obtain substantial (10–13-db) though not high-power gain in the amplitron [72], which is essentially a cylindrical magnetron with an interrupted circuit.

The idea of a crossed-field device with a linear circuit has remained attractive. Perhaps the most useful devices of this kind are voltage-tuned oscillators called M-type carcinotrons [46]; these are avilable at a range of frequencies for CW powers up to hundreds of watts.

Table V shows what sort of performance is commercially available in crossed-field tubes.

TABLE V
CROSSED-FIELD TUBES

Type	Frequency kMc	Peak Power Mw	Average Power kw	Efficiency Per cent	Gain db
Magnetron	1.3	10	17	30	—
Magnetron	2.8	4.5	4.5	50	—
Magnetron	9.4	1	0.875	46	—
Amplitron	2.9– 3.1	3	15	70	8
M Carcinotron	2.5– 3.3	CW	0.27	—	—
M Carcinotron	8.5–11.0	CW	0.20	—	—

REFERENCES

[1] H. Barkhausen and K. Kurz, "The shortest waves obtainable with vacuum tubes," *Z. Phys.*, vol. 21, pp. 1–6; 1920.
[2] "Radio communication very short waves," *Engineering*, vol. 131, pp., 402–483; April 10, 1931.
[3] G. R. Kilgore, "Magnetron oscillators for frequencies from 300–600 mc/s," PROC. IRE, vol. 24, pp. 1140–1157; August, 1936.
[4] C. E. Cleeton and N. H. Williams, *Phys. Rev.*, vol. 50, p. 1091; December 1, 1936.
[5] F. B. Llewellyn, "Diode Oscillator," U. S. Patent No. 2,190,668, filed July 31, 1937, issued February 20, 1940.
[6] A. V. Haeff, "Device for and Method of Controlling High Frequency Currents," U. S. Patent No. 2,064,496, filed October 23, 1933, issued December 15, 1936; "Device for and Method of Controlling High Frequency Currents," U. S. Patent No. 2,233,126, filed October 23, 1933, issued February 25, 1941.
[7] N. E. Lindenblad, "Electron Discharge Device," U. S. Patent No. 2,300,052, filed, May 4, 1940, issued October 27, 1942.
[8] F. B. Llewellyn, "Electron Discharge Apparatus," U. S. Patent No. 2,367,295, filed May 17, 1940, issued January 16, 1945.
[9] A. A. Heil and O. Heil, "A new method for the generation of short undamped electromagnetic waves of high intensity," *Z. Phys.*, vol. 95, pp. 752–773; July 12, 1935.
[10] J. Müller, "Electron oscillations in high vacuum," *Hochfrequenztech. u. Elektroakust.*, vol. 41, pp. 156–157; May, 1933.
[11] D. O. North, "Analysis of the effects of space charge on grid impedance," PROC. IRE, vol. 24, pp. 108–136; January, 1936.
[12] W. E. Benham, "A contribution to tube and amplifier theory," PROC. IRE, vol. 26, pp. 1093–1170; September, 1938.
[13] F. B. Llewellyn, "Electron Inertia Effects," Cambridge University Press, Cambridge, England; 1939.
[14] F. B. Llewellyn and L. C. Peterson, "Vacuum-tube networks," PROC. IRE, vol. 32, pp. 144–166; March, 1944.
[15] F. Gray, "Electron streams in a diode," *Bell Sys. Tech. J.*, vol. 30, pp. 830–845; October, 1951.
[16] J. R. Pierce, "Rectilinear flow in electron beams," *J. Appl. Phys.*, vol. 11, pp. 548–554; August, 1940.
[17] L. Brillouin, "A theorem of Larmor and its importance for electrons in magnetic fields," *Phys. Rev.*, vol. 67, pp. 260–266; April 1 and 15, 1945.
[18] C. C. Wang, "Electron beams in axially symmetrical electric and magnetic fields," PROC. IRE, vol. 38, pp. 135–147; February, 1950.
[19] J. R. Pierce, "Spatially alternating magnetic fields for focusing low-voltage electron beams," *J. Appl. Phys.*, vol. 24, p. 1247; September, 1953.
[20] E. Brüche and A. Recknagel, "On phase focusing of electrons moving in rapidly varying fields," *Z. Phys.*, vol. 108, pp. 459–482; March, 1938.
[21] R. H. Varian and S. F. Varian, "High-frequency oscillator and amplifier." *J. Appl. Phys.*, vol. 10, pp. 321–327; May, 1939.
[22] D. L. Webster, "Cathode ray bunching," *J. Appl. Phys.*, vol. 10, pp. 501–508; July, 1939.
[23] W. C. Hahn and G. F. Metcalf, "Velocity-modulated tubes," PROC. IRE, vol. 27, pp. 106–116; February, 1939.

[24] W. C. Hahn, "Small signal theory of velocity-modulated electron beams," *GE Record*, vol. 42, pp. 258–270; June, 1939.
[25] S. Ramo, "Space charge and field waves along an electron beam," *Phys. Rev.*, vol. 56, pp. 276–283; August, 1939.
[26] L. F. Broadway, C. J. Milner, D. R. Petrie, W. J. Scott and C. P. Wright, "Velocity-modulated valves," *J. IEE* (*London*), pt. 3 A, vol. 93, pp. 855–867; 1946.
[27] J. R. Pierce and W. G. Shepherd, "Reflex oscillators," *Bell Sys. Tech. J.*, vol. 26, pp. 460–681; July, 1947.
[28] R. Kompfner, "The traveling wave valve," *Wireless World*, vol. 52, pp. 369–379; November, 1946.
[29] J. R. Pierce, "Theory of the beam-type traveling-wave tube," Proc. IRE, vol. 35, pp. 111–123; January, 1947.
[30] L. J. Chu, presented at IRE Electron Devices Conference, University of New Hampshire, Durham, N. H.; June, 1951.
[31] W. H. Louisell and J. R. Pierce, "Power flow in electron beam devices," Proc. IRE, vol. 43, pp. 425–427; April, 1955.
[32] J. R. Pierce, "Coupling of modes of propagation," *J. Appl. Phys.*, vol. 25, pp. 179–183; February, 1954.
[33] W. E. Mathews, "Traveling-wave amplification by means of coupled transmission lines," Proc. IRE, vol. 39, pp. 1044–1051; September, 1951.
[34] C. K. Birdsall and J. R. Whinnery, "Waves in an electron stream with general admittance walls," *J. Appl. Phys.*, vol. 24, pp. 314–323; March, 1953.
[35] L. M. Field, P. K. Tien and D. A. Watkins, "Amplification by acceleration of a single-velocity stream," Proc. IRE (Correspondence), vol. 39, p. 194; February, 1951.
[36] T. G. Mihran, "Scalloped beam amplification," IRE Trans. on Electron Devices, vol. ED-3, pp. 3239; January, 1956.
[37] L. S. Nergaard, "Analysis of a simple model of a two-beam growing wave tube," *RCA Rev.*, vol. 9, pp. 585–601; December, 1948.
[38] Andrew V. Haeff, "The electron-wave tube," Proc. IRE, vol. 37, pp. 4–10; January, 1949.
[39] J. R. Pierce and W. B. Hebenstreit, "A new type of high-frequency amplifier," *Bell Sys. Tech. J.*, vol. 28, pp. 33–51; January, 1949.
[40] A. Nordsieck, "Theory of the large-signal behavior of traveling-wave amplifiers," Proc. IRE, vol. 41, pp. 630–637; May, 1953.
[41] P. K. Tien, L. R. Walker and V. M. Wolontis, "A large signal theory of traveling-wave amplifiers," Proc. IRE, vol. 43, pp. 260–277; March, 1955.
[42] J. E. Rowe, "Design information large-signal traveling-wave amplifiers," Proc. IRE, vol. 44, pp. 200–210; February, 1956.
[43] C. C. Cutler, "The nature of power saturation in traveling-wave tubes," *Bell Sys. Tech. J.*, vol. 35, pp. 841–876; July, 1956.
[44] Fritz Paschke, "Mikrowellen," Vortrage den Internationalen Tagung in Munchen, F. Vieweg und Sohn, Braunschweig, Germany; 1961.
[45] S. Millman, "A spatial-harmonic traveling-wave amplifier for six millimeter wavelength," Proc. IRE, vol. 39, pp. 1035–1043; September, 1951.
[46] P. Guenard, O. Doehler, B. Epszstein and R. Warnecke, "New oscillator tubes with wide electronic tuning band for ultra-high frequencies," *Compt. rend. Acad. Sci., Paris*, vol. 235, pp. 236–237; 1952.
[47] R. Kompfner and N. T. Williams, "Backward-wave tubes," Proc. IRE, vol. 51, pp. 1602–1611; November, 1953.
[48] C. C. Cutler and C. F. Quate, "Experimental verification of space-charge and transit-time reduction of noise in electron beams," *Phys. Rev.*, vol. 80, pp. 875–878; December 1, 1950.
[49] D. A. Watkins, "Traveling-wave tube noise figure," Proc. IRE, vol. 40, pp. 65–70; January, 1952.
[50] R. W. Peter, "Low-noise traveling-wave amplifier," *RCA Rev.*, vol. 13, pp. 344–368; September, 1952.
[51] J. R. Pierce, "A theorem concerning noise in electron streams," *J. Appl. Phys.*, vol. 25, pp. 931–933; August, 1954.
[52] H. A. Haus and F. N. H. Robinson, "The minimum noise figure of microwave amplifiers," Proc. IRE, vol. 43, pp. 981–991; August, 1955.
[53] G. E. Siegman, D. A. Watkins and H.-C. Hsieh, "Density-function calculations of noise propagation on an accelerated multivelocity electron beam," *J. Appl. Phys.*, vol. 28, pp. 1138–1148; October, 1957.
[54] M. R. Currie and D. F. Forster, "New mechanism of noise reduction in electron beams," *J. Appl. Phys.*, vol. 30, pp. 94–103; January, 1959.
[55] R. Adler, G. Hrbek and G. Wade, "The quadrupole amplifier, a low-noise parametric device," Proc. IRE, vol. 47, pp. 1713–1723; October, 1959.
[56] C. P. Lea-Wilson, R. Adler, G. Hrbek and G. Wade, "Some possible causes of noise in Adler tubes," Proc. IRE, vol. 48, pp. 255–256; February, 1960.
[57] A. Ashkin, "A low-noise microwave quadrupole amplifier," Proc. IRE, vol. 49, pp. 1016–1020; June, 1961.
[58] C. L. Cuccia, "The electron coupler: a development tube for amplitude modulation and power control at ultra-high frequencies," *RCA Rev.*, vol. 10, pp. 270–303; June, 1949.
[59] A. W. Hull, "The effect of a uniform magnetic field on the motion of electrons between coaxial cylinders," *Phys. Rev.*, vol. 18, pp. 31–57; July, 1921.
[60] E. Habann, "New generator tube," *Z. Hochfrequenz*, vol. 24, No. 5, pp. 115–120, No. 6, pp. 135–141; 1924.
[61] A. Zacek, "A method for generating short electromagnetic waves," *Casopis pro Pestovani Mathematiky a Fysiky* (*Prague*), vol. 53, p. 378; 1924.
[62] Kinjiro Okabe," On the short-wave limit of magnetron oscillators," Proc. IRE, vol. 17, pp. 652–659; April, 1929.
[63] K. Posthumus, "Oscillations in a split-anode magnetron," *Wireless Engr.*, vol. 12, pp. 126–132; March, 1935.
[64] A. L. Samuel, "Electron Discharge Device," U. S. Patent No. 2,063,342, filed December 8, 1934, issued December 8, 1936.
[65] N. T. Alekseev and D. E. Malainov, "Generation of high-power oscillations with a magnetron in the centimeter band," *J. Tech. Phys.* (*USSR*), vol. 10, No. 15, pp. 1297–1300; 1940.
[66] A. A. Boot and J. T. Randall, "The cavity magnetron," *J. IEE* (*London*), vol. 93, pt. 3 A, No. 5, pp. 928–938; 1946.
[67] J. B. Fisk, H. D. Hagstrum and P. L. Hartman, "The magnetron as a generator of centimeter waves," *Bell Sys. Tech. J.*, vol. 25, pp. 167–348; April, 1946.
[68] G. B. Collins, "Microwave Magnetrons," M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 6; 1948.
[69] J. Sayers, British Patent 588,916, filed October 3, 1941, issued June 16, 1947; U. S. Patent No. 2,546,870, filed February 9, 1945, issued March 27, 1951.
[70] P. L. Spencer, "High Efficiency Magnetron," U. S. Patent No. 2,408,235, filed December 31, 1941, issued September 24, 1946.
[71] J. Feinstein and R. J. Collier, "A magnetron controlled by a symmetrically coupled TE_{011} mode cavity," *Le Vide*, vol. 12, pp. 247–254; July–August, 1957.
[72] W. Brown, "The Platinatron: Amplitron and Stabilotron, in Crossed-Field Devices," vol. 2, E. Okress, Ed., Academic Press, Inc., New York, N. Y., 1961.
[73] J. R. Pierce, U. S. Patent No. 2,768,328, filed November 5, 1946, issued October 23, 1956.
[74] H. F. Webster, "Breakup of hollow electron beams," *J. Appl. Phys.*, vol. 26, pp. 1386–1387; November, 1955.

Current European Developments in Microwave Tubes*

A. H. W. BECK†, FELLOW, IRE

Summary—**The paper gives a brief description of the present state of the European microwave tube art. Developments on lines which differ from those common in the U.S.A. are stressed. After a few specialized tube types have been described, current research is surveyed, covering devices using electron beams in high vacuum, electron beam parametric amplifiers and microwave amplification in plasmas. Purely theoretical material has been omitted except where it is vital to the understanding of any device.**

Introduction

In looking back over the developments which have taken place in the microwave tube art since the end of World War II, one is conscious of both wonder that so many new devices should have been investigated, and of disappointment that the extent of their employment should have been so restricted, restricted not so much in numbers but in range of applications. If we leave aside purely military applications, the major European uses for microwave tubes are in marine and airfield surveillance radar and in high channel capacity microwave links, which mostly operate in the range between 2 and 7 kMc/sec.

In preparing this survey no attempt has been made to tabulate data on commercially available microwave valves because this information is readily available in manufacturers' catalogues. A few specialized types of valves are described but the major part of the article is devoted to research and possible future advances. It is perhaps necessary to emphasize that lack of space has forced the exclusion of any material on ancillary techniques such as: electron optical developments, cathode development and measurement techniques, all of which are vitally important in attaining the standards of performance and reliability vital in modern microwave valves. Moreover, purely theoretical work has also been omitted.

I. Some Tubes for Specialized Applications

A. Reflex Klystrons

A useful range of reflex klystrons is due to E.M.I.[1] who have produced two basic plug-in tubes of the external-resonator copper-glass disk-seal construction, one of which operates from 0.5 to 5.0 kMc and the other from 3.0 to 11.0 kMc [1]. These valves can be clamped into resonators whose design can be adjusted to optimize the performance for specific requirements. Using a $\frac{3}{4}\lambda$ radial cavity with a tuning device at the current node, a 30 per cent tuning range is combined with good efficiency [2]. Since the tubes themselves are relatively cheap, this system has considerable economic advantages. Other reflex klystrons which should be mentioned in view of their importance in research applications are the Philips 4-mm and 2-mm valves. Early samples of the 2-mm version gave 10 mw at 2.5 kv, 15 ma [3] but it is understood that considerably greater powers have now been achieved. Workers at S.T.L.,[2] using higher perveance guns, have achieved powers of 300 mw at 4 mm using voltages below 2.0 kv [4].

B. Traveling-Wave Tubes

In the traveling-wave tube field many developments are of interest. Several German laboratories have produced high-gain medium-power traveling-wave amplifiers. Of these, perhaps the most interesting is the Lorenz LWS4 [5] with a small signal gain of 55 db, giving a saturation output of over 5 w at about 45-db gain. This tube uses periodic permanent magnet focusing and an elegant package design has been developed.

Ferranti Ltd. has developed a grid-modulated 20-kw peak *X*-band tube with a minimum gain of 13 db. The grid is held at −250 v and is pulsed to about 1.6 kv to turn on the 26–30 kv, 4–6 a beam. The pulse power in the grid circuit is only about 0.1 per cent of that in the beam [6]. The same company has also worked on high-power CW tubes in which the helix was cooled by passing coolant through the helix under high pressure [7]. The main problems were technological and related to the pump and coolant arrangement. At *X* band an output of over 200 w for a 6.2-kv 450-ma beam with a gain of 20 db was achieved without apparently reaching the limit of the cooling system.

C. Backward-Wave Oscillators

Perhaps the most important series of tubes in the present category are the "M" (crossed field) and "O" (beam type) carcinotrons or backward-wave oscillators developed by C.S.F.[3] [8]. Recent developments in the high-power field are described in [9]. Very significant advances have been made in the millimeter region using "O" type tubes with finned-line (Millman-circuit) slow-wave structures. Five tubes cover the band 27.5–110kMc/sec. with a minimum power of 1 mw increasing

* Received by the IRE, November 13, 1961; revised manuscript received, December 14, 1961.

† Engineering Department, Cambridge University, Cambridge, England.

[1] Electrical and Musical Industries, Hayes, Middlesex, England.

[2] Standard Telecommunication Labs., Ltd., London, England.

[3] Compagnie Général de Télégraphie, Sons Fil, Levallois-Parret (Seine), France.

to over 4 mw below 50 kMc. The power output vs frequency curve is very peaky and the peak powers at selected frequencies are many times these minimum figures [10]. Narrow-band tubes with very high convergence electron guns and unspecified improved circuits using watercooling have given spectacular results, for example, 2 w in a 7 per cent band around 70 kMc [10]. Tubes for bands around 200 kMc and 270 kMc have been made and have given useful powers. An "M" type oscillator giving 10 w over the 27 to 36-kMc band has also been developed. These achievements, and those with reflex klystrons mentioned earlier, all of which relate to tubes which are available, at least for laboratory use, make it appear certain that it will not be necessary to resort to megavolt electronic devices to produce oscillations at wavelengths above a few tenths of a millimeter.

D. Power Klystrons and Traveling-Wave Tubes

A 10-kw peak *L*-band klystron which is pulsed by a simple apertured anode has been described [11]. Here, the pulse anode works at one-third of the main beam voltage and the pulse power required is only about 1/300 of the main beam power. Although the arrangement requires a little more power than gridded-pulse electrodes, the extreme mechanical simplicity is attractive.

In the high-power field S.E.R.L.[4] have been especially active, and they have described stagger-tuned multicavity klystrons and cloverleaf T.W.A.'s [12] and [15]. The circuit used in the latter is that developed at Stanford University. In France, klystrons giving 20-Mw peak, 20-kw mean, powers in the *S* band have been developed both by C.S.F. and C.F.T.H.[5] They differ from the original Stanford tubes in using guns of perveance 2×10^{-6} and therefore function at 250 kv. The maximum efficiency is about 45 per cent [13]–[14].

E. Magnetrons

Conventional pulsed magnetrons for the 4-mm band have been described by S.E.R.L. in England [16] and by Philips in Holland [17]. The S.E.R.L. tube gave 5-kw peak with about 7 per cent efficiency while the Philips 18-vane tube gave an output of 34 kw, and an efficiency of 19 per cent at a field of 1.8 Wb/m².

Some very light and cheap magnetrons, intended for radiosonde transmitters, have been described by the Mullard Co. [18]. These use an interdigital circuit attached to a Lecher line, in a manner used in some low-power German war-time magnetrons. *S* and *X* band models have been studied, that in *S* band giving 2.2-kw peak with an efficiency of 50 per cent at 4,500 gauss field.

[4] Services Electronics Research Laboratories, Baldock and Harlow, England.
[5] Compagnie Français, Thomson-Houston.

II. Current Research and Development on Microwave Valves

We have space only to mention briefly some of the more interesting projects recently reported.

Several versions of an electrostatically focused BWO using a novel focusing principle have been studied by the G.E.C.[6] [19]–[21]. These tubes, called "Ophitrons" use the inhomogeneous electrostatic fields set up between the delay line or lines (Figs. 1 and 2) and the external negative focusing plate to form a strip beam of sinuous shape. Tubes for *S* and *X* bands have been successful and the resulting oscillator is very small and light because of the absence of focusing magnet and seems most attractive for all low power applications of B.W.O.'s. W. E. Willshaw [21] reports an attempt at a 2-mm version with a line comprising 200 slots 0.004 in wide and 0.036 in long. The mechanical problems of such a tube are naturally formidable.

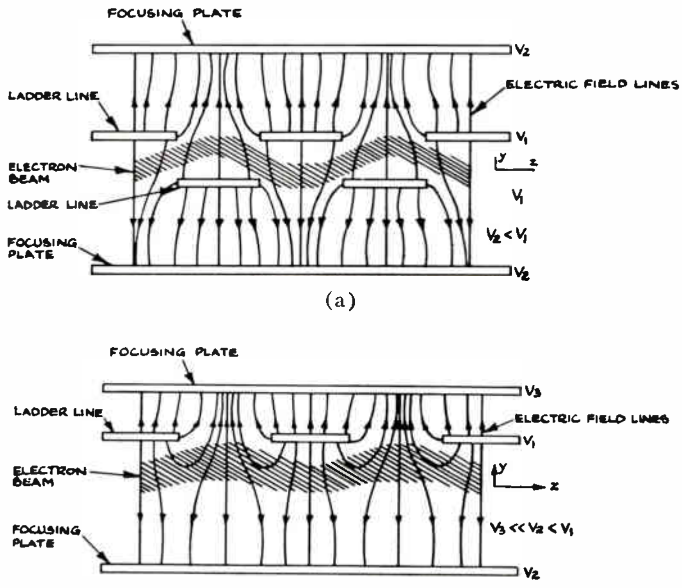


Fig. 1

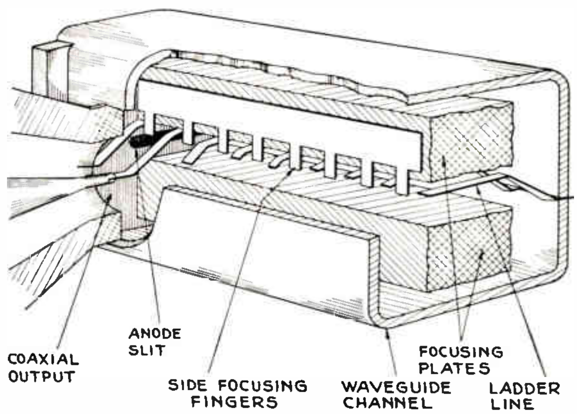


Fig. 2.—Cut away view of ophitron structure.

[6] General Electric Company Ltd., London, England.

A novel device, which appears to have been studied only in Russia, is the "Spiratron" (Fig. 3). Here, the slow wave structure has axial symmetry and is normally a helix. A focusing wire, at a different dc potential, runs along the axis of the helix. The injected beam is annular and the individual electrons perform helical trajectories whose pitch depends on the injection system [22]–[24]. The advantages of the Spiratron are claimed to be the following: First, with an annular beam, the change of coupling impedance with frequency is less than with a solid beam and broader bandwidths are obtained; second, the annular beam gives a bigger coupling impedance and therefore more gain and efficiency for a given current; third, the annular beam can transmit more current before unwanted dc radial potential variations become important; finally, under large signal conditions the beam is thought to conserve its cylindrical shape without radial displacement of charge and possible loss of efficiency.

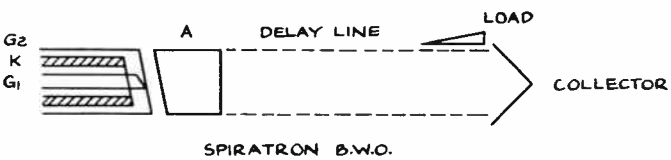


Fig. 3.

Most of these claims are borne out by the excellent results. For example, in amplifiers, bandwidths in excess of 3:1 have been achieved with 20-db gain. For the 12 to 20-cm band, outputs of 10 to 12.5-w C W were obtained with efficiencies of 25–30 per cent and a current of 25 ma, *i.e.*, at about 1.6 kv. Further results on Spiratron B.W.O.'s are given in [24]–[25]. Over the band 2.4–4.4 kMc/sec, requiring a voltage range of 250 v to 2.1 kv. output of up to 50-mw CW have been obtained for 12-ma beam current. The problem of the B.W.O. is more difficult than that of the amplifier, since, as the beam enters the interaction region where the RF fields are strongest, any tendency to RF beam instability has very serious consequences in beam interception, etc.

Russian investigators have studied some devices rather more thoroughly than has been the case in other European countries. One instance is the double stream-amplifier where at 3-kMc CW outputs of ≑ 5 w were achieved for beam currents of 40 ma when one stream was at 1550 v and the other at 1100 v. In this case the streams were very well mixed by making the second cathode of an emitting grid heated by interception of some of the beam from the first cathode [26]. Another group has used a reflex klystron as a frequency multiplier. Here, the reflector region was built into the low-frequency cavity while the ordinary cavity operated on the third harmonic. The device operated with a power drop of about 30 db from the 3.3 to 10-kMc band [27]. Also, in an intriguingly vague article, Golant [28] mentions a reflex klystron with 10–15 per cent electronic tuning range working at 150 v.

Of more importance are the recent achievements of cyclotron interaction between beams and waves on fast wave structures (Lecher lines) by the late S. I. Tetelbaum and his collaborators [29]–[30]. Such ideas had been discussed in Germany [31] but Tetelbaum was first to produce useful powers which he did with a very simple structure (Fig. 4). The electrons perform cycloidal paths when the magnetic induction is adjusted to the correct value $\omega_s = \eta B$. Theory shows that we may then consider that the electrons exhibit an effective phase velocity

$$p = \frac{u_0}{1 - \omega_c/\omega_s} \cdot \qquad (1)$$

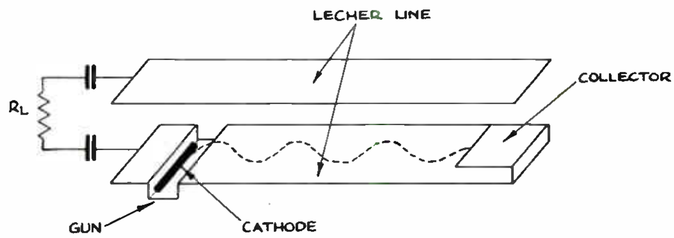


Fig. 4.

Thus, when $\omega_c = \eta B \doteqdot \omega_s$, the phase velocity is very large, and the beam will interact strongly with waves of high-phase velocity, which in Tetelbaum's case was the velocity of light c. His device was arranged as a B.W.O. and in spite of its simplicity gave 0.5 to 3-w CW over the band $\lambda = 6 - 18$ cm. This work is similar to the more refined experiments made at higher frequencies by investigators in the U. S. [32]–[33], and carried out with cylindrical beam devices. The investigation of magnetron-like devices was carried further in England [34]–[35].

Both Tetelbaum and Karp have given essentially steady-state analyses of the operation of the device. Tetelbaum's results agree fairly well with his measurements, but, of course, such a theory does not say anything about the starting current and buildup of oscillations. Pantell has given such a small-signal traveling-wave theory for his device which does predict these conditions.

Following on from the work of Karp, Reddish [36] has further examined the steady-state theory of the crossed-field version and has suggested means for improving the efficiency which are basically methods for operating with depressed collector potential. Experimental work has given promising results, when the electron optics of the gun were such that the cycloidal paths were not destroyed by space-charge forces.

Since these particular devices all use interaction at the cyclotron field, they require rather large magnetic fields, for example, about 13,000 gauss at $\lambda = 8$ mm, and the field increases linearly with frequency. Thus, it is not possible to go to very short wavelengths with existing means of producing magnetic fields (this should be qualified by noting recent successes in the production of 100-kilogauss fields with superconducting magnets). Since the fast wave structure is smooth, it has no mechanical problems comparable with those of slow wave structures and can certainly be accurately made even at very high frequencies. Also, the volume available for the electron beam is great and so large currents can be forced into the structure. These considerations made workers at Cambridge consider interactions between the beam and fields when angular frequency of the helical motion was adjusted to equal a submultiple of the cyclotron frequency. This is possible if a correct higher order waveguide mode is chosen. As an example, consider the case of half the cyclotron field. A spiralling electron beam is introduced into a waveguide propagating the TE_{11} mode. For square waveguide, the interaction is shown in Fig. 5. Consideration of the fields shows that the electrons always see fields which are accelerating or decelerating, depending only on the phase of entry. Thus, bunching takes place and the device behaves electronically like a traveling-wave tube in which the beam has been twisted into a spiral. Either forward-wave or backward-wave interactions can be used since the phase velocity is now

$$p = \frac{1}{1 - \eta B/n\omega_s}, \qquad n > 1.$$

Beck and Mayo [37] have outlined the theory of this device and are about to conduct an experimental test at 10 kMc/sec.

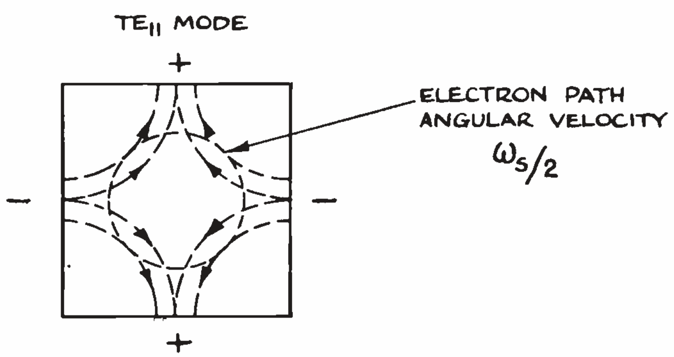


Fig. 5.

A novel device, whose practical embodiment is a little hard to visualize, has been described by Russian workers who call it the "Wavetron." In this a number of finite electron beams with differing velocities are considered to circulate in a waveguide of constant cross-section bent into a ring. Without detailed application of the theory it is difficult to say how this differs from linear multiple-beam tubes, but some solutions of the dispersion equation have been worked out [38] and of course regions of gain and instability appear.

The Siemens Company [39] have reported on the properties of a B.W.O. using a short interdigital line structure, which instead of being matched at both ends, is short circuited at the ends. Such a circuit exhibits a number of discrete resonances depending on the number of unit circuits in the particular line. Naturally, the resonance in the π mode has the greatest coupling impedance and will have the lowest starting current, but for higher beam currents, other modes may be excited at the appropriate beam velocity. In each mode only a small frequency bandwidth is expected. The experimental tube showed two modes, at 3.7 and 4.6 kMc, with tuning ranges of about $\frac{1}{2}$ per cent. The power output is not reported, but reasons for believing that this device can be developed for millimeter or submillimeter wavelengths are discussed.

III. Parametric Amplifiers

The successful development of the Adler type of low-noise parametric amplifier naturally aroused considerable interest in Europe and several developments were started. Publication in this field has been meagre but the English Electric Valve Co. has made Adler tubes for the 600-Mcs band.

Work at Cambridge [40] was directed towards investigating the question of whether the higher-order instabilities, which are predicted by the study of the unstable zones on Mathieu's equation, can be excited and if they give rise to useful values of gain. It has been shown both experimentally and theoretically that useful values of gain can be achieved in the particular unstable region which in the low-loss case starts at $\omega_p = 2\omega_s/3$. To obtain gain one must either pump the tube harder, or use a longer quadrupole, or improve the electronics over those of the standard $\omega_p = 2\omega_s$ device. Using a quadrupole about three times the normal length, Carroll has shown that over-all gains of over 10 db can be achieved in spite of circuit losses of order 3 db. The experimental tube used tuners sealed through the glass and worked at about 350 Mcs. The bandwidth was about 10 Mcs. When used at maximum gain the pump frequency was 254 Mcs, the shift from $\omega_p = 233.3$ Mcs $= 2\omega_s/3$ being predicted by the theory as the gain increases. Used conventionally, $\omega_p = 2\omega_s$ the tube gave $\doteqdot$ 33-db gain. Noise figures have not yet been measured. The work may be of value in exploiting parametric amplifiers at extreme frequencies where provision of the $2\omega_s$ pump would be difficult. It is also of theoretical interest in showing that the Adler tube can be described in terms of two coupled Mathieu equations.

Some work has been done by the G.E.C. [41] on the space-charge wave amplifier, orginally described by Louisell and Quate. The fundamental weakness of the simple form of this device is the small dispersion of ordinary space-charge waves and several methods of increasing the effective dispersion have been discussed.

The Philips Company [42] have made a comparison of the results obtained with an Adler tube operating at 550 Mcs/sec when the electron gun, of perveance 3×10^{-6}, was operated with the cathode first shielded (Brillouin flow) and then unshielded (confined flow). For the *Brillouin* gun the tube gave a noise temperature of 90°K, at 30-db gain while the saturation gain was 50-db. For the confined flow gun the noise temperature was 116°K at 22-db gain, saturation gain 30 db. The noise temperatures are a little higher than those of Adler's tube, probably because the cathode temperature is 1400°K. The gun has quite a low perveance (0.9×10^{-6}) in the cathode—1st anode region and the beam is then decelerated to the final 6-v value. The noise temperature at the first coupler is high, 5000°K, so that noise stripping is effective.

Workers at S.E.R.L. [43] have reworked the theory of the dc pumped quadruple tube to show that the noise temperature, previously thought to be high, can be made low by using beam cooling. A design with a theoretical noise figure of about 0.24 db is described.

IV. Microwave Plasma Phenomena

The world plasma is often misused in microwave tube literature to invoke the idea of a particular model of an electron beam in which the mean space-charge is neutralized by positive ions of infinite mass. Since this model involves none of the physical phenomena which occur in real plasmas it would be more logical to call it the ion neutralized beam. In this section, however, we are concerned with experimental work on real plasmas in which the collective interaction between the ion space-charge and the electronic space-charge is important and in fact gives rise to the observed phenomena.

This is not the place to review plasma theory but to make a description of the experiments intelligible one must say something about the propagation of waves in an infinite plasma. A characteristic frequency depending on the charge density is defined through

$$\omega_{pe}^2 = \frac{n_e e^2}{\epsilon_0 m_e}.$$

The theory then shows that if an electron beam density modulated at angular frequency ω is introduced into the plasma, gaining waves result for the frequency range $0<\omega<\omega_p$. For this range the plasma can be regarded as a dielectric medium with permittivity $\epsilon=\epsilon_0(1-\omega_{pe}^2/\omega^2)$, which is negative for the region under consideration. Maxwell's equations show that there is no ac space-charge in the medium and we may therefore, from a qualitative viewpoint, say that the requirement of charge neutrality has forced the plasma to assume a configuration giving gaining waves as in an inductive wall amplifier. The gain exhibits an infinite peak at the plasma frequency and then drops to zero. Above the plasma frequency ordinary transverse propagation of plane em waves takes place. For the noninfinite plasma, especially when confined in waveguides, the phenomena are complicated by the presence of the conductors. Above the plasma frequency, waves similar to the normal modes propagate, while below f_p waves with phase velocities less than that of light can propagate.

Even more complication is caused by adding a magnetic confining field because the plasma then becomes an anisotropic medium and new modes of propagation result.

The first successful attempts to confirm this picture were due to Boyd, Field and Gould [44]. They obtained amplification up to 4 kMc, with maximum gains at the plasma frequency. Russian workers [45], [46] have used increased plasma densities by providing magnetic ion traps at both ends of the interaction region and have thus been able to reach higher frequencies. A gain of 33 db at 3.0 cm is reported. Studies of the dependence of gain on discharge current, beam current and frequency have been made.

Other Russian groups are also active in the microwave plasma field as [47], [48] show. In Sweden, Agdur [49] has reported on amplification observed when a modulated electron beam is injected, by means of a helix coupler, into a gaseous atmosphere, the whole apparatus being immersed in a kilogauss magnetic field. In this case amplification of over 1 db/cm was observed, substantially at the electron cyclotron frequency. The qualitative explanation given is that slow electrons are created by the ionizing collisions and that the beam interacts with the slow wave carried by these electrons. The theory shows that gain should occur in the frequency range $\omega_c^2<\omega^2<\omega_c^2+\omega_r^2$ where ω_r = plasma frequency of the slow electrons. In the experiments, $\omega_r\ll\omega_c$ so that the range was small and near to ω_c.

Workers at Philips have obtained amplification of nearly 60 db in a low pressure mercury arc 3.0 cm long when working at 4.4 kMc [50].

In England, workers at the G.E.C. have started to study phenomena in plasma filled waveguides [51]. The theory has been extended [52] to cover the case of an electron beam penetrating the stationary magnetized plasma. If the slow wave velocity on the plasma column is nearly equal to the beam velocity gain results. Experiments on a hydrogen discharge gave partial confirmation of the theory and electronic gains were observed.

In the light of the experimental work so far performed, it is clear that very careful and prolonged series of measurements under carefully prescribed conditions will be necessary before the full details of all the interactions possible between electron beams and plasmas are understood. In particular, two difficulties have to be surmounted before operation at short millimetre wavelengths becomes possible. The first of these are the achievement, without excessive use of high-powered discharges, of sufficiently high-plasma densities and secondly, the solution of the problem of coupling to the plasma rather than to the beam, for, if beam coupling is used, most of the constructional difficulties of hard valves remain. Finally, one might guess that there are many engineering problems of stability and life to be solved before plasma devices take their place in microwave systems.

V. Conclusion

In this brief survey I have attempted to emphasize work not being done in the United States and by so doing I have to ask to be excused for ignoring a large amount of worthy research and development.

Perhaps the major conclusion one can draw is that conventional microwave tube types are now being developed for wavelengths in the region of 1 mm and that megavolt electronics does not appear necessary in this region, especially as the potentialities of cyclotron and subcyclotron fast wave tubes have, as yet, hardly been explored.

By contrast, it would appear that too much of the strictly limited European tube development effort is devoted to the production of competing types in well-known fields. It seems doubtful whether the microwave tube market will ever be sufficiently large to warrant the production of more than a few of the most efficient types.

Acknowledgment

The author wishes to acknowledge his indebtedness to Professor M. Chodorow for extending the hospitality of the Microwave Laboratory of Stanford University where most of the research for this survey was carried out. Acknowledgment is extended also to all the individuals and organizations who sent papers and catalogues to the author. Special thanks are due to W. H. Aldous and D. H. O. Allen.

References

[1] A. F. Pearce, K. H. Kreuchen, C. Baron, H. Houldin and S. Ratcliffe, "Plug-in reflex klystrons for microwaves," *J. Electronics and Control*, vol. 3, pp. 535–563; December, 1957.
[2] R. Mather, and J. Sharpe, "Wide-range tuning cavities for plug-in reflex klystrons," *Electronic Engrg.*, vol. 31, pp. 390–393; July, 1959.
[3] B. B. Van Iperen, "Klystrons reflex pour ondes millimetriques," *Le Vide*, vol. 65, pp. 264–266; 1956.
[4] Private communication.
[5] W. Klein and H. Neufischer, "Experimental medium-power high-gain traveling-wave tubes," *Proc. IEE*, vol. 105B, Suppl. 10, pp. 408–412; May, 1958.
[6] Data Sheet, Ferranti LY10 T.W.A. Ferranti Ltd., London.
[7] G. M. Clarke, "Experimental traveling-wave tubes," *Electronic Tech.*, vol. 38, pp. 38–45; February, 1961.
[8] R. Warnecke, "Sur quelques resultats recemment obtenus dans le domaine des tubes electroniques hyperfrequences," *Ann. de Radioelect.*, vol. 9, pp. 107–136; April, 1954.
[9] M. Favre, "Results obtained on crossed-field Carcinotrons under pulsed operation," *Proc. IEE*, vol. 105B, Suppl. 10, pp. 533–537; May, 1958.
[10] "Ondes Millimetriques—Nouvelles Perspectives," CSF, Paris, France; February, 1961.
[11] A. H. W. Beck and P. E. Deering, "A three-cavity L-band pulsed amplifier," *Proc. IEE*, vol. 105B, Suppl. 12, pp. 833–839; May, 1958.
[12] P. G. R. King, "A 5% bandwidth 2.5 mw S-band klystron," *Proc. IEE*, vol. 105B, Suppl. 12, pp. 813–820; May, 1958.
[13] C. Romiguiere, "Le klystron scelle 20 mw–20kw, TH2011," *Proc. Internatl. Congress on Microwave Tubes*, Munich, Germany, pp. 51–53; June 7, 1960.
[14] B. Palluel and G. Kantorowicz, "Pulse high perveance high power sealed-off klystrons for linear accelerators," *Proc. Internatl. Congress on Microwave Tubes*, Munich, Germany, pp. 73–77; June 7, 1960.
[15] M. O. Bryant, J. F. Gittins and F. Wray, "Improvements in a CW power traveling-wave tube," *J. Electronics and Control*, vol. 7, pp. 505–507; December, 1959.
[16] A. J. Monk, "A 3.8 mm wavelength pulsed magnetron," *Proc. IEE*, vol. 105B, Suppl. 10, pp. 440–443; May, 1958.
[17] J. Verweel and G. H. Plantinga, "A range of pulsed magnetrons for centimetre and millimetre waves," *Philips Tech. Rev.*, vol. 21, No. 1, pp. 1–10; 1959/1960.
[18] H. N. G. King, "Interdigital magnetrons," *Proc. Internatl. Congress on Microwave Tubes*, Munich, Germany, pp. 203–205; June 7, 1960.
[19] H. A. C. Hogg, "Periodic electrostatic beam focusing," *Proc. IEE*, vol. 105, Suppl. 12 pp. 1016–1020; May, 1958.
[20] R. B. Dyott, H. A. C. Hogg, M. A. Hulley and E. Kettlewell, "The ophitron—an electrostatically focused backward-wave oscillator," *Proc. Internatl. Congress on Microwave Tubes*, Munich, Germany, pp. 114–117; June 7, 1960.
[21] W. E. Willshaw, "Generation and amplification in the millimeter-wave field," *Proc. Internatl. Congress on Microwave Tubes*, Munich, Germany, pp. 6–17; June 7, 1960.
[22] Z. S. Chernov, *Radiotekh. Elektron.*, vol. 1, 1956.
[23] Z. S. Chernov, "Interaction of electromagnetic waves and electron beams in systems with centrifugal-electrostatic focusing," *Proc. IEE*, vol. 105B, Suppl. 10, pp. 594–595; May, 1958.
[24] G. A. Bernashevskii and T. A. Novskova, "Spiratron with backward-wave," *Radiotekh. Elektron.*, vol. 3, pp. 165–169; July, 1958.
[25] G. A. Bernashevskii and T. A. Novskova, "Experimental investigation of a backward-wave spiratron," *Radiotekh. Elektron.*, vol. 4, p. 135; September, 1959.
[26] G. A. Bernashevskii, P. S. Voronov, T. J. Isiamova and Z. S. Chernov, "The experimental investigation of double beam electron-wave systems," *Radiotekh. Elektron.*, vol. 4, pp. 165–172; October, 1959.
[27] S. A. Kornilov and V. A. Pamiatikhin, "Characteristics of a reflex klystron frequency multiplier," *Radiotekh. Elektron.*, vol. 4, pp. 261–264; November, 1959.
[28] M. B. Golant, "A reflex klystron with wide electronic tuning range," *Radiotekh. Elektron.*, vol. 3, pp. 97–105; June, 1958.
[29] S. I. Tetelbaum, "Return-wave phasochronous generators," *Radiotekh. Elektron.*, vol. 2, p. 45; January, 1957.
[30] K. I. Lizhdvoi, "Experimental investigations of non-slowed backward-wave oscillators," *Radiotekh. i. Elektron.*, vol. 4, pp. 78–86; February, 1959.
[31] K. Kleinwachter, *Elektrotechnische Zeitschrift*, vol. 24, p. 714; 1951.
[32] R. H. Pantell, "Backward-wave oscillations in an unloaded waveguide," Proc. IRE (Correspondence), vol. 47, p. 1146; June, 1959.
[33] R. M. Phillips, "The ubitron, a high-power traveling-wave tube based on a periodic beam interaction in unloaded waveguide," IRE Trans. on Electronic Devices, vol. ED-7, pp. 231–241; October, 1960.
[34] P. D. Dunn, "A Travelling-wave Magnetron Amplifier Operating in the Cyclotron Mode," Horwell, A.E.R.E. Rept. No. GB/M166; 1952.
[35] A. Karp, "Fast-wave interactions with an electron film at cyclotron resonance," *Proc. IEE*, vol. 105B, Suppl. 11, pp. 656–661; May, 1958.
[36] A. Reddish, "Fast-wave interactions using cyclotron resonance," *Proc. Internatl. Congress on Microwave Tubes*, Munich, Germany, pp. 235–240; June 7, 1960.
[37] A. H. W. Beck and R. F. Mayo, "Interactions between smooth-waveguides, etc.," *Proc. Int. Congress on Microwave Tubes*, Munich, Germany, pp. 240–242; June 7, 1960.
[38] V. M. Lopukhin, "Graphical investigation of a wavetron in the initial stage of oscillation," *Radiotekh. Elektron.*, vol. 4, pp. 681–687; April, 1959.
[39] H. Heynisch, "Der Resonanz-Rückwartswellen-Oszillator, ein Generator für Mikrowellen," *Arch. elekt. Übertragung*, vol. 14, pp. 531–538; 1960.
[40] J. E. Carroll, "Subharmonic parametric pumping of a quadrupole amplifier," submitted to *J. Electronics & Control.*
[41] C. R. Russell, "The space-charge wave beam parametric amplifier," *Proc. Internatl. Congress on Microwave Tubes*, Munich, Germany, pp. 348–350; June 7, 1960.
[42] P. A. H. Hart, and C. Weber, "Parametric amplifier electron gun," privately communicated and to be published.

[43] C. P. Lea-Wilson, T. J. Bridges and J. C. Vokes, "The application of beam cooling to quadrupole amplifiers," *J. Electronics & Control*, vol. 10, pp. 261–272; April, 1961.
[44] G. D. Boyd, L. M. Field and R. W. Gould, "Excitation of plasma oscillations and growing plasma waves," *Phys. Rev.*, vol. 109, p. 1393; 1958.
[45] E. V. Bogdanov, V. Ia Kislov and Z. S. Chernov, "Interaction between electron stream and plasma," *Radiotekhnika* (Moscow), vol. 5, pp. 75–90; February, 1960.
[46] Z. S. Chernov, and G. A. Bernashevskii, "Amplification of Microwaves by Means of Plasma," presented at Symp. on Electromagnetics and Fluid Dynamics, Brooklyn Polytechnic Inst., Brooklyn, N. Y.; 1961 (To be published).
[47] R. A. Demirkhanov, A. K. Gevorkov and A. P. Popov, "The Interaction of a Beam of Charged Particles with a Plasma," presented at Uppsala, Conf. Ionization Phenomena in Gases, North-Holland; September, 1960.
[48] I. F. Kharchenko, Ya. B. Fainberg, P. M. Nikolayev, E. A. Kornilov, E. A. Lutsenko and N. S. Pedenko, "Experimental and Theoretical Investigation of the Interaction of an Electron Beam with a Plasma," presented at Uppsala Conf. Ionization Phenomena in Gases, North-Holland; September, 1960.
[49] B. Agdur, "Oscillations in long electron beams," *Ericsson Tech.*, vol. 1, pp. 44–57; March, 1960.
[50] M. T. Vlaardingerbroek and U. Weimer, "Wisselwerking van Plasma en Elektronenbundel," presented at Meeting of N.N.V. Utrecht, The Netherlands, January, 1961.
[51] R. L. Ferrari and A. Reddish, "Propagation in a plasma filled waveguide," *Proc. Internatl. Congress on Microwave Tubes*, Munich, Germany, pp. 548–554; June 7, 1960.
[52] R. L. Ferrari, "Wave propagation and electron beam interaction in plasma filled waveguides," presented at 5th Conf. on Ionization Phenomena in Gases, Munich, Germany; September, 1961, to be published.

Beam-Deflection and Photo Devices*

K. SCHLESINGER†, FELLOW, IRE, AND E. G. RAMBERG‡, FELLOW, IRE

***Summary*—The development of practical beam-deflection and photo devices coincides with the lifetime of the Institute of Radio Engineers. The intimate association of these two electronic elements is illustrated by the development of the television camera tube. In more detail, both the historical development and the more significant recent advances are described for beam-deflection devices such as the cathode-ray oscillograph, storage tubes, and television display tubes. Similarly, advances are indicated in the field of photo devices such as photoemissive, photoconductive, and photovoltaic cells; multiplier photo-tubes; image tubes; and light amplifiers.**

I. General Introduction

A. Historical Background

FIFTY YEARS ago, at the time of the founding of the Institute of Radio Engineers, beam-deflection and photo devices were virtually nonexistent, although their basic principles were well-known. The cathode-ray oscilloscope for the study of the time variation of electron currents was developed by Ferdinand Braun [1] in 1897, the same year in which J. J. Thomson [2] measured the specific charge of the electron by its deflection in electric and magnetic fields. The photovoltaic effect was discovered by Edmond Becquerel [3] as early as 1839. In 1873, Willoughby Smith [4] first observed the photoconductivity of selenium. In 1887, finally, Heinrich Hertz [5] described the photoemissive effect of ultraviolet light on metal electrodes. In 1905, Albert Einstein [6] introduced the guiding theoretical principle for the photoelectric effect, showing how the laws of photoemission demonstrated the validity of Planck's quantum principle as applied to light.

Nevertheless, in 1912 there was not a single beam-deflection or photo device on the market. The reason is simple. In the absence of electronic amplifiers and oscillators, their utility was too limited to warrant commercial production.[1]

B. Purposes of Beam-Deflection and Photo Devices

Beam-deflection and photo devices serve essentially two purposes. These are:

1) Measurement: Thus, the cathode-ray oscillograph records the time variation of currents and voltages and the photoelectric cell measures light flux. As compared with the galvanometer, the cathode-ray oscillograph has the advantage of being able to represent extremely rapid variations of current and voltage. The same advantage of the representation of rapid variation of light intensities is realized by combining the photoemissive cell with the cathode-ray oscillograph.

2) Pictorial Representation: Thus, in television, instantaneous picture reproduction at remote points is made possible by a camera tube, a combined beam-deflection and photo device, and a beam-deflection device in the form of a viewing tube coupled by an electrical transmission channel. In radar and ultrasonic ranging, the pictorial reproduction utilizes a similar

* Received by the IRE, July 25, 1961; revised manuscript received, October 5, 1961. K. Schlesinger is primarily responsible for the section on beam-deflection devices, and E. G. Ramberg for those on television camera tubes and photo devices.
† Cathode Ray Tube Dept., General Electric Co., Syracuse, N. Y.
‡ RCA Laboratories, Princeton, N. J. (Presently on leave at Technische Hochschule, Darmstadt, Germany.)

[1] There were indeed a few uses which could have been satisfied without electronic auxiliary devices, exemplified by photographic exposure meters, relay-operated door openers and closers, and sealed-off cathode-ray oscillographs for measurements on rotating electrical machinery. The former had to await improvements in the sensitivity of photocells, while the latter, in the earlier stages of their development, could not compete with mechanical oscillographs in the low-frequency range of primary interest.

beam-deflection viewing tube. Storage tubes are beam-deflection devices which permit the delayed transmission of an electrical picture signal or its modification. Image tubes and image amplifiers are photo devices which translate an image formed by invisible or very-low-level radiation into an observable image.

A further use, which has achieved increasing prominence in the last few years, is the application of the photovoltaic cell as a solar energy converter. Still others, such as the use of photo devices as infrared signaling detectors, as scintillation counters, and as sensors in automatic industrial equipment are simply special examples of the measuring function.

C. Beam-Deflection and Photo Devices in Electronic Television

The association of beam-deflection and photo devices is best illustrated by the development of electronic television. Photo devices are essential in the conversion of the light values which compose the information content of the scene to be transmitted into electrical signals. Beam-deflection devices, on the other hand, were soon recognized as the most practical means for deriving a single time-dependent signal which described both the space variation *and* the time variation of the light distribution in the scene in sufficient detail to satisfy the ultimate viewer.

Even before this, the reconstitution of the picture from the electrical signal was found to be the special province of the suitably modified cathode-ray tube, the viewing tube or kinescope.

In addition, the needs of television have played a major role in the genesis of important photo devices such as the multiplier phototube and the image tube, as well as in the development of some types of storage tubes. Finally, the development of photoconductive materials has received important stimulation from the requirements of simplified camera tubes such as the vidicon, as well as from the precursor of a panel-type television display, the light amplifier.

II. Combined Beam-Deflection and Photo Devices: The Television Camera Tube

A. Basic Transmission Techniques

Television, or the instantaneous[2] electrical transmission of pictures, may employ any one of three basic techniques of signal generation (Fig. 1). Thus, it may utilize an array of photosensitive elements on which the picture to be transmitted is projected and employ individual transmission channels for every picture element to the picture reproducer [Fig. 1(a)]. This procedure was proposed as early as 1875 by G. R. Carey, but has survived only in animated cartoon advertising. The second procedure, reputedly first proposed by De Paiva

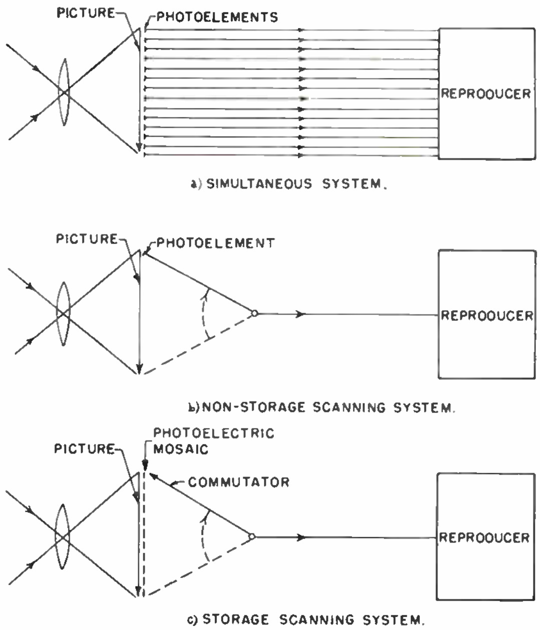


Fig. 1—Television picture transmission.

in 1878, employs a single photosensitive element which explores or scans the picture in a period comparable with the time of persistence of vision, transmitting the light value of every picture element once within this period over a single channel to the reproducer [Fig. 1(b)]. This procedure finds expression in all the various mechanical television pickup systems which played a leading role in television experimentation up to the early thirties. Its most famous realization is the Nipkow scanning disk. The fact that here, and in all except the earliest proposals, the photosensitive element is stationary and the picture element to be transmitted is selected by a great variety of scanning procedures is indifferent. The common features are 1) the single transmission channel; 2) a commutation system capable of sampling at a rate equal to the ratio of the number of picture elements to the time of persistence of vision—a quantity of the order of several million per second for pictures which are regarded subjectively as adequate; 3) reduction in the light available for signal generation (as compared with the simultaneous system) by a factor equal to the number of picture elements[3] and correspondingly increased requirements in sensitivity and lighting; and 4) the employment of photoelectric elements with signal lag smaller than "picture element time," or, approximately, the ratio of the time of persistence of vision to the number of picture elements. With present standards, picture element time is a fraction of a millionth of a second.

B. Electronic Nonstorage Systems

In brief, the simplification of a single transmission channel was purchased at the cost of imposing very stringent requirements on the signal generator. In the

[2] "Instantaneous" here has the practical meaning that delays between transmission and reception do not exceed the persistence of vision of the human eye.

[3] Except when the scene is scanned by the source of illumination, as in the flying-spot system.

early days of mechanical television, the second feature, *i.e.*, the high commutation rate, was regarded as particularly onerous, since it demanded the construction of rotating or reciprocating machinery operating at great speed and with high precision. This difficulty was eventually overcome by substituting the deflection of electron beams by alternating electric or magnetic fields for mechanical motion, as in Farnsworth's [7] image dissector [Fig. 2(a)] and in the flying-spot cathode-ray-tube pickup system [Fig. 2(b)]. In the former, the picture was projected on an extended photocathode. The emitted photoelectrons were accelerated and focused by a longitudinal magnetic field in a plane containing an aperture in front of a sensitive electron detector (*e.g.*, a secondary-emission multiplier), forming there an electron image of the light distribution on the photocathode. This electron image was swept across the aperture, a picture element in size, by alternating transverse magnetic fields synchronized with the deflection fields in the receiver. In the flying-spot system [8], [9], a scanning pattern is described by an electron beam in a simple cathode-ray tube on a short-persistence fluorescent screen. This pattern is imaged onto the scene to be transmitted, and the reflected or transmitted[4] light is collected by a phototube generating the picture signal.

While these electronic nonstorage scanning systems overcome the difficulty of high-speed commutation, they still suffered from the third characteristic, namely, high lighting requirements or inadequate sensitivity. These have limited their use to a few special-purpose applications. The required high-speed response of the photosensitive elements, on the other hand, has not proved a limitation, since the speed of response of photoemission in vacuum fully meets the need.

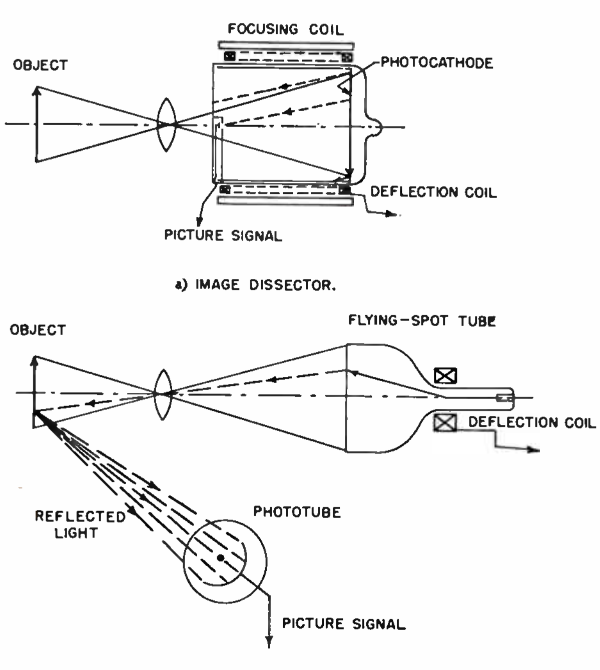


Fig. 2—Nonstorage scanning pickup systems.

[4] If the scene has been recorded on film or on a transparency.

C. Electronic Storage Systems

With the third procedure [Fig. 1(c)], the storage scanning system, the array of photosensitive elements is restored and the scanning takes place between this array and the transmission channel. The individual elements store charge by photoemission continuously[5] and release their charge for signal generation at the instance of commutation. Here, the single communication channel is obtained at the cost of only the high-speed commutation; the sensitivity and speed-of-response requirements for the photosensitive elements are the same as in the simultaneous system.

Mechanical commutation of the photosensitive elements at the required speeds was clearly out of question. Thus, all television storage scanning systems have employed camera tubes with electron beam scanning for commutation. The combination of electron beam scanning and storage was first proposed and carried into practice by V. K. Zworykin [8] in 1923—at a time preceding any other electronic pickup system. The original tube in which this was realized is shown in Fig. 3. The target was an aluminum film oxidized on one side, which was photosensitized with cesium vapor and faced a metal grill serving as collector for photoelectrons. The metal side, which served as signal plate, was scanned by a high-speed electron beam, which penetrated through the oxide layer, forming a temporary conducting path permitting the locally stored charge to flow off through the signal plate.

1) Iconoscope: While the tube as described was capable of transmitting only rudimentary patterns, it became the ancestor of the extended line of storage camera tubes which dominate all phases of television picture transmission. Their common features are electron beam scanning and a picture target with small transverse conductivity capable of storing charge released in response to light. The principal types are indicated in Fig. 4. The first practical storage camera tube was the iconoscope [11]. Here, the picture was projected on a mosaic of photosensitive elements capacitatively coupled to a signal plate. The mosaic was then scanned by a high-velocity beam, restoring the mosaic to a uniform potential and releasing photoelectrically stored charge for forming the picture signal. The secondary-emission ratio was greater than unity, so that the equilibrium potential under the beam was close to that of the electrodes facing the mosaic. The efficiency of the tube was low owing to the unfavorable field conditions for collecting the photoelectrons and secondary electrons; furthermore, redistribution of these electrons gave rise to shading errors in the picture.

[5] This requirement is not realized in the interesting early proposal for a cathode-ray camera tube with a mosaic target of A. A. Campbell-Swinton [10] (1911).

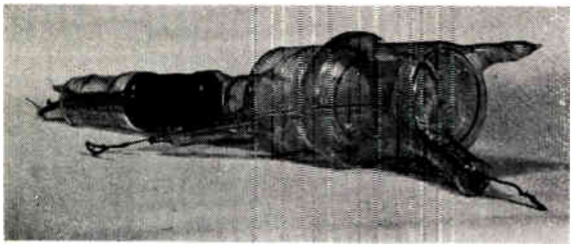


Fig. 3—Prototype of storage-type camera tube (Zworykin, 1923).

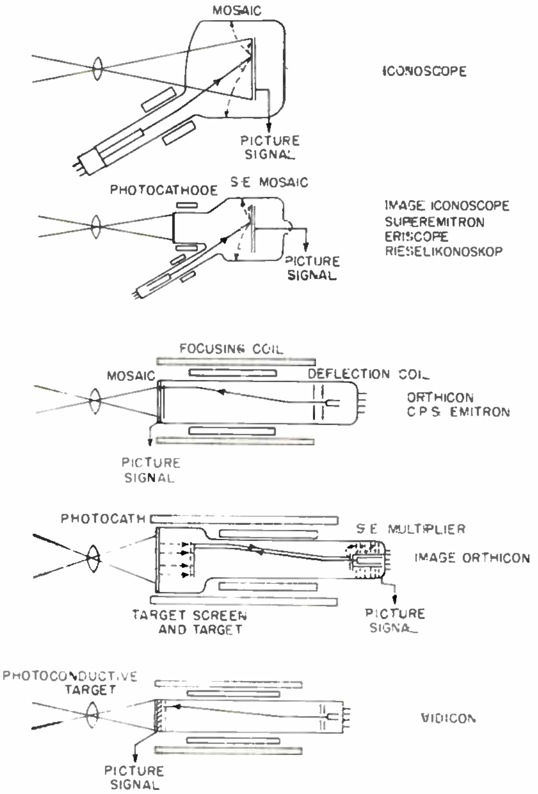


Fig. 4—Storage camera tubes.

2) *Image Iconoscope:* The first important improvement of the iconoscope consisted in the addition of an electron imaging stage [12], [13]. In the image iconoscope, the picture was imaged on a transparent photocathode, and the photoelectrons were then focused by an electron-optical accelerating and imaging system on the mosaic. This led to greater sensitivity of the camera tube as a result of the greater photoelectric sensitivity of the transparent photocathode as well as the secondary-emission gain at the mosaic. At a later date, a spray of photoelectrons from the photosensitized tube wall controlled by suitably distributed electrodes also served to eliminate shading [14]. Improved image iconoscopes of this type enjoy considerable popularity in European studios.

3) *Orthicon:* A more radical approach for improving sensitivity and correcting shading is represented by the orthicon [15], [16]. Here, the mosaic is scanned by the beam at low velocity. Since, then, the secondary-emission ratio is much less than 1, the beam drives the surface to a potential slightly below that of the cathode, so that no additional electrons can reach it. Collection is complete and there is no redistribution. At the same time, such low-velocity beams required new methods of beam focusing and deflection, making use of extended focusing and deflection fields. Earlier tubes of the orthicon type suffered from the disadvantage of instability at excessively high light levels.

4) *Image Orthicon:* In the image orthicon [17], this instability was overcome and the sensitivity of the camera tube was made to approach the theoretical limit for a pickup device with a photocathode of given sensitivity. The gain in sensitivity resulted from the addition of an electron-imaging stage ahead of the storage target and an electron multiplier for raising the signal level of the tube output. In this manner, sensitivity limitations imposed with earlier tubes by the camera amplifier were nullified. In addition, a target screen at a potential slightly above that of the gun cathode prevented charging of the target to a point at which unstable operation could result. Signal multiplication of the output signal in an electron multiplier was made possible by employing the beam returning from the target as signal carrier. The target itself was a thin glass foil with sufficient conductivity to permit charge equalization in a number of frame periods, or also a thin film of some other near-insulator, such as magnesium oxide.

5) *Vidicon:* With the image orthicon, the scope for material improvements in sensitivity was limited. On the other hand, there was considerable room for simplification, even at the cost of lower sensitivity. This was realized in the vidicon [18], a low-velocity tube with a photoconductive target. A potential difference of the order of 20 v is maintained, in darkness, between the front surface, in contact with the signal plate, and the back surface, brought to cathode potential by the scanning beam. Light liberates carriers within the photoconductive layer. These diminish the potential difference in the interval between successive scans. At the instant of scan, the beam deposits enough electrons on the back surface to restore the equilibrium condition and induces a corresponding charge displacement in the signal lead. In the dark, the photoconductor must be an insulator, with a resistivity of the order of 10^{12} ohm-cm, so that the relaxation time (the product of restivity and dielectric constant) may be large compared to the frame time, 1/30 sec. It is interesting to note that the decay of photoconductivity is, for all practical materials, far too slow to permit the use of photoconductors in any but storage systems. On the other hand, in such systems, their sensitivity can be made greater and their preparation is simpler than for insulating photoemissive mosaics in tubes of the orthicon type. This applies particularly for camera tubes of very small dimensions.

6) *Special-Purpose Tubes:* In addition to the standard camera tubes listed above, numerous camera tubes have been developed to serve special purposes. We may mention here in particular tubes which have been made

suitable for operation at extremely low light levels by placing multiple image amplifiers ahead of the storage target [19]; vidicons with special target materials for the infrared, and both vidicons and image orthicons for the ultraviolet [20], [21]; and photoconductive tubes for the reproduction of X-ray images [117].

III. Beam-Deflection Devices

A. The Cathode-Ray Oscillograph

1) Historical Background: The recording of time-varying electrical signals, or "waveform display," was the first technical application for the early cathode-ray tube. The turn of the century was marked by the fundamental work of Lord Kelvin and Heaviside on signal distortions in telegraph lines and cables, and by Ferranti's early experiences with transients and voltage surges on power transmission lines. This was also the era of Hertz and Marconi, the dawn of the "Radio Age."

All this contemporary work established a need for oscillographic instrumentation. It does not seem accidental, therefore, that it was a radio pioneer, Ferdinand Braun, who constructed the first cathode-ray oscillograph at the University of Strassbourg in 1897 [1].

Just like the early X-ray tube, the "Braun tube" used gas discharge phenomena for the emission and formation of an electron beam. Even after the introduction of a thermionic cathode into cathode-ray tubes by Wehnelt in 1905 [22], an argon atmosphere of about 10^{-3} mm Hg was still retained in commercial oscilloscopes [23], [24] for another 25 years [25]. The effect of ion focusing [26] facilitated the formation of long, filamentary, electron beams.

Electrostatic deflection does not perform ideally in an imperfect vacuum. On the other hand, the cathode-ray oscillograph must use electrostatic rather than magnetic deflection, since it is to handle wide-band signals and fast writing speeds with low deflection power. The development of the first high-vacuum tubes by Zworykin around 1930 [27], [28] marks, therefore, the date, for cathode-ray-tube design to be put on the firm scientific foundation of electron optics [29].

2) Oscilloscope Gun and Deflection System: A typical modern oscilloscope tube with intensifier is shown in Fig. 5. The electron gun is of the "zero current" focus type [30], implying negligible interception of beam current by the large-diameter focus cylinder *F*. Both pairs of deflection plates are flared to increase sensitivity. The first pair displays the signal waveform vertically with an average of about 16 v/cm·kv. The second pair provides the horizontal time base with about 20 v/cm·kv. Its contour is shaped to correct for pincushion distortion. There also exists an electrostatic yoke for deflection up to 60° from a common center [31]. This "Deflectron" is free from keystone distortion and has equal sensitivity in both axes. It has been used in industrial [33] and military [34] applications.

3) Post-Acceleration Techniques: Post-Deflection Acceleration (PDA) has become standard practice in modern oscilloscope tubes. It offers the immediate advantage of brightness intensification. In some designs, it also contributes to resolution by permitting to focus at a shorter effective throw. A disadvantage of PDA is its adverse effect on deflection sensitivity (for fixed beam velocity in the deflection plane) and scan linearity. In the simple dual-band intensifier shown in Fig. 5, these factors restrict the maximum usable PDA to 2:1.

Later designs of beam intensifiers are shown in Fig. 6. This includes the multiband intensifier [Fig. 6(a)] and also the helix-type [Fig. 6(b)]. These devices show much improvement in scan distortion since the ideal non-restoring, uniform field is more closely realized [35], [36]. As a result, very high screen voltages are feasible and this, in turn, permits the use of aluminized screens with their additional brightness gain. Even in a helix-type intensifier, a deflection loss still remains, amounting to 50 per cent at a PDA of 10. A configuration not beset by this difficulty is Allard's "ideal post-accelerator" [37] [Fig. 6(c)], which has been reduced to practice in the laboratory [32], [34].

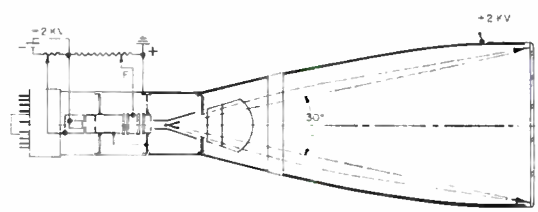


Fig. 5—Modern oscilloscope tube.

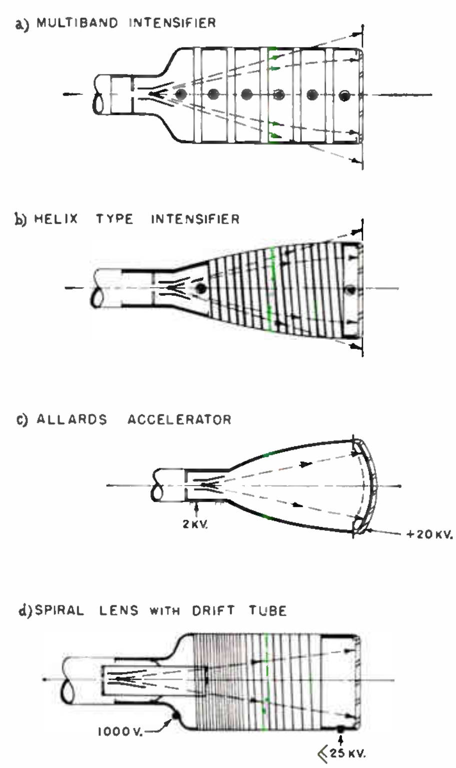


Fig. 6—Beam intensifiers for oscilloscopes.

Another more recent development in this field is the "spiral lens accelerator" of the General Electric Company [38] [Fig. 6(d)]. This design uses, in combination, a nonlinear spiral anode, acting as a negative lens, and an equally long drift tube after deflection at low voltage. With this combination, linear beam, propagation has been realized, resulting in a relatively high deflection factor (3.2 v/cm·kv). It has been found that the scanning width is practically stable over a screen voltage range from 1–22 kv.

4) Special Oscilloscope Tubes: Many special cathode-ray tubes for oscillography, including multibeam, switching, and low-voltage types, are listed in the recent reference book by Parr [25]. At present, advanced development is active mainly in these areas:

a) Ultra-high frequency
b) Ultra-high resolution
c) Storage-type oscilloscopes
d) Printing CRT's.

Space does not permit mention of more than a few examples.

a) and b) UHF oscilloscopes: Very broad-band oscilloscopes are needed for the observation of millimicrosecond pulses [39]. These tubes must have no deflection cutoff up to several hundred megacycles. Unfortunately, transit time limitations restrict the use of conventional "lumped" deflecting fields to frequencies not exceeding the lower VHF range [40]. Means to reduce transit time may include higher beam voltage (50 kv) and very short deflecting fields [41], [42]. Both approaches are hampered by loss of deflection sensitivity. Some relief is forthcoming from gun designs for ultra-fine resolution, whereby the deflection sensibility [43] is increased. As an example, a spot size of 0.0005 in, or 12 μ, was achieved by G. M. Lee in his "micro-oscilloscope" as early as 1946 [44]. Unfortunately, the use of optical [45] or electron-optical [46] magnification becomes necessary for observation of such microtraces.

In 1949, J. R. Pierce suggested the use of a distributed rather than lumped deflection system, whose individual segments were connected to a delay line with a phase velocity matching the beam velocity [47] (Fig. 7). This "traveling-wave" approach to CRT deflection was then further developed by Owaki [48], Smith [49], and others who actually incorporated a slow-wave structure into the cathode-ray tube. Germershausen [50] and Niklas [51] developed similar TW oscillographs for larger deflection angles. At present, the frequency range of the cathode-ray oscillograph has been extended through the entire UHF spectrum [46] and reaches far into the microwave range [52]. A recent report by Feinberg [53] shows the photograph of a magnetron oscillation at 9500 Mc. Similar results were also reported by Haines and Jervis in 1957 [52]. At present, the bandwidth capabilities of the CR oscillograph are far ahead of the availability of suitable signal amplifiers.

c) Storage-type oscilloscope: The introduction of picture retention is of great value in the oscillography of transients. With this in mind, several types of long-persistence screen materials have been developed [54]. These include the well-known cascaded double-layer screen [55], the "dark trace" KCL screens [55], [25], and also a new type of phosphors, whose persistence may be controlled by infrared light to reveal a latent image [56], [53]. A more ideal solution is offered by oscilloscope tubes with built-in electrostatic storage. Fig. 8 shows the "Memotron" of the Hughes Aircraft Company. This tube uses typical storage tube techniques (see Section B, below), including a dielectric mesh, for the transmission control of a flood beam. It writes at 100,000 inches per second, resolves 60 lines per inch, and retains scope traces at a brightness level of 50 foot-lamberts, until they are intentionally erased.

d) Printing with cathode-ray tubes: In most recent times, increasing attention is being given to the problem of obtaining lasting recordings by other than photographic methods. Recent approaches have included the development of several "printing" types of cathode-ray tubes, using either a thin mica window [57] or a metallic mosaic stripline target with over 60,000 wires per inch [58]. A high-resolution line scan in the tube will produce a charge image on moving paper outside. Pigmented plastic, charged tribo-electrically, can be used to "develop" this latent image. Other approaches have used xerographic methods [59] or paper with photoconductive surfaces [60]. Fiber optics [61] is being tried to improve optical coupling between tube and film. CRT's have been used to record on magnetic film [62] and also to detect such recordings [63], [64].

Thermoplastic recording (TPR) is a promising new method of recording with electron beams. This system

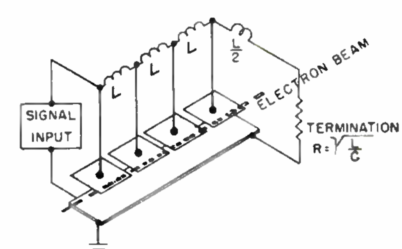


Fig. 7—Traveling-wave deflection (Pierce).

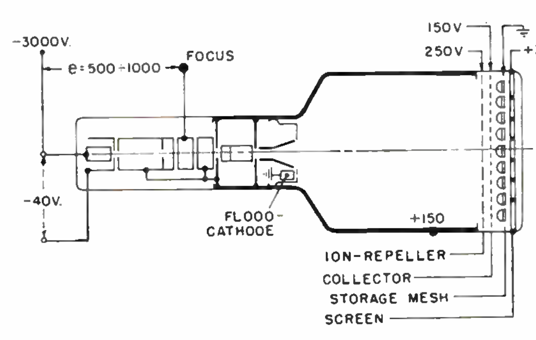


Fig. 8—Memotron (Hughes Aircraft Corp.).

has been described by W. E. Glenn of the General Electric Research Laboratory [65]. TPR uses a clear-film substrate charged by a microspot tube of ultra-fine resolution and read out by means of a "schlieren" optical system.

B. Storage Tubes

Storage tubes are multiple-beam tubes with built-in electrostatic memory [66], [67]. One beam writes the video signal into the storage target as fast as it is received. Another beam reads the charge-image in one or many recurrent frames, but not necessarily at the same time, nor at the same scanning rate (scan converter tubes).

In another type of storage tube, the stored charge image is made directly visible on a built-in viewing screen. These "direct view" storage tubes, or "bright displays," employ a nondeflected flood of slow electrons to project the entire stored image on a built-in viewing screen.

Historically, the addition of electrostatic memory to a cathode-ray tube was conceived long ago by some television pioneers, either in search of brighter pictures [68] or of bandwidth reduction [69]. However, the major development effort in storage tubes had to wait until the end of World War II [70]–[72]. At that time, radar developments had matured and had established a definite need for displays with practically no degeneration for the order of one second. Long-persistence phosphors could not meet this demand. Storage tubes have contributed, not only an improvement of three orders of magnitude in the brightness of radar displays [73], but also an apparent reduction in background noise [74], [75] and a new facility in the detection of moving targets [78].

1) Principles of Operation: The basic memory element in most storage tubes is the dielectric mesh. This is a fine, metallic screen with high transparency (about 50 per cent) and high resolution ($\frac{1}{4}$ to $\frac{1}{2}$ million holes per square inch). The first surface of the storage mesh is coated by a dielectric film with high secondary emission capabilities. Suitable materials include silica, anodized aluminum, the oxides or fluorides of magnesium, etc.

Fig. 9(a) shows a typical secondary emission characteristic of the storage mesh insulator. The two crossover points e_1 and e_2 for unity gain define three areas, *A*, *B*, and *C*, with three typical modes of operation, to be summarized below:

Mode A: Beam voltage below e_1 (approximately 100 v). At these slow velocities, the insulator receives more electrons than it can emit. The surface potential moves in a negative direction until it is "stabilized" at cathode potential. Mode A can be used for total, uniform erasure, and also for uniform "priming" of the target potential.

Mode B: Beam velocity is 1–2 kv, *i.e.*, intermediate between crossovers. High secondary emission causes the target potential to seek the potential of the collector grid.

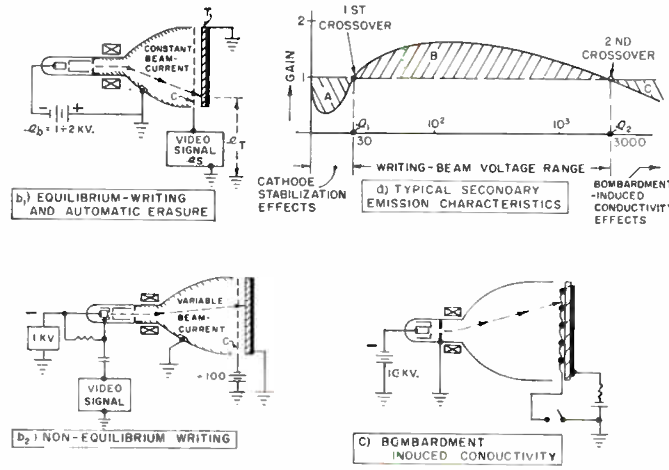


Fig. 9—Storage tube principles.

b_1: *Equilibrium writing:* Fig. 9(b_1) shows writing with constant current-variable collector voltage (*C* is on signal e_s). This process is bidirectional, has limited halftone capability and permits automatic erasure of previous information [70].

b_2: *Nonequilibrium writing:* Fig. 9(b_2) shows writing with constant voltage but variable current. The beam is signal-modulated. The target must be cathode stabilized, or "primed," prior to writing. Secondary emission drives the surface positive at a rate determined by instantaneous beam current. Nonequilibrium writing is the preferred mode of operation, since it offers a good halftone rendition and high speed.

Mode C: At very high beam voltages (6–10 kv), the beam interacts with the target material. Many hole-electron pairs are created in the dielectric by each primary electron. Bombardment induced conductivity (BIC) may cause currents in excess of beam current to flow as a field exists across the target [76].

BIC was first used by L. Pensak (RCA) in his "Graphecon" [77]. This pioneering tube used a film-target, rather than a dielectric mesh. The film was about 1μ thick and was scanned by a high-velocity (10-kv) beam from one side, and by a 1-kv beam from the other. The 10-kv beam was writing by relaxing a uniform charge through BIC. The 1-kv beam was used for reading and also for priming of the target, both in Mode B. The depth of penetration which is proportional to the square of beam voltage accounted for the difference in target behavior in response to the two beams.

2) Electrical Output Types: Of the large family of modern storage tubes, only four typical examples are shown in Figs. 10–13. Fig. 10 is the "Radechon" by A. S. Jensen (RCA, 1948) [78], [79], one of the earliest storage tubes for electrical output. Redistribution effects are reduced by the use of a "barrier grid," *i.e.*, a high-resolution mesh placed in direct contact with the storage surface. Writing can be done by either one of the two above modes in range B. Signal output is taken from the metallic target substrate. The number of readout frames is limited by incidental erasure (readout-decay).

Even so, the "Radechon" principle has found many important applications, mostly as a noise integrator [74] and also for the detection of moving targets by radar (MTI service). Here, the "Radechon" is used as a "frame-subtractor." It can be differentially connected so as to give zero output on identical frames [78]. Recent improvements of the "Radechon" include a high-resolution, 60° version [80], and also an adaptation to a high-speed, random-acess store for digital information [81].

Fig. 11 shows a modern scan converter of the Raytheon Company [82]. This tube is double-ended, with two high-resolution beams (1200 lines) scanning each side of a dielectric mesh. Sets of conductive bands on the envelope secure proper beam landing and deceleration. The reading beam is transmission-modulated and passes through the storage mesh at very low voltage (5–15 v). Thus, readout is practically nondestructive and up to 20,000 good halftone output frames can be read following a single writing frame. Simultaneous recording and reading without crosstalk is made possible by subcarrier modulation of the read-gun at 30 Mc. This separates the sidebands of the output signal from the video spectrum of the write-in signals.

3) *Visual Output Types:* Fig. 12 shows a display storage tube [83], [84]. In this particular version, the writing gun is on the tube axis and the floodbeam gun has a concentric, annular geometry. The low-voltage floodbeam is first collimated by a system of slotted anodes and accelerating rings, and is then transmission modulated by the storage mesh. With screen voltages of 8500 v, this tube can give 4000 foot-lamberts on a $2\frac{1}{2}$-inch screen. It is designed as a cockpit display in high ambient illumination.

Other developments in this area include viewing storage displays of large size [85], [86], a storage tube for the display of printed characters [87], and a large display for a stored tricolor image [88].

Fig. 13 shows a most recent development in the storage tube field, the "Multimode Tonotron" by N. H. Lehrer (Hughes Aircraft Company) [89]. This tube uses erasure by BIC. As Fig. 13 shows, the "Multimode Tonotron" has three electron guns, one for the low-voltage floodbeam, and the two others for nonequilibrium writing at 2 kv and for erasure by BIC at 6 kv. Since the erase beam is focused, erasure can be affected over selected areas, with high speed and with good uniformity of background. By varying the erase beam energy, BIC can be made to swamp secondary emission effects, or just to cancel them. At balance, the erase gun can be used to write nonstored information. At overbalance, dark-trace writing on a white background is practical. These many new capabilities, including a good, six-step, halftone rendition, mark the "Multimode Tonotron" as a major advance in storage tube technology.

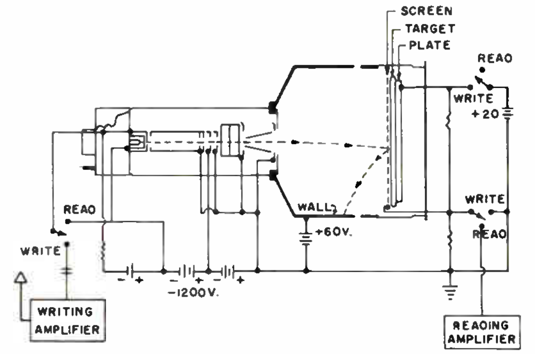


Fig. 10—Radechon tube (RCA).

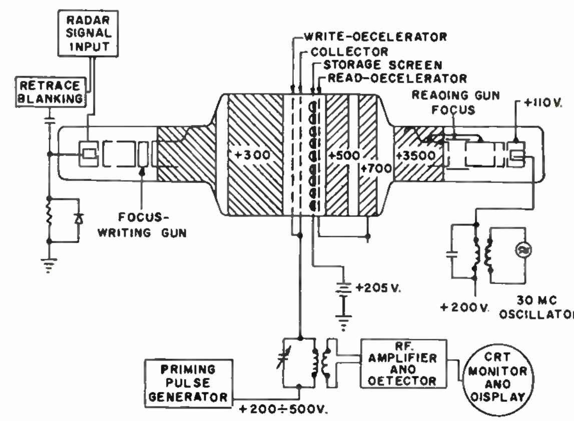


Fig. 11—Recording storage tube (Raytheon).

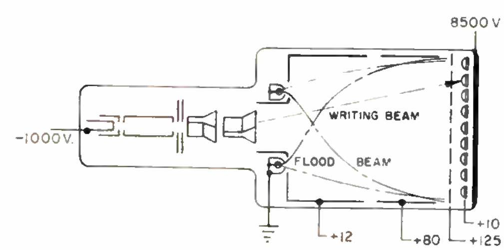


Fig. 12—Iatron storage display tube (ITT Labs.).

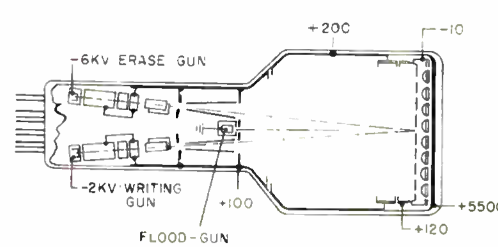


Fig. 13—Multimode Tonotron (Hughes Aircraft Corp.).

C. *Picture Display Tubes*

1) *Monochrome Tubes:* Modern cathode-ray tubes for picture display are characterized by relatively strong beam currents and by very wide deflection angles. Both features are a consequence of the large picture size and the shallow depth required by modern television sets.

The standard P-4 phosphor (ZnS-CdS, Ag) used in picture tubes has an average efficiency of 30 lumens per watt, or about five per cent quantum efficiency. On this basis, a simple calculation shows that the display of a 21-inch picture area at 200 foot-lamberts brightness requires 600 μa of focused beam current at 16 kv.

A highlight brightness three times as great has been achieved recently in picture tubes of the same size [102] by a corresponding increase in beam current. With other conditions equal, beam current capability of the gun should increase with picture area. The latter has grown by one order of magnitude during the past decade.

During the same period, similar progress was made in picture tube deflection techniques. The diagonal sweep angle rose from 52° to 110° and beyond. Such very wide scanning angles were realized in field-free envelopes, that is, without the use of post-acceleration, and also without any but magnetic methods of deflection.

The development of electron guns for picture tubes until about 1955 is very well outlined by Zworykin and Morton [90]. For progress since then, only a few typical examples can be presented here. Fig. 14(a) shows a kinescope gun as used in picture tubes with magnetic focus. This is a "tilted-barrel" type of gun which requires only a single, rather than a dual, magnetic ion trap [91]. A similar gun, but with adjustable spot size, is reported by Francken, *et al.* [92]. In these early tubes, in view of the limited effect of magnetic deflection on ion beams, ion traps were widely used to prevent ion spot formation on unaluminized screens and to protect the cathode from positive-ion bombardment.

However, later experience has shown that, as a result of improved aluminizing, exhaust, and gettering techniques, long life without ion-burn can be achieved in a straight gun, without an ion trap. This is all the more noteworthy, since magnetic focus has yielded, in the meantime, to focusing by electrostatic lenses, which are effective on ions.

The gun assembly of a modern picture tube is shown in Fig. 14(b). It features a unipotential focus lens of the "zero-voltage" type [93]. This type of lens obviates the need for a separate bleeder string across the high-voltage supply. To do so, manufacturing tolerances in the lens must be kept within 0.002 in [94].

Another feature of Fig. 14(b) is the re-entrant high-voltage barrel, which protrudes deeply into the Grid No. 2 cup. The resulting very low screen-grid voltage (+50 v) leads to less stringent cutoff and drive signal requirements. In this gun, a 30-v signal modulates 900 μa of beam current.

Circuit economy considerations as well as the trend towards transistorization have recently stimulated efforts to increase picture tube transconductance by electron-optical means [95]. Notable progress has been made in this direction, both here [96]–[99] and abroad [101]. Many of these low-drive guns employ planar grids made of extremely fine wire [100].

Fig. 14(c) shows a more recent type of high-transconductance gun which has been developed for the U. S. Signal Corps by the General Electric Company [102]. This gun uses a retarding-field type of hyperbolic electron lens to image a large cathode surface into a much smaller virtual point cathode, where the current density is greatly increased. Current from the virtual cathode is drawn forward through a spot defining aperture, or it is reflected back to the first anode. This process is called "Focus Reflex Modulation" (FRM). It modulates 1600-μa screen current by an 18-v signal, and maintains a more constant spot size than conventional guns.

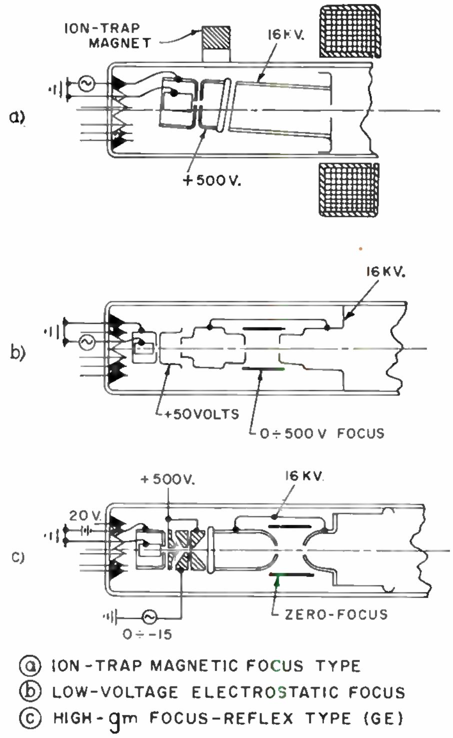


Fig. 14—Electron guns for picture tubes.

2) Flat Picture Tubes: There is, at present, a growing interest in developing "very-thin" picture tube displays [103]. Most of these approaches include radical departures from conventional scanning techniques. This is true of the Gabor tube [104] and also of the Aiken tube, [105] whose complex electron optics has been analyzed by Ramberg [106]. Magnetic mirrors for beam folding have also been proposed [107] and were demonstrated in the laboratory [103].

A very interesting new "flat tube" is the "Reflected-Beam Kinescope" by Law and Ramberg [108]. This tube gives a large picture (21-in diameter), has only 10-in depth, and permits a considerable saving in sweep power and space, particularly with large picture sizes.

3) Color Tubes: As a display device for color television, the cathode-ray tube has risen to the greatest challenge in its history! Early attempts to solve this problem are well-documented by E. W. Herold [109]. At present, the field is dominated by the triple-gun shadow mask tube [110] which produces an excellent color picture of large size (21 in) at medium brightness levels (25 foot-lamberts). In this tube, the three electron beams modulated by the red, green, and blue component picture signals strike a perforated mask with some 357,000 round apertures, which transmits only those electrons of every beam which are directed toward a phosphor dot of the corresponding color on the phosphor screen. The latter is deposited directly on the curved face-plate of the tube [111]. A photographic procedure for forming the screen, employing optical projection through the mask from three appropriately chosen light-source positions, assures precise alignment between the mask apertures, the phosphor dots, and the deflection centers of the scanning beams. An intensive development of manufacturing technology by RCA has resulted in color tubes characterized by consistently high color purity and simple adjustment [112a]–[112c].

The progressive improvement of the shadow mask tube has been paralleled by the investigation of other designs calculated to increase the picture brightness or to simplify the tube construction. Most attempts to realize high-intensity color have utilized the "post-focusing" principle, or the converging lens action at the apertures or slits in the mask structure, to increase the beam transmission of the mask [113a]–[113d], a principle applied by the late E. O. Lawrence [113e] among others. In a relatively recent demonstration by the General Electric Company, a brightness of 80 foot-lamberts was realized on a 22-in picture in this manner [114]. The attempt to eliminate the mask structure altogether by directing a very narrow scanning beam to the appropriate color phosphor elements with the aid of signals elicited by a guide beam and to simplify the tube construction in this manner has been carried forward in particular by Philco and has resulted in the so-called "Apple" system [115]. A still more radical departure from the conventional color tube is represented by the "Banana Tube" system recently announced by Mullard [116]. Here, a projected picture is formed employing a beam tube for scanning a color-line triplet and a rotating lens cylinder for frame scan.

IV. Photo Devices

A. *Photoemissive, Photoconductive, and Photovoltaic Cells*

The 50 years which have elapsed since the founding of the IRE lead from the introduction of the first useful vacuum phototubes, the hydrogenated alkali phototubes of Elster and Geitel [118] with maximum quantum efficiencies in the visible of the order of 1 per cent, to the multi-alkali phototubes with quantum efficiencies up to 40 per cent. The most important steps in this development are the introduction of silver-oxygen-cesium photocathodes by L. R. Koller [119] in 1929, sensitive in the entire visible spectrum as well as the near infrared; the preparation of alkali-antimony and alkali-bismuth films, typified by the antimony-cesium photocathode with quantum efficiencies up to 25 per cent, by P. Goerlich [120] in 1936; and finally, the introduction of the multi-alkali compounds by A. H. Sommer [121] in 1955.

Characteristic curves for the variation of the photoemission with the wavelength of the exciting light are shown, for various commercially important photocathodes, in Fig. 15. While the semicylindirical photocathode and rod anode remain common features of the simple phototube, the introduction of the semitransparent photocathode has aided the solution of special design problems. Gas amplification by collision-ionization of a low-pressure noble gas filling of the phototube, a process known since the beginning of the century, finds only limited application.

The developments in solid-state photocells in the 50 years which have elapsed since the founding of the IRE are perhaps even more striking. Though the discoveries of the photovoltaic and photoconductive effects date back to 1839 and 1873, respectively, the former came into wide use only after the discovery of the cuprous oxide cell by Grondahl [122] in 1926 and of the reintroduction of the selenium photovoltaic cell by Lange and others [123], [124] in 1931, principally as photographic exposure meters. Photoconductive cells have found, until recently, their major application in the detection of infrared radiation. The thalofide cell of Case [125], introduced over 40 years ago, was the pioneer in this field. It was followed by the lead salt detectors (PbS, PbSe, PbTe) first investigated in detail in Germany [126], the improved thallous sulfide cells of Cashman [127], and, more recently, by indium-antimonide and tellurium [128] as well as impurity-doped silicon-germanium cells [129]. Another photoconductive material, cadmium sulfide, initially introduced by Frerichs [130] in 1947, as well as the closely related cadmium selenide, have provided cells of very high sensitivity in the visible—10^5 that of ordinary phototubes—particularly well suited for light-operated relay applications [131].

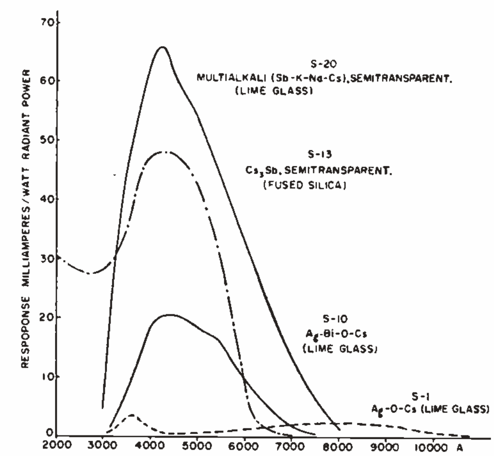


Fig. 15—Spectral response of typical commercial photocathodes. (After R. W. Engstrom, "Absolute spectral response characteristics of photosensitive devices," *RCA Rev.*, vol. 21, pp. 184–190; June, 1960.)

Apart from the photoconductive cell, single-crystal photovoltaic cells [132], commonly called photodiodes or phototransistors when provided with a bias in the blocking direction, and photoelectromagnetic (PEM) cells [133], provide effective radiation detectors. The principle of operation of the three types of cells is illustrated in Fig. 16. Neither photovoltaic cells nor PEM cells require an applied voltage. Negative electrons and positive holes generated by light are separated in the photovoltaic cell by the internal field of a *p-n*-junction, and in the PEM cell by a transverse magnetic field, and give rise to the cell potential or a current in an external circuit. Their output currents are relatively small (less than one electron per absorbed light quantum) and their response is rapid. In photoconductive cells, an applied field serves to separate the mobile charges. However, one of them (*e.g.*, the hole) is commonly trapped for a prolonged time in a recombination center and, by space-charge compensation, permits the passage of many electrons through the cell before its lifetime is terminated by recombination. Thus, photoconductive cells commonly deliver very large photocurrents and exhibit correspondingly slow response [134]. They are thus well suited for light relay operation. The high speed of response, favorable noise characteristics, and extraordinary compactness make the germanium photodiode valuable for sound pickup from film as well as for computer applications. Thus, cells of this type are employed in electronic route information translators in long-distance telephony [135].

Modern photovoltaic cells satisfy one other important objective: the conversion of solar energy with reasonable efficiency [136]–[138]. Thus, they provide an inexhaustible power source for the instrumentation of space craft. Whereas the usual exposure-meter selenium cell converts less than 1 per cent of the incident radiant energy into electrical power, single-crystal *p-n*-junction photocells have achieved conversion efficiencies of 14 per cent, not too far removed from the theoretical maximum of 24 per cent for a simple single-junction photocell.

B. Multiplier Phototubes

One of the most important by-products of television research is the multiplier phototube. When electrons of one or several hundred electron volts energy impinge on a suitably prepared conducting surface, they eject 4 to 10 low-velocity electrons, multiplying the initial current by a corresponding factor of 4 to 10. Repetition of this process leads to current multiplication by an arbitrarily high factor, practically without the addition of amplification noise. If the initial current is derived from a photocathode, the tube output reflects the variation of the light incident on the photocathode with a precision which depends only on the quantum efficiency of the cathode; with proper design, the dispersion in the transit time of the electrons from the cathode to the final collector can be held to quantities of the order of 10^{-10} second [139].

Several multiplier designs are sketched in Fig. 17. In the earliest effective multipliers (Zworykin, Morton,

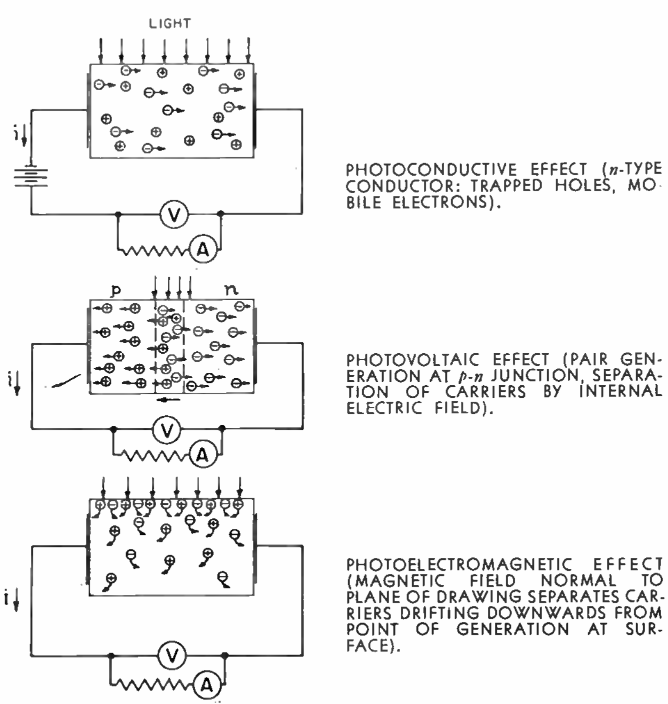


Fig. 16—Internal photoelectric effects.

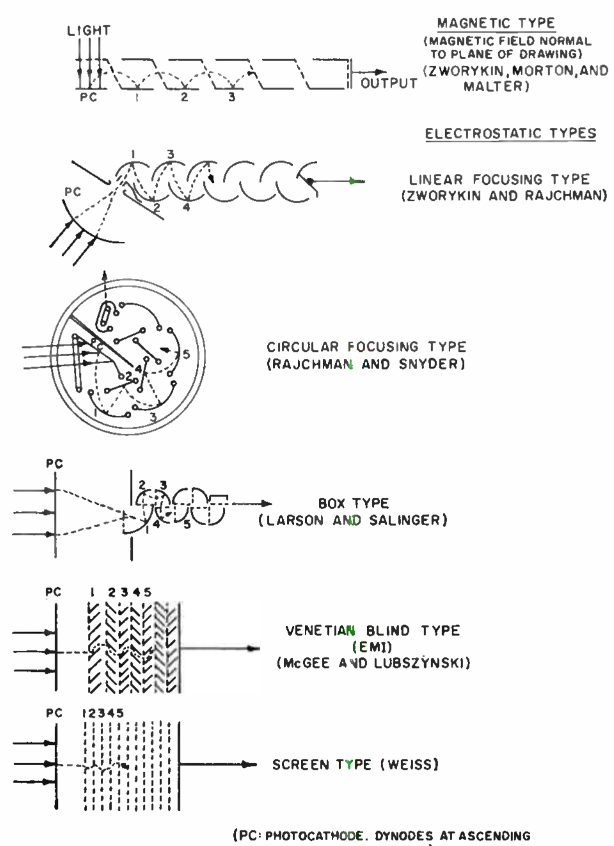


Fig. 17—Basic multiplier phototube structures.

and Malter, 1935) [140], the electrons were guided from dynode to dynode along an approximately cycloidal path by crossed electric and magnetic fields. Purely electrostatic focusing and acceleration sytems were developed subsequently by Zworykin and Rajchman [141] and Rajchman and Synder [142], as well as by Larson and Salinger [143]. These may be regarded as the prototypes of present-day multiplier phototubes of RCA and DuMont. The venetian-blind design utilized in the image orthicon is also employed by EMI and RCA for multiplier phototubes. Finally, the early and very simple screen multiplier of Weiss (1936) [144] does without focusing altogether, at the expense of materially lowered multiplication efficiency.

The multiplier phototube finds application wherever high sensitivity and high speed of response must be combined in the measurement of light values. It has become indispensable in the scintillation counter [145], [146] for high-energy particles and gamma rays. Here, the particle or quantum excites fluorescence in a crystalline or liquid medium, and as large a portion of this light as possible is made to fall on the photocathode of a multiplier phototube. Since every particle excites many light quanta in the phosphor, the multiplier phototube is not only capable of counting every absorbed particle, but provides in addition the possibility of classifying the particles according to energy as measured by the height of the output pulse. Finally, the high time resolution of the multiplier phototube permits the determination of coincidences.

C. Image Tubes

The image tube, also, owes its origin largely to the development of electronic television. It consists essentially of an extended semitransparent photocathode, an electron lens system for forming an image of the photocathode by means of the photoelectrons emitted by it, and a fluorescent screen for rendering this image visible. Tubes of this type, with combined electric and magnetic fields, were built in the middle thirties by Pohl [147] and Schaffernicht [148] in Germany and, with purely electrostatic lenses, by Zworykin and Morton [149] in the United States.[6] Most of the later image tubes have, also, been electrostatic. Image tubes have served principally four different functions:

1) The viewing of objects in illumination to which the eye is insensitive; apart from its obvious military applications [151], the infrared-sensitive image tube has been found useful in ophthalmic diagnosis [152] and in photographic film processing. Beyond the other end of the visible spectrum, the RCA Ultrascope image tube [153] converts the ultraviolet microscope into a direct-viewing instrument.
2) The utilization of the higher quantum efficiency of the photocathode as compared with that of a photographic emulsion in astronomic photography; this is accomplished for example in the direct photography on fine-grain plates within the image tube by A. Lallemand [154].
3) Brightness enhancement of low-level images. Thus, X-ray intensifier tubes [155], [156] introduced by J. W. Coltman and manufactured by Westinghouse, Philips, RCA, and others achieve, through electron acceleration and image size reduction, brightness enhancement by factors up to 1000, overcoming losses in detail recognition arising from the inefficient utilization of light by the human eye in normal viewing.
4) Electronic shutter action for high-speed photography [157], [158]. The provision of a gating grid in front of the photocathode permits photography with exposure times of the order of 10^{-8} sec.

In the examples cited above, single-stage image tubes have been used throughout. However, it has been found possible to greatly increase the brightness gain by staggering stages. Of the several types which are under investigation, the fluorescent-screen-photocathode sandwhich type (Morton and Ruedy [19] and Stoudenheimer [159]) and the transmission secondary-emission type as realized by W. L. Wilcock [160] have already attained excellent results. The former consists essentially of a sequence of complete image tubes. Good optical coupling is achieved by depositing the fluorescent screen of one stage and the photocathode of the next on the two sides of a thin glass membrane. In the

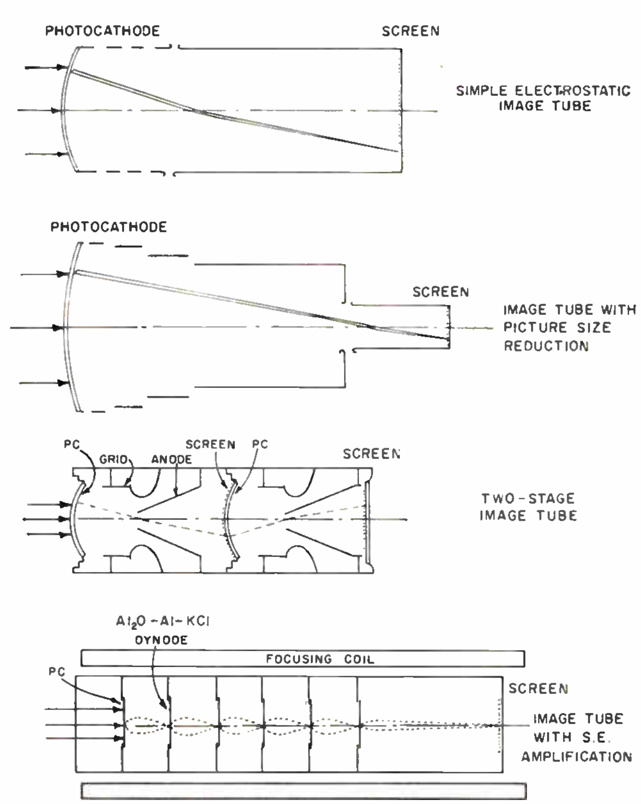


Fig. 18—Image tube structures.

[6] Image tubes without electron imaging were built by Holst, de Boer, Teves and Veenemans [150] in The Netherlands in 1934.

secondary-emission tube, a longitudinal magnetic field images the emission from one dynode, consisting of a thin aluminum—aluminum-oxide film coated with potassium chloride, to the next, achieving secondary-emission gains of the order of 5 with 1000 v between stages. Fig. 18 shows schematically a number of image tube structures.

D. Light Amplifiers

The solid-state counterpart of the image tube is the light amplifier [161]. In its simplest form, it consists of a sandwich of a powdered or sintered photoconductive layer and an electroluminescent layer separated by a light barrier to prevent feedback (Fig. 19). An alternating voltage is applied across the sandwich. In darkness, the potential drop across the electroluminescent layer is reduced by that across the photoconductive layer to a value below the threshold for effective light excitation. When the photoconductive layer is irradiated, its impedance drops, the potential difference across the electroluminescent layer increases, and light is emitted by it. With certain refinements, such light amplifier panels have yielded light gains in excess of 100. Furthermore, when employed for X-ray fluoroscopy [162], the image brightness of the panel could be made over 100 times as great as that of a conventional X-ray screen. Unfortunately, the slow response of the cadmium-sulfide photoconductor layers employed so far has limited the utility of this application. However, both experimental and theoretical work indicates that great improvements in this respect are obtainable.

Apart from the simple light amplifier, the combination of photoconductors and electroluminescent elements permits the construction of a wide range of storage and logical devices covered by the collective term optoelectronics [163]. This may well be one of the fields in which photo devices will demonstrate their expanding utility in the years to come.

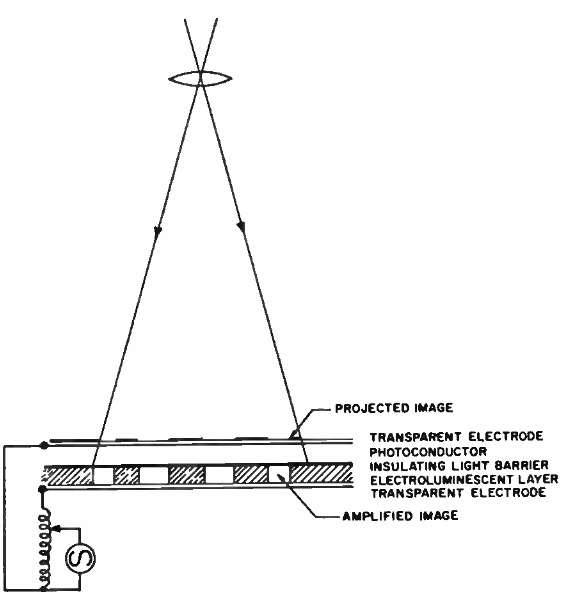


Fig. 19—Principle of the solid-state light amplifier.

Bibliography

[1] F. Braun, "On a method for the demonstration and study of currents varying with time," *Wiedemann's Ann.*, vol. 60, pp. 552–559; 1897.
[2] J. J. Thomson, "Cathode rays," *Phil Mag.*, vol. 4, pp. 293–316; 1897.
[3] E. Becquerel, "On electric effects under the influence of solar radiation," *Compt. rend. acad. sci.*, vol. 9, p. 561; 1839.
[4] W. Smith, "Effect of light on selenium during the passage of an electric current," *Am. J. Sci.*, vol. 5, p. 301; 1873.
[5] H. Hertz, "Ultraviolet light and electric discharge," *Ann. Physik*, vol. 33, pp. 301–312; 1888.
[6] A. Einstein, "A heuristic standpoint concerning the production and transformation of light," *Ann. Physik*, vol. 17, pp. 132–148; 1905.
[7] P. T. Farnsworth, "Television by electron image scanning," *J. Franklin Inst.*, vol. 218, pp. 411–444; October, 1934.
[8] V. K. Zworykin, U. S. Patent No. 2,141,059 (December 20, 1938); filed, December 29, 1923.
[9] M. von Ardenne, "The Braun tube as television receiver," *Fernsehen*, vol. 1, pp. 193–202, May, 1930; vol. 2, pp. 65–68, 173–178, April–July, 1931.
[10] A. A. Campbell-Swinton, "The possibilities of television," *Wireless World and Radio Rev.*, vol. 14, pp. 51–56; April 9, 1924.
[11] V. K. Zworykin, "The iconoscope—a modern version of the electric eye," Proc. IRE, vol. 22, pp. 16–32; January, 1934.
[12] J. D. McGee and H. D. Lubzynski, "EMI cathode-ray television tubes," *J. IEE*, vol. 48, pp. 468–475; 1937.
[13] H. Iams, G. A. Morton, and V. K. Zworykin, "The image iconoscope," Proc. IRE, vol. 27, pp. 541–547; September, 1939.
[14] J. E. Cope, L. W. Germany, and R. Theile, "Improvements in design and operation of image iconoscope type camera tubes," *J. Brit. IRE*, vol. 12, pp. 139–149; March, 1952.
[15] A. Rose and H. Iams, "The orthicon, a television pickup tube," *RCA Rev.*, vol. 4, pp. 186–199; October, 1939.
[16] J. D. McGee, "Distant electric vision," Proc. IRE, vol. 38, pp. 596–608; June, 1950.
[17] A. Rose, P. K. Weimer, and H. B. Law, "The image orthicon—a sensitive television pickup tube," Proc. IRE, vol. 34, pp. 424–432; July, 1946.
[18] P. K. Weimer, S. V. Forgue, and R. R. Goodrich, "The vidicon photoconductive camera tube," *Electronics*, vol. 23, pp. 70–73; May, 1950.
[19] G. A. Morton and J. E. Ruedy, "The low-level performance of the intensifier orthicon," in "Advances in Electronics and Electron Physics," L. Marton, Ed., Academic Press, Inc., New York, N. Y., vol. 12, pp. 183–193; 1960.
[20] —— and S. V. Forgue, "An infrared pickup tube," Proc. IRE, vol. 47, pp. 1607–1609; September, 1959.
[21] S. Taylor, "An infrared-sensitive television camera tube," in "Advances in Electronics and Electron Physics," L. Marton, Ed., Academic Press, Inc., New York, N. Y., vol. 12, pp. 263–275; 1960.
[22] A. Wehnelt, "The emission of negative ions from hot metal compounds and related phenomena," *Ann. Physik*, vol. 14, pp. 425–468; 1904.
[23] J. B. Johnson, "A low voltage cathode ray oscillograph," *Phys. Rev.*, vol. 17, pp. 420–421; March, 1921.
[24] M. von Ardenne, "Die Kathodenstrahl-Roehre," Julius Springer, Berlin, Germany; 1933.
[25] G. Parr and O. H. Davie, "The Cathode Ray Tube," Reinhold Publishing Co., New York, N. Y.; 1959.
[26] E. Bruche and A. Recknagel, "Elektronen-Geraete," Julius Springer, Berlin, Germany, 1941.
[27] V. K. Zworykin, "Television through a crystal globe," *Radio News*, vol. 11, pp. 905, 949, 954; April, 1930.
[28] V. K. Zworykin, "Television with cathode-ray tubes," *J. IEE*, vol. 73, pp. 437–451; October, 1933.
[29] H. Busch, "Calculation of the path of cathode rays in axially symmetric electromagnetic fields," *Ann. Physik*, vol. 81, pp. 974–993; December, 1926.
[30] L. E. Swedlund, "Improved electron gun for CR tubes," *Electronics*, vol. 18, pp. 122–124; March, 1945.
[31] K. Schlesinger, "Internal electrostatic deflection yokes," *Electronics*, vol. 25, pp. 105–109; July, 1952.
[32] ——, "Progress in the development of post-acceleration and electostatic deflection," Proc. IRE, vol. 44, pp. 659–667; May, 1956.
[33] H. Albertine, "Electrostatic Focus and Deflection Vidicon," presented at 1959 annual mtg. of the IRE-PGED, Washington, D. C.; October 29–31, 1959. Abstract: IRE Trans. on Electron Devices, vol. ED-7, p. 113; April, 1960.
[34] K. Schlesinger, B. Maggos, and A. F. Hogg, "An electrostatic character writing tube," 1957 IRE National Convention Record, pt. 3, pt. 3, pp. 160–168.

[35] G. C. Sponsler, "Trajectory traces study of helix and band type PDA," *J. Appl. Phys.*, vol. 26, pp. 676–682; June, 1955.
[36] E. Gundert and W. Schaffernicht, "New oscillograph tubes of high brightness," *Elektron. Rundschau*, vol. 11, no. 4, pp. 107–108; April, 1957.
[37] L. S. Allard, "An ideal post-deflexion accelerator CRT," *Electronic Engrg.*, vol. 22, pp. 461–463; November, 1950.
[38] K. Schlesinger, "Electron Trigomometry—a New Tool for Electron-Optical Design," presented at the Fall IRE-PGED Meeting, Washington, D. C.; October 27–28, 1960. Also, PROC. IRE, vol. 49, pp. 1538–1549; October, 1961.
[39] I. A. D. Lewis and F. H. Wells, "Millimicrosecond Pulse Techniques," Pergamon Press, New York, N. Y.; 1959.
[40] H. E. Hollmann, "CRT at UHF," *Wireless Engr.*, vol. 10, pp. 430–433; August, 1933.
[41] H. G. Rudenberg, "Deflection sensitivity of parallel wire lines," *J. Appl. Phys.*, vol. 16, pp. 279–284; May, 1945.
[42] T. H. Soller, M. Starr, and G. E. Valley, "Cathode Ray Tube Displays," M.I.T. Rad. Lab., Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 22, p. 74, Fig. 2; 1948.
[43] J. R. Pierce, "Theory and Design of Electron Beams," D. Van Nostrand Co., Inc., New York, N. Y., 2nd ed., sect. 8.7, pp. 137–143; 1954.
[44] G. M. Lee, "A three-beam oscillograph for recording at frequencies up to 10,000 megacycles," PROC. IRE, vol. 34, pp. 121W–127W; March, 1946.
[45] M. von Ardenne, "Elektronen-Mikro-Oscillograph," *Hochfrequenz-Tech. u. Elektro-Akustik*, vol. 54, pp. 181–188; December, 1939.
[46] D. J. Brangazzio, A. F. Dietrich, and J. W. Sullivan, "Broadband oscilloscope tube," *Bell Sys. Tech. J.*, vol. 37, pp. 447–460; March, 1958.
[47] J. R. Pierce, "Traveling wave oscilloscope," *Electronics*, vol. 22, pp. 97–99; November, 1949.
[48] K. Owaki, *et al.*, "The traveling-wave cathode-ray tube," PROC. IRE, vol. 38, pp. 1172–1180; October, 1950.
[49] S. T. Smith, R. V. Talbot, and C. H. Smith, "Cathode-ray tube for recording high-speed transients," PROC. IRE, vol. 40, pp. 297–302; March, 1952.
[50] K. J. Germershausen, S. Goldberg, and D. F. McDonald, "A high-sensibility cathode-ray tube for millimicrosecond transients," IRE TRANS. ON ELECTRON DEVICES, vol. ED-4, pp. 152–158; April, 1957.
[51] W. F. Niklas and J. Wimpffen, "Oscilloscope tube with traveling wave deflection system and large field of view," *J. Brit. IRE*, vol. 18, pp. 653–660; November, 1958.
[52] M. E. Haine and M. W. Jervis, "Ultimate performance of high speed oscillograph," *Proc. IEE*, pt. B, vol. 104, pp. 379–384; July, 1957.
[53] R. Feinberg, "New oscillograph tubes and their application in power and communication engineering," *Z. Elektro-Tech. u. Maschinenbau*, vol. 76, pp. 373–378; August, 1959.
[54] H. W. Leverenz, "Phosphors versus the periodic system of the elements," PROC. IRE, vol. 32, pp. 256–263; May, 1944.
[55] Soller, *et al.*, *op. cit.* [42], sect. 18, pp. 609–699.
[56] M. D. Dudley, "Phosphors for CRT's," *J. Brit. IRE*, vol. 19, pp. 485–489; August, 1959.
[57] R. G. Olden, "A thin window CRT for high speed printing with electrofax," *RCA Rev.*, vol. 18, pp. 343–350; September, 1957.
[58] "Printapix" Tube, Type L-4013/4016, Litton Industries, San Carlos, Calif.; 1960.
[59] R. M. Peterson and R. C. Ritchart, "Recent developments in shaped beam display and recording techniques," 1958 IRE NATIONAL CONVENTION RECORD, pt. 3, pp. 21–30.
[60] J. A. Amick, "A review of 'electrofax' behavior," *RCA Rev.*, vol. 20, pp. 753–769; December, 1959.
[61] N. S. Kapany, "Fibre optics," in "Concepts of Classical Optics," J. Strong, Ed., W. H. Freeman Co., San Francisco, Calif.; 1958.
[62] L. J. Mayer, "Electron mirror microscopy, etc.," *J. Appl. Phys.*, vol. 30, pp. 1101–1104; July, 1959.
[63] A. M. Skellet, L. E. Leveridge, and J. W. Gratian, "Electron beam head for magnetic tape playback," *Electronics*, vol. 26, pp. 168–171; October, 1953.
[64] M. M. Freundlich, *et al.*, "Experiments on magnetic tape readout with an electron beam," PROC. IRE, vol. 49, pp. 498–509; February, 1961.
[65] W. E. Glenn, "Thermoplastic recording," *J. Appl. Phys.*, vol. 30, pp. 1870–1873; December, 1959.
[66] M. Knoll and B. Kazan, "Storage Tubes," John Wiley and Sons, Inc., New York, N. Y.; 1952.
[67] J. Thomson and E. B. Callick, "Electron Physics and Technology," English Universities Press, London; 1959.
[68] H. Iams, U. S. Patent No. 2,259,507, filed, December 12, 1936; P. T. Farnsworth, U. S. Patent No. 2,355,212, filed, June 20, 1942; Loewe Radio, Brit. Patent No. 494,145; accepted, October 18, 1938.
[69] F. Schroeter, "Image storage in television," *Optik*, vol. 1, pp. 406–409; 1946.
[70] K. Schlesinger, "Storage Tube," U. S. Patent No. 2,532,339 (December 5, 1950); filed, May 9, 1946.
[71] A. V. Haeff, "A memory tube," *Electronics*, vol. 20, pp. 80–83; September, 1947.
[72] R. C. Hergenrother and B. C. Gardner, "The recording storage tube," PROC. IRE, vol. 38, pp. 740–747; July, 1950.
[73] M. Knoll and B. Kazan, "Viewing storage tubes," *Advances in Electronics*, vol. 8, pp. 447–492; 1956.
[74] J. V. Harrington and T. F. Rogers, "Signal-to-noise improvement through integration in a storage tube," PROC. IRE, vol. 38, pp. 1197–1203; October, 1950.
[75] R. C. Hergenrother and A. S. Luftman, "Performance of the recording tube," *Proc. NEC*, vol. 8, pp. 543–552; January, 1953.
[76] L. Pensak, "Conductivity induced by bombardment in thin films," *Phys. Rev.*, vol. 75, pp. 472–478; month, 1949.
[77] ——, "The 'graphecon,' a picture storage tube," *RCA Rev.*, vol. 10, pp. 59–73; March, 1949.
[78] A. S. Jensen, *et al.*, "Barrier grid storage tube," *RCA Rev.*, vol. 9, pp. 112–135; March, 1948.
[79] ——, "The 'radechon,' a barrier grid storage tube," *RCA Rev.*, vol. 16, pp. 197–215; June, 1955.
[80] M. Toohig and L. G. Wolfgang, "Improved Barrier Grid Storage Tube," presented at the Fall IRE-PGED Mtg., Washington, D. C.; October 29–31, 1959.
[81] T. S. Greenwood and R. E. Staehler, "A high speed barrier grid store," *Bell Sys. Tech. J.*, vol. 37, pp. 1195–1220; September, 1958.
[82] A. S. Luftman, "Survey of Recording Storage Tubes," presented at 7th Weather Radar Conf., Miami, Fla.; November, 1958.
[83] D. W. Davis, "Iatron storage tube," *Trans. AIEE*, vol. 29 (*Commun. and Electronics*), pp. 47–53; March, 1957.
[84] M. F. Toohig, "Iatron storage display tube," *Elec. Commun.*, vol. 36, No. 2, pp. 139–143; 1960.
[85] M. Knoll, P. Rudnick, and H. Hook, "Viewing storage tube with halftone display," *RCA Rev.*, vol. 14, pp. 492–501; December, 1953.
[86] T. Koda, N. Lehrer, and R. Ketchpel, "21-inch direct view storage tube," 1957 IRE WESCON CONVENTION RECORD, pt. 7, pp. 78–86.
[87] H. M. Smith, "The typotron," 1955 IRE CONVENTION RECORD, pt. 4, pp. 129–134.
[88] C. D. Beintema, S. T. Smith, and L. L. Vant-Hull, "Multicolor storage tube," IRE TRANS. ON ELECTRON DEVICES, vol. ED-4, pp. 303–309; October, 1957.
[89] N. H. Lehrer, "Selective erasure and nonstorage writing in direct-view half-tone storage tubes," PROC. IRE, vol. 49, pp. 567–573; March, 1961.
[90] V. K. Zworykin and G. A. Morton, "Television," John Wiley and Sons, Inc., New York, N. Y., 2nd ed., chs. 11 and 12; 1954.
[91] L. E. Swedlund and L. C. Wimpee, "The development of 110° television picture tubes having an ion trap gun," 1957 IRE NATIONAL CONVENTION RECORD, pt. 3, pp. 150–155.
[92] J. C. Francken, J. de Gier, and W. F. Nienhuis, "A pentode gun for TV picture tubes," *Philips Tech. Rev.*, vol. 18, pp. 73–81; September, 1956.
[93] H. Iams, "A fixed-focus electron gun for cathode-ray tubes," PROC. IRE, vol. 27, pp. 103–105; February, 1939.
[94] A. Y. Bentley, K. A. Hoagland, and H. W. Grossbohlin, "Self-focusing picture tube," *Electronics*, vol. 25, pp. 107–109; June, 1952.
[95] E. Gundert and H. Lotsch, "Steile Fernseh—Bildroehren für kleine modulations Spannungen," *Telefunken Zeitung*, vol. 33, pp. 58–65; March, 1960.
[96] J. W. Schwartz, "The annular geometry electron gun," 1958 IRE NATIONAL CONVENTION RECORD, pt. 3, pp. 13–20.
[97] ——, "A new high-transconductance electron gun for kinescopes," *RCA Rev.*, vol. 19, pp. 232–243; June, 1958.
[98] E. Atti, "High transconductance wideband television gun," 1958 IRE NATIONAL CONVENTION RECORD, pt. 3, pp. 3–12.
[99] P. H. Gleichauf, "Space-charge grid high-transconductance guns," PROC. IRE (Correspondence), vol. 46, p. 1542; August, 1958.
[100] K. Schlesinger, U. S. Patent No. 2,306,663 (December 29, 1942); application in Germany, March 2, 1938.
[101] E. Gundert and H. Lotsch, "Development of a television tube with high transconductance," *Telefunken Zeitung*, vol. 33, pp. 223–230; September, 1960.
[102] K. Schlesinger, "Focus reflex modulation of electron guns," 1960 IRE INTERNATIONAL CONVENTION RECORD, pt. 3, pp. 42–50. Also in IRE TRANS. ON ELECTRON DEVICES, vol. ED-8, pp. 224–230; May, 1961.
[103] J. J. Josephs, "A review of panel-type display devices," PROC. IRE, vol. 48, pp. 1380–1395; August, 1960.
[104] D. Gabor, P. R. Stuart, and P. G. Kalman, "A new cathode

ray tube for monochrome and color television," *Proc. IEE*, vol. 105B, pp. 581–606; May, 1958.
[105] W. R. Aiken, "A thin cathode-ray tube," Proc. IRE, vol. 45, pp. 1599–1604; December, 1957.
[106] E. G. Ramberg, "Electron-optical properties of a flat television picture tube," Proc. IRE, vol. 48, pp. 1952–1960; December, 1960.
[107] K. Schlesinger, U. S. Patent No. 2,935,643; filed, May 31, 1957. Also, see [34] .
[108] H. B. Law and E. G. Ramberg, "The reflected-beam kinescope," Proc. IRE, vol. 48, pp. 1409–1417; August, 1960.
[109] E. W. Herold, "Methods suitable for television color kinescopes," Proc. IRE, vol. 39, pp. 1177–1185; October, 1951.
[110] H. B. Law, "A three-gun shadow-mask color-kinescope," Proc. IRE, vol. 39, pp. 1186–1194; October, 1951.
[111] N. F. Fyler, W. E. Rowe, and C. W. Cain, "The CBS-colortron: a color picture tube of advanced design," Proc. IRE, vol. 42, pp. 326–334; January, 1954.
[112a] H. R. Seelen, *et al.*, "Development of a 21" metal-envelope color kinescope," *RCA Rev.*, vol. 16, pp. 122–139; March, 1955.
[112b] R. B. Janes, L. B. Headrick, and J. Evans, "Recent improvements in the 21AXP22 color kinescope," *RCA Rev.*, vol. 17, pp. 143–167; June, 1956.
[112c] C. P. Smith, A. M. Morrell, and R. C. Demmy, "Design and development of the RCA 21CYP22 21-inch glass color picture tube," *RCA Rev.*, vol. 19, pp. 334–348; September, 1958.
[113a] R. C. Hergenrother, "Design of lens-mask three-gun color television tubes," Proc. IRE, vol. 43, pp. 943–951; August, 1955.
[113b] E. G. Ramberg, *et al.*, "Focusing grill color kinescopes," 1956 IRE Convention Record, pt. 3, pp. 128–134.
[113c] E. F. De Haan and H. Zimmer, "Post-focusing color tubes," *Acta Electronica*, vol. 2, pp. 189–193; January–April, 1958.
[113d] M. E. Amdursky, R. G. Pohl and C. S. Szegho, "A new high-efficiency parallax mask color tube," Proc. IRE, vol. 43, pp. 936–943; August, 1955.
[113e] R. Dressler, "The PDF-chromatron—a single or multi-gun tri-color cathode-ray tube," Proc. IRE, vol. 41, pp. 851–858; July, 1953.
[114] C. G. Lob, "General Electric post acceleration color tube," 1956 IRE Convention Record, pt. 3, pp. 114–117.
[115] G. F. Barnett, *et al.*, "A beam-indexing color picture tube—the Apple tube," Proc. IRE, vol. 44, pp. 1115–1119; September, 1956.
[116] P. Schagen, "The banana-tube display system," *J. IEE*, vol. 108B, pp. 1–10; May, 1961 (Paper No. 3564). There are five companion papers: Nos. 3561–3566.
[117] J. E. Jacobs and H. Berger, "Large-area photoconductive x-ray pickup-tube performance," *Elec. Engrg.*, vol. 75, pp. 158–161; February, 1956.
[118] J. Elster and H. Geitel, "On colored hydrides of alkali metals and their photosensitivity," *Physik, Z.*, vol. 11, pp. 257–262, April, pp. 1082–1083, November, 1910; vol. 12, pp. 609–614, August, 1911.
[119] L. R. Koller, "Photoelectric emission from thin films of cesium," *Phys. Rev.*, vol. 36, pp. 1640–1647; December, 1930.
[120] P. Goerlich, "On composite transparent photocathodes," *Z. Physik*, vol. 101, pp. 335–342; 1936.
[121] A. H. Sommer, "New photoemissive cathodes of high sensitivity," *Rev. Sci. Instr.*, vol. 26, pp. 725–726; July, 1955.
[122] L. O. Grondahl, "The copper-cuprous oxide rectifier and photoelectric cell," *Rev. Mod. Phys.*, vol. 5, pp. 141–168; April, 1933.
[123] B. Lange and W. Eitel, "On a method of studying the absorption of mineralogical microspecimens with the aid of barrier-layer photocells," *Mineral. u. petrog. Mitt.*, vol. 41, pp. 435–452; 1931.
[124] L. Bergmann, "On a new selenium barrier layer photocell," *Physik. Z.*, vol. 32, pp. 286–288; April, 1931.
[125] T. W. Case, "Thalofide cell—a new photo-electric substance," *Phys. Rev.*, vol. 15, pp. 289–292; April, 1920.
[126] C. L. Oxley, "Characteristics of cooled lead sulfide photoconductive cells," *J. Opt. Soc. Am.*, vol. 36, p. 356; June, 1946.
[127] R. J. Cashman, "New photoconductive cells," *J. Opt. Soc. Am.* vol. 36, p. 356; June, 1946.
[128] F. F. Rieke, L. H. DeVaux, and A. J. Tuzzolino, "Single-crystal infrared detectors based upon intrinsic absorption," Proc. IRE, vol. 47, pp. 1475–1478; September, 1959.
[129] G. A. Morton, M. L. Schultz, and W. E. Harty, "Infrared photoconductive detectors using impurity-activated germanium-silicon alloys," *RCA Rev.*, vol. 20, pp. 599–634; December, 1959.
[130] R. Frerichs, "The photoconductivity of incomplete phosphors," *Phys. Rev.*, vol. 72, pp. 594–601; October, 1946.
[131] C. F. Hadley and E. Fisher, "Sintered cadmium sulfide photoconductive cells," *RCA Rev.*, vol. 20, pp. 635–647; 1959.
[132] J. N. Shive, "A new germanium photoresistance cell," *Phys. Rev.*, vol. 76, p. 575; August, 1949.
[133] I. K. Kikoin and M. M. Noskov, "A new photoelectric effect in cuprous oxide," *Phys. Zeits Sowjetunion*, vol. 5, pp. 586–596; 1934.
[134] A. Rose, "An outline of some photoconductive processes," *RCA Rev.*, vol. 12, pp. 362–414; September, 1951.
[135] J. N. Shive and P. Zuk, "Junction phototransistors," *Bell Labs. Record*, vol. 33, pp. 445–449; December, 1955.
[136] D. M. Chapin, C. S. Fuller, and G. L. Pearson, "A new silicon *p-n* junction photocell for converting solar radiation into electrical power," *J. Appl. Phys.*, vol. 25, pp. 676–677; May, 1954.
[137] P. Rappaport, "The photovoltaic effect and its utilization," *RCA Rev.*, vol. 20, pp. 373–397; September, 1959.
[138] M. Wolf, "Limitations and possibilities for improvement of photovoltaic solar energy converters. Part I: Considerations for earth's surface operation," Proc. IRE, vol. 48, pp. 1246–1263; July, 1960.
[139] G. A. Morton, R. M. Matheson, and M. H. Greenblatt, "Design of photomultipliers for the sub-millimicrosecond region," IRE Trans. on Nuclear Science, vol. NS-5, pp. 98–104; December, 1958.
[140] V. K. Zworykin, G. A. Morton, and L. Malter, "The secondary emission multiplier—a new electronic device," Proc. IRE, vol. 24, pp. 351–375; March, 1936.
[141] —— and J. A. Rajchman, "The electrostatic electron multiplier," Proc. IRE, vol. 27, pp. 558–566; September, 1939.
[142] J. A. Rajchman and R. L. Snyder, "An electrostatically focused multiplier phototube," *Electronics*, vol. 13, pp. 20–23, 58, 60; December, 1940.
[143] C. C. Larson and H. Salinger, "Photocell multiplier tube," *Rev. Sci. Instr.*, vol. 11, pp. 226–229; July, 1940.
[144] G. Weiss, "On secondary-emission multipliers," *Z. tech. Physik*, vol. 17, pp. 623–629; December, 1936.
[145] J. W. Coltman and F. H. Marshall, "Photo-multiplier radiation detector," *Phys. Rev.*, vol. 72, p. 528; September, 1947.
[146] G. A. Morton, "Recent developments in the scintillation counter field," IRE Trans. on Nuclear Science, vol. NS-3, pp. 122–135; November, 1956.
[147] J. Pohl, "Electron-optical imaging with photoelectrons," *Z. tech. Physik*, vol. 15, pp. 579–581; December, 1934.
[148] W. Schaffernicht, "On the transformation of light images into electron images," *Z. Physik*, vol. 93, pp. 762–768; February, 1935.
[149] V. K. Zworykin and G. A. Morton. "Applied electron optics," *J. Opt. Soc. Am.*, vol. 26, pp. 181–189; April, 1936.
[150] G. Holst, *et al.*, "An apparatus for the transformation of light of long wavelength into light of short wavelength," *Physica*, vol. 1, pp. 297–305; 1934.
[151] G. A. Morton and L. E. Flory, "An infrared image tube and its military applications," *RCA Rev.*, vol. 7, pp. 385–413; September, 1946.
[152] M. Kooman, R. Rousay, and H. A. Knoll, "An infrared pupillometer," *J. Opt. Soc. Am.*, vol. 38, pp. 719–722; August, 1948.
[153] "Ultrascope electron tube aid to medical and industrial research," *Elec. Engrg.*, vol. 78, pp. 395–396; April, 1959.
[154] A. Lallemand, M. Duchesne, and G. Wlerick, "Electron photography," in "Advances in Electronics and Electron Physics," L. Marton, Ed., Academic Press, New York, N. Y., vol. 12, pp. 5–16; 1960.
[155] J. W. Coltman, "Fluoroscopic image brightening by electronic means," *Radiology*, vol. 51, pp. 359–367; September, 1948.
[156] J. J. P. Hardenberg, "An apparatus for cinefluorography with 11-inch image intensifier," *Philips Tech. Rev.*, vol. 20, pp. 331–345; September, 1959.
[157] J. S. Courtney-Pratt, "A new method for the photographic study of fast transient phenomena," *Research*, vol. 2, pp. 287–293; June, 1949.
[158] R. G. Stoudenheimer and J. C. Moor, "An image-converter tube for high-speed photographic shutter service," *RCA Rev.*, vol. 18, pp. 322–331; September, 1957.
[159] ——, "Image intensifier development in the RCA Electron Tube Division," in "Advances in Electronics and Electron Physics," L. Marton, Ed., Academic Press, New York, N. Y., vol. 12, pp. 41–58; 1960.
[160] "British instrument magnifies light 50,000 times," *Elec. Engrg.*, vol. 79, p. 176; February, 1960.
[161] B. Kazan and F. H. Nicoll, "Solid-state light amplifier," *J. Opt. Soc. Am.*, vol. 47, pp. 887–894; October, 1957.
[162] ——, "A solid-state amplifying fluoroscope screen," *RCA Rev.*, vol. 19, pp. 19–34; March, 1958.
[163] E. E. Loebner, "Solid-state optoelectronics," *RCA Rev.*, vol. 20, pp. 715–743; December, 1959.

Semiconductor Devices*

J. M. EARLY†, FELLOW, IRE

Summary—**The semiconductor art began early in the nineteenth century. Gradual advances in materials, techniques, and understanding followed. The decisive event in creation of the modern semiconductor device field was the invention of the transistor in December, 1947 by Bardeen and Brattain. Understanding of minority carrier injection, invention of the junction diode and transistor, and major advances in materials and techniques then came in rapid sequence. In the last decade the semiconductor devices have grown prodigiously in variety and complexity and in speed, power, and reliability. Further growth will be shaped by the semiconductor materials and the constraints which they impose on speed, power, and impedance. The transistor and allied single-crystal *p-n* junction devices will probably be the key tools of electronic circuitry for many years.**

Introduction

The TRANSISTOR and related single-crystal *p-n* junction devices are the first of the modern solid-state devices. At present, the probability seems high that, in the electronics of the future, they will play many of the roles which de Forest's audion and its offspring played in creating the field. This paper includes in successive sections a brief history of the semiconductor technology, a short assessment of its influence on both engineering and science, and a summary of the present state and future prospects of single-crystal *p-n* junction devices. Emphasis is put on the key steps in the creation of the technology and on the work of the handful of pioneers who made the essential discoveries. Of necessity, many contributions of very considerable importance are passed over without mention. Easily ten times the space available here would be required to do justice to all those who by notable contributions have brought this complex technology to its present level of sophistication.

The growth of the semiconductor technology and industry has shown a variety of patterns, many of which are common to the evolution of other technical fields. Important among those which flowed from invention of the transistor are the discovery of new properties of matter, a subsequent rapid growth derived in large part from a penetrating theoretical understanding of the new properties and the device possibilities inherent in them, an accompanying transformation of the field from a frontier province of science to a central area of engineering, and the practical possibility of forecasting the broad future shape of the field from the properties of the materials.

Of these events, the recognition of a new property of matter, the injection into semiconductors of excess mobile minority charge carriers of useful lifetime, was the key. This phenomenon is essential to the transistor and to the junction diode and effectively determines their properties. Once this phenomenon was understood the general character of subsequent growth was predictable. Transistor technology therefore depends for its existence on a material property whose existence was not even suspected 15 years ago.

History

The history of transistor devices resembles that of other device technologies, particularly those which have followed it in the development of the solid-state field. It began in the nineteenth century and moved slowly but with gathering speed until the time of the invention of the transistor. From that point, which made a radical break with the past, major steps in device concepts and process technology followed each other in swift succession. Preceding the invention of the transistor and essential to it were improvements which had been made in semiconductor materials throughout the preceding decade or more. The transistor and other device inventions which followed created powerful new stimuli for further improvements of the materials.

Semiconductors can be characterized by the increase in charge-carrier concentration which accompanies increase of temperature. This effect was first noticed as a decrease of resistance with rise of temperature in Faraday's experiments on silver sulphide. Rectification was observed some 40 years later, and the Hall effect was first seen in 1879. Various advances both in technology and understanding were made with the years. The discovery of the "cat's-whisker" detector for radio, and the later development of copper-oxide and selenium power rectifiers stimulated both scientific and engineering advances. This work, however, gave little, if any, hint of the revolution that was to come.

The improvements in materials which were to lay the foundation for the discovery of injected carrier action began with a new emphasis on silicon point-contact diodes as microwave radio detectors. Silicon and germanium were studied extensively at the Bell Telephone Laboratories and Purdue University, respectively. Study of germanium was stimulated by its similarity to silicon and its lower melting point. Nearly a decade of work in the purification of germanium and innumerable ingenious experiments preceded the invention of the transistor by Bardeen and Brattain in December, 1947.[1]

* Received by the IRE, January 19, 1962.
† Bell Telephone Laboratories, Inc., Murray Hill, N. J.

[1] J. Bardeen and W. H. Brattain, "The transistor, a semiconductor triode," *Phys. Rev.*, vol. 74, pp. 230–231, July 15, 1948; "Physical principles involved in transistor action," *Phys. Rev.*, vol. 75, pp. 203–231, July 15, 1949.

Their work was soon followed by the suggestion that excess charge carriers opposite in sign to those characterizing the material at thermal equilibrium could be injected into germanium and passed from electrode to electrode. The validity of this concept was underscored by the work of Shockley, Pearson, and Haynes, and will no doubt be demonstrated to coming generations of students by their study of and repetition of the classic Haynes-Shockley experiment.[2] This experiment determines clearly and quantitatively the mobility, lifetime, and effective sign of charge for the injected minority carriers.

Implicit in the new understanding was the fact that the charge-carrier densities could be characterized by a Fermi-Dirac distribution in energy and were described to good accuracy by the familiar Boltzmann distribution. This meant that the injected excess carrier densities increased by the factor e when the injection potential increased by kT/q. In theory, then, injected carrier densities and therefore currents could be increased by a factor of 10 for each increase of approximately 60 mv in injection voltage, thus allowing very large forward currents at low forward voltages.

Still other consequences of importance were implicit in these results. The two oppositely-charged types of carrier, the conduction electron and the hole, permitted most of the semiconductor bulk to remain electrically neutral and thereby circumvented significant space-charge limitation of current density. The carrier mobilities and large carrier concentrations obtainable led to the possibility of current densities of many thousands of amperes per square centimeter in materials with dielectric constants of 12 to 16. At the same time, the absence of need for a heated filament and the exponential dependence of current on voltage opened the opposite possibility of operation at minuscule currents and therefore powers. The subsequent history of transistor devices is the story of the realization of these marvelous possibilities.

Most of these possibilities were delineated explicitly or implicitly in Shockley's synthesis of device theory which closely followed the recognition of the existence of excess minority carriers.[3] From here, theory and experiment fed upon each other with steady improvement and with spectacular advance in the control of the properties of germanium and silicon leading the way. This work on crystals[1] by demonstrating dramatically the importance of purity and crystallinity in studies of solid materials, has probably made semiconductor technology's most important indirect contribution to scientific and technological progress.

A tremendously improved control of the composition and structure of the materials, and thereby of devices, is the essence of this advance. Perfection of crystal-pulling techniques,[4] the invention of zone-leveling,[5] plus a wide variety of supporting techniques in the preparation of raw materials and the subsequent handling of the purified ones reduced undesired constituents of the new single crystals by factors of 1000 or more. "High back-voltage germanium" with perhaps 10^{15} impurities per cubic centimeter was replaced by intrinsic germanium containing 10^{12} or fewer impurities per cubic centimeter. Control of crystallinity and of chemical-purity concentration was followed by techniques for similarly reducing the concentration of other defects, such as dislocations. Hand in hand with the perfection of means for removing or avoiding undesired chemical or physical defects came techniques for introducing many of these same impurities or defects in a highly controlled fashion. Substantially all of the present techniques for control of the form and characteristics of junctions are the outcome of such work.

The technological possibilities of the new materials and the transistor devices were obviously large, and the problems of development and manufacture soon accompanied those of research. Control and understanding of the structure and behavior of the junction surfaces lagged far behind the advances in bulk properties. Even today most changes in the behavior of diodes and transistors during normal use result from changes at the junction surfaces. A fully satisfying description of surface structure and behavior is still evolving, but in the most stable devices yet developed surface changes have been minimized by application of a simple rule for design and manufacture: "seal the junctions clean, dry, and very hot." The lesson to be learned in the advances made through refinement of materials has been repeated in this technique of improving reliability. Control of composition and structure is the common factor in these parallel developments.

The decade since measurements on junction diodes and junction transistors were first reported[6,7] has witnessed major advances in device concepts, process techniques, materials properties, and in the device performance sought through these improvements. The levels of performance reached in today's multitude of devices are surveyed in the closing section of this paper. Quite aside from the specifics of today's device realities and tomorrow's device possibilities are influences which the transistor and its technology have had on engineering and science. These are the subjects of the next section.

[2] W. Shockley G. L. Pearson, and J. R. Haynes, "Hole injection in germanium-quantitative studies and filamentary transistors," *Bell Sys. Tech. J.*, vol. 28, pp. 344–366; July, 1949. Also see J. R. Haynes and W. Shockley, "The mobility and life of injected holes and electrons in germanium," *Phys. Rev.*, vol. 81, pp. 835–843; March 1, 1951.

[3] W. Shockley, "The theory of *p-n* junctions in semiconductors and *p-n* junction transistors," *Bell Sys. Tech. J.*, vol. 28, pp. 435–489; July, 1949.

[4] G. K. Teal and J. B. Little, "Growth of germanium single crystals," *Phys. Rev.*, vol. 78, p. 647; June 1, 1950.

[5] W. G. Pfann, "Principles of zone refining," *Trans. Am. Inst. Mining and Metallurgical Engrs.*, vol. 194, pp. 747–753; July, 1952.

[6] F. S. Goucher, G. L. Pearson, M. Sparks, G. K. Teal, and W. Shockley, "Theory and experiment for a germanium *p-n* junction," *Phys. Rev.*, vol. 81, no. 4, pp. 637–638; 1951.

[7] W. Shockley, M. Sparks, and G. K. Teal, "*p-n* junction transistors," *Phys Rev.*, vol. 83, pp. 151–162; July 1, 1951.

Impact on Technology and Science

The swift advance of semiconductor-device technology has had significant influence, both direct and indirect, on many technical fields. Particularly important for both technology and science have been the lessons taught concerning purification of materials and control of their composition and structure. Also of broad significance in many areas are the improvements in digital computers arising from the transistor work. The speed, power, reliability, cost, and size advantages of semiconductor devices have also modified innumerable other applications. Effects found in single-crystal junction devices have become in some cases important tools in studying the structure of solids. Finally, the highly rational theory of semiconductor devices has not only re-emphasized the importance of such theory but has stimulated development of unified theories covering devices of many types.

The close relation between materials improvements and the transistor technology showed the power of single-crystal studies and deliberate control of crystal composition. It stimulated attempts at purifying and controlling other crystalline solids, both directly by providing techniques, such as zone-refining, and indirectly by showing the advantage of such an approach. At the same time, the emphasis on purity and control was extended in the semiconductor technology from bulk properties to surface conditions. There can be little doubt that semiconductor successes have underlined most distinctly the importance of understanding and controlling the composition and therefore the properties of materials of many kinds.

Of the various semiconductor-device applications, use in computers will almost certainly have the widest influence on other areas of technology. Although the digital computer was already highly successful, the transistor brought increases in computer speed, complexity, and reliability which touched off an explosion in computer manufacture and use. Engineering and science, like the rest of our society, have felt only the beginnings of the changes which will result from the new machines. Other semiconductor-device applications range from dc clamp-on ammeters to low-noise microwave amplifiers. This applications revolution seems sure to grow in diversity and complexity. The new devices, until superseded by major new invention, seem certain to rule the electronics of the future.

The Esaki diode is an example of a new and useful junction device which is simultaneously a tool for the study of solids. Already used in several applications and under study for a variety of others, it provides a beautiful electronic instance of quantum-mechanical tunneling through a potential barrier. Careful study of the details of the voltage-current characteristic of this device, particularly at low temperatures, has given considerable data on phonon spectra in the electronic semiconductors.

Beginning with the publication of the first *p-n* junction paper, theory has charted much of the course in device development. Even the first junction transistors showed remarkable agreement with the predictions. This pattern has continued. Transistor theory has evolved in two directions, both towards more detailed understanding and prediction, and toward coarser and simpler pictures of transistor operation. The latter trend was instrumental in evoking the charge-control theory which describes in a common language[8] the fundamentals of devices so diverse as photoconductors, electron tubes, and the transistor. This generalization of device theory may be extended still further with the development of a gain-band product given by the reciprocal of relaxation time multiplied by the ratio of two energies.[9] Still another form of a gain-band product can be found as the quotient of a change in controlled power divided by the change in control energy. These unifications of viewpoint have been a welcome advance among, and byproduct of, increasing specialization. This is perhaps only an expression of the trend toward a more unified viewpoint in engineering which has accompanied the closer approach of engineering to the basic sciences.

Status and Future

Semiconductor technology has become the basis of a rapidly maturing industry making a wide variety of devices by use of many complex processes and techniques. The skills of the factory and the market are largely supplanting those of the research and development laboratories. Beyond these easy generalizations other directions of progress common to most device technologies can be seen. Present trends in materials, structures, performance, and function are rooted in the past and provide the raw material for an outline of the future.

Materials limit present devices and will determine those to come. Concepts of structure and techniques for achieving them must work within the limits set by nature. Pressure for improvement of devices is based on the insatiable demands of users for more speed, power, and reliability, and less cost. For the most part, even the search for lower noise is but the opposite face of the demand for greater power. At the same time, sharpened perception of applications needs and of device possibilities stimulates development of numerous new and specialized devices. This section discusses present and future prospects in materials, then surveys the present and future of some of the principal devices with regard to speed, power, and reliability, and ends with speculation on what lies still further beyond.

Among the materials, germanium and silicon have carved out major and secure uses and contend with each other and with some of the secondary materials for a

[8] E. O. Johnson and A. Rose, "Simple general analysis of amplifier devices with emitter, control, and collector functions," Proc. IRE, vol. 47, pp. 407–418; March, 1959.

[9] E. O. Johnson, private communication, September, 1960.

variety of special applications. Various of the 3-5 compounds have already proved their worth where some superior property offers a performance gain to balance the increased difficulty in material refining and device fabrication. To date germanium has been the usual first choice when cost per device was the prime consideration. Progress in silicon fabrication may change this balance. In low-frequency power transistors, germanium appears to have a permanent niche because of the relatively large currents—up to 100 amperes or more—achievable in rather simple devices. In diodes, gallium arsenide is finding uses based on its very high carrier mobilities and good high-temperature performance. Indium antimonide is the major choice for Hall effect devices, and gallium antimonide for low-noise Esaki tunnel diodes, again because of special characteristics. As technology evolves, these materials will become purer and cheaper, as have germanium and silicon. The 3-5's will become relatively more important, but the older materials seem certain to dominate the field.

Perhaps the chief factor insuring leadership for the elemental semiconductors is their relatively simple metallurgy. In formation of junctions in or contacts to the compound materials, reaction of the alloying metal with each of the elements of the compound material must be considered. This restricts the choice of alloying metals severely. In addition, virtually every 3-5 compound has a high vapor pressure, and crystal growth is thereby complicated. It is sometimes necessary to carry out alloying operations in pressure chambers for the same reason. By contrast, despite difficulties, the processing of germanium and silicon offers no problems.

Device structure, as well as materials, determines performance. Among the major performance characteristics, reliability tends to be the one most nearly the same for transistors and diodes. In both silicon and germanium devices, the first reliability goal, that of device parameters permanently stable during normal use in ambients near room temperature, appears to be in sight. Clean junctions, sealed vacuum-tight at high temperatures in the absence of water vapor, seem common to all successful processes so far developed, including the very important one of a surface oxide on silicon. The future will bring use of such techniques on ordinary as well as premium devices. A further goal is the raising of the temperatures at which the devices are permanently stable. These will probably go as high as the temperatures at which the devices are electrically useful—about 150°C for germanium and near 300°C for silicon.

Power and speed of devices are determined by the device material and structure and are to some degree mutually interdependent. The diode is the most fundamental of the new devices and shows the widest ranges in speed and power-handling ability. This superiority results almost directly from the simplicity of diode structure—in ideal principle all diode currents flow at right angles to the junction and none flow parallel. It is precisely the flow of base current parallel to the transistor junctions which allows the transistor to perform its complex functions and at the same time restricts its speed power because of power losses in lateral current flow through layers of substantial sheet resistance.

The power-handling ability of diodes and transistors is set by both thermal and electrical considerations. Use of diode-junction areas of a square centimeter or more, together with the provision of forced cooling, has resulted in units with peak current ratings measured in thousands of amperes. At the same time, use of very thick junctions fabricated of high-resistivity silicon has produced devices with peak voltage ratings of thousands of volts. Individual diodes rectify tens of kilowatts and rectifier combinations for conversion of a megawatt have been built. The very high rectification efficiencies obtainable with germanium and silicon rectifiers promise continued effort to raise the maximum ratings of such rectifier combinations as well as of the individual units. In contrast to these maximum powers, reverse currents of silicon diodes are frequently less than 10^{-10} amperes and gallium-arsenide diodes have shown reverse currents in the 10^{-13}-ampere range.

P-N junction devices are subject to the usual frequency-scaling laws of electronics and must be reduced in size as their speed is increased. The evolution of high-speed junction diodes brings them ever closer in gross form to the point-contact diodes which preceded them. This general parallelism of structure provides a fine example of the evolutionary principle that form tends to follow function, producing a convergence in general appearance between members of different evolutionary lines specializing to occupy similar ecological niches. Silicon-computer diodes (now operate in the nanosecond range) and still faster gallium arsenide diodes have been reported. Structures in the tenth-nanosecond class will almost certainly evolve.

The varactor diode, however, provides the present outer bounds in both speed and power of semiconductor device performance at high frequencies.[10] Cutoff frequencies ($Q=1$) above 200 Gc have been reached in germanium, silicon, and gallium arsenide. Gallium-arsenide devices operating at liquid-nitrogen temperature have given parametric amplification in the degenerate mode with noise temperatures as low as 30°K thus setting the lowest minimum so far in the power-handling ability of semiconductor devices. At the other end of the power spectrum, varactor diodes of larger area, used as harmonic multipliers, have produced output powers of a few watts at a few Gc.

Although the advance in diode speed and power will continue, increase of power at high frequency must encounter limitations imposed by frequency-scaling laws. If, in these devices, linear dimensions are made inversely proportional to operating frequency, device impedance

[10] This, of course, excludes the optical maser, which is not a *p-n* junction device.

at operating frequency[11] remains fixed, as does the device Q at operating frequency. However, the junction area then decreases inversely as the square of frequency and, because power per unit area is nearly an electrical design constant, power-handling ability at a fixed impedance level also varies as the inverse square of operating frequency. This limitation applies to diodes and transistors alike. The circumvention of its effects by development of successful means for operation at very low impedance levels and by series connection of units is a major challenge to the years to come.

Advances in the speed and power of transistors have paralleled those in diodes. Cutoff frequencies have increased from the few tens of megacycles seen in the first point-contact transistors and in early grown junction transistors to the several gigacycles of today's fastest diffused-base units. Major steps in this evolution to greater speed were the concept of the thin base-region, thick collector-barrier transistor, and the development of diffused-base technology to realize this structure.

Again, as in diodes, the primary tool in achieving increase of speed has been the reduction of dimensions. Minority-carrier transit distances were cut by diffusion fabrication of thin base layers. A wide variety of techniques have been developed for achieving small electrode areas. These range from ball-alloying through jet-plating plus alloying, and evaporation plus alloying to the use of photolithography.

As indicated previously, the power-handling ability of transistors has stayed well inside that of diodes and seems certain to do so in the future. The difficulties and challenges to the expansion of power and speed capabilities of transistors are closely similar to those in diodes. By way of comparison, transistors have controlled powers of several kilowatts at low frequencies, have given outputs of a few watts at a few hundred megacycles, and work is underway on units to provide a watt or more in the low gigacycle range.

The *p-n-p-n* diodes and triodes now appear to be the most important of the many devices related to the diode and transistor. Their major impact has been in the use of *p-n-p-n* triodes as controlled rectifiers and pulse generators. Devices switching tens of amperes at many hundreds of volts in a few nanoseconds have already been put into service. It appears that the *p-n-p-n* structures will remain intermediate in power capabilities between transistors and diodes and will be restricted to speeds lower than those of transistors.

The transistor art has spawned many other devices. Many of these have more limited design ranges in speed and power and justify their existence by utility in special applications. Various thermistors, Hall-effect devices, microwave protective limiters, and the field-effect transistor are other offshoots from the principal line of development. Many of these do not depend on excess injected minority carriers for their operation, but nonetheless are derived from transistor technology. The Esaki diode, for example, which depends for its operation on a quantum-mechanical tunneling effect not imagined in or relevant to the transistor theory, is a single-crystal *p-n* junction device built by transistor techniques from transistor techniques. It appears to have established a niche in low-noise amplification around 1 Gc, as well as in some level-setting and triggering applications. Despite difficulties, active study continues on Esaki units for computers, frequency conversion, and microwave-oscillator service.

The broad outlines of the future can be seen. The technology of the more complex and difficult compound semiconductors will gradually be mastered, and their importance will therefore increase. Germanium and silicon will almost surely remain the basic semiconductor materials. Reliability will increase, thus allowing higher device temperatures. Scaling laws of frequency, power, and impedance dictate the use of ever lower impedances and smaller dimensions in the evolution toward higher power and greater speed. Expression of these limitations in terms of specific powers and speeds seems dangerous. Forecasts of this kind have almost always been regarded as extravagant at the time and as ridiculously conservative a generation later.

Accompanying the growth in power, speed, and reliability will be an increase in device complexity. This includes not only the development of more refined designs for particular service and numerous specialized types of devices, but also the development and use of multiple devices and integrated circuits. The use of several diodes and several transistors in a single encapsulation provides economies of manufacture and use. Combining of these with passive components provides additional advantages in some instances. Multiple devices, integrated circuits, and individual semiconductor devices will all play important roles in the future of electronics.

In the far future, displacement of the diode, transistor, and allied components by now unimagined devices seems possible. However, such new devices will almost certainly find competition with diodes and transistors more difficult than these devices have found the competition with their high vacuum and gaseous predecessors. The requirements for effective competition can be seen in the limitations of the semiconductor devices. The diode and transistor, useful as they are, show objectionable temperature dependence. Breakdown fields and relatively low charge-carrier velocities restrict their powers and speeds. Any successful major competitor to semiconductor devices must almost certainly show marked advantage in one or more of these characteristics, and at the same time must equal or surpass them in simplicity and cost. Perfection of new solid-state triodes might bring an advantage in speed, but so far the trend in such devices has been away from the originally-conceived schemes and toward the transistor materials and structures, and therefore toward acceptance of their limitations.

[11] J. M. Early, "Speed in semiconductor devices," to be presented at the 1962 IRE International Convention, New York, N. Y.

Solid-State Devices Other than Semiconductors*

B. LAX† AND J. G. MAVROIDES†

Summary—**The solid-state properties of matter are of interest not only to the research scientist but also to the engineer who has translated the findings of basic research into endeavors of great technological significance. The research in ferromagnetism has led to the magnetic amplifier, which for certain applications is superior to amplifiers using vacuum tubes and to the development of new materials such as ferrites, which are now finding use at radio frequencies in pulse transformers, radio frequency inductors, antenna rods, magnetic core delay lines, computer elements, and magnetic recording media. Ferrites and antiferromagnetic materials are now also widely used at microwave frequencies and millimeter wavelengths as nonreciprocal phase shifters, load isolators, filters, ferrite switches, ferrod radiators, modulators and power limiters. Such phenomena as piezoelectricity, ferroelectricity and magnetostriction have resulted in the use of ultrasonic waves for the study of phonon-electron and phonon-spin interactions as well for such practical devices as the detection of imperfections and faults in solids, electromechanical transducers, resonators, filters, delay lines, computer memories, ultrasonic soldering, cleaning, drilling and cutting, and strain and acceleration gauges. Dielectric materials have found applications as prisms, polarizers, restrahlen plates and more recently in dielectric wave guides and fiber optics. Advances in low-temperature techniques have facilitated the study of resonance phenomena in paramagnetic materials, and this has led to such exciting new devices as the solid-state maser and the laser; from a study of superconductivity at these low temperatures there have emerged such new devices as the cryotron for computers, frictionless magnetic bearings for gyroscopes and other applications, and infrared detectors. The future promises many applications employing as yet unexploited phenomena in solids involving electrical, mechanical, magnetic and optical effects.**

HISTORICALLY, solid-state components are as old as technology itself. We may classify these devices in terms of the physical phenomena which are involved. Thus the phenomenon may be dielectric, optical, electrical, mechanical or magnetic, the latter of which in turn may be subdivided further into paramagnetic, ferromagnetic and ferrimagnetic. In this broad classification a variety of materials and their special properties may be included, and in some cases in more than one of the above categories. Thus, for example, we are not only interested in the dielectric properties of materials for their role in electrical circuits but in optical components as well. It is the latter property, indeed, that was one of the earliest uses of solid-state phenomena, namely, the spectrometer by Newton, the telescope by Galileo, and the microscope by Leuwenhoek. Another device, one as old if not older, is the magnetic compass. In this review, however, we shall not try to bridge the gap between early history and modern times, but shall restrict ourselves to the twentieth century, and particularly to the last 50 years. We shall emphasize the modern developments which exploit the complex properties of solid state materials involving piezoelectric, magnetostrictive, ferroelectric, superconducting, ferro- and ferrimagnetic, and resonance as well as optical phenomena. These have led to such devices as nonreciprocal circuit elements, computer components and such exciting new developments as microwave and optical masers.

* Received by the IRE, July 25, 1961.

† Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Mass., operated with support from the U. S. Army, Navy, and Air Force.

Ferromagnetic Devices

Magnetic Amplifier

Magnetic amplifiers and related devices whose operation depends on magnetic saturation are among the oldest solid-state electronic devices [1]. They were first studied before the turn of the century. However, they did not become an important technological factor until fairly recent years when improved magnetic materials [2] with rectangular hysteresis loops such as super-malloy, mu-metal, hipersil, etc., and efficient solid-state rectifiers enabled them to compete successfully with vacuum tubes and rotating amplifiers in the performance of a wide variety of functions.

The basic principal of operation is shown in Fig. 1, where the flow of ac current in the output circuit is controlled by adjusting the dc current in the input which determines the degree of saturation of a magnetic core, as shown by the hysteresis loop in Fig. 2. When the core is saturated, all the ac voltage appears across the output. For nonsaturation, very little voltage appears; thus a small dc power can control a much larger ac output power—hence the term magnetic amplifier.

Magnetic amplifiers have been used primarily to control ac machinery and power supplies at 60 and 400 cycles. More recently, amplifiers have been used for controlling power from 10 to 100 Kc using metallic cores, and to 10 Mc using ferrite cores [3]. At the lower frequencies magnetic amplifiers can control powers ranging from a microwatt to over a million watts [4] and with an amplification factor in some cases much greater than 10^4.

High-Frequency Ferrite Components

As the frequency of electronic circuitry increased in many applications prior to World War II, it was necessary to make thin metallic ribbons and tapes for transformers and powdered iron for core material in order to

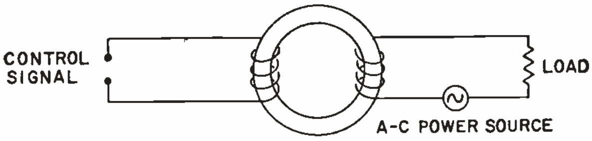


Fig. 1—Basic magnetic amplifier circuit. The direct current through the control winding determines the degree of saturation of the core. When the core is completely saturated, the full ac power appears across the load.

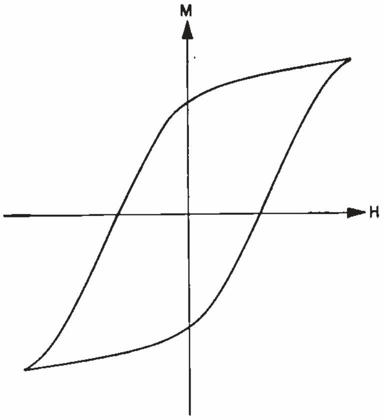


Fig. 2—Ferromagnetic hysteresis loop. Here *H* is the magnetic driving field and *M* is the resultant magnetization.

reduce the high-frequency eddy current losses. The solution to this problem was offered by the use of ferrite materials which are magnetic insulators with very high resistivity and therefore very low eddy current losses, which of course increase as the square of the frequency. The first proposal for high-frequency application was made in Germany in 1909 by Hilpert, who first synthesized such ferrites [5]. However, the practical development and exploration of these materials did not come about until Snoek [6] from the Phillips Laboratories in Holland carried out extensive investigations on the high-frequency properties of such materials as manganese and nickel ferrite well into the UHF region. Today we have a large variety of ferrites, often of mixed composition, with copper, magnesium, manganese, nickel, zinc, cobalt, as the divalent ion in the spinel structure for many of these applications. The specific applications which are of greatest practical interest are pulse transformers, television deflection transformers and yokes, radio frequency inductors of fixed and variable variety, ferrite antenna rods, and magnetic core delay lines.

Although the ferrites have smaller hysteresis losses than the magnetic metals, the saturation magnetization which ranges from about 3000 to 5000 gauss is much lower (by almost an order of magnitude), so that at lower frequencies the prohibitive size makes their use less desirable. Another property of ferrites which is utilized in the application of the antenna rod is their large permeability at radio frequencies, which may be of the order of 1000 or more. This permits the size of the antenna rod to be very small. Furthermore, the magnetic antenna has a better signal-to-noise response than the electrical counterpart since the electrical component of noise is usually larger than the magnetic.

Magnetic Recording

Perhaps one of the most popular applications of remanent or permanent magnetization of materials is in the tape recorder. Although most of us think of the tape recorder as a modern invention, of the postwar era, the original patent was obtained by Poulsen at the turn of the century [7]. However, it did not realize its full potential until relatively recently.

The basic phenomenon of remanent or permanent magnetization is of equal importance for commercial and military use of magnetic drums in computers. It is known that many magnetic materials possess hysteresis loops; this property permits them to be set in a permanently magnetized state by appropriate inducement. The amplitude of the magnetization will depend upon the strength of the magnetic induction, which may be applied by a current through a coil or by the poles of a small electromagnet. For example, tape moving past the recording head when a signal is applied is left in a magnetized state, with a remanent magnetization proportional to the signal (see the Schematic drawing in Fig. 3). The inverse process requires that the tape induce a weak signal, proportional to the magnetization found at any point on the tape, in a magnetic circuit, without seriously affecting the state of magnetization found in the tape. Usually there is a small loss of signal, although the fidelity is not affected by successive runs. The direction of magnetization of the tape may be longitudinal, *i.e.*, parallel to the direction of motion, perpendicular, *i.e.*, at right angles to the surface of the tape and to the direction of motion; or transverse, *i.e.*, parallel to the surface of the tape but perpendicular to its motion.

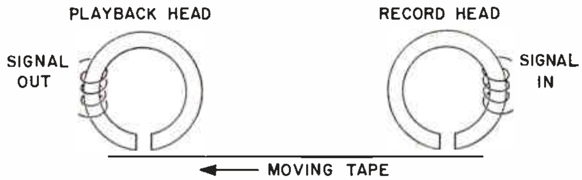


Fig. 3—Schematic representation of a magnetic recording and playback system.

The medium utilized for recording may have one of several geometric forms—wire, tape, or a rigid frame such as that in a magnetic drum. Metallic or nonmetallic material may be used. Metallic material in the form of a carbon steel wire was first used. Since then tungsten steel and stainless steel, whose magnetic properties are obtained by appropriately cold working and heat treatment, have been utilized [8]. The nonmetallic materials most commonly used are iron oxide powder, in small particles of a micron or less, incorporated in a nonmetallic binder. These particles are usually in the form of needles, which are often oriented by a magnetic field within the coating, usually flexible film of acetate or mylar. The most widely used oxide powders are γFe_2O_3 (red oxide) and, less extensively, Fe_3O_4 (magnetite or black oxide) [9]. The two properties that are important in the application for a recording medium are coercive force and magnetization.

For metallic materials the coercive force may range from about 40 to several hundred oersteds, and the remanent magnetization may be as high as 10 to 12 Kg, whereas for nonmetallic oxides the coercivity may have typical values of from 200 to 300 oersteds, and the remanence from 600 to 1200 gauss.

Fidelity of reproduction is one of the primary aims of recording. The nonlinearity of the magnetization curve however results in serious distortions. A linear transfer characteristic, and therefore, high-fidelity reproduction is achieved by mixing a high-frequency bias current with the signal to be recorded.

Electromechanical Devices

Propagation of mechanical waves in solids has been considered theoretically for well over a century, and has been the major topic of two volumes by Lord Raleigh [10]. From a fundamental viewpoint the existence of such waves, called *phonons*, is of interest to physicists because of their interaction with electrons in semiconductors and metals, or with magnetic spins in other solids. From a device viewpoint phonons acquired more practical interest when electro-mechanical interactions were observed as early as 1847 with Joule's discovery of magnetostriction [11], and in 1880 with the Curie brothers' discovery of piezoelectricity [12]. Today ultrasonic waves are used for a great many purposes in solids, including the detection of imperfections and faults in solids, and the composition of polycrystalline metals, as well as in more practical processes such as ultrasonic drilling and cutting, and in transducers, resonators, and filters [13].

Piezoelectricity

It has been found that when certain crystals are subjected to pressure, electrical charges appear on their faces; it has also been found, inversely, that voltages applied to the faces of a crystal produce corresponding changes in the sample dimensions. This phenomenon is called the piezoelectric effect. It exists only in crystalline materials which do not have a center of symmetry. Thus, when a stress is applied, the center of gravity of the positive and negative charges can be displaced relative to one another so that the permanent electric dipole moment is altered. There are a number of crystals which exhibit this property. Among the more common materials are quartz, rochelle salt, ammonium and potassium dihydrogen phosphate, and tourmaline.

Piezoelectricity is used as an electromechanical transducer in which electrical energy is converted to mechanical energy and vice versa, and also as a resonator for frequency control of oscillators, time standards and highly selective filters. In the latter application, which was introduced by Cady [14], the piezoelectric effect drives the crystal into mechanical vibration, which in turn reacts back to control the electrical impedance of the circuit. Quartz has been used almost exclusively as a resonator, since it is highly stable, has very low internal losses, and can be cut so that its properties are independent of temperature. In the laboratory, Q's as high as 6×10^6 have been measured.

However, it was as a transducer that this phenomenon first found use as a device. During World War I, in order to combat the submarine menace at the request of the French government, Langevin [15] began research which led to the development of a sonar system which employed quartz piezoelectric elements. Since then, this transducer action has been used for other devices such as phonograph pickups, ultrasonic delay lines, ultrasonic soldering and cleaning, and strain and acceleration gauges. In electromechanical transducers the maximum sound generation is obtained when the frequency of the applied electrical voltage coincides with one of the modes of vibration of the crystal element. In general, the smaller the dimensions of the element the higher the natural frequency ν. For example, for free-free thickness resonant vibrations of a slab, $\nu = nv/2t$ where v is the velocity of sound in the medium, t is the thickness of the slab, and $n=1, 2, 3$, etc. The dimensions thus set an upper limit to the highest frequency possible for sound generation. Using very thin slices of quartz fundamental frequencies just under 100 Mc are obtainable. With harmonic operation this technique sets an upper limit for ultrasonic wave generation of about 1000 Mc. More recently, new techniques for piezoelectric sound generation have been developed. Ultrasonic waves have been excited at the optically flat surface of quartz rods by placing the surface at the high electric field region of a microwave cavity [16]. These waves then propagate into the interior of the rod. The frequency of these waves has been extended to about 30 kMc, and it appears that the availability of electromagnetic driving fields now determines the highest frequency of ultrasound [17]. These microwave ultrasonic frequencies have been used in spin resonance experiments which have provided basic information on the spin-phonon interactions and selection rules in paramagnetic host crystals [18] such as MgO with Fe^{3+} and Mn^{2+} and ruby with Cr^{3+}. Another new type of generator employs semiconducting intermetallic compounds, such as InSb and GaAs, which have a zinc blende structure with no inversion symmetry and which are piezoelectric [19]. A junction is produced by depositing a thin metallic film on the semiconductor. The piezoelectric action takes place in the back-biased junction depletion area, which sustains all the applied voltage and behaves as a piezoelectric crystal. Since this depletion area is very thin, its frequency of operation is very high; by varying the back bias, the thickness of the depletion area and the resonant frequency of the transducer are varied accordingly. Depletion layers of the order of 0.1 μ, then, could provide an operating frequency which could exceed 10 kMc, with efficiencies about 100 times greater than any other known technique. At the present time a working model at Bell Telephone Laboratories operates at 600 Mc with a 5 per cent (30-Mc) bandwidth, which should provide storage of large amounts of information for computers [20].

Ferroelectricity

Ferroelectricity, the electrical analogue of ferromagnetism, involves the spontaneous alignment of electric dipole moments below the Curie point by material interaction. Ferroelectric materials are characterized by hysteresis loops such as that shown in Fig. 2, except in that the electric field E replaces the magnetic field H and the electric polarization P replaces the magnetic polarization M. Thus, by subjecting the material to an electric field a remanent polarization will persist, even after the removal of this field; and when thus polarized the substance exhibits piezoelectric behavior. Therefore, ferroelectric materials can be used as transducers and resonators in much the same way as piezoelectric crystals. Ferroelectric substances such as barium titanate have the highest known electromechanical coupling coefficients and are frequently used, consequently, as underwater sound transducers and other types of pickups and delay lines. However, because of the relatively large hysteresis losses, temperature variations and aging effects of ferroelectrics, piezoelectric materials are usually preferred for high-Q resonators. These ferroelectrics may be employed either as single crystals or as rugged ceramic material which may be cast in any desired shape for special applications. By using single crystal rather than polycrystalline ceramic ferroelectrics, the squareness of the corners of the hysteresis loop can be improved. Such square hysteresis loops have a potentially important use in information storage devices. Ferroelectricity was first discovered by Valasek [21] in 1921 in Rochelle salt, which is piezoelectric above its Curie point; however, the most important ferroelectric material today is barium titanate, which was discovered in 1942 and developed shortly thereafter at M.I.T. [22] for ferroelectric applications. Since then, other materials of the Perovskite structure and of other structures have been found to be ferroelectric [23].

More recent applications [24] of the electromechanical property of piezoelectric materials are the development of ferroelectric voltage transformers and ferroelectric electromechanical filters. The voltage transformer is a simple device in which the driver sets the ferroelectric specimen into vibration at one of its natural modes. The voltage output is taken off at the other end of the crystal, as shown in Fig. 4, at a lower current level but at a much higher voltage. The output voltage, which can be as large as 1000 volts rms and which can be generated with a bar of the order of 6 inches, is limited by the electrical breakdown or depolarization effects that are due to heating of the rod. These types of devices are much simpler and smaller than the more conventional high-voltage coil or magnetic transformers.

The other interesting device is the ferroelectric filter, which is essentially another use of the transformer element just discussed. The filtering action comes about from the high Q of the mechanical resonance required for transforming the input to an output signal. For many applications these filters may very likely replace large single crystals of quartz, since they are easily molded and fabricated. Additional flexibility is possible from composite filters made of both ferroelectric and magnetostrictive materials.

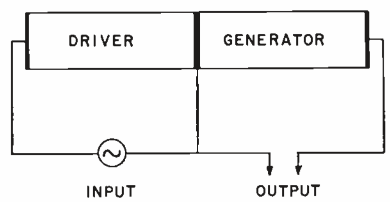


Fig. 4—Ferroelectric electromechanical voltage transformer. For maximum output voltage, the over-all length of the rod is made mechanically resonant at the electrical driving frequency.

Magnetostriction

Magnetostriction is a well-known phenomenon [11] in which the mechanical dimension of a magnetic material is altered as the magnetization is varied; and in which, conversely, the magnetization is altered as the dimension is changed. Thus, an alternating current applied through a coil wrapped around a specimen can induce mechanical vibrations in it; and, alternately, mechanical vibrations set up in such a specimen can transform the mechanical energy into electrical energy in a coil wound around it. The ideal material for such devices would have a large magnetostrictive constant, high permeability, low losses and a low Young's modulus. The material that fulfills these requirements most closely is nickel, which has been widely used for magnetostrictive transducers in applications very similar to those of piezoelectric devices. Ferrites also appear to have promise in this respect [25], although their electromechanical coupling is smaller; however, they are suitable for high-frequency filters in the Mc range, whereas nickel magnetostrictive devices operate best in the 20 to 50 kc range. A unique advantage of the magnetostrictive as well as ferroelectric mechanical filter is the characteristic (shown in Fig. 5) of remarkably steep sides with a flat top, which makes such filters superior for many applications to the quartz filter. The magnetostrictive properties of magnetic ma-

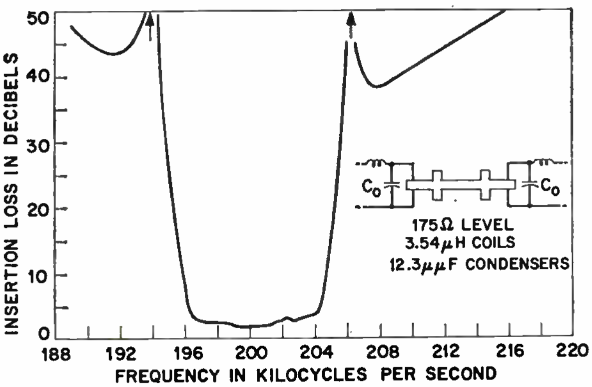


Fig. 5—Measured attenuation for a two-section mechanical filter made from barium titanate ceramic. [*After* W. P. Mason, "Physical Acoustics and the Properties of Solids," D. Van Nostrand Co., Inc., Princeton, N. J., p. 146; 1958.]

terials have also been used to construct delay lines, particularly for low-speed computer applications.

Dielectric Materials

Optical and Infrared Devices

The more conventional optical devices, such as lenses and prisms, are well known to most scientists. However, in recent years large single crystals of quartz and various alkali halides have provided selective dispersion in different regions of the infrared spectrum; for example, NaCl is most useful from 8 to 15 μ, KBr from about 12 to 24 μ and CsBr and CsI up to 40 or 50 μ, respectively. In the infrared spectrum, elementary semiconductors such as Ge and Si have been used as lenses and as polarizers, since at wavelengths corresponding to energies below the energy gap these materials behave as conventional low-loss dielectrics. Polarizers of such materials, with light incident at the Brewster angle, have been successfully used in the wavelength region from a few μ to 20 μ. Another useful effect in these materials is associated with the anomalies in the dielectric at long wavelengths well into the far-infrared region of the spectrum. These are the so-called restrahlen frequencies, which are associated with resonances of the vibrations of ions in the lattice. At these resonances the electromagnetic wave transmitted through a thin crystal shows a peak absorption. At a slightly different wavelength the reflection shows a relatively narrow peak known as the residual ray, or restrahl. This reflection peak is utilized as a selector for narrow-band radiation to permit selected far-infrared radiation from an incandescent source. Thus NaCl gives a peak reflection at 52 μ, KCl at 63 μ and KI at 94 μ, providing sources of "monochromatic radiation" in a very simple fashion; often several of these plates are used in series to narrow the bandwidth to perhaps 5 per cent. Such restrahlen plates are also used as filters for the far infrared.

Fiber Optics

Fiber optics is a very exciting development which utilizes the dielectric properties of glass. These fibers are essentially electromagnetic dielectric waveguides analogous to those found at microwave frequencies, except in that the electromagnetic radiation at optical frequencies is reflected internally by a well-known phenomenon in optics. The angle of incidence of the internally reflected wave is well below the critical angle, known as the Brewster angle, so that there is total internal reflection. It is possible to analyze the modes mathematically and to obtain the field configuration as a consequence. In practice, when light is transmitted through such fibers the mode configuration can be optically displayed as shown in Fig 6, and readily identified in terms of the mathematical analysis [26]. Optical fibers used in bundles are finding important applications as display devices, as a snooper scope, and for illuminating and examining the internal organs of patients. It is conceivable that fibers containing optically active impurities may lead to the development of traveling-wave optical masers.

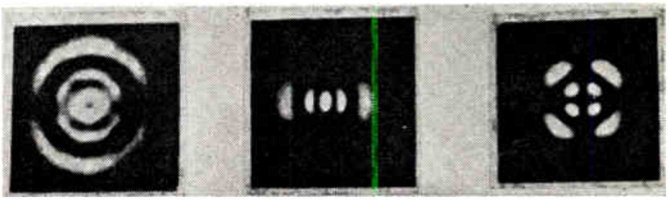


Fig. 6—Dielectric waveguide modes in the visible region. [*After* E. Snitzer, and A. Osterberg, "Observed dielectric waveguide modes in the visible spectrum," *JOSA*, vol. 51, pp. 499–505; May, 1961.

Of course, dielectric waveguides [27] are not unique to optical frequencies, but have been used at microwave and millimeter waves for some time. They operate on the principle that the electromagnetic energy is concentrated in the dielectric rod which acts as a waveguide, and that electromagnetic energy beyond the air-dielectric interface falls off exponentially. Thus the electromagnetic energy is confined within a small region surrounding the dielectric waveguide. This principle has been effectively used at millimeter wavelengths to build relatively low-loss components, including such devices as directional couplers [28].

Computer Elements

Magnetic Memories

The last decade has seen the advent of modern electronic computers, which are likely to reshape our modern technology perhaps more than any other single invention. At the heart of this modern computer lies a tiny ferrite core which serves as its memory element. The ferrite, as we have mentioned, is a magnetic insulator which in an appropriate composition possesses a rectangular hysteresis loop on magnetization. This makes it an excellent bistable magnetic memory element. A current I through a wire interleaving the memory core can magnetize it clockwise or counterclockwise, representing the binary digits 0 or 1, respectively. Usually this is done by two sets of wires which interleave a two-dimensional array of cores such as that shown in Fig. 7. Each wire carries a current $I/2$, thus magnetizing the core in one or the other direction when both wires are carrying this current in the same sense. This is known as current coincidence. In practice, a large number of these ferrite memories is utilized in a modern computer, exceeding a million in a few very large ones, such as the TX-2 Lincoln Laboratory computer and/or the IBM Stretch computer. Small cores of magnesium manganese ferrite have been engineered to give a minimum response time of 10^{-7} seconds, which is limited primarily by the domain wall motion. However, the random-access times of such systems range from 1 to 10 μsec. Therefore, in order to increase the time response and ultimately the capacity of modern computers, new memory and switching elements have been sought.

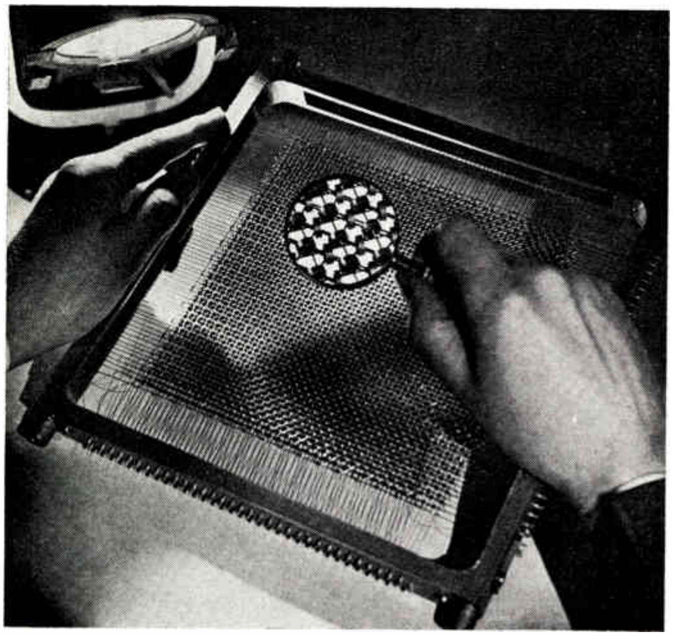


Fig. 7—Memory matrix of ferrite cores. [*After* W. N. Papian, M.I.T. Lincoln Laboratory, Lexington, Mass., private communication.]

Two approaches have been considered, one using thin magnetic films [29] and the other using cryogenic components such as the cryotron [30] and cryosar [31]. The principle of operation of thin magnetic films is quite different from that of the toroidal ferrite core, in which the two states of magnetization were in opposite directions in the toroidal plane. In the thin film, an internal uniaxial anisotropy is introduced by annealing in a magnetic field. The thin film of approximately 1000 Å is is deposited on a glass substrate. The material of 80–20 permalloy, which was first reported by Blois [32], showed a characteristic rectangular hysteresis loop. Following this work, these films have been fabricated with switching speeds of the order of mμsec, or 300 times faster than the ferrite cores in a comparable driving field of about 1 oersted. The magnetic film memories, such as that in Fig. 8, can be made as small or smaller

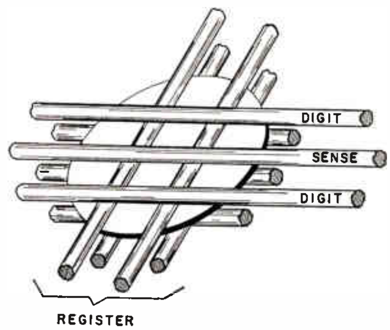


Fig. 8—Single magnetic film memory cell. The film which is supported on a glass substrate is about ¼ inch in diameter and 1000 Å thick. For high-speed switching drive lines, simulate transmission lines. [*After* J. I. Raffel, M.I.T. Lincoln Laboratory, Lexington, Mass., private communication.]

than the ferrite cores; furthermore, transmission-line techniques for magnetizing the films will permit higher-speed computers with access time well below the μsec range.

Cryogenic Elements

A superconducting switching element was first examined by de Haas and Casimir-Jonker [33] in 1935; however, it was not until 1955 that Buck demonstrated a practical device which he called the cryotron [30]. The basic principle of the cryotron depends on the existence of a critical magnetic field above which the superconducting metal becomes a normal conductor. The original cryotron utilized a small tantalum rod wound with a niobium wire. At liquid helium temperatures the tantalum wire has a critical field of the order of several hundred gauss, whereas that of niobium is of the order of 2000 gauss. Consequently, when the niobium wire is pulsed with a suitable current, the magnetic field that it creates is sufficient to destroy the superconductivity in the tantalum but not in itself. The current in the niobium wire can be smaller than that in the tantalum wire so that a small current can control a larger one, thus producing a current gain in the device. The more recent cryotron elements, such as that shown in Fig. 9, utilize small superconducting strips which cross one another and are separated by thin insulators (hundreds of angstroms thick) in which one superconducting strip controls the current in the other [34]. Usually the narrower film with the greater critical field controls a larger current in the thicker film. Cryotron loops performing the function of a memory can be made in very small configurations, possibly 100 to 1000 per cubic inch with extremely low power dissipation such that a network of a million cryotron loops would require about 1 watt of power. The switching time is now competitive with that of the thin magnetic films and is of the order of 10 mμsec. Although no computer as yet has been built utilizing the cryotron, the possibility in the near future looks promising.

Even though it is a semiconducting device, another computer element that should be mentioned in this context is the cryosar [31]. Germanium containing impurities can be made conducting at liquid helium temperatures by an avalanche process. Localized breakdown can be produced by very low fields of order of a few v/cm, permitting switching junctions of crossed conducting strips on either side of a thin germanium slab such that as many as 1000 per cm² can be fabricated on a thin wafer. Thus as many as 10^4 elements in a cubic centimeter are possible. The breakdown phenomenon is extremely fast and the time involved is shorter than that of other computing elements mentioned. A compound cryosar has been developed by McWhorter and Rediker which provides bistable elements in a coincident-voltage memory for computer applications.

At present, the difficulty with cryogenic computer elements involves the necessity of providing many leads from the low-temperature Dewar to the driving transistors at room temperature. The large number of wires required results in an undesirable heat leak. However, it appears that, in the future, low-gap semiconductors and semimetals may provide transistors and transistor-like devices [35] making it possible to contain the entire computer at liquid helium temperatures except for the output and input terminals.

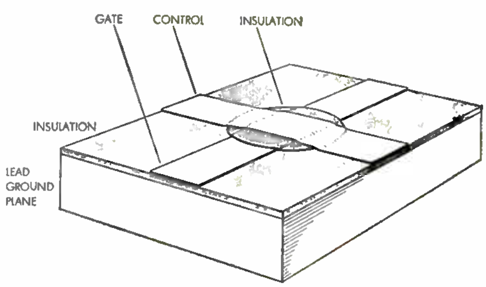


Fig. 9—Film cryotron. [*After* W. B. Ittner and C. J. Craus, "Superconducting computers," *Sci. Am.*, pp. 124–136; July, 1961.]

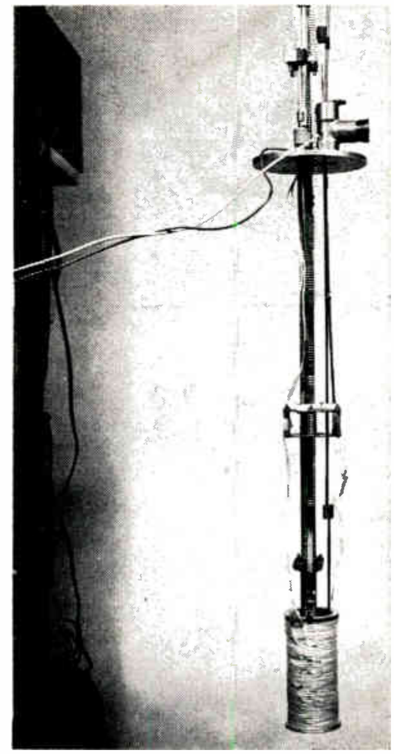


Fig. 10—Tunable *L*-band maser using superconducting solenoid to provide the magnetic field. This particular solenoid can supply up to 6500 gauss with a 1½-volt dry cell and a variable resistor. A closed superconducting circuit may be set up, after which a persistent current maintains the magnetic field and the dry cell may be removed. [*After* S. H. Autler, M.I.T. Lincoln Laboratory, Lexington, Mass., private communication.]

Other Superconducting Applications

That several other superconducting devices have been developed, in addition to the cryotron, makes this phenomenon more than an intellectual curiosity. Perhaps the superconducting magnet is the one that is likely to have the greatest technological impact. It is known that Onnes experimented with small coils of superconducting lead for making solenoids capable of producing a few hundred gauss [36]. A few years ago this idea was extended by Yntema [37], who made a coil of niobium wire to produce fields of the order of several thousand gauss. However, the first serious application was that of Autler [38], who used niobium to make superconducting magnets for use with microwave masers, as shown in Fig. 10. He wound wire which was stressed such that he was able to reach approximately 8 to 10 Kg, and took advantage of the persistent current to produce very stable fields by shorting out the battery supply. He also developed small electromagnets immersed in a helium Dewar in which the fringing field around the superconducting coils was below the critical value. However, the fields within the pole tips were about 25 Kg in a gap$\sim\frac{1}{4}$ inch. The major breakthrough was made very recently by Kunzler, *et al.* [39], at Bell Telephone Laboratories, when they fabricated niobium tin wires with critical fields in the vicinity of 200 Kg. Subsequently, niobium zirconium wires have been made by this group with smaller current capacity and critical fields of the order of 80 Kg and current densities of 10^4/cm². With appropriate technological developments it appears that superconducting magnets with fields of between 50–100 Kg should be feasible in the near future.

A number of devices based on the Meissner effect [40] and its related properties in superconductors are receiving serious consideration at this time. It is known that a superconductor below the critical magnetic field is essentially a perfect diamagnetic material; *i.e.*, the magnetic field does not penetrate the superconductor, which behaves as a magnetic mirror. Thus it is possible to suspend a magnet above a spherical lead mirror immersed in a liquid helium bath. Alternatively, a magnetic flux can be used to suspend a superconducting sphere or plate in space, thereby providing a frictionless bearing. This provides the basis for a superconducting gyroscope with possibilities for application in inertial guidance systems. Other possible applications which would utilize this type of phenomenon are superconducting motors and magnetic focusing for increasing the resolution of electron microscopes.

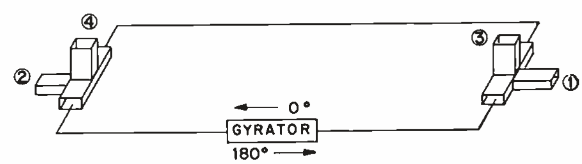
The bridge between fundamental theoretical discovery and its application to technology is beautifully illustrated by the tunneling phenomenon observed in superconductors. Approximately four years ago the far-infrared experiments of Tinkham and Glover [41] gave evidence of an energy gap in superconductors. At about that time a quantum theory of superconductivity developed by Bardeen, Cooper and Schrieffer [42] quantitatively predicted this energy gap. Perhaps the most elegant demonstration of the existence of the energy gap and the related details of the density of states was provided by experiments of Giavaer [43], in which the tunneling between superconductors separated by a thin insulating aluminum oxide film was observed. Undoubtably this phenomenon will be exploited for a number of practical devices. Already it has been suggested that such sandwiches can be utilized as sensitive photoconducting detectors in the far infrared [44].

Microwave Ferrites and Ferrimagnetics

Nonreciprocal and Reciprocal Devices

The last decade has seen a great many practical developments in solid state, but perhaps one of the most far reaching was that involving the microwave application of ferrites. A great stimulus for this work was the demonstration in 1952 by Hogan [45] of the microwave circulator utilizing the nonreciprocal Faraday rotation in a circular waveguide. Soon thereafter, Kales, Chait and Sakiotis observed nonreciprocal phase shift and loss in ferrite loaded rectangular waveguides [46]; this led to a number of devices, soon exploited commercially. The phenomenon involved in the rectangular waveguide nonreciprocal components is rather straightforward. It is known that an electromagnetic wave traveling in a rectangular waveguide possesses an RF magnetic field component which is circularly polarized at a point between the center and the side of the waveguide. The sense of circular polarization is opposite, depending on the direction of propagation. Consequently, if a ferrite specimen in the shape of a long flat bar magnetized transversely is inserted in this region of the waveguide, the polarization of the magnetic field within the ferrite remains circular. For one direction of propagation the spin dipoles of the ferrite precess in the same sense as the rotation of the RF magnetic field, thereby reacting strongly with the electromagnetic wave. For the opposite direction of propagation the precession is opposed to the rotation of the RF field, so that the interaction with the electromagnetic field is much weaker. Consequently, the phase shift is said to be nonreciprocal, since the change is much greater for the first direction of propagation then for the second. When the magnetic field is such that the frequency of spin precession differs substantially from the frequency of the electromagnetic wave, the interaction is essentially reactive. When the magnetic field is adjusted so that the frequency of precession is equal to that of the microwave frequency, then for the positive direction of propagation (i.e., when the rotation and precession are in the same sense), the ferrite absorbs the electromagnetic energy rather strongly. In the opposite direction the absorption is weak. This latter device is known as the resonance isolator.

The performance of these devices, when appropriately designed, in microwave circuits agrees well with the figure of merit predicted theoretically, based upon the loss properties of the ferrites [47]. The nonreciprocal phase shift can be used in a microwave circuit in which it forms one branch of a four-terminal network, which may consist of two magic T's as shown in Fig. 11(a). The ferrite phase shifter provides the same phase shift in the two arms for one direction of propagation but a difference of 180° for the other direction of propagation. If the microwave energy is then properly traced through such a circuit, it can be shown that it becomes a circulator such that if energy is fed into arm 1 it comes out in arm 2, if fed in arm 2 it comes out in arm 3, as indicated in the circuit symbol of Fig. 11(b). More recently, a



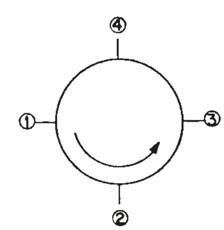


Fig. 11—Ferrite circulator. (a) Schematic representation indicating microwave circuit with either a nonreciprocal phase shifter or gyrator in one arm and magic T's in the other two arms. (b) Symbolic representation.

much simpler circulator known as the three-port or Y circulator has been developed. It is a planar junction with a 120° symmetry, and contains a ferrite rod or post at its center such that each arm is perfectly matched. It has been shown theoretically and demonstrated experimentally that under these conditions energy fed into arm 1 goes out arm 2, energy from arm 2 proceeds through arm 3, energy through arm 3 goes to arm 1, and so on. Circulators of the latter type have been constructed from the UHF frequencies of several hundred Mc well into the mm region as high as 150 kMc [48].

Isolators use a variety of ferrites, from substitutional low magnetization aluminates which are useful in the UHF region, to the barium ferrites or magneto-plumbites with high anisotropy which extend into the millimeter region to about 70kMc [49]. Beyond this range the isolators become impractical because the internal magnetic field of the highly anisotropic magneto-plumbite is inadequate to provide sufficiently high magnetic fields or resonance. This has been overcome by the use of antiferromagnetic materials at low temperatures, in which resonance occurs in internal fields of the order of 50,000 oersteds or more. Heller, *et al.*, using Cr_2O_3, have built a resonance isolator operable at 150 kMc [50].

Nonreciprocal ferrite devices have been engineered to handle high peak powers of Mw, as well as average power in the kilowatt range for use in communications and radar systems to isolate transmitters, and also receiving circuits, at low-power levels. A number of other nonreciprocal and reciprocal devices at microwave frequencies have been constructed in a variety of configurations. Dielectric-loaded coaxial line is used to reduce the size of the devices at lower frequencies [51]. The dielectric distorts the normal TEM mode and generates a longitudinal component of RF magnetic field to provide circular polarization for nonreciprocal devices. Similar techniques have permitted the use of ferrites in strip lines as well.

Another nonreciprocal device which has received a fair amount of attention is the field displacement isolator [52], which depends on the different electric field configurations which exist in the ferrite for the two opposing directions of propagation in a ferrite-loaded rectangular waveguide. When the parameters of the ferrite are suitably adjusted there is an electric field null at a ferrite-to-air interface for one direction of propagation, whereas a large field component exists for the other. By coating the interface with a lossy material a field displacement isolator with a large figure of merit can be obtained. These are usually low-power devices.

Applications using the reciprocal properties of ferrites and waveguides have also been developed. Among these perhaps the most important is the ferrite phase shifter located at the center of a rectangular waveguide and magnetized longitudinally as shown in the inset of Fig. 12. This device [53] operates on the principle that electromagnetic energy is concentrated primarily in the ferrite, due to its relatively large dielectric constant. The RF permeability is increased or altered substantially as the ferrite rod or slab is magnetized, thereby producing a large reciprocal phase shift. The principle of operation has been theoretically accounted for on this basis [54]. Another device that has been utilized in an analogous manner is a cavity resonator which can be used either as a filter or as a variable tuning device where frequency is altered by the application of an externally applied magnetic field which magnetizes the ferrite rod in the cavity [55].

Other passive microwave ferrite devices which have been developed are the single-sideband modulator, which utilizes the birefringant properties of a transversely magnetized ferrite; ferrite switches; ferrod radiators using ferrite rods often in an array [56]; and nonreciprocal isolating elements in traveling-wave tubes and traveling-wave masers.

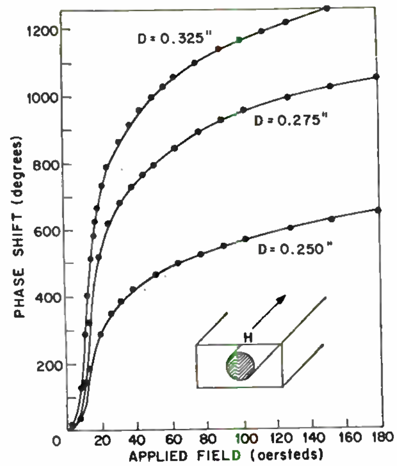


Fig. 12—Reciprocal phase shift for a longitudinally magnetized ferrite rod at the center of a rectangular waveguide as a function of applied dc magnetic field with rod diameter as a parameter. The phase shift increases sharply at the threshold thickness of 0.2 inch at *X*-band wavelength. [*After* F. Reggia and E. G. Spencer, "A new technique in ferrite phase shifting for beam scanning of microwave antennas," PROC. IRE, vol. 45, pp. 1510–1517; November, 1957.]

Nonlinear Devices

From theoretical considerations of the equation of motion of a magnetic dipole in a large RF magnetic field, it is evident that the motion exhibits a nonlinear behavior. The first exploitation of this phenomenon was by Ayres, Vartanian and Melchor [57], who designed a harmonic generator by simply driving a ferrite specimen with a large linear RF magnetic field transverse to the applied dc field. At the high-power levels obtained from a magnetron they were able to get quite effective frequency conversion which now has been extended into the mm region and capable of producing peak powers of the order of 50 watts at 150 kMc. This phenomenon obviously can be utilized as a nonlinear mixer as well.

Shortly after this development, the theoretical work of Suhl [58] on the nonlinear behavior of ferrimagnetic materials at high-power levels led him to propose a new device known as the ferromagnetic amplifier. This amplifier operates on the principle of a parametric device in which the spin system is pumped at the resonant frequency and the signal frequency is at a lower value such that it is equal to the pump frequency reduced by the idler frequency. In the original amplifier (shown in Fig. 13 constructed by Weiss [59], the idler and the signal frequencies corresponded to resonant modes of the microwave cavity and were operated with a pulsed signal. This is known as the electromagnetic mode of operation. Since then a variety of other combinations using different modes of resonance in the ferrites have been demonstrated [60]; however, the particular mode that proved to be most successful is the magnetostatic mode of operation which in addition used longitudinal pumping of the spin system. This device utilized an RF magnetic field at the pump frequency parallel to the dc magnetic field and equal to twice the ferromagnetic resonant frequency. The idler and the signal frequencies,

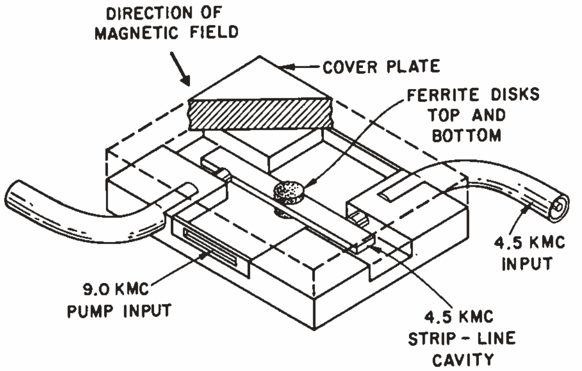


Fig. 13—Cavity configuration for microwave ferrite amplifier. [*After* M. T. Weiss, "A solid state microwave amplifier and oscillator using ferrites," *Phys. Rev.*, vol. 107, p. 317; July, 1957.]

straddled on either side of the ferromagnetic resonance, employed long wave length spin-wave modes of the ferrimagnet, fulfilling the frequency requirement of a parametric device. The amazing result was that the device developed by Denton [61] operated as a CW resonator and amplifier with a relatively low level of pump power of the order of 500 mw. The gain of the device was about 25 db, with a bandwidth of 100 kc and a noise figure of approximately 12 db.

Still another important application of the nonlinear behavior of spin systems in ferrimagnets is the microwave power limiter. According to Suhl, as the power into the spin system is increased, the ferrimagnet exhibits greater loss corresponding to the increased saturation of the spin system. Thus, a waveguide circuit containing such a ferrite can behave in a manner analogous to a gas T-R switch. Unfortunately, this high-power device exhibits a spike in the initial phases of the pulse which prohibits its use as an effective device for crystal protection in the receiving circuit. A low-power application, developed by DeGrasse [62], utilizing the narrow resonance line of yttrium iron garnet (YIG) takes advantage of the substantial decrease in susceptibility at relatively low threshold power in two transversely crossed microwave lines which become decoupled when the susceptibility of the garnet is reduced. Although the nonlinear properties were first discovered in ferrites at high microwave powers [63], the significant exploitation of these devices followed the discovery of the yttrium iron garnet whose resonance linewidth under suitable conditions was found to be one to two orders of magnitude narrower than that of ferrite crystals. In addition, the garnet materials provided an ideal medium for studying the fundamental properties of ferromagnetic resonance.

Masers

Microwave

The word *maser* was first coined by Gordon, Zeiger and Townes [64] to describe the microwave amplifier which operated by removing the energy from an inverted population of an electrically focused ammonia beam by stimulating the emission of these atoms from the excited to a lower energy state. The maser really became a much more practical device after Bloembergen [65] proposed the three-level scheme using the resonance phenomena in paramagnetic materials. He suggested that microwave energy from a source excites spins from the lowest energy level 1 to the highest energy level 3 such that at low temperatures and at relatively high microwave power the number of spins occupying state 3 would exceed considerably that at the intermediate level state 2. With an inverted population, if a microwave cavity were tuned to the frequency corresponding to the energy separation $\epsilon_{23}=h\nu_{23}$, then maser action would ensue at this frequency by stimulated emission. The device would oscillate if the total losses, including those of the cavity and of the internal circuits coupled to it, were exceeded by the energy stimulated within the crystal. However, if the stimulated energy is sufficient to overcome the losses of the cavity but not those of the total system, then the device operates as an amplifier. Indeed, such a maser was built at Bell Telephone Laboratories [66], and it demonstrated this principle unequivocally. Subsequently, using such materials as potassium chromicyanide [67] and ruby [68], microwave amplifiers with operation over a wide frequency range have been constructed. The dual frequency cavity for one of these masers is shown in Fig. 14. These masers were shown to have the low noise figure predicted theoretically [69], which is essentially equal to that of the crystal itself plus the losses encountered in the microwave circuitry associated with the maser. The initial noise figures corresponded to temperatures of the order of 20°K, but figures of the order of 17°K have been obtained subsequently [70]. These masers have since been used in radioastronomy for obtaining radar as well as radiometric signals from Venus and for the observation of the hydrogen line from external galactic systems [71]. Significant progress has also been made in the design and construction of traveling-wave masers, which has permitted the increase of the gain bandwidth product of such a device to the order of 13 db with a bandwidth of 67 Mc and tunable over a 350-Mc range with a pumping power of 100 mw at a frequency of about 19 kMc [72]. The three-level maser principle has now been extended by the group at Lincoln Laboratory [73] as far as 70 kMc, utilizing the levels of Fe^{3+} impurity in rutile TiO_3 such that the frequency range from 26 to 76 kMc can be achieved in two separate cavities. Because of the high dielectric constant of this material, many resonant modes of operation were obtained in this interval by merely rotating the magnetic field relative to the crystal axes. The pump power for the lower frequency range was obtained from a 4-mm klystron, whereas for the higher frequency range the pump was a 2-mm carcinatron, and both were at a power level of the order of about 5mw.

In addition to the three-level masers, a number of two-level devices have also been developed which

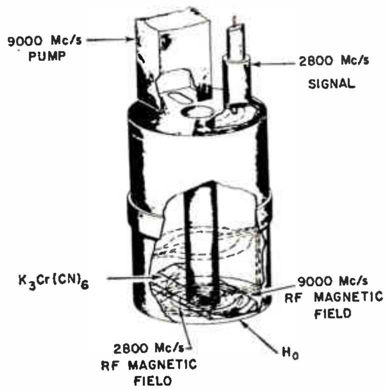


Fig. 14—Three-level maser cavity for 2.8 kMc operation with *X*-band pumping. [*After* A. L. McWhorter and J. W. Meyer, "Solid-state maser amplifiers," *Phys. Rev.*, vol. 109, pp. 312–318; January 15, 1958.]

operate on a pulse basis. The first of these two-level schemes [74] utilized adiabatic fast passage in Si, and although unsuccessful, it was actually the precursor of the first such successful solid-state maser which was operated by Feher, *et al.*, at Bell Telephone Laboratories [75]. Another maser employed a two-level electron-spin system in single crystals of quartz and magnesium oxide containing paramagnetic levels introduced by neutron irradiation, and was operated by the group at Westinghouse [76], who also obtained inversion of the electron population by adiabatic rapid passage in which the magnetic field was swept through resonance. A unique maser using pulse techniques was that developed by Foner [77], in which two sets of levels in ruby were inverted by pumping at 9 kMc, then pulsing a magnetic field in a period of the order of 100 μsec, a time which is short compared to that of the relaxation of the inverted spin system. By this technique the energy separation between two inverted levels made it possible to obtain stimulated emission as high as 70 kMc. Two interesting features of this system were that the signal frequency was higher than the pump frequency, and that energy was supplied not only by the pump power but also by the magnetic field.

Optical and Infrared Masers

Although the initial proposal of optical masers by Schawlow and Townes was centered around well-known spectral transitions in alkali vapors [78], the first optical maser was a solid-state device, developed by Maiman [79], who used a ruby rod containing 0.05 per cent trivalent Cr^{3+} impurities surrounded by a helical xenon flashtube which excited electrons into a band of levels as shown in Fig. 15. These electrons then quickly emptied into a lower set of doublet levels whose lifetime relative to the ground state was of the order of msec. During the pulse, which was of the order of about 100 μsec, it was possible to excite more than half the electrons from the ground state into the *R* levels, thus inverting the population momentarily. The silver-plated

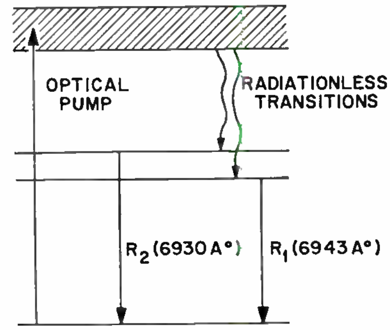


Fig. 15—Energy levels of dilute ruby (Cr^{3+} in Al_2O_3) utilized in the ruby *R*-line maser.

mirror ends of the rod provided an interferometer which acted as the resonant cavity of the microwave maser, thus producing stimulated emission from the *R* levels to the ground state, resulting in maser action at 6943 Å. Subsequently the coherence properties of the ruby maser were demonstrated by the Bell Telephone Laboratories group [80], who observed diffraction patterns produced by the beam of the maser. Masers similar to that of the ruby but effectively using four levels instead of three were developed by Sorokin and Stevenson, who initially employed uranium and then later samarium in CaF_2 host crystals [81]. These have the property that the ground state is split; therefore, at low temperatures the Boltzmann factor reduces the equilibrium population of the higher of the two pertinent ground states, thus requiring less energy for inverting the excited states relative to the higher ground state. Indeed, their maser required pumping power one order of magnitude below that of the ruby maser. The four-level optical maser appears more promising than the three-level as a candidate for CW operation, although it has not yet been achieved. The only CW device is the gas maser of Javan [82], which operates in the infrared at 2.1 μ. The red ruby maser emits lines at 6943 Å, but another ruby maser [83] with high concentration of impurities 0.7 per cent Cr^{3+} produces a splitting of the ground level into two satellite lines due to pairing of chromium ions so that emission at 7040 and 7010 Å has been observed at 77°K. One other solid-state maser that should be mentioned in this context is the cyclotron resonance maser [84], which is essentially a three-level device utilizing transitions of electrons from the quantized magnetic levels in the valence band to those of the conduction band of the same quantum number. Then cyclotron resonance transitions between neighboring levels within the valence band produce stimulated emission for maser action. Although this maser has yet not been built, theory indicates that it is feasible if an optical ruby maser or similar device with a pulse output well in excess of 100 watts is utilized and if magnetic fields of the order of 100,000 gauss are imposed at liquid helium temperatures on crystal wafers of Ge. This maser should be magnetically tunable in the far infrared and submillimeter region of the electromagnetic spectrum.

Conclusions

From the variety of applications discussed in this paper, it is apparent that solid state has become an extremely important technological field for the future. Areas of research in this science which were viewed as intellectual curiosities until recently have now become endeavors of great practical significance. Superconductivity is probably the most notable example. Just within the last two years it has emerged from the cocoon of basic physics to a practical prominence which may well revolutionize many electrical and electronic devices. Even semimetals, which are closely related to semiconductors, have future possibilities at low temperatures for electronic devices similar to those of the transistor and the tunnel diode.

The magnetic materials of the postwar era, ferrites and garnets, have already proved of extreme value for computer and microwave applications. In addition, ferrimagnetic materials and their related antiferrimagnetic counterparts show natural resonances in the submillimeter and far-infrared region which undoubtably will be exploited in a practical way when the technological art in this range of the frequency spectrum is sufficiently advanced. Such advances do not seem too far away. Masers, of one variety or another, should provide coherent monochromatic sources much more energetic and efficient than the black-body radiation or harmonic generators used at the present time. Low-temperature detectors using the solid-state phenomena of superconductivity or maser amplifiers combined with coherent sources should make this energy-starved region as fruitful for the future as the microwave region has been for the last two decades.

Without a doubt, unexploited phenomena in dielectrics, in alkali halides and in metals involving electrical, mechanical, magnetic and optical effects, which are today in the province of the research physicist tomorrow will become the bases of new inventions by the applied scientist.

References

[1] E. F. W. Alexanderson and S. P. Nixdorff, "A magnetic amplifier for radio telephony," Proc. IRE, vol. 4, pp. 101–129; April, 1916.
L. Kuhn, "Uber ein neues radiotelephonisches System," *Drah. Teleg. Teleph.*, vol. 9, pp. 502–534; 1914.

[2] F. Benjamin, "Improvements extend magnetic amplifier amplifications," *Electronics*, vol. 25, pp. 119–123; June, 1952.

[3] R. E. Morgan, "The amplistat—a magnetic amplifier," *Elec. Eng.*, vol. 68, pp. 663–667; August, 1949.

[4] E. L. Harder, "Power control with magnetic amplifiers," *Electronics*, vol. 25, pp. 115–118; October, 1952.

[5] G. Hilpert, "Genetische und konstitutive Zusammenhänge in den magnetischen Eigenschaften bei Ferriten und Eisenoxyden," *Berichte deutsch. chemisch. Gesell.*, vol. 42, pp. 2248–2261; 1909.

[6] J. L. Snoek, "New Developments in Ferromagnetic Materials," Elsevier Publishing Co., New York, N. Y.; 1947.

[7] V. Poulsen, U. S. Patent No. 661,619; November 13, 1900.

[8] S. J. Begun, "Magnetic Recording," Murray Hill Books, Inc., New York, N. Y.; 1949.

[9] J. Tall, "Techniques of Magnetic Recording," The Macmillan Co., Inc., New York, N. Y.; 1950.

[10] Lord Raleigh, "The Theory of Sound," Dover Publications, Inc., New York, N. Y.; 1945.

[11] J. P. Joule, "On the effects of magnetism upon the dimensions of iron and steel bars," *Phil. Mag.*, vol. 30, pp. 226–241; April, 1847.

[12] J. Curie and P. Curie, "Developpement, par pression, de l'electricite polaire dans les cristaux hemiedres a faces inclinees," *Comp. Rend.*, vol. 91, pp. 294–295; July–December, 1880.

[13] W. P. Mason, "Physical Acoustics and the Properties of Solids," D. Van Nostrand Co., Inc., Princeton, N. J.; 1958.

[14] W. G. Cady, "The piezoelectric resonator," *Phys. Rev.*, vol. 17, p. 531; April, 1921.

[15] M. Langevin, "Echo sounding," *Nature*, pp. 689–690; May, 1925. French Patent No. 505,903; 1918.

[16] K. N. Baranskii, "The excitation of hypersonic frequencies in Quartz," *Soviet Phys. Dokl.*, vol. 2, pp. 237–238; May–June, 1957.
H. Bommel and K. Dransfield, "Excitation of very-high-frequency sound in quartz," *Phys. Rev. Lett.*, vol. 1, pp. 234–236; October 1, 1958.
E. H. Jacobson, "Piezoelectric Production of Microwave Phonons," *Phys. Rev. Lett.*, vol. 2, pp. 249–250; March 15, 1959.

[17] ——, private communication.

[18] N. S. Shiren and E. B. Tucker, "Selection rule for the interaction of microwave ultrasonics with spins," *Phys. Rev. Lett.*, vol. 6, pp. 105–106; February 1, 1961.
E. B. Tucker, "Attenuation of longitudinal ultrasonic vibrations by spin-phonon coupling in ruby," *Phys. Rev. Lett.*, vol. 6, p. 183; February 15, 1961.

[19] J. H. Wasilik and R. B. Flippin, "Piezoelectric effect in indium antimonide," *Phys. Rev. Lett.*, vol. 1, pp. 233–234; October 1, 1958.

[20] D. L. White, "Depletion layer ultrasonic transducer—a new high frequency transducer," 1961 IRE International Convention Record, pt. 6, pp. 304–309. Abstract in Proc. IRE, vol. 49, p. 681; March, 1961.

[21] J. Valasek, "Piezoelectric and allied phenomena in rochelle salt," *Phys. Rev.*, vol. 17, pp. 475–481; April, 1921.

[22] E. Wainer and A. N. Salomon, Titanium Alloy Manufacturing Co., Elec. Rept., vol. 8, and vol. 10, 1943.
A. von Hippel, R. G. Breckenridge, F. G. Chesley and L. Tisza, "High dielectric constant ceramics," *Ind. Engrg. Chem.*, vol. 38, pp. 1097–1109; November, 1946.

[23] B. Matthias, "Phase transitions in ferroelectrics," in "Phase Transformations in Solids," R. Smoluckowski, Ed., John Wiley and Sons, Inc., New York, N. Y.; 1951.

[24] C. A. Rosen, "Ceramic Transformers and Filters," Syracuse University, Syracuse, N. Y., Ph.D. dissertation; 1956. Also in "Solid State Magnetic and Dielectric Devices," H. W. Katz, Ed., John Wiley and Sons, Inc., New York, N. Y., Ch. 8; 1959.

[25] W. van B. Roberts, "Some applications of permanently magnetized ferrite magnetostrictive resonators," *RCA Rev.*, vol. 14, pp. 3–16; March, 1953.

[26] E. Snitzer, "Cylindrical dielectric waveguide modes," *JOSA*, vol. 51, pp. 491–498; May, 1961.

[27] S. A. Schelkunoff, "Electromagnetic Waves," D. Van Nostrand Co., Inc., New York, N. Y., pp. 425; 1943.

[28] S. E. Miller, "Coupled-wave theory and waveguide applications," *Bell Syst. Tech. J.*, vol. 33, pp. 661–719; March, 1954.

[29] A. V. Pohm and E. N. Mitchell, "Magnetic film memories—a survey," IRE Trans. on Electronic Computers, vol. EC-9, pp. 308–314; September, 1960.
D. O. Smith, "Static and dynamic behavior of permalloy film," *J. Appl. Phys.*, vol. 29, pp. 264–273; March, 1958.
—— and G. P. Weiss, "Steady state and pulse measurement techniques for thin magnetic films in the vhf-uhf range," *J. Appl. Phys.*, vol. 29, pp. 290–291; March, 1958.

[30] D. Buck, "The cryotron—a superconductive computer component," Proc. IRE, vol. 44, pp. 482–493; April, 1956.

[31] A. L. McWhorter and R. H. Rediker, "The cryosar—a new low temperature computer element," Proc. IRE, vol. 47, pp. 1207–1213; July, 1959.

[32] M. S. Blois, Jr., "Preparation of thin films and their properties," *J. Appl. Phys.*, vol. 26, pp. 975–980; August, 1955.

[33] W. J. DeHass and J. M. Casimir-Jonker, "Penetration of a magnetic field into superconductive alloy," *Commun. Phys. Lab. Leiden*, no. 233c, pp. 1–7; 1935.

[34] V. L. Newhouse, J. W. Bremer, and H. H. Edwards, "The crossed-film cryotron and its application to digital computer circuits," *Proc. Eastern Joint Computer Conf.*, pp. 255–260; December, 1959.

[35] B. Lax, "Semi-metals a new horizon," *Bull. Am. Phys. Soc.*, Ser. 2, vol. 6, pp. 58; February, 1961.
R. H. Rediker, A. R. Calawa, and B. Lax, "Magnetotunneling—Possible Four Terminal Device," presented at Stanford University Solid State Device Conf., Palo Alto, Calif.; June, 1961.

[36] H. K. Onnes, "Further experiments with liquid helium C. On the change of electrical resistance of pure metals at very low temperatures, etc. IV. The resistance of pure mercury at helium

temperatures," *Commun. Phys. Lab. Leiden*, no. 120b, pp. 17–19; April 28, 1911.

——, "Further experiments with liquid helium J. The imitation of an ampere molecular current or of a permanent magnet by means of a supra-conductor," *Commun. Phys. Lab. Leiden*, No. 140b, pp. 9–18; April 24, 1914.

[37] G. B. Yntema, "Superconducting winding for electromagnet," *Bull. Am. Phys. Soc.*, vol. 30, No. 1, p. 58; January, 1955.

[38] S. H. Autler, "Superconducting electromagnetics," *Bull. Am. Phys. Soc.* ser. 2, vol. 4, pp. 413; November, 1959. Also in *Rev. Sci. Instr.*, vol. 31, pp. 369–373; April, 1960.

—— and L. J. Donadieu, "Ten-kilogauss superconducting solenoid; resistive transition in superconducting niobium," *Bull. Am. Phys. Soc.*, ser. 2, vol. 6, pp. 64; February, 1961.

[39] J. E. Kunzler, E. Buehler, F. S. L. Hsu, and J. H. Wernick, "Super-conductivity in Nb_3Sn at high current density in a magnetic field of 88 Kgauss," *Phys. Rev. Lett.*, vol. 6, pp. 89–91; February 1, 1961.

[40] W. Meissner and R. Ochsenfeld, "Ein neuer Effekt bei Eintritt der Supraleitfahigkeit," *Naturwissenschaften*, vol. 21, pp. 787–788; March 11, 1933.

[41] R. E. Glover and M. Tinkham, "Conductivity of superconducting films for photon energies between 0.3 and 40 kT_c," *Phys. Rev.*, vol. 108, pp. 243–256; October 15, 1957.

[42] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, "Theory of Superconductivity," *Phys. Rev.*, vol. 108, pp. 1175–1204; December 1, 1957.

[43] L. Giavaer, "Energy gap in superconductors measured by electron tunneling," *Phys. Rev. Lett.*, vol. 5, pp. 147–148; August 15, 1960.

——, "Electron tunneling between two superconductors," *Phys. Rev. Lett.*, vol. 5, pp. 464–466; November 15, 1960.

[44] E. Burstein, D. N. Langenberg, and B. N. Taylor, "Superconductors as quantum detectors for microwave and sub-millimeter-wave radiation," *Phys. Rev. Lett.*, vol. 5, pp. 92–94; February 1, 1961.

[45] C. L. Hogan, "The ferromagnetic Faraday effect at microwave frequencies and its applications," *Bell Syst. Tech. J.*, vol. 31, pp. 1–31; January, 1952.

[46] M. L. Kales, H. N. Chait, and N. G. Sakiotis, "A nonreciprocal microwave component," *J. Appl. Phys.*, vol. 24, pp. 816–817; June, 1953.

B. Lax, K. J. Button, and L. M. Roth, "Ferrite phase shifters in rectangular waveguide, *J. Appl. Phys.*, vol. 25, pp. 1414–1421; November, 1954.

[47] B. Lax, "Figure of merit of microwave ferrites at low and high frequencies," *J. Appl. Phys.*, vol. 26, p. 919; July, 1955.

——, "Frequency and loss characteristics of microwave ferrite devices," PROC. IRE, vol. 44, pp. 1368–1385; October, 1956.

[48] J. B. Thaxter and G. S. Heller, "Circulators at 70 and 140 kMc," PROC. IRE (Correspondence), vol. 48, p. 110; January, 1960.

H. N. Chait and T. R. Curry, "Y circulator," *J. Appl. Phys.*, vol. 30, pp. 152S–153S; April, 1959.

L. Davis, Jr., U. Milano, and J. Saunders, "A strip-line *L*-band circulator," PROC. IRE (Correspondence), vol. 48, pp. 115–116; January, 1960.

[49] E. O. Schulz-Dubois, G. J. Wheeler, and M. H. Sirvetz, "Development of a high-power *L*-band resonance isolator," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-6, pp. 423–428; October, 1958.

M. T. Weiss and P. W. Anderson, "Ferromagnetic resonance in ferroxdure," *Phys. Rev.*, vol. 98, pp. 925–926; May 15, 1955.

F. K. du Pre, D. J. DeBitetto and F. G. Brockman, "Magnetic materials for use at high microwave frequencies," *J. Appl. Phys.*, vol. 29, pp. 1127–1128; July, 1958.

M. T. Weiss and F. A. Dunn, "A 5 mm resonance isolator," IRE TRANS ON MICROWAVE THEORY AND TECHNIQUES (Correspondence), vol. MTT-6, p. 331; July, 1958.

L. C. Kravitz and G. S. Heller, "Resonance isolators at 70 kMc," PROC. IRE, vol. 47, pp. 331; February, 1959.

[50] G. S. Heller and J. J. Stickler, "Antiferromagnetic devices for the millimeter and submillimeter region," Proc. Annual IRE-PGMTT Conf.; May 9, 1960. See also, *J. Appl. Phys.*, vol. 32, pp. 307S–312S; March, 1961.

[51] M. Sucher and H. J. Carlin, "Coaxial line nonreciprocal phase shifter," *J. Appl. Phys.*, vol. 28, pp. 921–922; August, 1957.

B. J. Duncan, L. Swern, K. Tomiyasu, and J. Hannwacker, "Design considerations for broad band ferrite coaxial line isolator," PROC. IRE, vol. 45, pp. 483–490; April, 1957.

K. J. Button, "Theory of nonreciprocal ferrite phase shifters in dilectric loaded coaxial line," *J. Appl. Phys.*, vol. 29, pp. 998–1000; June, 1958.

[52] A. G. Fox, S. E. Miller, and M. T. Weiss, "The behavior and applications of ferrites in the microwave region," *Bell Syst. Tech. J.*, vol. 34, pp. 5–103; January, 1955.

S. Weisbaum and H. Seidel, "The field displacement isolator," *Bell Syst. Tech. J.*, vol. 35, pp. 877–898; July, 1956.

[53] D. Bush, "Discussion on Microwave Apparatus," *Proc. IEE*, vol. 104, pt. B, Suppl. No. 6, pp. 368–369; November, 1956.

F. Reggia and E. G. Spencer, "A new technique in ferrite phase shifting for beam scanning of microwave antennas," PROC. IRE, vol. 45, pp. 1510–1517; November, 1957.

[54] J. A. Weiss, "The Reggia-Spencer microwave phase shifter," *J. Appl. Phys.*, vol. 30, pp. 153S–154S; April, 1959.

——, "A phenomenological theory of the Reggia-Spencer phase shifter" PROC. IRE, vol. 47, pp. 1130–1137; June, 1959.

K. J. Button and B. Lax, "Perturbation theory of the reciprocal ferrite phase shifter using perturbation theory," *Proc. IEE*, vol. 108, pt. B; 1961 (to be published).

[55] G. S. Heller, "Ferrite loaded cavity resonators," *L'Onde Electrique*, vol. 38, Special Suppl. No. 376, vol. 2, pp. 588–593; August, 1958.

[56] F. Reggia, E. G. Spencer, R. D. Hatcher, and J. E. Tompkins, "Ferrod Radiator Systems," Diamond Ordnance Fuze Labs., Washington, D. C., Tech. Rept. No. 357; June, 1956.

[57] W. P. Ayres, P. H. Vartanian, and J. L. Melchor, "Frequency doubling in ferrites," *J. Appl. Phys.*, vol. 27, pp. 188–189; February, 1956.

J. L. Melchor, W. P. Ayres, and P. H. Vartanian, "Microwave frequency doubling from 9 to 18 kMc in ferrites," PROC. IRE, vol. 45, pp. 643–646; May, 1957.

J. E. Pippin, "Frequency doubling and mixing in ferrites," PROC. IRE, vol. 44, pp. 1054–1055; August, 1956.

R. W. Roberts, W. P. Ayres, and P. H. Vartanian, "Nonlinear effects in ferrites," in "Quantum Electronics," C. H. Townes, Ed., Columbia University Press, New York, N. Y., pp. 314–323; 1960.

[58] H. Suhl, "The nonlinear behavior of ferrites at high microwave signal levels," PROC. IRE, vol. 44, pp. 1270–1284; October, 1956.

——, "The theory of ferromagnetic resonance at high signal powers," *J. Phys. Chem. Solids*, vol. 1, pp. 209–227; January, 1957.

——, "Proposal for a ferromagnetic amplifier in the microwave range," *Phys. Rev.*, vol. 106, pp. 384–385; April, 1957.

——, "Theory of the ferromagnetic microwave amplifier," *J. Appl. Phys.*, vol. 28, pp. 1225–1236; November, 1957.

——, "The ferromagnetic microwave amplifier," *Phys. Today*, vol. 11, No. 9, pp. 28–30; September, 1958.

[59] M. T. Weiss, "A solid-state microwave amplifier and oscillator using ferrites," *Phys. Rev.*, vol. 107, p. 317; July, 1957.

——, "Solid-state microwave amplifier and oscillator using ferrites," *J. Appl. Phys.*, vol. 29, p. 421; March, 1958.

[60] A. D. Berk, L. Kleinman, and C. E. Nelson, "Modified semi-static ferrite amplifier," 1958 IRE WESCON CONVENTION RECORD, pt. 3, pp. 9–12.

W. L. Whirry and F. B. Wang, "Experimental study of the modified semistatic ferrite amplifier," *J. Appl. Phys.*, vol. 30, pp. 150S-151S; April, 1959.

[61] R. T. Denton, "A ferromagnetic amplifier using longitudinal pumping," 1960 IRE-PGMTT Natl. Symp., Coronado, Calif.; May 9, 1960. Also in PROC. IRE (Correspondence), vol. 48, pp. 937–938; May, 1960.

[62] R. W. DeGrasse, "Low-loss gyromagnetic coupling through single crystal garnets," *J. Appl. Phys.*, vol. 30, pp. 155S–156S; April, 1959.

[63] R. W. Damon, "Relaxation effects in ferromagnetic resonance," *Rev. Modern Phys.*, vol. 25, pp. 239–245; January, 1953.

N. Bloembergen and S. Wang, "Relaxation effects in para-magnetic and ferromagnetic resonance," *Phys. Rev.*, vol. 93, pp. 72–83; January, 1954.

N. G. Sakiotis, H. N. Chait, and M. L. Kales, "Nonlinearity of microwave ferrite media," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-4, pp. 111–115; April, 1956.

[64] J. P. Gordon, H. J. Zeiger, and C. H. Townes, "The maser—New type of microwave amplifier, frequency standard and spectrometer," *Phys. Rev.*, vol. 99, pp. 1264–1274; August 15, 1955.

——, *et al.*, "Molecular microwave oscillator and new hyperfine structure in the microwave spectrum of NH_3" *Phys. Rev.*, vol. 95, pp. 282–284; July 1, 1954.

[65] N. Bloembergen, "Proposal for a new type solid state maser," *Phys. Rev.*, vol. 104, pp. 324–327; October 15, 1956.

[66] H. E. D. Scovil, G. Feher, and H. Seidel, "Operation of a solid state maser," *Phys. Rev.*, vol. 105, pp. 762–763; January 15, 1957.

[67] A. L. McWhorter and J. W. Meyer, "Solid-state maser amplifier," *Phys. Rev.*, vol. 109, pp. 312–318; January 15, 1958.

R. H. Kingston, "A UHF solid-state maser," PROC. IRE (Correspondence), vol. 46, p. 916; May, 1958.

[68] C. Makhov, C. Kikuchi, J. Lambe, and R. W. Terhune, "Maser action in ruby," *Phys. Rev.*, vol. 109, pp. 1399–1400; February 15, 1958.

[69] A. L. McWhorter, J. W. Meyer, and P. D. Strum, "Noise temperature measurement on a solid state maser," *Phys. Rev.*, vol. 108, pp. 1642–1644; December 15, 1958.

[70] R. W. DeGrasse, D. C. Hogg, E. A. Ohm, and H. E. Scovil, "Ultra Low Noise Antenna and Receiver Combination for Satellite or Space Communication," Bell Telephone Co., Murray Hill, N. J., Tech. Publ. Mono. No. 3624; 1960.

[71] R. Price, P. E. Green, Jr., T. J. Goblick, Jr., R. H. Kingston, L. G. Kraft, Jr., G. H. Pettengill, R. Silver, and W. B. Smith, "Radar Echoes from Venus," *Science*, vol. 129, pp. 751–753; March 20, 1959.
J. A. Giordmaine, "Centimeter wavelength radio astronomy including observations using the maser," *Proc. Natl. Acad. Sci.*, vol. 46, pp. 267–276; March, 1960.
F. V. Jelley and B. F. C. Cooper, "The Operational 21-cm Maser at Harvard College Observatory," presented at 2nd Internatl. Conf. on Quantum Electronics, Berkeley, Calif.; March, 1961.

[72] R. W. DeGrasse, "Slow structures for unilateral solid-state maser amplifiers," IRE WESCON CONVENTION RECORD, pt. 3, pp. 29–35; August 19, 1958.

[73] S. Foner, L. R. Momo, J. B. Thaxter, G. S. Heller, and R. M. White, "Millimeter Wave Masers," presented at 2nd Internatl. Conf. on Quantum Electronics, Berkeley, Calif.; March, 1961.

[74] J. Combrission, A. Honig, and C. H. Townes, "Utilisation de la resonance de spins electroniques pour realiser un oscillateur ou un amplificateur en hyperfrequences," *Comp. Rend.*, vol. 242, pp. 2451–2453; May 14, 1956.

[75] G. Feher, J. P. Gordon, E. Buchler, E. A. Gere, and C. D. Thurmond, "Spontaneous emission of radiation from an electron spin system," *Phys. Rev.*, vol. 109, pp. 221–222; January 1, 1958.

[76] P. F. Chester, P. E. Wagner, and J. G. Castle, "Two level solid-state maser," *Phys. Rev.*, vol. 110, pp. 281–282; April 1, 1958.

[77] S. Foner, L. R. Momo, A Mayer, and R. A. Myers, "Pulsed field millimeter wave maser," in "Quantum Electronics," C. H. Townes, Ed., Columbia University Press, New York, N. Y., pp. 487–498; 1960.

[78] A. L. Schawlow and C. H. Townes, "Infrared and optical masers," *Phys. Rev.*, vol. 112, pp. 1940–1949; December 15, 1958.

[79] T. H. Maiman, "Stimulated optical radiation in ruby," *Nature*, vol. 187, pp. 493–494; August 6, 1960. Also in *British Commun. and Electronics*, vol. 7, pp. 674–675; 1960.
——, "Optical and microwave-optical experiments in ruby," *Phys. Rev. Lett.*, vol. 4, pp. 564–566; June 1, 1960.

[80] R. J. Collins, D. F. Nelson, A. L. Schawlow, W. L. Bond, C. G. B. Garrett, and W. K. Kaiser, "Coherence, narrowing, directionality and relaxation oscillations in the light emission from ruby," *Phys. Rev. Lett.*, vol. 5, pp. 303–305; October 1, 1960.

[81] P. P. Sorokin and M. J. Stevenson, "Stimulated infrared emission from trivalent uranium," *Phys. Rev. Lett.*, vol. 5, pp. 557–559; December 15, 1960.
—— and ——, "Solid state optical maser using divalent samarium in calcium fluoride," *IBM J. Res. and Dev.*, vol. 5, pp. 56–58; 1961.

[82] A. Javan, "Possible production of negative temperature in gas discharges," *Phys. Rev. Lett.*, vol. 3, pp. 87–89; July 15, 1959.
——, "Possibility of obtaining negative temperature in atoms by electron impact," in "Quantum Electronics," C. H. Townes, Ed., Columbia University Press, New York, N. Y., pp. 564–571; 1960.
——, W. R. Bennett, Jr., and D. R. Herriott, "Population inversion and continuous optical maser oscillation in a gas discharge," *Phys. Rev. Lett.*, vol. 6, pp. 106–110; February 1, 1961.

[83] I. Wieder and L. R. Sarles, "Stimulated optical emission from exchange-coupled ions of Cr^{+++} in Al_2O_3," *Phys. Rev. Lett.*, vol. 6, pp. 95–96; February 1, 1961.
A. L. Schawlow and G. E. Devlin, "Simultaneous optical maser action in two ruby satellite lines," *Phys. Rev. Lett.*, vol. 6, pp. 96–98; February 1, 1961.

[84] B. Lax, "Cyclotron resonance and impurity levels in semiconductors," in "Quantum Electronics," C. H. Townes, Ed., Columbia University Press, New York, N. Y., pp. 428–449; 1960.
——, "Cyclotron Resonance Maser," presented at 2nd Internatl. Conf. on Quantum Electronics, Berkeley, Calif.; March, 1961.

For more extensive references on solid-state devices, the reader may refer to the following recent publications:

A. A. R. von Hippel, Ed., "Molecular Science and Molecular Engineering," M.I.T. Technology Press, Cambridge, Mass., and John Wiley and Sons, Inc., New York, N. Y.; 1959.

B. H. W. Katz, Ed., "Solid State Magnetic and Dielectric Devices," John Wiley and Sons, Inc., New York, N. Y.; 1959.

C. W. P. Mason, "Physical Acoustics and the Properties of Solids," D. Van Nostrand Co., Inc., Princeton, N. J.; 1958.

D. PROC. IRE, Special Issue on Computers; January, 1961.

E. W. B. Ittner and C. J. Kraus, "Superconducting Computers," *Sci. Am.*, vol. 203, pp. 124–136; July, 1961.
T. A. Buchhold, "Applications of Superconductivity," *Sci. Am.*, vol. 202, pp. 74–82; March, 1960.

F. J. R. Singer, "Masers," John Wiley and Sons, Inc., New York, N. Y.; 1959.

G. C. H. Townes, Ed., "Quantum Electronics," Columbia University Press, New York, N. Y.; 1960.

H. A. L. Schawlow, "Infrared and Optical Masers," *Solid State J.*, vol. 2, pp. 21–29; June, 1961.

Contributions of Materials Technology to Semiconductor Devices*

RICHARD L. PETRITZ†, SENIOR MEMBER, IRE

Summary—**Two basic disciplines underlying the design and development of solid-state electronic devices are: 1) the design and analysis of the device from the electron physics standpoint, and 2) the materials technology by which suitable structures in solids are prepared to fulfill the conditions specified in the device design. This article considers the second discipline, and reviews the basic steps of purification of materials, crystal growth, and processes such as diffusion and alloying for fabricating structures. Germanium, silicon, and compound semiconductors are discussed. An historical review is given, and directions of future developments are indicated.**

THERE are two basic disciplines underlying the design and development of solid-state electronic devices. One is the design and analysis of the device from the electron physics standpoint. This involves an understanding of the transport properties of electrons and holes in the presence of *p-n* junctions and other energy band structures, independent of the technique used to fabricate the device. The second, which we shall call materials technology, undertakes to make suitable structures in solids by various techniques to fulfill the conditions specified in the device design.

Within this materials technology article, we shall consider the basic steps of purification of materials, crystal growth, and processes such as diffusion and alloying for fabricating semiconductor devices.

Semiconductor technology progress can be divided in two broad categories. We group and review those contributions related to the purification and control of germanium and silicon single crystals and associated device technology in the first category. The second considers work related to compound semiconductors.

Materials Research on Germanium and Silicon

Research Prior to 1948 [1]–[3].

The history of semiconductor technology can mark its beginning with the crystal rectifier which was used as the detector in the early radio receivers. Typically, a detector was made by soldering or clamping a piece of the crystal in a receptacle, and a flexible wire "cat whisker" held in light contact with the crystal made the rectifying contact. This arrangement allowed good rectification only from "sensitive" spots on the crystal and with frequent adjustments of the contact point.

Until the 1940's the crystal rectifier served as a laboratory device to detect and monitor UHF power, since the thermionic tubes had replaced its utility in radio receivers. For this work the combination of a silicon crystal and a tungsten or molybdenum whisker was used because of the sensitive characteristic provided.

With the advent of radar in the late thirties, renewed attention was given to crystal rectifiers because mixer tubes did not perform well at radar frequencies. The first mixer crystals made by the British Thomson-Houston, Ltd., used commercial silicon of about 98 per cent purity. They exhibited the usual sensitive spots, but showed considerable variation in sensitivity. Similar results were obtained at the Radiation Laboratory of Massachusetts Institute of Technology. The General Electric Company, Ltd. of England took the first steps toward using silicon of higher purity than commercially available. Their crystals were obtained from melts of highly purified silicon powder with a fraction of a per-cent of aluminum and beryllium added.

The theory of semiconduction predicted that the conductivity and, hence, the rectifying properties of silicon would be affected by small amounts of impurities. This, together with the success of the General Electric Company, Ltd. rectifier process, stimulated research along similar lines in the United States.

F. Seitz of the University of Pennsylvania initiated a program in connection with E. I. duPont de Nemours & Company to develop a method to produce high purity silicon. Reduction of silicon tetrachloride with zinc yielded silicon with a spectroscopic purity better than 99.9 per cent. Other impurities not detectable by spectroscopic analysis, such as carbon, were present in larger amounts. Ingots made with this silicon, with an appropriate impurity added, had improved conductivity and rectifying properties over prior materials.

K. Lark-Horovitz of Purdue University recommended that the Eagle-Pitcher Company develop techniques to prepare high purity germanium. Their work led to the germanium oxide process for the preparation of high purity germanium.

J. H. Scaff and H. C. Theuerer at BTL investigated metallurgical methods for preparing ingots of silicon and germanium for use in point contact diode studies. A directional freezing technique was developed which produced polycrystalline ingots. Because of the segregation of impurities occuring between the liquid and solid phases, this process provided ingots of improved purity.

The role of Group III and V elements as *p*- and *n*-

* Received by the IRE, January 2, 1962.
† Semiconductor Research Laboratory, Texas Instruments Incorporated, Dallas, Tex.

type dopants respectively for germanium and silicon was studied during the World War II period. The Purdue group made systematic studies of various impurities added to germanium. A similar, but not as complete, study was made on silicon at the University of Pennsylvania. The similarity of effects of impurities in the two materials was recognized. In spite of these efforts, we note that even as late as 1948 rectifiers were made with tin-doped germanium which had very high back resistance for inverse voltages of the order of 100 volts, and it was thought that tin was the doping element. We now know that impurities in tin, probably antimony or arsenic, were responsible for this action and that tin itself does not affect germanium electrically. Thus, impurities in impurities led to confusion during the early forties. However, the research during 1940–1948 resulted in considerable progress in the purification of germanium and silicon and in the controlled addition of impurities to them.

The success of the silicon and germanium diodes developed during this period further stimulated scientists to seek methods for making an amplifying device from a solid. The discovery of the point contact transistor by Bardeen and Brattain [4] at BTL was made in 1948, giving impetus to the over-all program.

Germanium Materials Research—1948 to 1952

In the latter part of 1948 G. K. Teal and J. B. Little of BTL began experiments to grow germanium single crystals, selecting the pulling technique [5]. They succeeded in growing large single crystals [6] of germanium of high structural perfection (Fig. 1). They also improved the impurity of the material by repeated recrystalization methods. The advance in our understanding of basic semiconductor properties has come largely from experiments on single crystals. Furthermore, single crystal technology has become a cornerstone of the transistor industry.

In 1949, W. L. Shockley showed the possibility of obtaining both rectifying and transistor action using *p-n* junctions in bulk material [7]. This important theoretical step stimulated research toward preparing *p-n* junctions.

At BTL Teal, working with M. Sparks, devised a unique method for preparing *p-n* junctions by modifying his crystal-pulling apparatus (Fig. 2), to allow controlled addition of impurities during crystal growth. Using ingots they prepared single crystals containing *p-n* junctions [8], and soon afterward *n-p-n* grown-junction transistors [9] (Fig. 3) which had many of the properties predicted by Shockley [7]. Thus the BTL achieved both the theoretical prediction of junction transistors and the distinction of first solving the physical and metallurgical problems of fabricating a useful device. These transistors were first prepared in 1950, but their public announcement was not made until July, 1951 [9].

At General Electric R. N. Hall and W. C. Dunlap

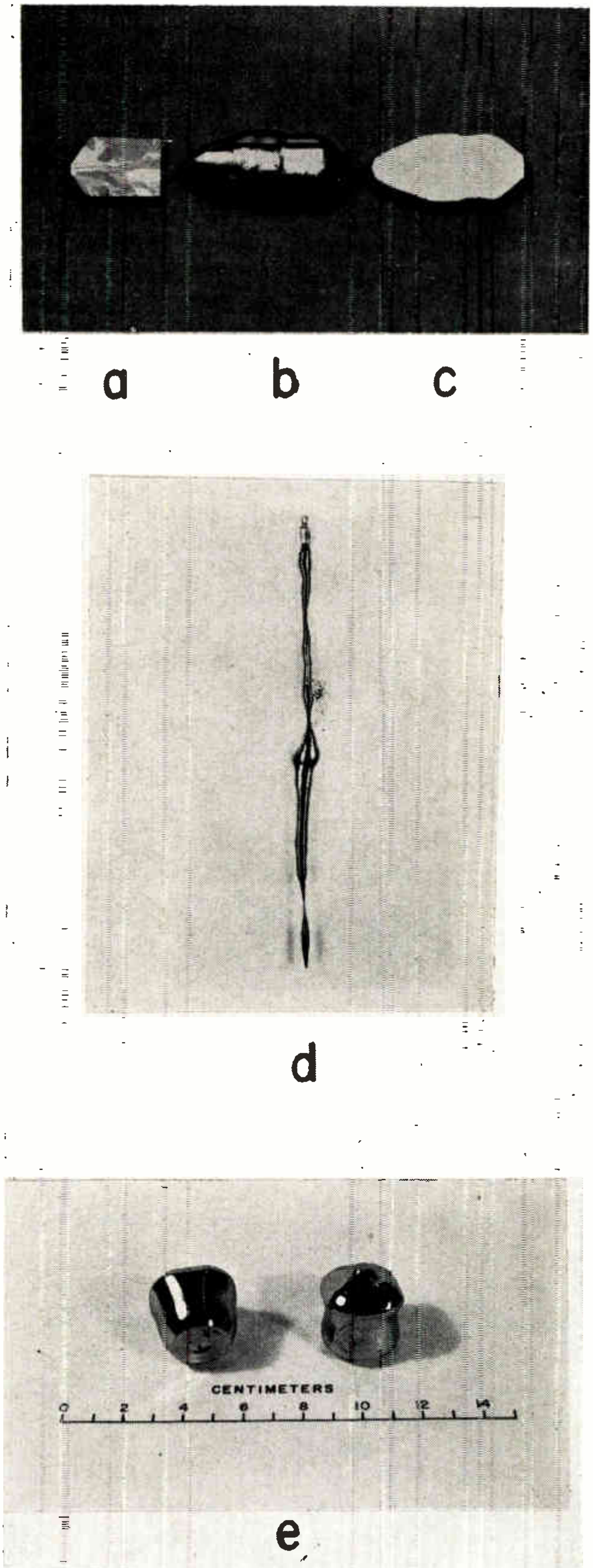


Fig. 1—Polycrystalline and single crystal germanium. A shows a cross section of a polycrystalline germanium ingot revealing a large number of grain boundaries. B shows a single crystal grown by the Teal-Little method. C is a cross section of such a crystal showing the absence of grain boundaries. D is one of the early germanium crystals grown by the Teal-Little technique, while E is a recent germanium single crystal grown by the same method.

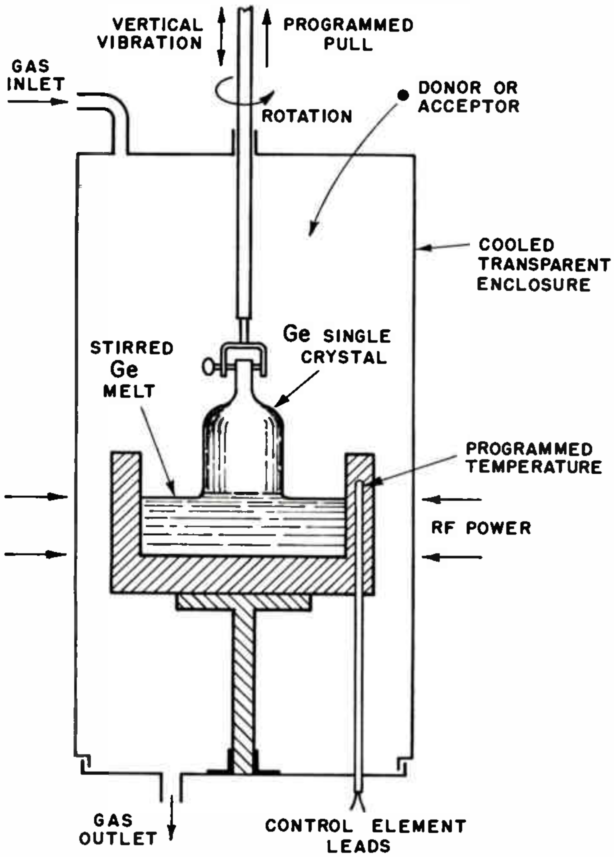


Fig. 2—Cross-section of germanium crystal pulling equipment for growing single crystals of controlled composition. (See Fig. 3 caption for a description of how this equipment is used in the preparation of *n-p-n* transistors.)

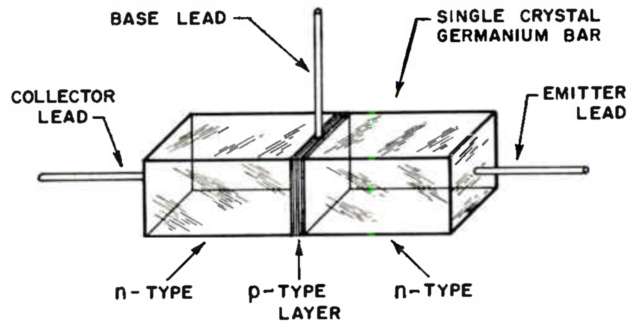


Fig. 3—The structure of an *n-p-n* germanium grown-junction transistor. Using the equipment of Fig. 2, a single crystal is drawn from an appropriately doped *n*-type melt (phosphorus or arsenic, typically), aided by a seed crystal, and the collector region is formed. At a suitable time, a *p*-type impurity such as boron or gallium is added to the melt to dope the base. Subsequently, an additional donor impurity is added to the melt to form the strongly doped *n*-type emitter. The crystal is then withdrawn, cut into bars, and leads attached. Several hundred transistors are obtained from a single crystal.

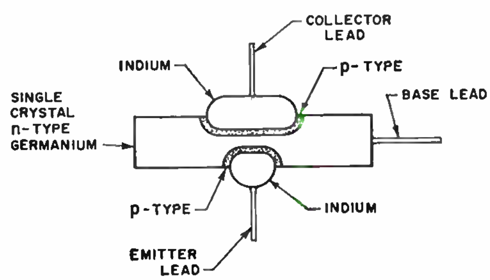


Fig. 4—A *p-n-p* alloyed germanium junction transistor. Two small dots of indium are alloyed into a single crystal *n*-type germanium wafer at a temperature of approximately 500°C. At each indium-germanium interface the surface of the germanium wafer dissolves as indicated in the drawing and upon cooling the assembly, the indium becomes supersaturated with germanium, and a recrystallization layer of germanium heavily doped with indium is formed upon the wafer. These recrystallized layers then form the emitter and the collector sections of a *p-n-p* junction transistor with the original *n*-type wafer playing the part of the base. The leads to the three sections are connected to the indium dots and to the base as shown.

succeeded in preparing single *p-n* junctions by an alloying technique [10]. Later J. S. Saby [11] prepared *p-n-p* transistors by the alloy technique (Fig. 4). It is interesting to note that *n-p-n* grown junction (Shockley, Sparks, and Teal) and *p-n-p* alloy transistors (Saby were discussed publicly for the first time at the same informal technical meeting [12].

Rapid developments soon followed in materials technology. W. G. Pfann [13] discovered a simple method for repeating the action of normal melting and freezing, which avoided handling the material between each operation (Fig. 5). This resulted in material of extremely high purity which was then grown into single crystals by the pulling technique. Pfann also developed the zone leveling technique, which distributes impurities uniformly through a rod. He grew single crystals in his zone leveling apparatus using seeding techniques. The combination of zone leveling and horizontal growth of single crystals has become the standard technique used in today's transistor manufacturing operations [14], [15].

In 1952 the Bell Telephone Laboratories released the technological details of the grown-junction transistor process to a large number of licensed manufacturers. At a historic symposium, details of the crystal growing and related experiments were disclosed, and a set of articles was written in the form of a two-volume publication

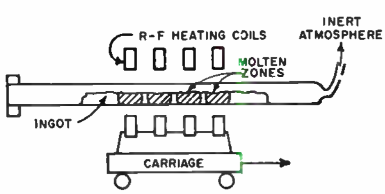


Fig. 5—Cross section of zone refining equipment. The RF heating coils cause melting to occur in the narrow zones. The zones are caused to move from one end of the semiconductor bar to the other by moving the carriage. Segregation at the liquid-solid interfaces causes impurities to move toward the ends of the bar. Several stages of recrystallization can be obtained without handling the material by using several zones and making repeated passes, so an extremely high degree of purification can be achieved.

[16] for the licensees, some 28 companies. Later they were made available to the public. A revised version has been published recently [17]. Fig. 6, a reproduction of page 4 of "Transistor Technology" shows the stages through which germanium materials technology had proceeded at BTL by 1952. The left side of the figure shows the early stages in which chemical purification (GeO_2) was used along with normal freezing (N.F.) of ingots to producer point contact varistor material. In the center of the figure the phase is shown where normal frozen ingots were used in the grown junction apparatus to produce single crystals, grown-junction transistors and diodes. The right side indicates that zone refining had replaced normal freezing for purification of the ingot and that the pulling technique was used for preparation of single crystals, diodes, and transistors. The dotted line shows the zone leveling single crystal technique as being under development.

The IRE provided a needed service by devoting an entire issue of PROCEEDINGS [18] to transistors. One of the many useful purposes this issue served was to disclose details of processes being used for transistor fabrication by a number of industrial concerns.

Silicon Materials Research—1951 to 1954

Large single crystals of silicon and silicon *p-n* junctions were prepared by Teal and Buehler [19] by an extension of the pulling technique developed for germanium. These crystals were used by G. L. Pearson [20] to prepare *p-n* junction diodes by the alloy method.

Silicon, while similar to germanium in many respects, presented certain problems which delayed the development of transistors for several years. Because it has a higher melting point and higher reactivity, a completely satisfactory crucible to hold a silicon melt has not been developed even today. The most satisfactory crucible material is fused quartz; however, it tends to be dissolved slowly by molten silicon. Impurities present in the crucible, such as boron, are transferred to the silicon. Also, oxygen dissolves in the melt and is subsequently incorporated into the growing crystal.

Horizontal zone refining of silicon is also complicated by the crucible problem. Because molten silicon wets the quartz boat, and since silicon and quartz have quite different coefficients of thermal expansion, cracking occurs on cooling. Because of the difficulty in applying zone refining to silicon, more reliance was placed on chemical purification methods [21].

Attempts to fabricate silicon transistors proceeded along the lines of grown-junction and alloy techniques. Control over alloying to the degree necessary to get effective transistor action proved difficult with the material available at that time, while the grown-junction technique was successfully used to fabricate silicon transistors [22], [23]. Relatively few industries adapted the grown-junction technique (even for germanium), and silicon transistors were produced commercially by a single company, Texas Instruments Incorporated, for several years.

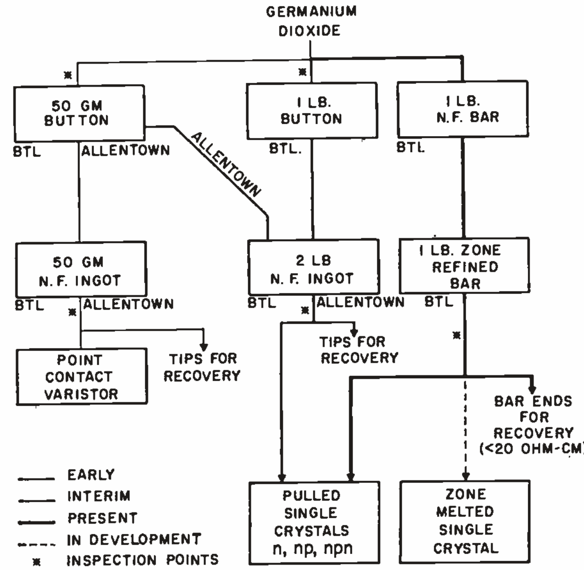


Fig. 6—Stages in the preparation of germanium for varistors and transistors at the Bell Telephone Labs. in 1952. The left side of the figures shows the early stages in which chemical purification (GeO_2) was used along with normal freezing (N.F.) of ingots to produce point contact varistors. The interim phase, where one goes from the normal frozen ingot to the grown-junction technique for production of transistors and diodes, is shown in the middle of the figure. The right side indicates how zone refining replaces normal freezing for the purification of the ingot, and how the pulled crystals are made for the transistors. In the dotted line, the zone leveling single crystal technique is also indicated as being under development.

An improved silicon purification technique was developed which produced material of sufficient quality that alloy silicon transistors could be fabricated with good yields. This was a novel variation of zone refining called "floating zone refining" developed by P. H. Keck [24] and independently by R. Emeis and H. C. Theurer [25]. This operation employs a vertical system and uses surface tension to support a stable liquid zone formed by induction heating. Hence, the crucible is completely eliminated. Silicon with thousands of ohm-cm resistivity and minority carrier lifetimes of greater than 100 μsec can be produced by this method. This material has been used to prepare silicon transistors by the alloy method. However, diffusion processes were developed about the same time, so that alloy techniques have been relatively less important for silicon than for germanium transistors.

Further Developments in Process Technology

While the grown-junction and alloy processes served to give the transistor industry two basic technologies with which practical devices could be manufactured, their inherent limitations were soon recognized. Chief of these were the inability to produce extremely narrow base regions, and to fabricate small area junctions, both of which were needed for performance above a few hundred kilocycles.

The double-doping technique of Teal did not have sufficient control to fabricate base regions narrower than about 0.5 mils. Hall's rate-growing technique [26] offered improvement and transistors of good performance at intermediate radio frequencies were manufactured at good yields.

The jet etching technique [27] was developed at the Philco Corporation to achieve narrow base regions (Fig. 7) and small area junctions; this produced improved high-frequency performance. When metal contacts are plated to both sides of the thin section [28], the surface barrier transistor (SBT) [29] is produced. The micro-alloy (MAT) [30] process developed as a result of forming alloyed junctions on each side of the etched region.

The next major advance in device technology was the diffusion process (Fig. 8). Research on the diffusion of III–V impurities into germanium and silicon by Fuller at the BTL, and by Dunlap at GE laid the foundation for transistor fabrication using diffusion as a key process step [31]. The BTL was the first to fully integrate these results into germanium and silicon transistors [32], [33].

Diffusion techniques have proved to be one of the best controlled methods for preparing *p-n* junctions (Fig. 8). Because the common doping impurities diffuse very slowly in semiconductors at rates which can be varied by adjusting temperatures, close control and reproducibility of the impurity distributions can be achieved. Hence, control over the electrical parameters of the resulting devices may be maintained. The ability to form base regions only a fraction of a micron thick allows very high-frequency transistors to be fabricated.

Diffusion technology, in addition to offering an entirely new technique for processing transistors, was adaptable to many existing processes [34]. The grown-junction technique was quickly modified to utilize diffusion. The result was the grown-diffused transistor [35] which had significantly improved high-frequency performance for both germanium and silicon transistors.

The built-in electrical field which results from a diffusion process was found to give higher-frequency performance for the basic alloy transistor, as predicted by Krömer [36]. This transistor, called a "drift transistor" by RCA has higher frequency response than the conventional alloy transistor (which has uniform doping in the base region).

The addition of a diffused base to the jet etching procedure has resulted in the Philco micro-alloy diffused transistor (MADT) [37], which has excellent high-frequency response, fast switching speeds, and low collector series resistance.

The diffusion technology was not only adapted to existing processes, but it also led to new device structures, the first of which was the mesa transistor (Fig. 9) [32], [33]. For germanium the base was formed by diffusion and the emitter by evaporation and alloying. Precise, small area emitters and narrow base regions

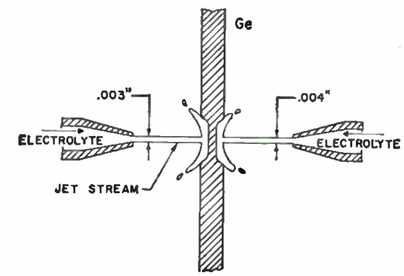


Fig. 7—Jet etching a germanium transistor blank. The wafer is jet etched from both sides until a relatively thin section remains. When metal contacts are plated [28] to both sides of the thin section, the surface barrier transistor (SBT) [29] is produced. If the structure is heated and the metals alloy to the semiconductor the microalloy transistor (MAT) [30] is formed.

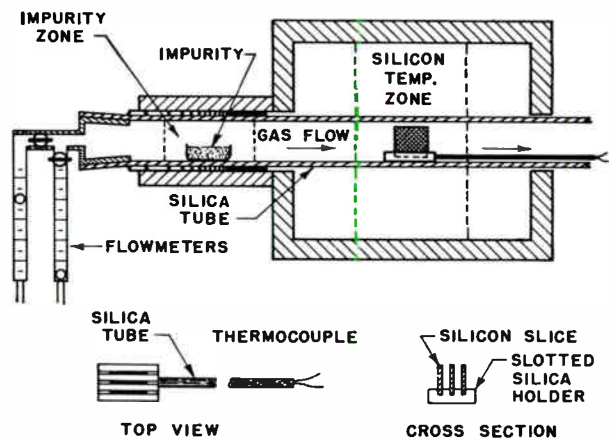


Fig. 8—Two-zone open-tube diffusion apparatus for silicon. The impurity zone temperature can be maintained independent of the silicon temperature to control the surface concentrations. An inert gas flow transfers the impurity to the silicon slices. Generally, the impurity sources are compounds—for example, P_2O_5 and BBr_3. A chemical reaction occurs on the silicon surface resulting in the phosphorus or boron being deposited on the silicon. Diffusion then occurs into the silicon. Other diffusion techniques include the closed tube, and the semi-closed tube (box) systems.

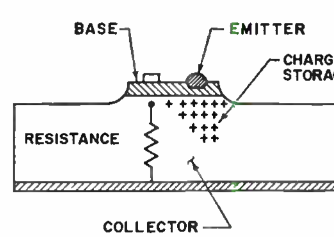


Fig. 9—Cross section of a diffused base, alloyed emitter, germanium mesa transistor. The starting material, which becomes the collector region, is *p*-type single crystal germanium with a typical resistivity of a few ohm-cm. An *n*-type impurity such as arsenic or antimony is diffused to form the base region and base-collector junction. The emitter is formed by evaporating and alloying a metal stripe containing a *p*-type impurity such as indium. An ohmic contact to the base region is similarly made using an *n*-type impurity in the stripe. Photomasking and chemical etching techniques form the mesas. The slice is scribed and broken into several hundred individual transistor wafers which are assembled onto headers.

were possible. Chemical etching was used to remove unwanted diffused material so that small area emitter and collector junctions were obtained. The frequency response of germanium was extended into hundreds of megacycles. These techniques allowed fabrication of hundreds of transistors on a single wafer. The chief limitations of this transistor are the high collector series resistance and charge storage in the collector.

A variation of the combination of alloy and diffusion processes is the alloy-diffused process [38], where the emitter is first alloyed and the base impurity diffused from the emitter region.

For silicon mesa transistors [33], both the base and emitter were formed by diffusion, and etching was again used to define a small collector area. These processes also allowed hundreds of transistors to be fabricated on a single slice, with performance far exceeding that achieved previously in silicon.

Another important technological advance in this period was the development of oxide masking for silicon by C. J. Frosch of BTL [39]. He observed that a thermally grown oxide on silicon impeded the diffusion of certain impurities, including boron and phosphorus. This technique, coupled with photographic masking against etching, provides a powerful tool for silicon processing, and will be illustrated by example in a later section.

Physical imperfections, such as dislocations, are also important in the physics and chemistry of solids, and in achieving good process control. W. C. Dash has made important contributions to this subject. First he decorated dislocations in silicon by precipitating copper and then observed them directly with an infrared imaging microscope [40]. He then developed methods for growing silicon single crystals with essentially no dislocations [41]. This technique has also been used to grow dislocation-free germanium single crystals, and the etch pit method for observing dislocations has been used to verify this.

Surface Technology

With the announcement of the grown-junction transistor in 1951, enthusiasm ran high for a device with performance determined principally by bulk properties. Unfortunately, this optimism was not justified because developments soon indicated that while the primary transistor action was a bulk phenomenon, the surface determined such important properties as leakage currents and over-all stability of the device.

Even today the surface remains a major technological problem. While progress has been slow in this field, there are encouraging signs that improved methods of stabilizing the surface are being developed.

Basic research in the field of surfaces has been oriented toward obtaining a fundamental understanding of surface states and how they are related to various chemical and physical processes [42]–[44]. A second, more empirical study has been to develop techniques for stabilizing the surface to achieve long-life, stable transistors. Stabilization techniques have been investigated ranging from the extreme of a bare transistor sealed in a hermetic can with a getter, to surfaces with various protecting agents put or grown on them.

Thermal oxidation is proving to be an important technique, and M. M. Atalla [45] of BTL has been a significant contributor. Much of this research has been directed toward growing the oxide on the device after fabrication. One difficulty is that high temperatures (1100–1200°C) must be used to grow good oxides on silicon; at these temperatures the impurity distributions will change because of diffusion. For diodes, where the *p-n* junction is deep in the material and its position is not critical, post oxide passivation [46] works well and devices are commercially available. For transistors, where the diffused junctions are near the surface and in close proximity to each other, the redistribution of impurities that occurs during growth of the oxide seriously degrades the device.

Another important approach to this problem is the planar process [47], where the oxide is grown on the silicon slice before the base and emitter diffusions are performed. The oxide serves as a mask in the diffusion processes and protects the *p-n* junctions which are formed under the oxide (Fig. 10). The over-all planar process is described in a later section. Silicon planar transistors have extremely low leakage currents (in the nanoampere range), improved current gain (β) at low currents, and generally good stability and reliability characteristics. Commercial devices made by this method were first introduced by the Fairchild Transistor Corporation.

Epitaxial Crystal Growth

Until 1960 the semiconductor industry followed a pattern of starting with a crystal as pure as needed in the initial stage, and each step added impurities in a controlled manner. In June, 1960, the Bell Telephone Laboratories [48] announced a new method of fabricating transistors using epitaxial single crystals grown from the gas phase with controlled impurity levels. This technique (Fig. 11) had been studied at a number of laboratories [49], but it was not until the 1960 announcement that the potential was fully grasped by the semiconductor industry. Its unique advantage is the ability to grow very thin regions of controlled purity.

The specific process employed by BTL for a germanium mesa transistor utilizes the epitaxial technique to grow a thin, high resistivity *p* collector region, on a low resistivity ($p+$) substrate crystal. The transistor is then fabricated by conventional diffusion and alloy techniques (Fig. 12) in the thin epitaxial layer. The device retains the advantages of the mesa structure, but undesirable aspects (high collector series resistance and minority carrier storage time) are appreciably reduced.

While the principal application to date of the epitaxial method has been to grow *n* regions on $n+$ silicon

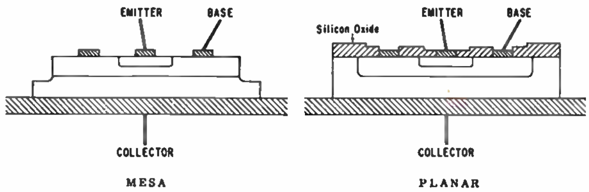


Fig. 10—Silicon mesa and planar transistors. On the left is a mesa transistor where the base-collector junction area is defined by the mesa etching process and the emitter-base junction by the oxide masking process (the oxide subsequently removed). In the planar structure at the right, both junctions are defined by the oxide masking process. The silicon oxide also serves to protect the junctions of planar devices. Further details of the mesa etching process and of the oxide masking process are given under Figs. 9 and 13, respectively.

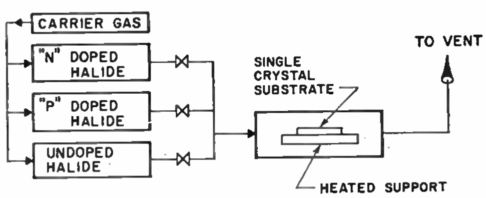


Fig. 11—Schematic of apparatus for epitaxial growth of single crystal semiconductors of controlled composition. A typical silicon process is as follows: The single crystal substrate slice of $n+$ material is placed on the support and heated to 1100°C. The undoped halide ($SiCl_4$) is carried to the silicon slice by a carrier gas (H_2). A chemical reaction occurs at the silicon surface

$$SiCl_4 + Si + 2H_2 \rightarrow Si + 4HCl$$

resulting in the growth of the crystal. Doping can be achieved by mixing n- or p-doped halides (BCl_3 or PCl_3) with the $SiCl_4$ by the valve system. The boron is deposited along with the silicon by the surface reaction

$$2BCl_3 + Si + 3H_2 \rightarrow \underbrace{2B + Si} + 6HCl.$$

The degree of doping is controlled by the gas mixtures and flow rates. Because of the high temperatures involved, diffusion of impurities occurs during the growing period, and account of this must be taken in the device process design.

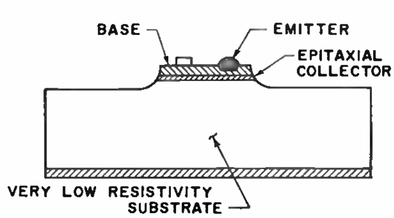


Fig. 12—Cross section of an epitaxial, diffused base, alloyed emitter mesa transistor. Typical fabrication steps for a p-n-p germanium transistor are similar to those described under Fig. 9, with the important difference being that the starting slice is one on which a high-resistivity p region has been epitaxially grown on a low-resistivity $p+$ slice. The fabrication steps of Fig. 9 then occur in the p region, the $p+$ region serving mainly as a mechanical support and a low-resistance electrical connection to the collector.

substrates, and p regions on $p+$ germanium substrates for use with mesa-planar transistors, its potential is much broader because of its flexibility. This technique may well be the most powerful yet developed for the semiconductor industry. A comprehensive review of this subject has not yet been written, but J. Sigler and S. V. Watelski [50] provided a good introduction.

Semiconductor Process Technology—1962

The resultant effect of semiconductor technology developments is the transistor industry, which daily works with crystals of chemical, physical, and structural perfection orders of magnitude better than in any other industry. The processes used for manufacturing devices are sophisticated combinations of diverse technologies. The successful manufacturers are those who have best integrated the various technologies to manufacture specific devices, with economies often being a deciding factor. The so-called "obsolete" technologies, grown) junction and alloy, in their modern version, still account for an appreciable share of the semiconductor market.

The silicon transistor of today combines many aspects of advanced technology and exemplifies the heights to which process technology has progressed. To give the reader an appreciation of these strides, we shall outline the steps in manufacturing the epitaxial, planar, double-diffused silicon transistor produced by most manufacturers today.

The discussion will refer to Fig. 13, where the key process steps are shown. Step a shows the starting slice, a single crystal of silicon of low resistivity ($n+$) 5–10 mils thick (figure not to scale). Step b is to grow an epitaxial layer of n-type material with a resistivity of a few ohm-cm and a thickness of a few tenths mil. Next (step c), an oxide layer a few microns thick is thermally grown on the surface. Then (step d), a photosensitive compound similar to that used in printed circuit technology is applied over the oxide. A suitable mask covers the defined regions for base diffusion, and the remainder is exposed to light. Chemical development removes the unexposed photomasking over the base area as shown in step e. The oxide over the base region is then removed by an acid such as HF, the photomasking not being attacked (step f). Next, the photomasking is removed by a suitable solvent (step g).

The base diffusion is performed using a boron compound and the oxide restricts diffusion to the exposed silicon surface (step h). The boron diffuses laterally under the oxide into the silicon as well as forward into the material, and an oxide is grown over the base region while the diffusion is in process. Note that the base collector junction is formed under the oxide, minimizing possibilities of contamination.

The emitter area is defined by a second photomasking and etching process (similar to d through f), resulting in step j. The emitter diffusion, using a phosphorus compound, is carried out with the oxide again masking

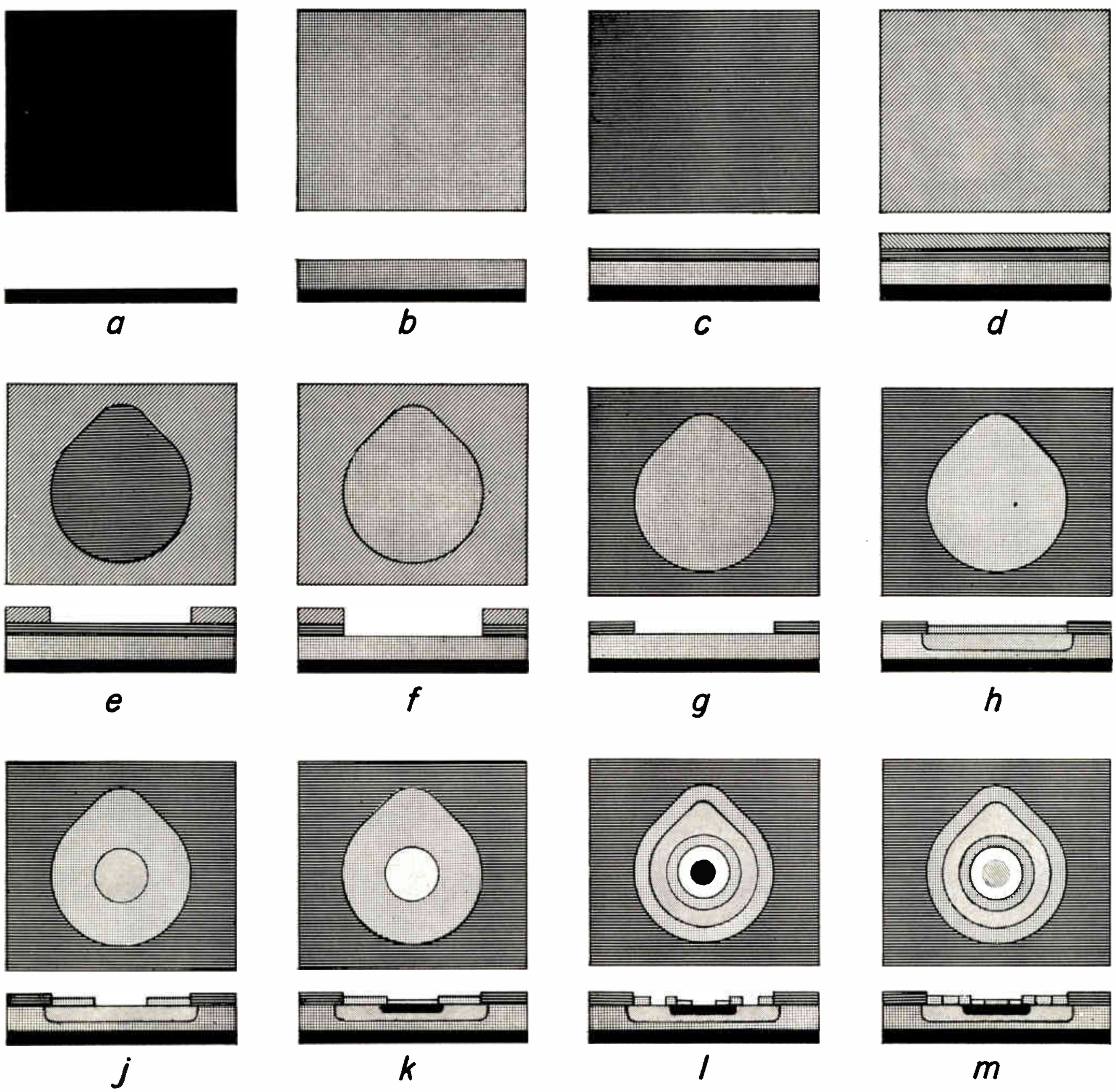


Fig. 13—Steps in fabricating a silicon epitaxial planar transistor. The operations are described in the text.

all but the desired region. The emitter-base junction is formed under the oxide, and another oxide is grown over the emitter region while the diffusion is in process (step k).

Another photomasking-etching operation defines the base and collector contact regions (step l). Aluminum or a similar contact material is evaporated over the slice and another photomasking step removes the aluminum from all but the base and emitter contact regions. The aluminum is alloyed to form a good ohmic contact to the silicon (step m).

Several hundred transistors are being formed in a single slice during these operations (Fig. 14). The slice is next scribed and cut into individual transistor wafers. These are assembled onto headers.

Hopefully, this illustration, while somewhat laborious, will give the reader some feeling for the subtleties and sophistication that the materials technology developments have brought into the semiconductor industry over the past twenty years. It is an achievement hardly envisioned in 1948, when there was little hope for single crystals ever becoming important in a practical sense, since it was generally thought that such a technique would be too expensive.

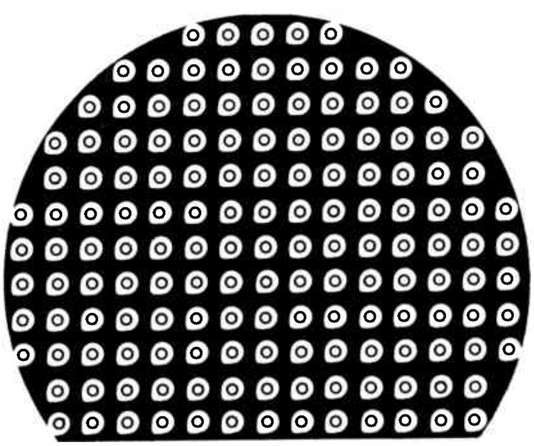


Fig. 14—Sketch showing a typical silicon planar transistor slice prior to scribing and breaking into individual units.

Compound Semiconductor Developments

New Compound Semiconductors and the Chemical Bond Theory

A second major development in the field of materials technology during the past 20 years was the discovery and classification of new compound semiconductors. Prior to 1940 there had been some research on naturally occurring compound semiconductors such as cadmium sulfide, zinc oxide, and others, but no systematic approach to the discovery of new compounds was made.

As early as 1925, E. Friederich [51] speculated that high electrical conductivity occurs in compounds in which not all of the valence electrons are involved in the bonding. He also noted the absence of an appreciable electronic conduction in molecular crystals. In 1944, W. Myer [52] considered electronic conduction in solid compounds and found that Friederich's basic question of the necessary and sufficient conditions for electronic conduction in compounds was still not rigorously answered. Systematizing the wealth of accumulated data from many investigations, Myer concluded the following:

1) The electrical conductivity of a solid compound is related to the electron configurations of the component atoms.
2) Semiconductors and insulators are normal valence compounds; compounds which do not adhere to the valence rules and compounds with compositions deviating from stoichiometry are good (metallic) conductors.

Myers' conclusions could very well have been a guidepost to search for new semiconducting materials. However, by the end of World War II, the scientists' attention was focused on the two Group IVB elements: germanium and silicon. The explosive growth of solid-state devices which followed the discovery of the transistor created a desire for new semiconducting materials suitable for various applications.

A major advance occurred in 1952 when H. Welker [53] recognized the isoelectronic similarity between the Group III-V compounds and the Group IVB (germanium, silicon, and gray tin) elements, and predicted that the former would also be semiconductors. Subsequent investigations verified Welker's prediction.

Welker reasoned that the eight electrons which make up the four covalent bonds for each atom in the diamond lattice (Fig. 15), would, in the case of a III–V compound such as gallium arsenide, be furnished by the five and three valence electrons of arsenic and gallium respectively. Thus Welker predicted a whole new class of semiconductors, which should have properties similar to, but distinct from those of the Group IVB elements. He predicted that they should have the zinc blende crystal structure, which is analogous to the diamond structure as shown in Fig. 15.

Welker went further and attempted to predict the electrical properties of the III-V compounds. For example, he argued that the bonding of the elemental semiconductors—germanium, silicon, and gray tin— is purely covalent, but that in the corresponding isoelectronic compounds—gallium arsenide, aluminum phosphide, and indium antimonide—the bonds should have an ionic as well as a covalent component. Resonance between the covalent and ionic structures leads to an increased bond strength; therefore, the band gaps of the compounds should be larger than those of the corresponding elements. Since higher bond strength lowers the amplitudes of thermal vibrations of the atoms, Welker predicted higher charge carrier mobilities for the compounds. Experiments verified some of these predictions for several of the III-V compounds. However, these arguments are not conclusive, and band structure

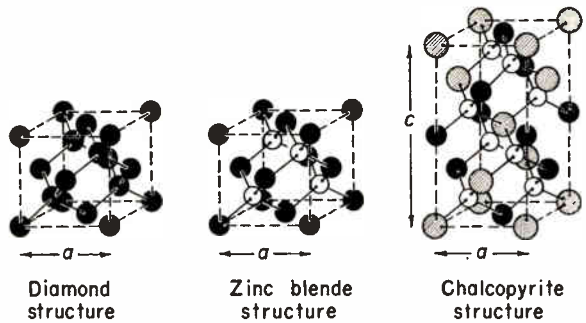


Fig. 15—Relationship between the crystal structures of some elemental and compound semiconductors. Elemental semiconductors such as silicon and germanium possess the diamond structure. Each atom is surrounded by four identical atoms arranged tetrahedrally. The zinc blende structure, characteristic of the III–V compounds such as gallium arsenide, is very similar. The atomic positions are unchanged; however, every atom is surrounded by four atoms of the opposite element. In the chalcopyrite structure, characteristic of compounds such as $ZnGeAs_2$, each atom is bonded to atoms of two different elements causing the tetrahedra to be distorted. Although the composition of these compounds differs, the similarity in structures and electronic configurations causes them all to be semiconductors.

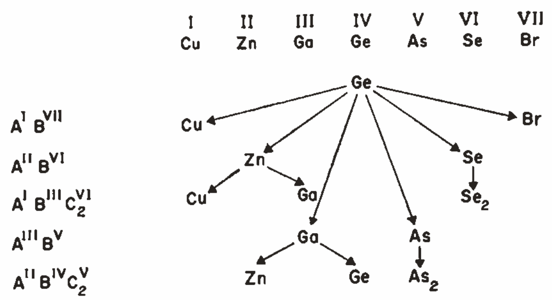


Fig. 16—Relation between some elemental and compound semiconductors. Examples of five types of semiconducting compounds related isoelectronically to germanium are shown. The lattice structure and properties are similar as shown in Fig. 15 and Table I.

TABLE I

LATTICE CONSTANTS OF SEMICONDUCTING COMPOUNDS OF COMPOSITION $A^{II}B^{IV}C_2^{V}$

$ZnSiP_2$	$a = 5.39_8$	$\frac{c}{a} = 1.93_4$
$ZnGeP_2$	$a = 5.46$	$\frac{c}{a} = 1.97$
$ZnGeAs_2$	$a = 5.67_0$	$\frac{c}{a} = 1.96_7$
$CdGeAs_2$	$a = 5.94_2$	$\frac{c}{a} = 1.88_9$
$CdSnAs_2$	$a = 6.09_2$	$\frac{c}{a} = 1.95_7$

calculations are needed for a more complete understanding of the transport properties.

Other classes of compound semiconductors were discovered using the isoelectronic argument (Fig. 16). C. H. L. Goodman and R. W. Douglas [54] predicted and found semiconducting properties in the chalcopyrites of composition $A^{I}B^{III}C_2^{VI}$, which crystallizes in a tetrahedrally distorted superlattice of the zinc blende structure (Fig. 15). O. G. Folberth and H. Pfister [55] later predicted the existence, the structure, and the semiconducting properties of the series of compounds with composition $A^{II}B^{IV}C_2^{V}$. Figs. 15 and 16 and Table I show some of the interesting correspondence in the properties of these three classes of isoelectronic semiconductors.

These discoveries led to the attempt to classify semiconductors in terms of a valence bond description analogous to L. Pauling's [56] work in the theory of the chemical bond. Basically, the difference between the metallic state and the semiconducting state is that the former is characterized by covalent bonds which lead to partial filling of the valence orbitals of the atoms. By sharing electrons in forming the bonds, all atoms in elemental semiconductors acquire filled valence subshells. Thus, the difference between the two types of electronic conductors—metals and semiconductors—is attributed to the difference in the degree of filling the valence shell of the component atoms.

The chemical bond theory is still in development; a current review is E. Mooser and W. B. Pearson [57]. It does afford a chemical criterion analogue to A. H. Wilson's [58] physical criterion of filled bands, permitting a distinction to be made between metals and semiconductors. However, only little progress has been made in predicting and calculating such fundamental properties as energy band gaps and charge carrier mobilities of semiconductors from this theory.

Energy band calculations [59] are a more powerful tool for predicting and understanding the detailed properties of a particular semiconductor. The complex band structure of germanium and silicon (Figs. 17 and 18) that has been revealed by cyclotron resonance [60] and other experiments indicates that transport properties depend in an intricate way upon the detailed structure of the conduction and valence bands [61]. Thus, it is doubtful that the chemical bond approach can predict the subtleties of the true electron band picture.

In the author's opinion, the chemical bond approach is valuable in predicting and classifying new compound semiconductor families, but band theory calculations are needed to predict the detailed properties of a particular semiconductor. Conversely, band theory calculations are too complex to be of much value in predicting and classifying new semiconductor compounds, and it is in this area that the chemical bond approach is especially important. The two approaches necessarily complement each other since they are different means of describing the same physical phenomena.

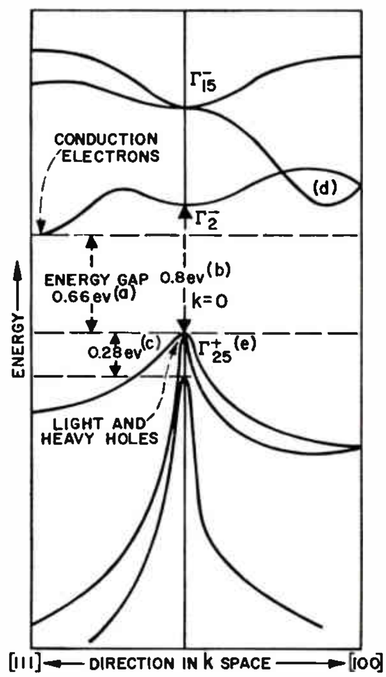


Fig. 17—Energy as a function of wave vector k for the [111] and [100] directions in germanium.

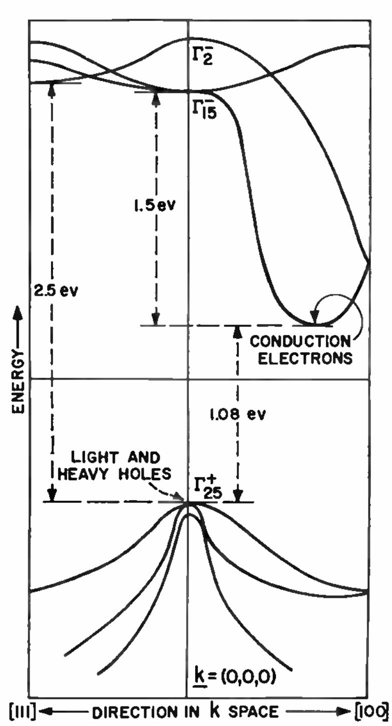


Fig. 18—Energy as a function of wave vector k for the [111] and [100] directions in silicon.

Technology of Compound Semiconductors

Technology problems of compound semiconductors are similar, but in general more complex than those discussed for germanium and silicon, and many techniques that have been developed for the elements are applicable. Progress has been relatively slow compared to germanium and silicon, one reason being that the wealth of materials has made a concentrated effort difficult. Another reason is the heavy concentration of industrial research and development on silicon and germanium. However, in certain areas considerable progress has been made. For example, indium antimonide [62] can be considered the third best-known semiconductor in terms of physical and chemical properties. Since both indium and antimony have rather low vapor pressures at the melting point, indium antimonide has proved relatively easy to prepare and purify by zone refining, and has received considerable study. This research has produced impurity concentrations of less than $10^{14}/cm^3$ and collision times sufficiently long that cyclotron resonance has revealed its band structure.

On the other hand, the arsenides and phosphides of the Group III elements have appreciable vapor pressures at the melting point, which complicates purification. It has been necessary, therefore, to construct sealed furnaces with heating arrangements to maintain the entire system at a high temperature to prevent condensation of the constituents. This requires control of the gas phase as well as the solid and liquid phases. However, progress is being made and application of chemical and metallurgical techniques is bringing better control of these new materials. Reviews of compound semiconductor physics and technology can be found in references [62]–[67].

Implications of Compound Semiconductors for Device Electronics

The discovery and systematic investigation of compound semiconductors afford the device electronics man an expanded degree of freedom for his work. For example, the properties of gallium arsenide make it interesting for transistors [68] (Table II and Fig. 19). This compound has potentially higher-frequency and higher-temperature performance than silicon or germanium. The latter property means better power handling capabilities for the same active area. Transistors have been constructed from gallium arsenide which demonstrate their feasibility [69]. Moreover, the improved electrical parameters expected from the physical properties can actually be realized.

New materials are also being sought in the field of thermoelectricity [70]. For a highly efficient thermoelectric generator or heat pump, a material with high electric conductivity, low thermal conductivity, and high thermoelectric power is needed. No simple semiconducting material satisfies all these requirements, so compounds are required. Using the chemical bond approach, the search for better materials is not so random

TABLE II
ELEMENTAL III-V COMPOUND SEMICONDUCTOR PROPERTIES

Semiconductor	Band Gap (ev)	Electron Mobility cm^2/v sec	Hole Mobility cm^2/v sec	Max. Temp. (Usable °C)
Ge	0.67	3,900	1,900	85–100
Si	1.11	1,500	500	150
C (diamond)	6.70	1,800	1,200	>1,000
InSb	0.18	85,000	3,000	−196
InAs	0.33	33,000	450	~30
InP	1.25	4,800	150	~300
GaAs	1.40	9,000	450	~400
GaP	2.25	>100	>20	~650
AlSb	1.52	460	460	~450

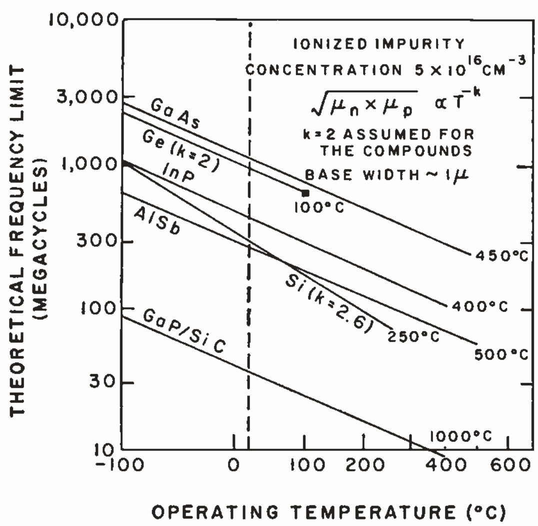


Fig. 19—Theoretical frequency limit vs operating temperature of bipolar compound semiconductor transistors.

as it might otherwise be. Materials may be tailored to achieve certain properties needed for specific devices. For example, alloying two similar materials such as bismuth telluride and bismuth selenide reduces thermal conductivity without materially affecting electrical conductivity.

Infrared detectors are another example of semiconductor devices which depend on new materials. For detectors to operate over a broad range of spectral response, materials with intrinsic energy gaps of different magnitudes, and/or extrinsic impurity energy levels of great variety, are needed. Indium antimonide is already of commercial importance because of its low intrinsic energy gap, which allows detection of infrared radiation to approximately 5 microns [71]. The lead sulfide, lead selenide, and lead telluride family of compounds yields spectral response from 2.7 to over 5 microns [72], [63]. Alloys of these materials allow response to any specified value in this range. Similarly, alloys of indium arsenide and indium antimonide can be used to adjust the spectral response between 3 and 5 microns.

Impurity levels in both elemental and compound semiconductors are important for infrared detectors. Already devices using the gold level in germanium [73] have been brought to the commercial stage, and more recently copper-doped germanium detectors with response to 20 microns have been developed. We thus see that the availability of a variety of impurities, along with a variety of host lattices, offers considerable flexibility to the approach to infrared detection [74].

The discovery of the Group III-V semiconductors along with the systematic development of the theory of the chemical bond for solids has brought forth many new compound semiconductors to be explored for a variety of device applications. No longer is the device man restricted to germanium and silicon, but he now has a broad, highly variable class of materials with which to work.

FUTURE TRENDS OF MATERIALS SCIENCE AND TECHNOLOGY

It is appropriate in an issue dedicated to the 50th Anniversary of the IRE to discuss trends in materials technology which may be expected to take place in the next 50 years.

Four areas will be of particular importance:

1) Improved materials and process technology of germanium and silicon for *p-n* junction-type devices.
2) Investigations of compound semiconductors for electronic devices.
3) New device principles which will require new materials technology for their implementation to a practical device.
4) The full development of the integrated circuit concept, based on a high degree of sophistication in materials and process technology.

The first is necessary to further improve the performance and reliability of *p-n* junction devices of germanium and silicon, as well as to improve manufacturing yields so that prices will be further reduced. New techniques will be discovered and the technologies discussed above will be further refined; surface studies can be expected to receive particular emphasis. New ways of using advanced technologies will result in devices which, while fundamentally of the *p-n* junction type, will have higher performance and reliability and lower cost than present devices.

More attention will be focused on the technology of compound semiconductors, and we can expect the universities to be instrumental in exploring the basic properties of these materials. Compound semiconductors will be especially important in developing new types of devices where some particular properties, not found optimum in germanium or silicon, are utilized.

Category three already contains several new solid-state devices which involve principles different from those of *p-n* junction devices, but which are limited by materials techniques. An example is the metal-oxide-metal tunneling [75], [76] device, where the device function depends on quantum-mechanical tunneling and hot electrons. While the theoretical understanding of this device is quite advanced, the practical development is lagging because of lack of technique in the field of thin films. This example illustrates the general point; namely, that new device principles will be discovered, which will in turn place new demands on the materials technologist in order for the device to reach a stage of practical development.

The last category will place the most stringent demands of all upon the materials technologist. One is attempting to make "devices" which perform complex functions, and the need for control of materials processes is obvious—if it is not obtained the cost of manufacturing such "devices" will not be brought low enough to compete with more conventional approaches, except in special areas. We are confident that the necessary advances will be made, and that integrated circuits will be extremely important in the future.

While the developments in solid-state electronics have already been of spectacular proportions, we believe that the field is still in its infancy; that many exciting developments will take place in the years ahead; and that materials technology will be a most important factor in this continued development. We highly recommend the field of solid-state electronics to our young engineers and scientists, and encourage them to fortify their challenging future with a concentrated study of chemistry, metallurgy, and physics.

References

[1] Karl Lark-Horovitz, "The new electronics," from "The Present State of Physics," Publ. of the Amer. Assn. for the Advancement of Sci. (N.D.). (A survey of semiconductor research prior to 1940 is included in this article.)

[2] H. C. Torrey and C. A. Whitmer, Eds., "Crystal Rectifiers," McGraw-Hill Book Company, Inc., New York, N. Y.; 1948. (This is an excellent summary of the semiconductor research up to the end of World War II.)

[3] G. L. Pearson and W. H. Brattain, "History of semiconductor research," Proc. IRE, vol. 43, pp. 1794–1806; December 1955.

[4] J. Bardeen and W. H. Brattain, "Physical principles involved in transistor action," *Phys. Rev.*, vol. vol. 75, pp. 1208–1225; April, 1949.

[5] J. Czochralski, "Measuring the velocity of crystallization of metals," *Z. Phys. Chem.*, vol. 92, pp. 219–221; April, 1917.

[6] G. K. Teal and J. B. Little, "Growth of germanium single crystals," *Phys. Rev.*, vol. 78, p. 647; June, 1950.

[7] W. L. Shockley, "Theory of *p-n* junctions in semiconductors and *p-n* junction transistors," *Bell Sys. Tech. J.*, vol. 28, pp. 435–489; July, 1949.

[8] G. K. Teal, M. Sparks, and E. Buehler, "Growth of germanium single crystals containing *p-n* junctions," *Phys. Rev.*, vol. 81, p. 637; February, 1951.

[9] W. Shockley, M. Sparks, and G. K. Teal, "*p-n* junction transistors," *Phys. Rev.*, vol. 83, p. 151–162; July, 1951.

[10] R. N. Hall and W. C. Dunlap, "*P-N* junctions prepared by impurity diffusion," *Phys. Rev.*, vol. 80, pp. 467–468; November, 1950.

[11] J. S. Saby, "Recent developments in transistors and related devices," *Tele-Tech.*, vol. 32, p. 10; 1951. (This is the first published account of Saby's work. A more complete report is Saby, "Fused impurity *P-N-P* junction transistors," Proc. IRE, vol. 40, pp. 1358–1360; November, 1952.

[12] IRE-AIEE Electronic Devices Conf., Durham, N.H.; June, 1951.

[13] W. G. Pfann, "Segregation of two solutes, with particular reference to semiconductors," *J. Metals*, vol. 4, pp. 861–865; August, 1952.

[14] W. G. Pfann, "Techniques of zone melting and crystal growing," in "Solid State Physics," F. Seitz and D. Turnbull, Eds., Academic Press, Inc., New York, N. Y., vol. 4, pp. 423–521; 1957.

[15] C. D. Thurmond, "Control of composition in semiconductors by freezing methods," in"Semiconductors," N. B. Hannay, Ed., Reinhold Publishing Corp., New York, N. Y., ch. 4; 1959.

[16] "Transistor Technology," Bell Telephone Labs., Inc. and Western Electric Co., Inc., vols. I and II; 1952.

[17] "Transistor Technology," (vol. I—H. E. Bridgers, J. H. Scaff, and J. N. Shive, Eds.; vols. II and III—F. J. Biondi, Ed.), D. Van Nostrand Co., Inc., Princeton, N. J., 1958.

[18] Proc. IRE, vol. 40; November, 1952. (Special Issue on Transistors.)

[19] G. K. Teal and E. Buehler, "Growth of silicon single crystals and of single crystal silicon *p-n* junctions," *Phys. Rev.*, vol. 87, p. 190; July, 1952.

[20] G. L. Pearson and P. W. Foy, "Silicon *p-n* junction diodes prepared by the alloying process," *Phys. Rev.*, vol. 87, p. 190; July, 1952.

G. L. Pearson and B. Sawyer, "Silicon *P-N* junction alloy diodes," Proc. IRE, vol. 40, pp. 1348–1351; November, 1952.

[21] J. M. Wilson, "The chemical purification of germanium and silicon," in "Progress in Semiconductors," Alan F. Gibson, Ed., John Wiley & Sons, Inc., New York, N. Y., vol. 3, pp. 27–51; 1958.

[22] G. K. Teal, "Some Recent Developments in Silicon and Germanium Materials and Devices," presented at Nat'l Conf. on Airborne Electronics, Dayton, Ohio; May 10, 1954.

[23] W. A. Adcock, M. E. Jones, J. W. Thornhill, and E. D. Jackson, "Silicon transistor," Proc. IRE (Correspondence), vol. 42, p. 1192; July, 1954.

[24] P. H. Keck and M. J. E. Golay, "Crystallization of silicon from a floating liquid zone," *Phys. Rev.*, vol. 89, p. 1297; March, 1953.

[25] R. Emeis, "Growing silicon single crystals without a crucible," *Z. Naturforsch*, vol. 9a, pp. 67; January, 1954.

H. C. Theurer, "Removal of boron from silicon by hydrogen water vapor treatment," *J. Metals*, vol. 8, pp. 1316–1319; October, 1956.

[26] R. N. Hall, "*p-n* junctions produced by rate growth variation," *Phys. Rev.*, vol. 88, p. 139; October, 1952.

[27] J. W. Tiley and R. A. Williams, "Electrochemical techniques for fabrication of surface-barrier transistors," Proc. IRE, vol. 41, pp. 1706–1708; December, 1953.

[28] R. F. Schwarz and J. F. Walsh, "The properties of metals to semiconductor contacts," Proc. IRE, vol. 41, pp. 1715–1720; December, 1953.

[29] W. E. Bradley, "Principles of the surface-barrier transistor," Proc. IRE, vol. 41, pp. 1702–1706; December, 1953.

[30] A. D. Rittmann, G. C. Messenger, R. A. Williams, and E. Zimmerman, "Microalloy transistor," IRE Trans. on Electron Devices, vol. ED-5, pp. 49–54; April, 1958.

[31] H. Reiss and C. S. Fuller, "Diffusion processes in germanium and silicon," ch. 6 of ref. [15]. (This contains an excellent bibliography for diffusion.)

[32] Charles A. Lee, "A high-frequency diffused base germanium transistor," *Bell Sys. Tech. J.*, vol. 35, pp. 23–34; January, 1956.

[33] M. Tanenbaum and D. E. Thomas, "Diffused emitter and base silicon transistors," *Ibid.*, pp. 1–22.

[34] F. M. Smits, "Formation of junction structures by solid-state diffusion," Proc. IRE, vol. 46, pp. 1049–1061; June, 1958.

[35] B. Cornelison and W. A. Adcock, "Transistors by Grown-Diffused Technique," presented WESCON, Semiocnductor Devices Section, San Francisco, Calif.; August 20–23, 1957.

[36] H. Krömer, "The drift transistor," *Naturwiss*, vol. 40, pp. 578–579; November, 1953; "On the theory of the diffusion and the drift transistor," *Arch. Elekt. Übertr.*, vol. 8, pp. 223–228, May, 1954.

[37] C. G. Thornton and J. B. Angell, "Technology of micro-alloy diffused transistors," Proc. IRE, vol. 46, pp. 1166–1176; June, 1958.

[38] P. J. W. Jochems, O. W. Memelink, and L. J. Tummers, "Construction and electrical properties of a germanium alloy-diffused transistor," Proc. IRE, vol. 46, pp. 1161–1165; June, 1958.

J. R. A. Beale, "Alloy-diffusion: a process for making diffused-base junction transistors," *Proc. Phys. Soc.* (*London*), vol. B70, pp. 1087–1089; November, 1957.

[39] C. J. Frosch and L. Derick, "Surface protection and selective masking during diffusion in silicon," *J. Electrochem. Soc.*, vol. 104, pp. 547–552; September, 1957.

[40] W. C. Dash, "Copper precipitation on dislocations in silicon," *J. Appl. Phys.*, vol. 27, pp. 1193–1195; October, 1956.

[41] W. C. Dash, "Silicon crystals free of dislocations," *J. Appl. Phys.*, vol. 29, pp. 736–737; April, 1958.

[42] Jay N. Zemel, Ed., "Semiconductor Physics," *Proc. of the 2nd Conf. on the Physics of Semiconductor Surfaces*, U.S. Naval Ordnance Lab. and Office of Naval Res., Pergamon Press, Oxford, England; 1960.

[43] R. H. Kingston, Ed., "Semiconductor Surface Physics," *Proc. of the Conf. on the Physics of Semiconductor Surfaces*, Office of Naval Res., University of Pennsylvania, Mass. Inst. of Tech.,

Philadelphia, Pa., June, 1956, University of Pennsylvania Press, Philadelphia, Pa.; 1957.
[44] T. B. Watkins, "The electrical properties of semiconductor surfaces," in "Progress in Semiconductors," A. F. Gibson, Ed., John Wiley & Sons Inc., New York, N. Y., vol. 5, pp. 1–52; 1960.
[45] M. M. Atalla, E. Tannenbaum, and E. J. Scheibner, "Stabilization of silicon surfaces by thermally grown oxides, "*Bell Sys. Tech. J.*, vol. 38, pp. 749–783; May, 1959.
[46] C. R. Fuller and S. S. Baird, "Oxidation Induced Degradation of Reverse Characteristics of Silicon Diodes," presented at the Electrochem. Soc. Meeting, Detroit, Mich.; October, 1961.
[47] Jean A. Hoerni, "Planar Silicon Transistors and Diodes," presented at the 1960 Electron Devices Meeting, Washington, D. C.; October, 1960.
[48] H. H. Loar, H. Christensen, J. J. Kleimack, and H. C. Theuerer, "New Advances in Diffused Devices," presented at the IRE-AIEE Solid-State Device Res. Conf., Pittsburgh, Pa.; June, 1960.
[49] J. C. Marinace, "Epitaxial vapor growth of Ge single crystals in a closed-cycle process," *IBM J. Res. & Dev.*, vol. 4, pp. 248–255; 1960.
H. Christensen and G. K. Teal, "Method of fabricating germanium bodies," Patent No. 2,692,839; October 26, 1954.
R. C. Sangster, E. F. Maverick, and M. L. Croutch, "Growth of silicon crystals by a vapor phase pyrolytic deposition method," *J. Electrochem Soc.*, vol. 104, pp. 317–319; May, 1957.
[50] J. Sigler and S. V. Watelski, "Epitaxial techniques in semiconductor devices," *Solid State J.*, vol. 2, pp. 33–37; March, 1961.
[51] E. Friederich, "Some new properties of simple compounds and different varieties of the solid state," *Z. Phys.*, vol. 31, pp. 813–826, 1925; vol. 34, p. 637, 1925. (Same title; elaboration of same material.)
[52] W. Meyer, "Electron conduction in solid chemical compounds," *Z. Elektrochem.*, vol. 50, pp. 274–290; 1944.
[53] H. Welker, "Semiconducting compounds," *Z. Naturf.*, vol. 7a, pp. 744–749; November, 1952.
[54] C. H. L. Goodman and R. W. Douglas, "New semiconducting compounds of diamond type structure," *Physica*, vol. 20, pp. 1107–1109; November, 1954.
[55] O. G. Folberth and H. Pfister, "The new semiconductor compounds with chalcopyrite structure," in "Semiconductors and Phosphors," M. Schön and H. Welker, Eds., Friedr. Vieweg and Sohn, Berlin, Germany, vol. 5, pp. 474–476; 1958.
[56] L. Pauling, "The Nature of the Chemical Bond," Cornell University Press, Ithaca, N. Y., 2nd ed.; 1945.
[57] E. Mooser and W. B. Pearson, "The chemical bond in semiconductors," in "Progress in Semiconductors." *op. cit.*, Gibson, Ed., vol. 5, pp. 103–139.
[58] A. H. Wilson, "Theory of electronic semiconductors," *Proc. Royal Soc.*, vol. A133, pp. 458–491; 1931.
[59] F. Herman, "Theoretical investigation of the electronic energy band structure of solids," *Rev. Mod. Phys.*, vol. 30, pp. 102–121; January, 1958.
Joseph Callaway, "Electron Energy Bands in Solids," in "Solid State Physics," F. Seitz and D. Turnbull, Eds., Academic Press Inc., New York, N. Y., vol. 7, pp. 99–212; 1958.
H. Y. Fan, "Valence Semiconductors, Germanium and Silicon," in "Solid State Physics," *Ibid.*, vol. 1, pp. 283–365; 1955.
[60] B. Lax and J. G. Mavroides, "Cyclotron resonance," in "Solid State Physics," *Ibid.*, vol. 11, pp. 261–400; 1960.
[61] T. H. Geballe, "Group IV semiconductors," ch. 8 of Ref. [15].
[62] T. S. Moss, "Indium antimonide," in "Progress in Semiconductors," *op. cit.* [21], vol. 5, pp. 189–220.
[63] W. W. Scanlon, "Polar semiconductors," in "Solid State Physics," *op. cit.* [14], vol. 9, pp. 83–137; 1959.
[64] R. A. Smith, "The electronic and optical properties of the lead sulfide group of semiconductors," *Proc. of the Internat'l Conf. on Semiconductors*, pp. 910–929, Amsterdam, Netherlands; June, 1954.
[65] J. M. Whelan, "Properties of some covalent semiconductors," ch. 9. of Ref. [15].
[66] Harry C. Gatos, "Properties of Elemental and Compound Semiconductors," Proc. Mettallurgical Soc. Confs. Interscience Publ., Inc., New York, N. Y., vol. 5; 1959.
[67] H. Welker and H. Weiss, "Group III–Group V compounds," in "Solid State Physics," *op. cit.* [14], vol. 3, pp. 1–78; 1956.
[68] D. A. Jenny, "The status of transistor research in compound semiconductors," PROC. IRE, vol. 46, pp. 959–968; June, 1958.
[69] M. E. Jones and E. C. Wurst, Jr., "Recent advances in gallium arsenide transistors," 1961 IRE INTERNATIONAL CONVENTION RECORD, pt. 3, pp. 26–30.
[70] R. R. Heikes and R. W. Ure, Jr., "Thermoelectricity: Science and Engineering," Interscience Publ., Inc., New York, N. Y.; 1961.
[71] F. F. Rieke, L. H. DeVaux, and A. J. Tuzzolind, "Single-crystal infrared detectors based upon intrinsic absorption," PROC. PROC. IRE, vol. 47, pp. 1475–1478; September, 1959.
G. R. Pruett and R. L. Petritz, "Detectivity and preamplifier considerations for indium antimonide photovoltaic detectors," *Ibid.*, pp. 1524–1529.
[72] R. J. Cashman, "Film-type infrared photoconductors," *Ibid.*, pp. 1471–1475.
[73] Henry Levenstein, "Impurity photoconductivity in germanium," *Ibid.*, pp. 1478–1481.
[74] PROC. IRE, vol. 47; September, 1959. (Special Issue on Infrared.)
[75] J. P. Spratt, R. F. Schwarz, and W. M. Kane, "Hot electrons in metal films: injection and collection," *Phys. Rev. Letters*, vol. 6, pp. 341–342; April, 1961.
[76] C. A. Mead, "The tunnel-emission amplifier," PROC. IRE, vol. 48, pp. 359–361; March, 1960.

Section 12
ELECTRONIC COMPUTERS

Organized with the assistance of the IRE Professional Group on Electronic Computers

The Evolution of Computing Machines and Systems*

R. SERRELL†, SENIOR MEMBER, IRE, M. M. ASTRAHAN‡, SENIOR MEMBER, IRE,
G. W. PATTERSON||, FELLOW, IRE, AND I. B. PYNE§, MEMBER, IRE

Summary—**The development of digital computing machines is described and illustrated, from the mechanical devices of the seventeenth century to the electronic systems of today, in three chronological sets of sketches. Only a few typical ones of the modern systems are treated and not all early machines are discussed. A bibliography is appended.**

Introduction

IN APRIL, 1944, during the third year of the second World War, a huge electromechanical computing machine formally named the "IBM Automatic Sequence Controlled Calculator" and familiarly known as the "Mark I," was put in operation in the basement of Harvard University's Cruft Laboratory in Cambridge, 24 hours a day, 7 days a week. By the time it had been moved, late in 1946, to its permanent location in the Harvard Computation Laboratory, the Mark I had fairly revolutionized the calculation of mathematical tables.

During the same period, early in 1946, the eighteen-thousand-tube "Electronic Numerical Integrator and Computer"—called "ENIAC" for short—was put to work at the Moore School of the University of Pennsylvania in Philadelphia, to perform such tasks as the computation of ballistic tables. The ENIAC could compute the trajectory of a projectile in less than the time of flight of the projectile.

Thus, large-scale automatic computation began. A scant fifteen years later, automatic electronic computation had become a billion-dollar-a-year industry of unsurpassed vigor, supplying solutions to scientific problems of almost inconceivable complexity and performing data-processing tasks of almost unimaginable extent.

The purpose of this paper is to describe panoramically—in a chronological series of glimpses—the evolution of electronic digital computing machines from early con-

* Received by the IRE, January 26, 1962.
† Consultant, 39 Lovers Lane, Princeton, N. J.
‡ Advanced Systems Development Division, International Business Machine Corporation, San Jose, Calif.
|| Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.
§ Princeton Computer Center, Princeton University, Princeton, N. J.

cepts of over three-hundred years ago, through a transitional period in the decade following World War II, to the present era of modern computing systems.[1] Only *digital* machines and systems are described; the important class of *analog* machines are not treated here.[2] (A digital computing machine is one in which numbers are represented by discrete symbols—discrete physical states of the machinery used—forming a definite number system. In an analog machine, numbers are represented by continuous physical entities, such as voltages.) Because of space limitations, only a few of the modern digital computing machines (and only some of the earlier ones) can be discussed. The mere identification of all the distinct digital machine types now in use would require more than the entire space allocated to this paper [142]. For the same reason, foreign modern machines are not explicitly mentioned. As the whole evolution of electronic digital computers must be covered in a paper of reasonable length, which should, as George Sarton said of his history of Hellenistic science, "enlighten the reader without overwhelming him," it is obvious that the authors cannot deal with every part of the subject.

In the sense in which the term is used here, an automatic computer is essentially an assemblage of units designed to recognize information (usually but not exclusively numerical) supplied at the computer input, store information (both to be used and to be supplied by the machine), perform arithmetic and other operations, form definite sequences of these operations under appropriate control, and supply the results of the computation to the computer output. A digital automatic computer is one in which numbers are represented in terms of a base (such as the base 10 of decimal numeration or the base 2 of binary numeration, etc.) as the discrete physical positions or states of the computing mechanism. For example, a wheel which can be driven to any one of ten discrete angular positions can represent a decimal digit 0 to 9; a two-tube toggle circuit, or *flip-flop*, in which only one or the other of the tubes is conducting, can represent a binary digit 0-1, and so on. The first digital computers, like the well-known hand adding machines, were entirely mechanical. Then electrical (solenoid or motor) actuation and sliding contacts appeared and digital computers became electromechanical. Later, electronic conducting states, as in the vacuum-tube or transistor flip-flop, were used to represent numbers. Of course, definite variations of these states could be, and were, used to perform operations on the numbers as well. Physical motion of parts was no longer an essential feature of a computer. And the action became so fast that it made no sense to build an electronic digital computer that was not also an automatic digital computer. So the word "automatic" was gradually dropped.

A mechanical digital computer can add two digits, for example the two decimal digits 5 and 8, by counting first up to 5 positions and then 8 positions more with a ten-position counting wheel, generating a carry of 1 into the next higher decimal *place* (also called *column*, or *stage*) as the wheel passes through 9. An electronic digital computer can add in the same manner with an electronic counter. Many adders based on this principle have been built, including most of the early and all of the mechanical ones. But, in electronic machines, it was soon found more economical and much faster to use various kinds of *logic circuits* which generate the sum digit and the carry digit directly from the given augend and addend digits (5 and 8 in our example) by means of the elementary operations of logic.[3]

Subtraction has often been performed by means of complementary numbers.[4] In electronic machines, more direct techniques have been used as well. Multiplication can be performed, of course, simply as repeated addition. The early digital computers (and some not-so-early ones) multiplied in just this manner. In some machines, multiplication was performed by means of a built-in multiplication table. This, however, required large amounts of "hardware"; more efficient techniques were developed later. Not all computers could divide explicitly. Division has usually been performed as a specially-designed combination of the other three operations; (many such combinations exist). For all four operations, specially-designed logic circuits, together with specially-designed number systems, have played a large part in the development of electronic digital computers. In some of them, the *parallel* computers, computation proceeds simultaneously with all digits of the given numbers, as in a desk calculator. In other machines, the *serial* computers, computation is performed one digit at a time, as with pencil and paper.

The automatic character of an automatic digital computer consists in its ability to perform definite sequences of arithmetic and other operations (such as finding square roots, looking up entries in a table, etc.) without manual intervention. The operations in such a sequence are specified by a corresponding sequence of computer *instructions*. In early machines, these instructions were communicated one at a time to a computer by means of punched cards or punched tape read by the machine as

[1] The chronology is based upon the dates at which the various machines were put in operation. However, since the operativeness of an evolving machine is difficult to define, many of the dates given are only approximate.

[2] See, G. A. Korn, "The impact of hybrid techniques on the analog computer art," Proc. IRE, this issue, pp. 1077–1086.

[3] These elementary operations of logic are usually defined as operations in Boolean Algebra (named for George Boole, English mathematician, 1815–1864). They include the operation of union (A or B or both), intersection (A and B simultaneously) and complementation (not A), in which A, B, etc., are sets of binary conditions or states.

[4] "The complement on ten of a number is that second number which must be added to the first in order to obtain a power of ten. This complement on ten may be read off from left to right by taking the complement on nine of each successive digit except the last on the right, of which the complement on ten must be taken." (Quoted from Aiken and Hopper [4].)

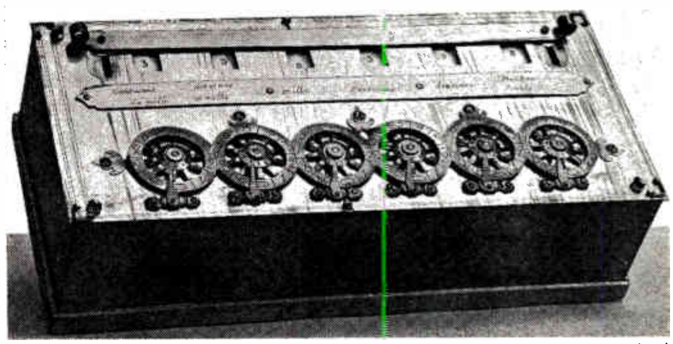
its work went on. Later, the whole *program* (as a set of computer instructions is called) was stored in the internal *memory* of the computer along with the numbers to be used in the computation. Today, most large digital computers (and many small ones) are "stored-program" machines which utilize this idea. Since each instruction usually includes one or more *addresses, i.e.*, storage locations in the memory, a stored-program machine can be instructed to perform calculations on the instructions themselves. For example, it can add a number to an address. In this manner, such a machine can modify its own original program. More precisely, the machine is able to make explicit those parts of its own program that it determines are needed as the computation proceeds. In this ability resides much of the tremendous power of the modern electronic digital computer.

Early Concepts

Some aids to computation, like the abacus, are quite ancient. However, the invention of the first mechanical device capable of addition and subtraction in a digital manner has been generally credited to Pascal, who built his first machine in 1642. This claim has been contested on the basis of letters sent to Kepler in 1623 and 1624 by Wilhelm Schickhardt of Tübingen, in which the latter describes the construction of a calculator [121]. Pascal, who at the age of 19 had wearied of adding long columns of figures in his father's tax office in Rouen, made a number of calculators, some of which are still preserved in museums. His machines had number wheels with parallel, horizontal axes. The positions of these wheels could be observed and sums read through windows in their covers. Numbers were entered by means of horizontal telephone-dial-like wheels which were coupled to the number wheels by pin gearing. Most of the number wheels were geared for decimal reckoning but the two wheels on the extreme right had twenty and twelve divisions, respectively for sous and deniers. A carry ratchet coupled each wheel to the next higher place. The stylus-operated pocket adding machines now widely used are descendents of Pascal's machine (Figs. 1 and 2).

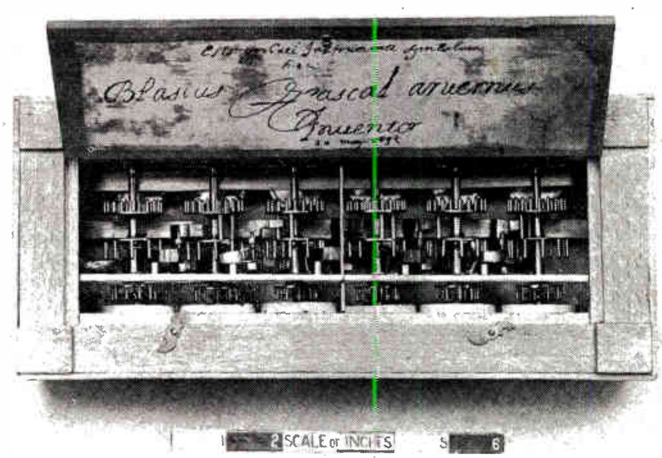
In 1671, Leibniz proposed a machine which could multiply by rapid, repeated addition. One was built in 1694, but it was not reliable (Fig. 3). The first machine to perform all four basic arithmetic operations well enough for commercial manufacture was the Arithmometer of C. X. Thomas, built in 1820. However, only a small number of Thomas's machines were constructed; real commercial exploitation of mechanical calculators did not take place until the last two decades of the nineteenth century. Two further inventions deserve mention as forerunners of automatic calculators. The first is the Jacquard loom punched-card system devised to control the automatic weaving of complex patterns. Jacquard's loom, which came into wide use during the decade following 1804, was the first successful application of principles of punched tape and card control demonstrated originally between 1725 and 1745 by Bouchon, Falcon, and Jacques.

The second invention is the difference engine, a device for automatically calculating mathematical tables of functions whose higher-order differences are constant. Such a machine requires a register for each order of difference and a means for successively adding the contents of each register to those of the next lower-order register. The difference engine was conceived by Müller in 1786. Construction was first attempted by Charles Babbage,



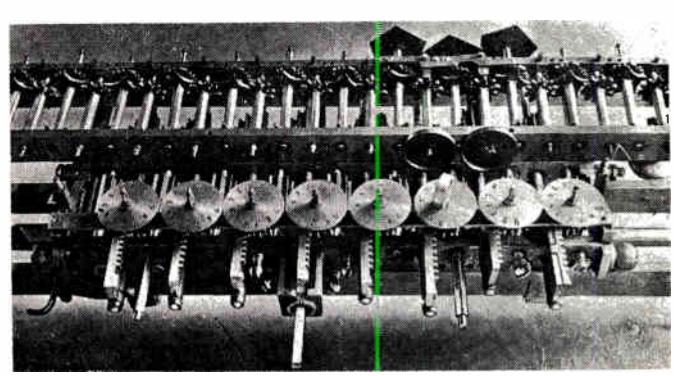
(*British Crown Copyright, Science Museum, London*)

Fig. 1—Replica of Pascal's calculating machine.



(*British Crown Copyright, Science Museum, London*)

Fig. 2—Replica of Pascal's calculating machine (mechanism).



(*Photo. Science Museum, London*)

Fig. 3—Leibnitz calculating machine (close-up of mechanism).

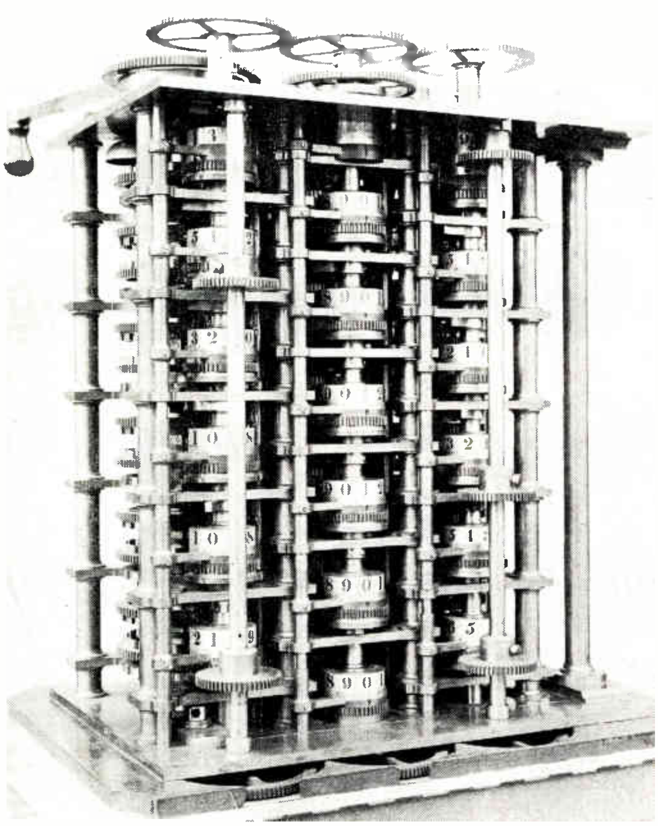
who, between 1812 and 1822, designed and built a small working model with three registers of six digits each [36] (Fig. 4). It was characteristic of the man that he immediately proposed to built the biggest difference engine that might ever be needed. It was to have seven 20-digit registers and printed output. With the backing of the Royal Society he obtained government support and began work in 1823. However, the engineering of the time was not up to such a machine and Babbage had to invent techniques for engineering drawing and for precision construction. The machine was only partially completed in 1833 when government support ended and work stopped. A five-register sixteen-digit machine was later built in Sweden by Scheutz and demonstrated in England in 1854 [1].

In 1833, Babbage conceived his analytical engine, the first design for a universal automatic calculator. He worked on it with his own money until his death in 1871. Babbage's design had all the elements of a modern general-purpose digital computer; namely: memory, control, arithmetic unit, and input/output. The memory was to hold 1000 words of 50 digits each, all in counting wheels. Control was to be by means of sequences of Jacquard punched cards. The very important ability to modify the course of a calculation according to the intermediate results obtained—now called conditional branching—was to be incorporated in the form of a procedure for skipping forward or backward a specified number of cards. As in modern computer practice, the branch was to be performed or not depending upon the algebraic sign of a designated number. The arithmetic unit, Babbage supposed, would perform addition or subtraction in one second while a 50×50 multiplication would take about one minute. Babbage spent many years developing a mechanical method of achieving simultaneous propagation of carries during addition to eliminate the need for fifty successive carry cycles. Input to the machine was to be by individual punched cards and manual setting of the memory counters; output was to be punched cards, printed copy, or stereotype molds. When random access to tables of functions—stored on cards—was required, the machine would ring a bell and display the identity of the card needed. Although Babbage prepared thousands of detailed drawings for his machine, only a few parts were ever completed (Fig. 6).

The description of Babbage's ideas would not be ade-

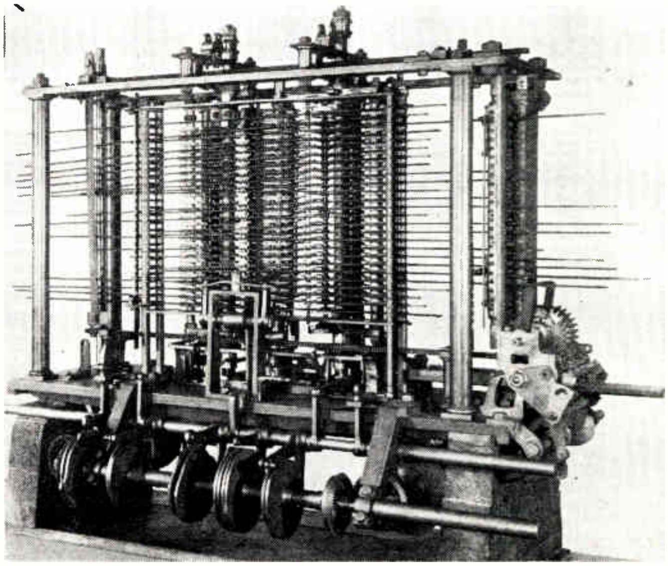
(British Crown Copyright, Science Museum, London)

Fig. 5—Charles Babbage (1792–1871).



(British Crown Copyright, Science Museum, London)

Fig. 4—Portion of Babbage's difference engine.



(British Crown Copyright, Science Museum, London)

Fig. 6—Portion of Babbage's analytical engine.

quate without a mention of Lady Ada Augusta, Countess of Lovelace, who was acquainted with Babbage and his work. Her writings have helped us to understand this work and contain the first descriptions of programming techniques.

Less than twenty years after Babbage's death, H. Hollerith conceived and developed the idea of machine-readable unit-record documents. He introduced electromechanical sensing means and apparatus for entering, classifying, distributing and recording data on punched cards. Hollerith's machines were used during the compilation of the 1890 Census reports. The development of many types of electromechanical accounting machines and early computers was based on Hollerith's inventions.

The development by J. W. Bryce and his associates of devices and circuitry which would transfer data between registers, or from registers to recording devices, was an important step in the evolution of electromechanical computers. One result was a machine developed by the International Business Machines Corporation for Columbia University during 1929 which solved mechanical problems that had thwarted Babbage. The inventions which had made this machine possible strongly influenced the development of the 600 series of IBM multiplying machines introduced in 1931 and were used later in the Mark I.

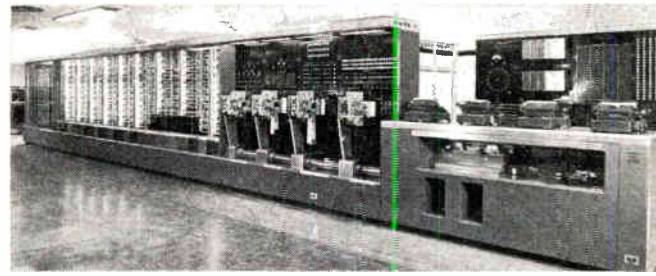
Another significant series of developments were the Bell Telephone Laboratories relay computers, which were based initially on the work of G. Stibitz in 1938. These biquinary (using both the base 2 and the base 5) and binary-coded decimal machines included paper tape input, program control, branching, self checking and many other features later incorporated in electronic computers. The Model II and III machines were built for the National Defense Research Council and were placed in operation in 1943 and 1944. Later, Model V was built for the National Advisory Committee on Aeronautics at Langley Field, Va., and for the Ballistic Research Laboratory at Aberdeen, Md. [10].

Transitional Machines

Mark I

Utilizing twentieth-century advances in mechanical and electrical engineering, the Automatic Sequence Controlled Calculator, or Mark I, brought Babbage's ideas into being, giving concrete existence to much more at the same time. The Mark I, an electromechanical calculator 51 feet long and 8 feet high, was built by the International Business Machines Corporation between 1939 and 1944, when it was formally presented to Harvard University by T. J. Watson, IBM's President. The Mark I was developed jointly by Howard Aiken, Professor of Applied Mathematics at Harvard, (later, also Commander, U.S.N.R.) and his staff, and by IBM engineers (Figs. 7 and 8).

The machine was a parallel, synchronous calculator, using a number length of 23 decimal digits plus alge-



(Computation Laboratory of Harvard University)

Fig. 7—The IBM Automatic Sequence Controlled Calculator ("Mark I"). *Left to right:* Constant registers, storage counters, multiplying-dividing unit, functional counters, interpolators, sequence control, electric typewriters and punched-card equipment.

(Computation Laboratory of Harvard University)

Fig. 8—Howard Aiken and part of the Mark I calculator.

braic sign. Calculations could also be done with 46 digits. It could perform any specified sequence of five fundamental operations, addition, subtraction, multiplication, division and reference to tables of previously computed results. The operation of the entire calculator was governed by an automatic sequence mechanism. The machine consisted of 60 registers for constants, 72 adding storage registers, a central multiplying and dividing unit, means of computing the elementary transcendental functions $\log_{10} x$, 10^x and $\sin x$, and three interpolators reading functions coded in perforated paper tapes. The input was in the form of punched cards and switch positions. The output was either punched into cards or printed by electric typewriters.

Each of the 60 constant registers consisted of 24 ten-position switches (for the 23 digits and the sign). Each of the 72 adding storage registers was composed of 24 electromechanical counter wheels capable of adding numbers as well as storing them. Each of these was essentially a ten-position switch (with additional contacts for carries) actuated through a magnetic clutch. Subtraction was effected in the same counter wheels by means of complementary numbers. The multiply-divide unit multiplied by first forming and storing the 9 integer multiples of the multiplicand; then selecting the multiples indicated by each digit of the multiplier and shifting and adding them in the proper columnar position. The position of the decimal point was fixed and determined by a plugboard. Division was carried out with the same unit in much the same manner: the 9 integer multiples of the divisor were first formed and stored. These multiples were then compared with the dividend and the largest multiple smaller than the dividend was selected; it was subtracted from the dividend and a digit defining it was entered in the quotient register. Logarithms, antilogarithms and sines were calculated automatically by means of special registers and appropriate series expansions of the functions. Each of the three interpolators carried an endless function tape containing values of the function for equidistant arguments together with appropriate interpolation coefficients. The tape was first positioned by moving it automatically in the direction of shortest travel; then the function was read by the machine and the interpolated value was calculated.

The sequence mechanism consisted of a sprocket drum over which ran a perforated paper tape, the "control" tape. Each transverse line of the tape had space for 24 equidistant holes in three groups, A, B, and C, of 8 holes each. Each line of holes constituted the instruction: "Take the number out of unit A, deliver it to B, start operation C." The sequence and interpolator mechanisms and the counter wheels were all synchronized and driven by a gear-connected mechanical system and a 5-HP motor. The fundamental cycle was of 300 msec. A typical multiplication required about 3 sec [4], [8], [14].

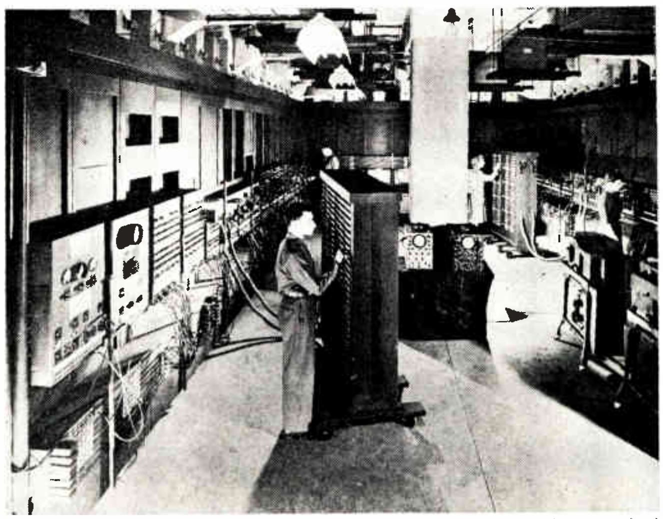
The Automatic Sequence Controlled Calculator was in operation for more than fifteen years. Results of its work are used in almost every computation laboratory.

ENIAC

The *ENIAC* was developed and built at the Moore School of Electrical Engineering of the University of Pennsylvania beginning in 1943, and was completed in 1946. As evidenced by its full name, "Electronic Numerical Integrator and Computer," it was not intended to be a general-purpose machine. Its principal objective was the computation of firing and ballistic tables for the Aberdeen Proving Ground of the U. S. Army Ordnance Department. This computation required the integration of a simple system of ordinary differential equations involving arbitrary functions.

Physically, the ENIAC was impressive. It consisted of 40 relay racks of equipment, arranged in a U-shaped assembly, with 4 relay racks forming a module, 4 modules on each side and 2 at the rear, as seen from the inside in Fig. 9. Half of the racks contained accumulators (registers with facilities for addition), one accumulator per rack, and the other half included the following:

1) Master programmer, cycling unit and initiating unit. These emitted the control impulses, cleared the machine to zero (4 racks),
2) Function-table panels (6 racks). These supplied on demand the functional values associated with a given argument. There were 3 such tables, mounted on heavy casters. The values were settable in advance on the rotary switches,
3) Divider and square root unit (2 racks),
4) Multiplier (3 racks),
5) Constant transmitter (3 racks),
6) Printer (2 racks).



(Remington Rand UNIVAC Division, Sperry Rand Corporation)

Fig. 9—The ENIAC.

This bulky equipment occupied a space 30×50 feet and contained 18,000 vacuum tubes. It required 130 kw of power. The computing elements consisted largely of decade rings, flip-flops and pentode gates. These were direct-coupled, requiring 18 power-supply voltages and consequent heater-cathode potential differences of over 100 volts. The input-output system consisted of modified IBM card readers and punches.

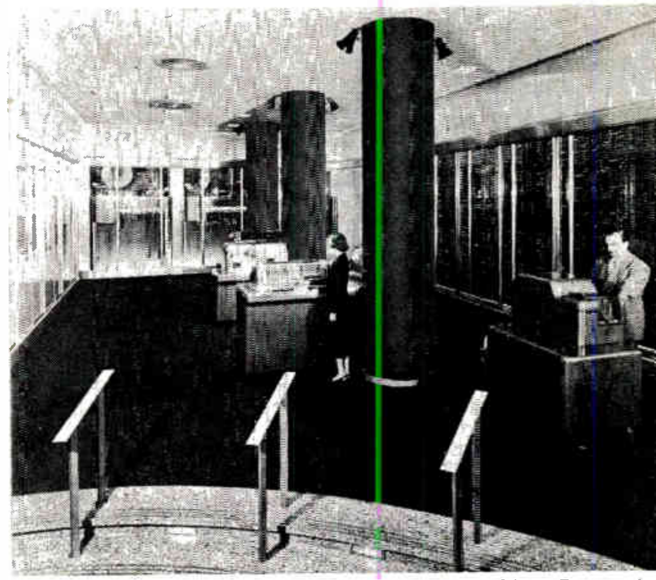
Numbers were of ten decimal digits, with negative numbers represented by complements. A basic pulse rate of 100 kilopulses per second was used. The machine carried out its calculations in the parallel mode by transmitting numbers additively or subtractively from one accumulator to another, with references to the multiplier and divider, table look-ups from the function tables and new numbers from the constant transmitter. Addition required 20 pulse times (0.2 msec) and multiplication required 14 addition times (2.8 msec). All this required elaborate preparation of the so-called background wiring of the machine. This wiring—plugged connections and switches—effected the cycling of addition, multiplication, etc., under control of the cycling unit. The connections required from 30 minutes to a full working day to set up and constituted a serious limitation on the flexibility of the system.

Even though it was not a serious obstacle to the calculation of ballistic and firing tables, this limitation was eased in the late 1940's when a more flexible wired programming system was devised as a result of suggestions of John von Neumann. The background wiring was made permanent and the function tables were used to store instructions rather than values of functions. This was made possible by extensive changes in the cycling unit and associated equipment. The instructions did not then constitute an internally-stored program, however, since they were manually set on the function-table switches [6], [7], [9], [11], [14].

The ENIAC is no longer a productive computer. The units composing it have been separated and given to museums and other institutions where they are accessible to the public. There, they serve as reminders of a pioneering effort in extending the scope of man's ability.

IBM 604 Electronic Calculator

The IBM 604 Electronic Calculator was a general-purpose electronic digital computer produced in quantity. Development began late in 1946 and the first deliveries took place at the end of 1948. This machine remained in production until 1960, and more than 4000 were built. The IBM 604, a descendent of the IBM 603 (a punched-card, six-digit electronic multiplier) employs miniature tubes in one-tube pluggable circuit packages for economy in assembly and maintenance. The 604 contains over 1400 tubes and operates at 50 kilopulses per second. The basic machine has a capacity of 60 program steps, which are set up on a plugboard. All steps,



(International Business Machines Corporation)

Fig. 10—The IBM Selective Sequence Electronic Calculator (SSEC), a 13,000-tube, 23,000-relay machine, went into operation at IBM New York Headquarters in December 1947, and was used until August, 1952. The SSEC was capable of conditional branching, permitting the selection of alternative program sequences according to intermediate results.

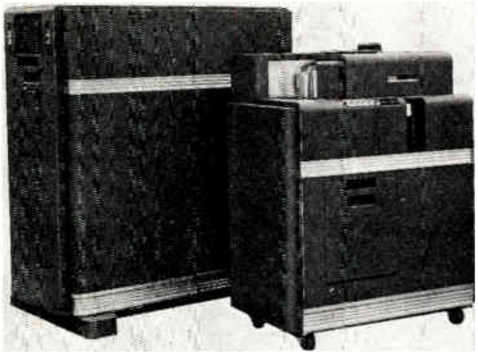
except those calling for multiplication or division, are executed in 0.5 msec (Fig. 11).

An important feature of the IBM 604 was the inclusion of a zero-test and a sign-test on the contents of the accumulator. These tests permitted a conditional program branch by allowing suppression of some steps of the normal program sequence. Additional flexibility was provided by the use of relay selectors. The calculator has a 13-decimal-digit accumulator and 37 digits of additional storage in vacuum-tube registers. Input and output are on punched cards. Marginal checking (*i.e.*, determination of performance margins) by clock frequency variation is provided [142].

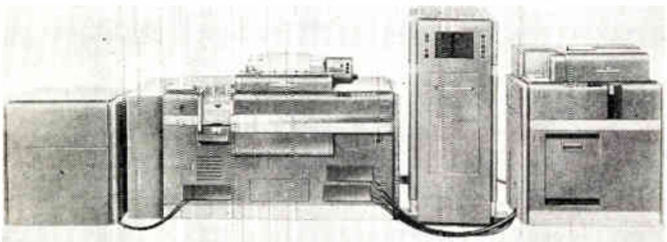
IBM Card-Programmed Calculator (CPC)

The IBM Card-Programmed Calculator (CPC) line descended from the connection, made for Northrop Aviation in 1947–1948, of a multiplier to an accounting machine tabulator. First an IBM 601 Multiplier (mechanical) and then an IBM 603 Electronic Multiplier was connected to an IBM 405 Alphabetic Accounting Machine for solution of linear differential equations. In 1948–1949, an IBM group produced the models I and II Card-Programmed Calculator by connecting an IBM 604 to an IBM 402 Alphabetic Accounting Machine. This system provided a parallel printer and 80 digits of counter wheel storage. Sixteen additional ten-digit-plus-sign words could be stored in counter wheels in each of three separate IBM 941 Auxiliary Storage Units.[5] (Fig. 12.)

[5] A "word" is an ordered set of digits (or characters) used together.



(International Business Machines Corporation)
Fig. 11—IBM Type 604 Calculator.



(International Business Machines Corporation)
Fig. 12—IBM Card-Programmed Electronic Calculator. *Left to right:* 941 auxiliary store, tabulator-printer, 605 computer, card reader-punch.

The Card-Programmed Calculator took its instructions from cards which could call for programs stored in a plugboard. Conditional branches were effected by switching to alternate sequences in the cards. The arithmetic speed of the IBM 604 System was lower than those of the large-scale machines then in operation and under construction, and its internal storage capacity was more limited. But access of the Calculator to an unlimited program on cards eliminated housekeeping operations (so time-consuming on the early stored-program machines) required to manipulate a stored program, and enabled the CPC's to outperform more sophisticated equipment in many applications. At one time about 200 Card-Programmed Calculators were simultaneously in use [142].

Electronic Discrete Variable Automatic Computer (EDVAC)

The Electronic Discrete Variable Automatic Computer, or EDVAC, was built at the Moore School (University of Pennsylvania) between 1947 and 1950 for the Ballistic Research Laboratory at the Aberdeen Proving Ground. It is a serial, synchronous machine in which all pulses are timed by a master clock operating at 1 megapulse per second. It contains some 5900 vacuum tubes, about 12,000 semi-conductor diodes and utilizes the binary number system with a word length of 44 binary digits (Fig. 13).

(Ballistic Research Laboratory)
Fig. 13—The EDVAC.

The EDVAC was built as a stored-program machine. The instructions, each of which consists of four ten-bit addresses and a four-bit operation code,[6] are stored in an internal memory together with the numbers to be used in the computation. This internal memory is composed of 128 thermostatically-controlled acoustic delay lines, each storing 384 bits (and accommodating 8 words) as sound waves in mercury. The information constantly circulates through the line. $128\times 8=1024$ words of rapid-access storage are thus available, with an access time of 48 to 384 μsec.

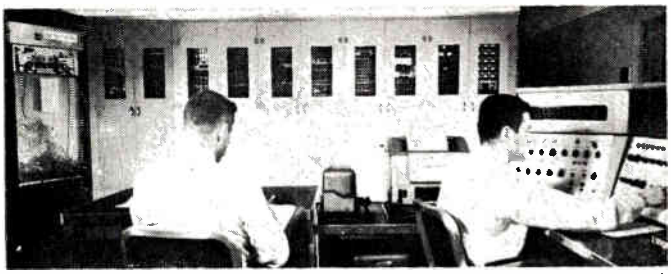
Twelve distinct operation codes are used, out of the possible sixteen, including double-precision operations (*i.e.*, with double-length numbers) and a branch instruction. Diode logic circuits are used. Subtraction is done directly, not through complements. The time required for the average addition is 864 μsec, for the average multiplication, 2900 μsec. Input-output means include paper tape, teletypewriter and punched cards. Equipment additions have been made to the machine from time to time, including particularly magnetic-drum and magnetic-tape storage.

Important features of the EDVAC have been used in other machines, notably the UNIVAC (built by Eckert and Mauchly) and the SEAC (National Bureau of Standards). The EDVAC was still in operation at Aberdeen at the time this paper was written [17], [142].

Standards Electronic Automatic Computer (SEAC)

The Standards Electronic Automatic Computer, or SEAC was built by the staff of the Electronic Computer Laboratory of the National Bureau of Standards. The design began in June, 1948 and the machine was put in operation in May, 1950. It was built under the sponsorship of the Office of the Air Comptroller, Department of the Air Force, principally to carry out mathematical investigations of techniques for solving large logistics programming problems (Fig. 14).

[6] "Bit" is a common abbreviation for "binary digit."



(*National Bureau of Standards*)

Fig. 14—The SEAC.

The SEAC was built as a stored-program computer. Its design was based upon that of the EDVAC. The SEAC is a serial, synchronous, binary, fixed-point machine operating at a one-megapulse-per-second rate with a word length of 44 bits plus sign. It was planned and constructed as a nucleus to which improvements and subassemblies could be added to increase its problem-solving capabilities. With some 750 tubes and 10,500 diodes, the original installation consisted mainly of the computing and control circuitry, a memory of 64 mercury acoustic delay lines providing for the storage of 512 forty-five bit words, a manual keyboard for direct input and a teletype printer for direct output. Punched paper tape was used to provide indirect operation. There were seven basic four-address instructions including "logical transfer" or branching.

Equipment additions have been made to the machine from time to time. One of the first was that of a cathode-ray-tube electrostatic memory of a type developed by Williams at the University of Manchester [18]. This electrostatic memory was later improved and enlarged and acoustic-delay-line memory capacity was doubled. The machine now has a memory capacity of 1024 words in acoustic delay lines and 1024 words in CR tubes. It contains some 2300 vacuum tubes and 24,000 germanium diodes. A "Flexowriter" (an electric typewriter punching paper tape) and magnetic tape are used for input and output as well as the original means. Facilities have been provided for the use of three-address instructions as an alternative to the four-address structure.

The National Bureau of Standards has operated the SEAC as one of its own facilities since January, 1954. The machine has been working nearly 12 years, not only solving innumerable problems for government agencies, but also permitting studies in the data processing field and providing means of developing and testing new computer components and devices [40], [44], [66], [142].

Whirlwind I

An assignment to build a real-time aircraft simulator was given in 1945 to the Digital Computer Laboratory of the Massachusetts Institute of Technology, at that time a part of the Servomechanisms Laboratory of M.I.T. Beginning in 1947, the major part of the effort was devoted to the design and construction of the electronic digital computer known as Whirlwind I. The project was sponsored by the Office of Naval Research and the United States Air Force. The machine was put in operation in March, 1951 (Fig. 15).

Whirlwind I was a parallel, synchronous, fixed-point computer utilizing a number length of 15 binary digits plus sign (16 binary digits in all). Physically, it was a large machine containing some 5000 vacuum tubes (mostly single pentodes) and some 11,000 semiconductor diodes. It consisted of an arithmetic "element" including three registers; a control element including central control, storage control, arithmetic control and input-output control; a program counter; a source of synchronizing pulses, or master clock, supplying 2 megapulses per second to the arithmetic element and 1 megapulse per second to the other circuits; an internal storage element, or memory; terminal equipment; and extensive test and marginal checking equipment.

The arithmetic element included three registers designated AC, A and B and a counter. The first register was a shifting accumulator, formed of an adding register and a carry register; it held the product during multiplication. The A register held the multiplicand during multiplication. The B register was a shifting register holding the multiplier. Special fast-carry circuits were provided for addition. Subtraction was by nine's complements and "end-around" carry.[7] Multiplication was by successive addition and division by successive subtraction. (In binary arithmetic, multiplication and division are carried out by addition, or subtraction, and shifting.) Complete addition required 3 μsec. The average multiplication required 16 μsec.

Whirlwind I was a stored-program computer. The instructions, each of which consisted of one ten-bit address and one five-bit operation code, were stored in the internal memory of the machine. There were 32 distinct operation codes, of which about 27 were assigned, including addition, subtraction, multiplication, division, shifting, transfer of all or parts of words, branching and conditional branching. There were terminal-equipment control instructions and also special instructions to facilitate double-length and floating-point operations.

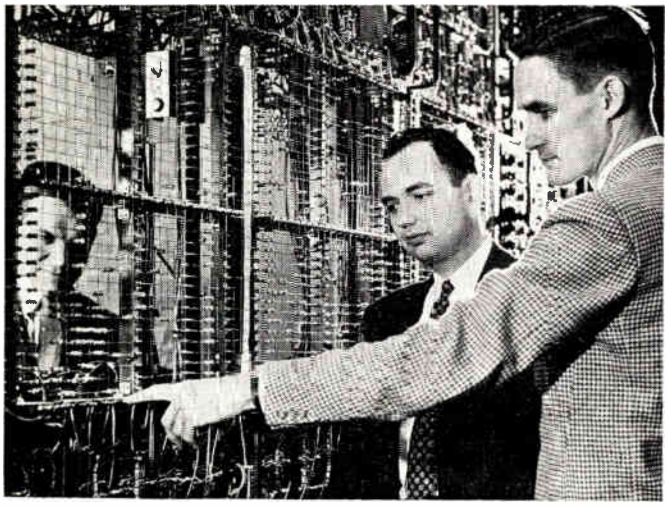
The central control of Whirlwind I was arranged in the form of two "crystal matrices" or diode arrays (Fig. 16). The first of these was a 32 position control switch to decode the order part of the instruction while the second was an operation control matrix in which one horizontal selection made a large number of vertical selections. Both arrays were driven at a rate of one megapulse per second.

The memory of Whirlwind I consisted at first mainly of 16 specially-designed electrostatic storage tubes, each with an array of $32 \times 32 = 1024$ charged spots, providing 1024 registers of storage [20]. The M.I.T. electrostatic storage tube differed from the Williams CR storage tube

[7] From the highest to the lowest column of the machine.

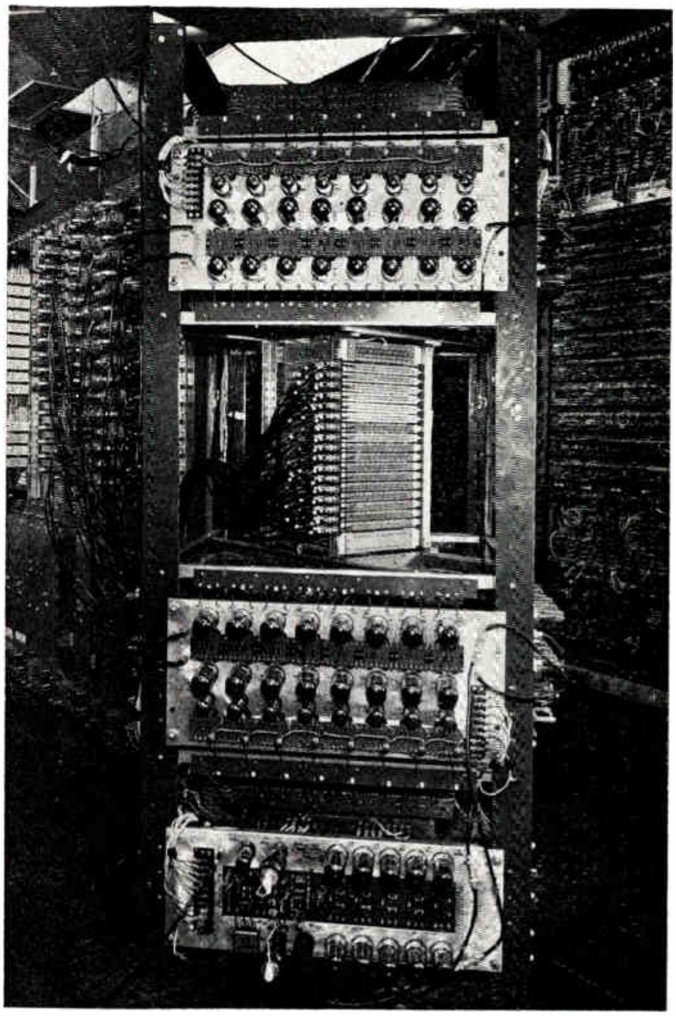
(*Massachusetts Institute of Technology*)

Fig. 15—Control room of the Whirlwind I Computer.



(*Massachusetts Institute of Technology*)

Fig. 16—Diode switching array of the Whirlwind I Computer.



(*Massachusetts Institute of Technology*)

Fig. 17—Coincident-current magnetic-core storage array of the Whirlwind I Computer.

developed at the University of Manchester [18] in that a second electron gun was used to maintain the charge, instead of cyclic replacement, and the storage depended on a change of the potential of the charged spot, rather than on a change of the geometry of the spot. Later, a coincident-current magnetic-core memory was used. The cores were made of a material possessing a rectangular hysteresis loop [30], [50]. Two banks, each of 32×32×17 cores were in operation in the machine in 1953 (Fig. 17). They provided 2048 registers of storage, each of 16 binary digits and a parity-check digit.[8]

The terminal equipment included:

1) A photoelectric paper tape reader,
2) Electromechanical paper tape readers and punches,
3) Electromechanical typewriters,
4) CR oscilloscope displays, including a computer-controlled scope camera,
5) Inputs from various analog equipments needed for control studies,
6) Outputs to analog equipment.

[8] A parity-check binary digit is used to make the number of binary "ones" of constant parity, either odd or even. Since a single error changes the parity, the parity check permits the detection of any single error.

Magnetic-tape units and magnetic drums were added later.

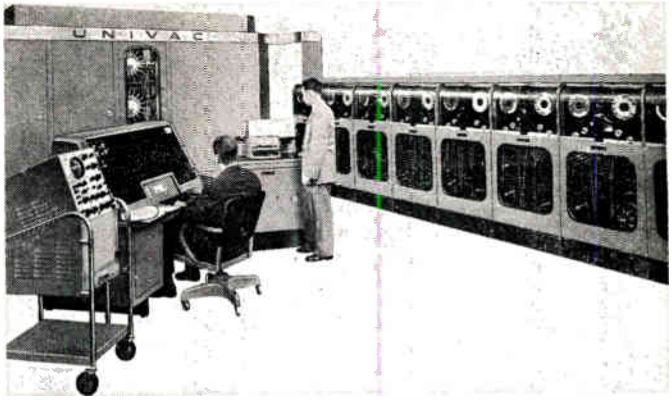
Preventive maintenance techniques called "marginal checking" were used on the Whirlwind I to establish the life expectancy of various components, particularly vacuum tubes and diodes, by determining performance margins. Variations of operating voltages and of pulse times were used, together with special pulse sources, to detect imminent failures. This was done automatically by means of telephone switching apparatus. The marginal checking of the whole computer, which could be carried out in 15 minutes, helped to bring its operational reliability to comparatively high levels [25].

Ideas embodied in the Whirlwind I are used in most of the modern digital computers [28], [33].

Universal Automatic Computer (UNIVAC)

The development of the Universal Automatic Computer, or UNIVAC was started about 1947 by Presper Eckert and John Mauchly who founded the Eckert-Mauchly Computer Corporation in December of that year. The first UNIVAC I was built for the U. S. Bureau of the Census and was put in operation in the Spring of 1951. (The Eckert-Mauchly Corporation later became a subsidiary of Remington Rand, forming the organization which is now the Remington Rand UNIVAC Division of the Sperry Rand Corporation.)

UNIVAC I was a direct descendent of the ENIAC and of the EDVAC, in the development of which Eckert and Mauchly had both had an important part at the University of Pennsylvania. It was a serial, synchronous machine operating at a rate of 2.25 megapulses per second. It contained some 5000 vacuum tubes and several times as many semi-conductor diodes in logic and clamp circuits. One-hundred mercury delay lines provided 1000 twelve-decimal-digit words of internal storage. Twelve additional delay lines were used as input-output registers. Aside from console switches and an electric typewriter providing small amounts of information, the input-output medium was metal-base magnetic tape. 48 UNIVAC I machines were built (Fig. 18).



(Remington Rand UNIVAC Division, Sperry Rand Corporation)

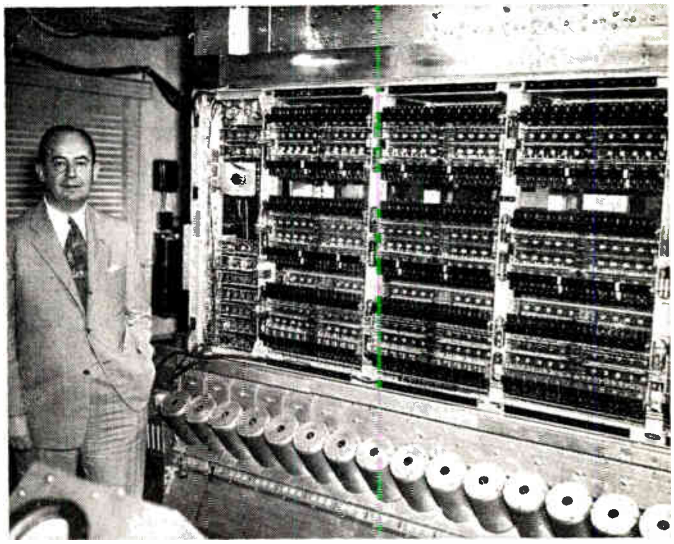
Fig. 18—UNIVAC I.

UNIVAC I was built as a stored-program machine. Through the use of a six-bit code, alphabetic characters, punctuation, control symbols, etc., could be handled by the machine as well as decimal digits. Within this binary-coded representation, decimal digits were represented in the "excess-three" code.[9] The code included a seventh bit for a parity check of every "alphanumerical" (and other) character. There were some 45 distinct instructions, including rounded and unrounded arithmetic, conditional transfers and input-output control. Each instruction consisted of a one-alphabetic-character operation code and of a three-decimal-digit address. There were two instructions per word. Subtraction was direct. Addition required 0.5 msec; about 2.5 msec were required for multiplication.

Not long after UNIVAC I was put in operation, "automatic programming" techniques were developed for the machine by its builders. These techniques consist essentially of means to instruct a computer to prepare that sequence of its own instructions which is needed to perform a given task. Sets of such means have become extensive programming languages that are important for the effective use of the modern digital computers[10] [27], [45], [142].

Institute for Advanced Study Computer

The design of the Institute for Advanced Study Computer was started in 1946 and the machine was completed in June, 1952, by a staff under the direction of Prof. John von Neumann, IAS. Its development and construction were initially sponsored by the Ordnance Corps, U. S. Army; later, the project was also sponsored by the ONR, the USAF and the AEC (Fig. 19).



(Institute for Advanced Study)

Fig. 19—John von Neumann (1903–1957) and the Institute for Advanced Study Computer.

The IAS Computer was an electronic, parallel, asynchronous, fixed-point machine using numbers consisting of a sign and 39 binary digits. It was of unusually small size, 2×8×8 feet (not including input-output equipment). It contained some 2300 vacuum tubes (mostly double triodes) and consisted essentially of 3 registers designated as RI, RII and RIII, a vacuum-tube parallel adder, a control unit and an electrostatic memory composed of 40 cathode-ray tubes. A magnetic drum was added to the machine later. Input and output utilized punched cards and teletype tape.

RI was an accumulator, a device capable of holding a number and of adding another number to it. It was a shifting register holding the augend and then the sum during addition. RII was called the "arithmetic register." It also was a shifting register and was designed to hold the multiplier and the least significant digits of the partial products during multiplication, and the quotient during division. RIII was a nonshifting "memory

[9] In the "excess-three" code each decimal digit is represented by the (four-bit) pure binary equivalent of the decimal digit plus three. In this code, binary complementation yields nines complements.

[10] See R. D. Elbourn and W. H. Ware, "The evolution of computing techniques and concepts," PROC. IRE, this issue, pp. 1059–1066.

register" associated with 40 complementing gates.[11] It contained the addend during addition, the multiplicand during multiplication and the divisor during division. The adder circuit contained 4 tubes (per digit) which propagated a carry to the next stage and which reduced the 8 possible inputs to 4 possible voltage levels. These 4 voltage levels were then resolved to 2 levels, representing a "1" or a "0," with additional circuitry. The control unit contained shift and address counters and other control devices. The electrostatic memory was of the type developed by Williams at the University of Manchester [18]. It had a capacity of 1024 forty-bit words and consisted of an array of $32\times32=1024$ charged spots on the phosphor surface of each of forty cathode-ray tubes.

Negative numbers were represented by complements. Multiplication was performed by successive addition and division by successive subtraction. The IAS computer was a stored-program machine, the instructions being stored in the CRT memory. Each instruction was composed of a ten-bit address and of a ten-bit operation code. There were thus a pair of instructions in each instruction word. The basic "machine language" or set of basic instructions, included: addition (or subtraction, executed in 62 μsec), multiplication (720 to 990 μsec), division (1100 μsec), left shift, right shift, load in RII, store from RI, input, output, unconditional transfer and conditional transfer (branching). This last instruction was executed when the number in the accumulator was positive, ignored when it was negative.

The IAS Computer was the prototype for various subsequent machine developments, including three for the AEC and one for the U. S. Army Ordnance Corps, Ballistic Research Laboratory, at the Aberdeen Proving Ground [5], [41].

IBM 701

The development of the IBM 701 Data Processing System began at the end of 1950. A model was operating late in 1951 and the first production machine was delivered at the end of 1952. The heart of the IBM 701 system was a 36-bit single-address, binary, parallel, synchronous processor employing vacuum-tube flip-flops and diode logic at a rate of one megapulse per second. Multitube pluggable circuit packages were used. The arithmetic registers employed a recirculating-pulse bit-storage circuit, developed for the NORC [67], in which a combination of diode gating and pulse delay made it possible to store, shift right, or shift left with one triode per bit. Computation was governed by a single-address stored program of two 18-bit instructions per 36-bit word (Fig. 20).

For the high-speed memory, a parallel-mode cathode-ray-tube system—patterned after the serial CRT store developed by Williams [18] was devised. Seventy-two CR tubes, storing 1024 bits each, provided a 2048-word memory with a 12-μsec memory cycle. Most instructions took 5 cycles, 3 of which were for regeneration of the CRT memory. Multiplication and division took 38 cycles, during which a reserve of regeneration cycles was built up that the succeeding 12 instructions could draw upon for faster operation.

Input to the processor was on punched cards; a 150-line-per-minute alphanumeric printer and a 100-card-per-minute card punch provided output. Magnetic-tape units using plastic tape operated at a rate of 7500 six-bit groups per second (100 bits per inch). Vacuum columns were used to buffer tape between the reels and the rapid start-stop pinch-roller capstan. Both clock-frequency and supply-voltage marginal checking were used. Nineteen IBM 701 Systems were completed [37], [42], [52], [142].

IBM 650

The IBM 650, an intermediate-size, vacuum-tube computer was considered a workhorse of the industry during the late 1950's. Development began in 1949 and the first installation was made late in 1954. Over a thousand 650's have been in service since then (Fig. 21).

The 650 operates serially by character on words of 10 decimal digits plus sign. A 2-out-of-5 decimal representation in storage is translated into a biquinary code in the operating registers, allowing a fixed-count check to detect the presence of more or less than 2 bits per character. The main store of the 650 is a 12,500-rpm, 2000-word magnetic drum. A two-address instruction format accommodates, as part of each instruction, the

(International Business Machines Corporation)

Fig. 20—IBM 701 Data-Processing System. *Left to right:* Drums, CRT store, cpu and card reader, power control, tapes, card punch, printer.

(International Business Machines Corporation)

Fig. 21—IBM 650 (right, rear) with magnetic tape and ramac file.

[11] In computer terminology, a "gate" is a logic circuit designed to supply an output when and only when a specified set of conditions exist at its (several) inputs. See Ref. [75].

location of the next program step. This format allows the programmer to place instructions anywhere on the program drum, and makes it possible for him to minimize access times to successive instructions.

The 650 has a basic clock rate of 125 kilopulses per second, an 800-μsec add time, and a 24- to 200-msec multiply time. A three-triode latch—basically a double inverter with cathode-follower feedback—is employed for bit storage. Checking includes a character validity check and control timing ring start and stop. A buffer area on the drum allows simultaneous card input-output and processing. The original punched-card input-output devices are currently supplemented by paper tape, magnetic-tape units, a line printer, and remote inquiry stations. Additional supplementary equipment subsequently available for use with the 650 includes random-access disk files, a 60-word core memory, and the logic necessary for index registers, floating-point operations, and the handling of alphabetic characters using two digits per character [65], [142].

Modern Systems

The transitional machines differed markedly from each other, especially in the beginning. But modern electronic digital computers generally possess many attributes in common. They are usually built in several units, only one of which is a computer, or "processor." The other units are control, storage or input and output devices. Thus, the modern machine is more properly called a computing system than a computer. Most modern digital computing systems use semi-conductors (diodes and transistors) rather than tubes. Most include magnetic-core and magnetic-tape storage. Many contain circuits to perform floating-point arithmetic directly; and, in the later ones especially, sizes and amounts of equipment can be varied to meet specific needs.

The computers of modern digital systems are all stored-program machines and all include automatic checking. The modern systems nearly all provide alphabetic as well as numerical input and output. In many of them, especially the later ones, arithmetic is performed concurrently with other operations.

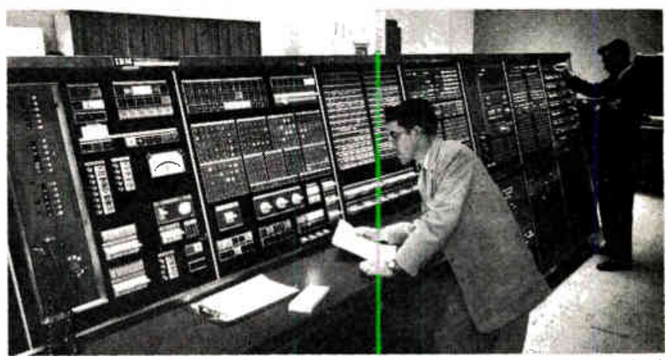
Since the engineering and other technical aspects of the modern digital computing systems are discussed in accompanying papers, we shall describe only the highlights of a few typical systems among the larger ones.

Air Defense Computers

Descendants of MIT's Whirlwind I and of the IBM 701, the AN/FSQ-7 air defense computers for the SAGE System began, in 1952, as a cooperative IBM-M.I.T. Lincoln Laboratories effort based upon previous studies and specifications by Lincoln Laboratories. SAGE, a real-time communication-based digital computer control system, accepts radar data over phone lines, processes, displays information for operator decisions and guides interception weapons. The first engineering model of the computer was delivered by IBM in 1955, and production deliveries began in June, 1956 (Figs. 22 and 23).

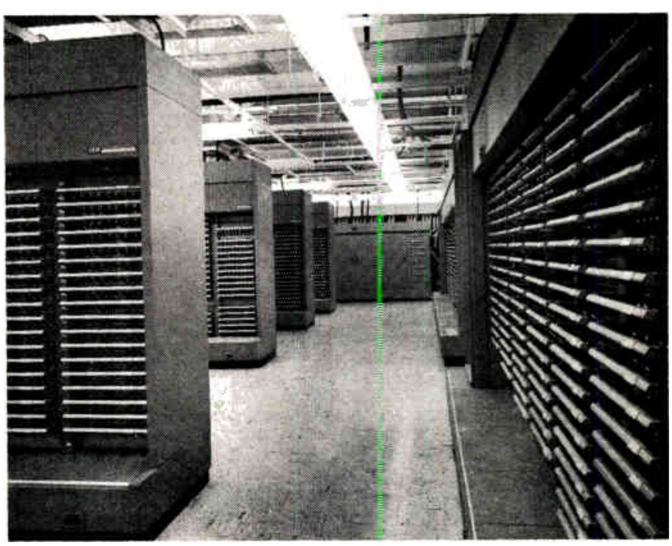
Each computer of this duplex system is a vacuum-tube and germanium-diode machine using diode and multigrid gate-tube logic, flip-flop bit storage, and pulses through pulse transformers for bit transmission. The computer originally incorporated an 8192-word core memory (6-μsec cycle), since increased to 69,632 words, and provides over 150,000 words of drum storage. The Q-7 is single-address, binary, parallel and synchronous. Words are 32 bits plus a parity bit in length. The arithmetic unit is "dual," *i.e.*, capable of operating simultaneously on two separate 16-bit numbers for efficiency in handling two-coordinate calculations. Other noteworthy features include instruction indexing, computation in parallel with input-output, high-speed multiplication (16.5-μsec total execution time for 16-bit factors), and buffer drums to couple the processor to many slow-speed inputs and outputs.

An extensive console system with keyboard inputs provides alphanumeric and pictorial display information to system operators. The processor, drums and tape



(International Business Machines Corporation)

Fig. 22—AN/FSQ-7 SAGE Computer (operating console).



(International Business Machines Corporation)

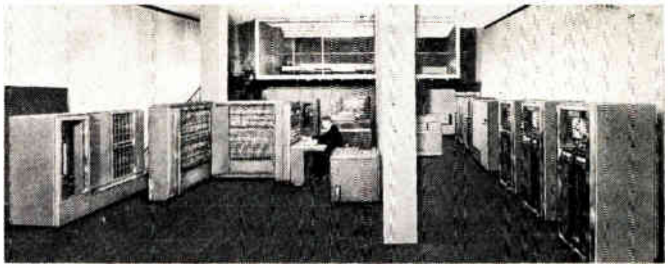
Fig. 23—AN/FSQ-7 SAGE Computer (computer frames).

units are duplexed for real-time reliability, while backup units are ready at all times to assume input-output and display functions. [72], [81], [142].

IBM 704, 709 and 7090

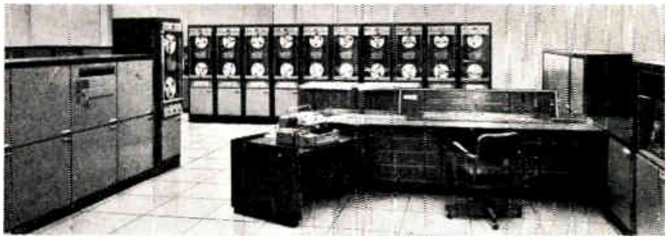
The development of the IBM 704, descended from the 701, began in November, 1953, and the first system was delivered in January, 1956. The 704 featured higher speed, magnetic-core memory, floating point and indexing (Fig. 24). It was followed in 1958 by the 709, featuring simultaneous read, write and compute by means of Data Synchronizer units that allowed the input-output channels to operate independently, as well as several special operations, including a table look-up instruction and indirect addressing. A 32,768-word core memory was installed on a 704 in April, 1957.

The IBM 7090, the first units of which were delivered in 1959, is a transistorized system compatible with the 709. About 5-times faster than the 709 on typical problems, the 7090 incorporates a 2.18-μsec core memory and improved magnetic-tape units [142].



(International Business Machines Corporation)

Fig. 24—IBM 704.



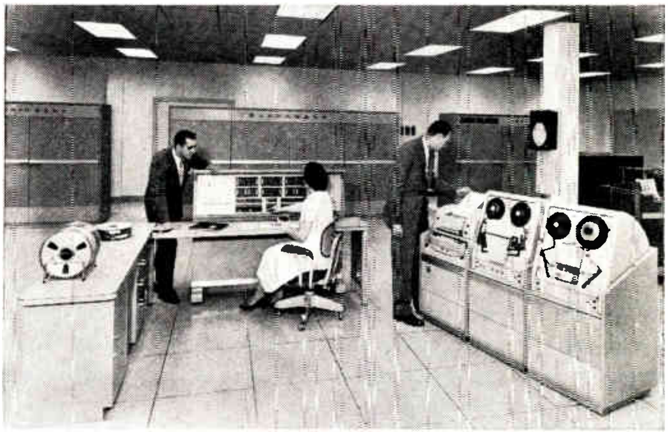
(Philco Corporation)

Fig. 25—Philco 2000 System.

Philco 2000

The Philco 2000 is an all-purpose data-processing system. The first installation was made in November, 1958. The machine's design is based upon direct-coupled, asynchronous logic circuits using surface barrier transistors (Fig. 25).

A typical Philco 2000 system contains 450 tubes, 1200 diodes and 56,000 transistors. The computer utilizes binary, binary-coded-decimal and alphanumeric internal number systems with a 48-bit word length. There are two 24-bit, one-address instructions per word; 59 of the 225 distinct instructions are floating-point instructions. As many as 32 index registers can be provided for instruction modification, relative addressing, etc. Arithmetic is performed in the parallel asynchronous mode concurrently with other operations. Addition requires 1.7 μsec, multiplication, 40.3 μsec. 4096 to 32,768 words of magnetic-core storage and 32,768 to 1,048, 576 words of magnetic-drum storage are provided. Input, output utilize punched cards, magnetic tape (up to 256 units) and a 900-lines-per-minute printer. Among the units available with the system is an input-output processor which connects 16 channels to the central computer, multiplexing any 4 channels [142].



(Burroughs Corporation)

Fig. 26—Burroughs 220.

Burroughs 220

The Burroughs 220 is a general-purpose data-processing system operating in the series-parallel mode (Fig. 26). The first installation was made in December, 1958.

The central processor contains 1800 vacuum tubes. The internal number system is binary-coded decimal, with a word length of 10 decimal digits plus sign. There are 93 instructions, each containing one four-decimal-digit address, a two-digit operation code and four "control digits" for instruction modification, relative addressing, etc. Arithmetic is performed in the serial, synchronous mode, with fixed or floating point. Addition requires 185 μsec, multiplication 2055 μsec. 10,000 words of magnetic-core storage are provided. Input-output equipment includes paper tape, punched cards, keyboard and monitor printer and up to 10 units of magnetic tape. Output includes a 1500-lines-per-minute printer [142].

National 304

The National 304, built by the National Cash Register Company, is a general-purpose data-processing system (Fig. 27). The first installation was made in November, 1959.

The central processor contains some 8000 diodes and 4000 transistors. The internal number system is binary-coded decimal. Word length is either of 10 alphanumeric characters or of 5 to 60 characters. 83 one-address "scientific" instructions or 37 three-address "business" instructions (with variations) can be used. Special opera-

(The National Cash Register Company)

Fig. 27—National 304.

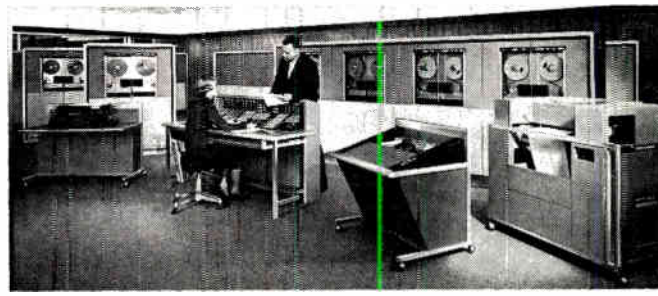
tion codes, relative addressing and index registers are provided. Arithmetic is serial by word, parallel by character, with either floating point or fixed point and proceeds concurrently with other operations. With single-address instructions, addition requires 60 μsec, multiplication 1260 μsec. 2400 to 4800 ten-character words of magnetic-core storage can be used. Input-output equipment includes punched cards, punched paper tape and magnetic tape. Magnetic tape has no gaps between records. Up to 64 units can be used. Output also includes an 850–1200 lines-per-minute printer [142].

RCA 501

The RCA 501 is a general-purpose data-processing system using "building block" construction and completely variable word length (Fig. 28). The first installation was made in December, 1959.

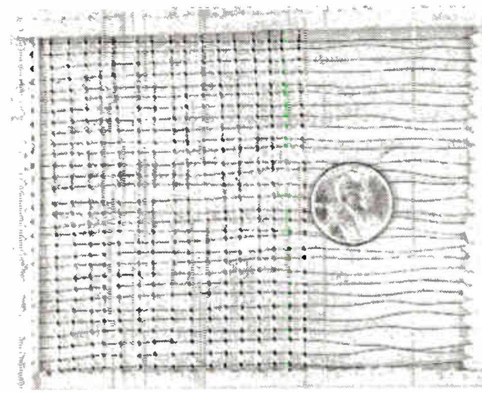
The RCA 501 utilizes transistor and diode logic circuits. The internal number system is binary. A word consists of all the characters between special symbols defining its beginning and its end. Forty-nine two-address instructions are used, each including a one-character (six-bit) operation code, three characters for each of the two addresses and a one-character address modifier. This address modifier permits adding the contents of any one or two of seven memory locations or other registers to the instruction addresses. Arithmetic is performed with fixed point in the serial, synchronous mode concurrently with other operations. Typical addition requires 0.24 to 0.42 msec, typical multiplication 1.9 to 9.6 msec.

Up to 262,144 octal digits (base 8) of magnetic-core storage can be used. Decimal information is automatically stored as its octal equivalent. A magnetic drum storing 1.5 million characters is optional equipment. Input equipment includes keyboard, paper tape, punched cards and magnetic tape. (Up to 62 tape units can be connected on line. For accuracy, all information on magnetic tape is automatically recorded twice.) In addition to paper and magnetic tapes, output equipment includes a 900-lines-per-minute printer and a monitor typewriter [98], [142].



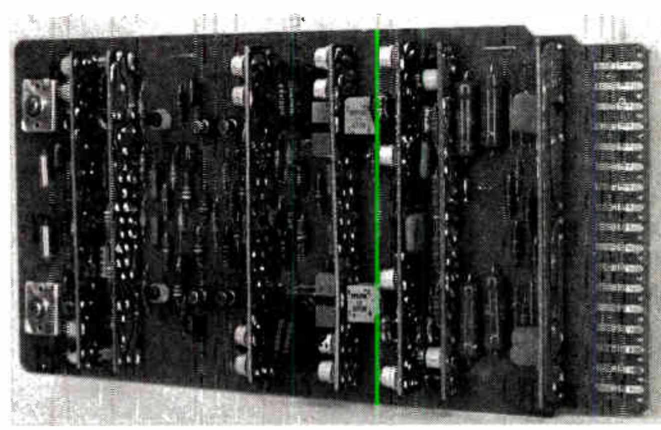
(Radio Corporation of America)

Fig. 28—RCA 501 System.



(Radio Corporation of America)

Fig. 29—Small-magnetic-core array.

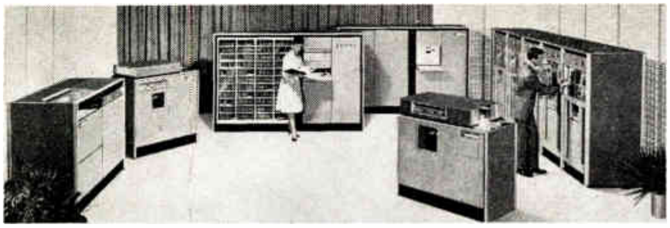


(Radio Corporation of America)

Fig. 30—Multiple printed-circuit card assembly.

UNIVAC Solid State 80/90

The UNIVAC Solid State 80/90 was designed as a medium-size data-processing system. The term "solid-state" refers to the use of "Ferractor" magnetic amplifiers and transistors (Fig. 31). The system consists of a central processor, a read-punch unit, a 450 card-per-minute card reader and a 600-lines-per-minute printer. The card equipment can be obtained for either the 80-

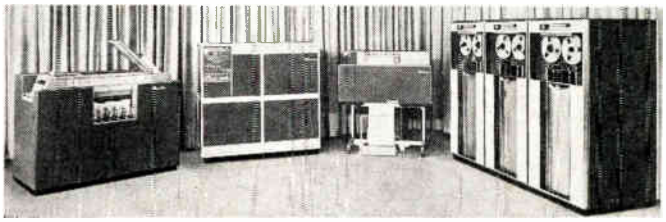


(Remington Rand UNIVAC Division, Sperry Rand Corporation)

Fig. 31—UNIVAC Solid-State Computer, with tapes.

(Control Data Corporation)

Fig. 32—Control Data 1604.



(International Business Machines Corporation)

Fig. 33—IBM 1401 System.

column or the 90-column punched-card systems. The first installation was made in January, 1960.

The system contains 215 tubes, 900 transistors and 36,500 diodes. The internal number system is biquinary-coded decimal. Word length is 10 decimal digits plus sign. 53 two-address, ten-decimal-digit instructions are used; (first address, operand; second address, next instruction). Arithmetic is performed with fixed point, in the synchronous mode, serially by digit, in parallel by bit, concurrently with other operations. Addition requires 1.36 msec, multiplication 1.275 msec, including storage access. The internal memory of the machine consists of one or more magnetic drums providing for 4000 words of storage with 1.7 msec average access and 1000 words with 0.425 msec average access. Up to 10 units of magnetic tape are available. The card-punching printer provides the ability to punch a card and print on both sides of the same card [142].

CDC 1604

The CDC 1604 is a general-purpose data-processing system manufactured by the Control Data Corporation (Fig. 32). The first installation was made in January, 1960.

The entire system includes some 100,000 diodes and 25,000 transistors. The internal number system is the binary, with a word length of 48 bits. There are 62 24-bit one-address instructions (2 per word) each including a six-bit operation code, a three-bit index and 15 bits for the address. Indirect addressing is built-in and six index registers are provided. Arithmetic is performed with fixed or floating point in the parallel synchronous mode concurrently with other operations. Addition requires 4.8 to 9.6 μsec, multiplication 25.2 to 63.6 μsec, including storage access. 32,768 words of magnetic-core storage are provided. Input-output equipment includes paper-tape typewriter, punched cards, magnetic tape (up to 24 units) and a 667–1000 lines-per-minute printer [142].

IBM 1401

The IBM 1401, an intermediate-size system, was installed initially in September, 1960, upon completion of a 3-year development program (Fig. 33). Employing transistor-diode logic, and a serial-by-character core memory, the 1401 has an 11.5-μsec character cycle. A word-mark bit, additional to the 1401's six-bit-plus-parity-bit alphanumeric character code, allows variable length data and instruction words. Add-to-storage logic replaces the conventional accumulator. The 1401's card reader-punch reads 800 cards per minute and punches up to 200 cards per minute, while the system's chain printer delivers 600 lines per minute [142].

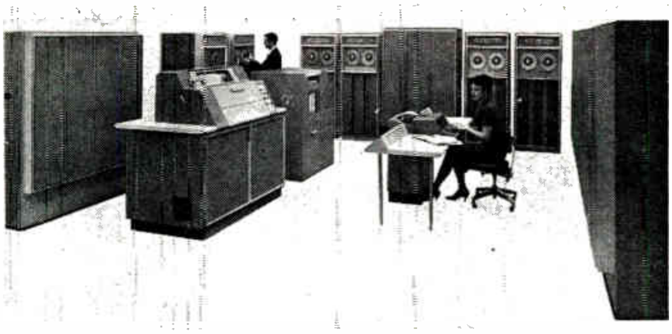
Honeywell 800

The Honeywell 800 is a general-purpose data-processing system capable of running as many as 8 distinct programs simultaneously without special instructions (Fig. 34). The first installation was made in December, 1960.

The system includes 30,000 diodes and 6000 transistors, excluding peripheral equipment. The internal number structure is binary and binary-coded decimal, with a word length of 48 bits, or 12 decimal digits. These 48 bits are assignable to numerical, alphanumerical or pure binary information. There are 59 basic instructions, each consisting of a twelve-bit operation code and 3 twelve-bit addresses. Eight index registers are available for each of the 8 programs which can be run concurrently. Other special-purpose registers are available. Arithmetic is performed in a synchronous parallel-serial-parallel mode, concurrently with other operations. Addition requires 24 μsec, multiplication 162 μsec, including storage access. Up to 32,000 words of magnetic-core storage can be used. Input-output equipment includes punched cards, paper tape and a 900 lines-per-minute printer. Up to 64 magnetic tape units can be connected to the system [142].

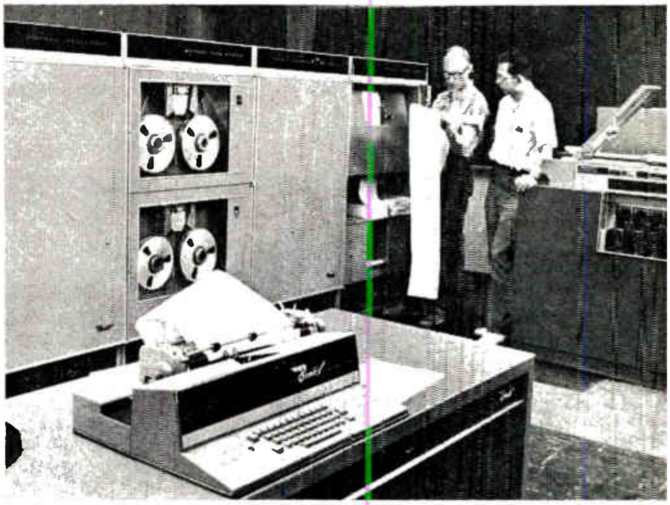
Bendix G-20

The Bendix G-20 is a general-purpose data-processing system of completely modular construction (Fig.



(Minneapolis Honeywell Regulator Company)

Fig. 34—Honeywell 800.



(Bendix Computer Division, Bendix Aviation Corporation)

Fig. 35—Bendix G-20 Computing System at Carnegie Institute of Technology.

35). The first installation was made in April, 1961, at the Humble Research Center, Houston, Tex., where the machine is used for oil-reservoir studies, including the solution of problems requiring eight to ten billion arithmetic operations.

The Bendix G-20 system typically contains 240 tubes, 38,000 diodes and 8900 transistors. The internal number system is binary, with a word length of 32 bits. One-address instructions are used, with provisions for indirect addressing. Arithmetic is performed with floating point in the parallel, synchronous mode, concurrently with other operations. Addition requires 13 μsec, multiplication, 56 μsec. 4096 to 32,768 words of magnetic-core storage are provided. Input-output is by paper tape, punched cards and magnetic tape (any number of units). Output includes also a 600-lines-per-minute printer [142].

Scribe

Digital computers primarily built to compute—to do common arithmetic—form a major part of the machinery used in the field. But there are rapidly increasing numbers of important digital data-processing machines that are not primarily "computers" in this sense. Examples are machine-tool control devices, such as those which control milling machines from information on tape; and "information-retrieval" machines, in which large files of indexed information (on magnetic tape or in some photographic form) can be systematically interrogated to supply designated subsets of the information stored. The Scribe test-scoring machines built in 1960 and 1961 by the Norden Division of United Aircraft Corporation for the Educational Testing Service form another example which is fairly typical of the class (Fig. 36).

(Norden Division, United Aircraft Corporation)

Fig. 36—The Scribe Scoring-Machine System.

The Scribe scoring-machine system consists of a paper handler in which pencil marks on $8\frac{1}{2}$ inch × 11 inch test "answer sheets" are sensed photoelectrically at the rate of 100 sheets per minute; a processor containing a magnetic-core memory, a magnetic drum and some 500 plug-in "cards" of transistor and diode circuits; and two output devices, an IBM type 514 card punch and an RCA type 581 magnetic-tape unit. The processing of up to 2280 coded marks, supplying up to 530 alphanumerical characters of several distinct kinds of information from each sheet, is performed serially, at a synchronous rate of 144 kilopulses per second, under the control of a program initially stored in the core memory. To score the "answers" on any one sheet, the machine can select as many as any six scoring "keys" from 24 such keys initially stored on the drum, supplying up to 54 part and total scores for each sheet. The system can modify its own program as may be required as the work proceeds [142].

Conclusion

Within the past fifteen years or so, digital computers have progressed from relays and rotating switches, as computing elements, to vacuum-tube and to semi-conductor logic circuits; from electromechanical, delay-line and cathode-ray-tube storage to magnetic-drum, magnetic-tape and magnetic-core storage. They have progressed from input-output speeds of one or two dozen characters per second to tens of thousands of characters per second; and from computation speeds of a few operations per second to thousands and tens of thousands of operations per second. Through the use of self-checking codes and by other means, they have reached high levels of reliability.

But the progress of the electronic computer art could not be so rapid if it did not concern more than machines. By programming its operation in some new way, almost every digital computer built has been found capable of doing more than it was originally designed to do. This has often led to significant improvements of the computing machinery. And then the process was repeated.

Early digital computers were built to solve a few important but relatively small scientific problems. Electronic digital computing machines and systems now at work

perform calculations for millions of pay checks,

bank accounts, insurance policies every day,

prepare weather forecasts,

solve scientific and engineering problems requiring from a few hundred to billions of arithmetic operations,

perform calculations for the design of almost every product of advanced technology,

translate English into Braille,

process data for the production, inventory control and transportation of millions of products,

perform the calculations needed in many diverse fields for the effective defense of the nation,

advance medical research by finding new patterns of diagnosis, by bringing new understanding,

make commercial airline passenger reservations,

perform calculations for satellite launching, orbiting and tracking,

keep records for billions of items of technical and patent information,

doing many other things as well, constantly accomplishing untold new tasks and demonstrating abundantly their usefulness to mankind.

Acknowledgment

The authors acknowledge with thanks the help of many colleagues and friends who generously supplied information and photographs for this paper.

Bibliography

References to papers and books are arranged chronologically (by year only) and alphabetically (by last name of first author) within each year.

[1] E. M. Horsburgh, Ed., "Modern Instruments and Methods of Calculation," G. Bell & Sons, Ltd., London, Eng., 343 pp.; 1914.

[2] J. A. V. Turck, "Origin of Modern Calculating Machines," The Western Society of Engineers, Chicago, Ill.; 1921.

[3] C. E. Shannon, "A symbolic analysis of relay and switching circuits," *Trans. AIEE*, vol. 57, pp. 713–723; 1938.

1946

[4] H. H. Aiken and G. M. Hopper, "The automatic sequence controlled calculator," *Elec. Engrg.*, vol. 65, pp. 384–391, August–September, 1946; pp. 449–454, October, 1946; pp. 522–528, November, 1946.

[5] A. W. Burks, H. H. Goldstine and J. von Neumann, "Preliminary Discussion of the Logical Design of an Electronic Computing Instrument," report prepared for the Research and Development Service, Ordnance Dept., U. S. Army; June, 1946.

[6] H. H. Goldstine and A. Goldstine, "The electronic numerical integrator and computer (ENIAC)," *Mathematical Tables and other Aids to Computation*, vol. 2, pp. 97–110; July, 1946.

[7] D. R. Hartree, "The ENIAC, an electronic computing machine," *Nature*, vol. 158, pp. 500–506; October 12, 1946.

[8] Staff of the Computation Laboratory of Harvard University, "A Manual of Operation for the Automatic Sequence Controlled Calculator," Annals of the Computation Laboratory, vol. 1, Harvard University Press, Cambridge, Mass., 561 pp.; 1946.

1947

[9] A. W. Burks, "Electronic computing circuits of the ENIAC," Proc. IRE, vol. 35, pp. 756–767; August, 1947.

[10] S. B. Williams, "A relay computer for general application," *Bell Lab. Record*, New York, N. Y., vol. 25, pp. 49–54; February, 1947.

1948

[11] J. G. Brainerd and T. K. Sharpless, "The ENIAC," *Elec. Engrg.*, vol. 67, pp. 163–172; February, 1948.

[12] W. J. Eckert, "The IBM pluggable sequence relay calculator," *Mathematical Tables & Other Aids to Computation*, vol. 3, pp. 149–161; July, 1948.

[13] F. J. Murray, "The Theory of Mathematical Machines," King's Crown Press, New York, N. Y., 139 pp.; 2nd ed.; 1948.

[14] "Proceedings of a Symposium on Large-Scale Digital Calculating Machinery," Annals of the Computation Laboratory of Harvard University, vol. 16; Harvard University Press, Cambridge, Mass., 302 pp.; 1948.

1949

[15] D. R. Brown and N. Rochester, "Rectifier networks for multiposition switching," Proc. IRE, vol. 37, pp. 139–147; February, 1949.

[16] D. R. Hartree, "Calculating Instruments and Machines," University of Illinois Press, Urbana, Ill., 138 pp.; 1949.

[17] G. W. Patterson, *et al.*, "A Functional Description of the EDVAC," The Moore School of Elec. Engrg., University of Pennsylvania, Philadelphia, Pa., Res. Div. Rept. 50-9; November, 1949.

[18] F. C. Williams and T. Kilburn, "A storage system for use with binary-digital computing machines," *Proc. IEE* (London, Eng.), pt. 3, vol. 96, pp. 81–100; March, 1949 (same paper in pt. 2, vol. 96, pp. 183–202; April, 1949).

1950

[19] A. Cohen, "Magnetic drum storage for digital information processing systems," *Mathematical Tables and Other Aids to Computation*, vol. 4, pp. 31–39; January, 1950.

[20] S. H. Dodd, H. Klemperer and P. Youtz, "Electrostatic storage tube," *Elec. Engrg.*, vol. 69, pp. 990–995; November, 1950.

[21] R. W. Hamming, "Error detecting and error correcting codes," *Bell Sys. Tech. J.*, vol. 29, pp. 147–160; April, 1950.

[22] M. Rees, "U. S. computing machine program," *Science*, vol. 112, pp. 731–736; December, 1950.

[23] R. F. Shaw, "Arithmetic operations in a binary computer," *Rev. Sci. Instr.*, vol. 21, pp. 687–693; August, 1950.

[24] Staff of Engineering Research Associates, Inc., "High-Speed Computing Devices," McGraw-Hill Book Co., Inc., New York, N. Y., 451 pp.; 1950.

[25] N. H. Taylor, "Marginal checking as an aid to computer reliability," Proc. IRE, vol. 38, pp. 1418–1421; December, 1950.

[26] A. Wang and W. D. Woo, "Static magnetic storage and delay line," *J. Appl. Phys.*, vol. 21, pp. 49–54; January, 1950.

1951

[27] J. P. Eckert, Jr., J. R. Weiner, H. F. Welsh and H. F. Mitchell, "The UNIVAC system," *Rev. of Electronic Digital Computers, Joint AIEE-IRE Computer Conf.*, Philadelphia, Pa., pp. 6–14; December 10–12, 1951.

[28] R. R. Everett, "The Whirlwind I Computer," *Rev. Electronic Digital Computers, Joint AIEE-IRE Computer Conf.*, Philadelphia, Pa., pp. 70–74; December 10–12, 1951.

[29] J. H. Felker, "The transistor as a digital computer component," *Rev. of Electronic Digital Computers, Joint AIEE-IRE Computer Conf.*, Philadelphia, Pa., pp. 105–108; December 10–12, 1951.

[30] J. W. Forrester, "Digital information storage in three dimensions using magnetic cores," *J. Appl. Phys.*, vol. 22, pp. 44–48; January, 1951.

[31] W. Keister, A. E. Ritchie and S. H. Washburn, "The Design of Switching Circuits," D. Van Nostrand Co., Inc. New York, N. Y., 556 pp.; 1951.

[32] Staff of the Computation Lab. of Harvard University, "Synthesis of Electronic Computing and Control Circuits," Annals of the Computation Lab., vol. 27, Harvard University Press, Cambridge, Mass., 278 pp.; 1951.

[33] N. H. Taylor, "Evaluation of the engineering aspects of Whirlwind I," *Rev. of Electronic Digital Computers, Joint AIEE-IRE Computer Conf.*, Philadelphia, Pa., pp. 75–78; December 10–12, 1951.

1952

[34] W. N. Papian, "A coincident-current magnetic memory cell for the storage of digital information," Proc. IRE, vol. 40, pp. 475–478; April, 1952.

[35] J. A. Rajchman, "Static magnetic matrix memory and switching circuits," *RCA Rev.*, vol. 13, pp. 183–201; June, 1952.

1953

[36] B. V. Bowden, Ed., "Faster than Thought," Pitman & Sons, Ltd., London, Eng., 416 pp.; 1953.

[37] W. Buchholz, "The system design of the IBM Type 701 Computer," Proc. IRE, vol. 41, pp. 1262–1275; October, 1953.

[38] L. P. Crosman, "The Remington Rand Type 409-2 Electronic Computer," Proc. IRE, vol. 41, pp. 1332–1340; October, 1953.

[39] J. P. Eckert, Jr., "A survey of digital computer memory systems," Proc. IRE, vol. 41, pp. 1393–1406; October, 1953.

[40] R. D. Elbourn and R. P. Witt, "Dynamic circuit techniques used in SEAC and DYSEAC," Proc. IRE, vol. 41, pp. 1380–1387; October, 1953.

[41] G. Estrin, "The Electronic Computer at the Institute for Advanced Study," *Mathematical Tables and other Aids to Computation*, vol. 7, pp. 108–114; April, 1953.

[42] C. E. Frizzell, "Engineering description of the IBM Type 701 Computer," Proc. IRE, vol. 41, pp. 1275–1287; October, 1953.

[43] S. E. Gluck, H. J. Gray, Jr., C. T. Leondes and M. Rubinoff, "The design of logical OR-AND-OR pyramids for digital computers," Proc. IRE, vol. 41, pp. 1388–1392; October, 1953.

[44] S. Greenwald, R. C. Haueter and S. N. Alexander, "SEAC," Proc. IRE, vol. 41, pp. 1300–1313; October, 1953.

[45] G. M. Hopper and J. W. Mauchly, "Influence of programming techniques on the design of computers," Proc. IRE, vol. 41, pp. 1250–1254; October, 1953.

[46] H. D. Huskey, R. Thorensen, B. F. Ambrosio and E. C. Yowell, "The SWAC—Design features and operating experience," Proc. IRE, vol. 41, pp. 1294–1299; October, 1953.

[47] G. W. King, G. W. Brown and L. N. Ridenour, "Photographic techniques for information storage," Proc. IRE, vol. 41, pp. 1421–1428; October, 1953.

[48] J. H. McGuigan, "Combined reading and writing on a magnetic drum," Proc. IRE, vol. 41, pp. 1438–1444; October, 1953.

[49] M. Palevsky, "The design of the Bendix Digital Differential Analyzer," Proc. IRE, vol. 41, pp. 1352–1356; October, 1953.

[50] W. N. Papian, "The MIT magnetic-core memory," *Proc. of the Eastern Joint Computer Conf.*, Washington, D. C.; December 8–10, 1953.

[51] J. A. Rajchman, "A myriabit magnetic-core matrix memory," Proc. IRE, vol. 41, pp. 1407–1421; October, 1953.

[52] H. D. Ross, Jr., "The arithmetic element of the IBM Type 701 Computer," Proc. IRE, vol. 41, pp. 1287–1294; October, 1953.

[53] A. L. Samuel, "Computing bit by bit, or digital computers made easy," Proc. IRE, vol. 41, pp. 1223–1230; October, 1953.

[54] R. Serrell, "Elements of Boolean Algebra for the study of information-handling systems," Proc. IRE, vol. 41, pp. 1366–1380; October, 1953.

[55] C. E. Shannon, "Computers and automata," Proc. IRE, vol. 41, pp. 1234–1241; October, 1953.

[56] C. E. Shannon and E. F. Moore, "Machine aid for switching circuit design," Proc. IRE, vol. 41, pp. 1348–1351; October, 1953.

[57] P. C. Sherertz, "Electronic circuits of the NAREC Computer," Proc. IRE, vol. 41, pp. 1313–1320; October, 1953.

[58] J. H. Vogelsong, "A transistor pulse amplifier using external regeneration," Proc. IRE, vol. 41, pp. 1444–1450; October, 1953.

[59] W. H. Ware, "The logical principles of a new kind of binary counter," Proc. IRE, vol. 41, pp. 1429–1437; October, 1953.

[60] D. J. Wheeler and J. E. Robertson, "Diagnostic Programs for the ILLIAC," Proc. IRE, vol. 41, pp. 1320–1325; October, 1953.

[61] G. S. White, "Coded decimal number systems for digital computers," Proc. IRE, vol. 41, pp. 1450–1452; October, 1953.

1954

[62] "Automatic digital computation." *Proc. Symp. held at the National Physical Laboratory*, March 25–28, 1953, Her Majesty's Stationery Office, London, Eng.; 1954.

[63] C. J. Bashe, W. Buchholz and N. Rochester, "The IBM Type 702, An electronic data processing machine for business," *J. Association for Computing Machinery*, vol. 1, pp. 149–169; October, 1954.

[64] C. C. Elgot, "On single vs triple address computing machines," *J. Association for Computing Machinery*, vol. 1, pp. 118–123; July, 1954.

[65] F. E. Hamilton and E. C. Kubie, "The IBM Magnetic Drum Calculator Type 650," *J. Association for Computing Machinery*, vol. 1, pp. 13–20; January, 1954.

[66] A. L. Leiner, W. A. Notz, J. L. Smith and A. Weinberger, "System design of the SEAC and DYSEAC," IRE Trans. on Electronic Computers, vol. EC-3, pp. 8–23; June, 1954.

1955

[67] W. J. Eckert and R. Jones, "Faster, Faster," International Business Machines Corporation, New York, N. Y., 160 pp.; 1955.

[68] B. Gilchrist, J. H. Pomerene and S. Y. Wong, "Fast carry logic for digital computers," IRE Trans. on Electronic Computers, vol. EC-4, pp. 133–136; December, 1955.

[69] J. A. Rajchman and A. W. Lo, "The Transfluxor—A magnetic gate with stored variable setting," *RCA Rev.*, vol. 16, pp. 303–311; June, 1955.

[70] R. K. Richards, "Arithmetic Operations in Digital Computers," D. Van Nostrand Co., Inc., New York, N. Y., 397 pp.; 1955.

1956

[71] J. R. Anderson, "A new type of ferroelectric shift register," IRE Trans. on Electronic Computers, vol. EC-5, pp. 184–191; December, 1956.

[72] M. M. Astrahan, H. F. Heath, Jr., R. A. Nienburg and L. R. Walters, "Reliability of an air defense computing system," IRE Trans. on Electronic Computers, vol. EC-5, pp. 224–237; December, 1956.

[73] D. A. Buck, "The Cryotron—A superconductive computer component," Proc. IRE, vol. 44, pp. 482–493; April, 1956.

[74] W. K. Halstead, J. W. Leas, J. N. Marshall and E. E. Minett, "Purpose and application of the RCA BIZMAC System," *Proc. Western Joint Computer Conf.*, San Francisco, Calif., pp. 119–123; February 7–9, 1956.

[75] "IRE Standards on Computers," Proc. IRE, vol. 44, pp. 1167–1173; September, 1956.

[76] G. J. Prom and R. L. Crosby, "Junction transistor switching circuits for high-speed digital computer applications," IRE Trans. on Electronic Computers, vol. EC-5, pp. 192–196; December, 1956.

[77] J. A. Rajchman and A. W. Lo, "The transfluxor," Proc. IRE, vol. 44, pp. 321–332; March, 1956.

[78] C. E. Shannon and J. McCarthy, Eds., "Automata Studies," Annals of Mathematics Studies, no. 34, Princeton University Press, Princeton, N. J., 285 pp.; 1956.

[79] A. Weinberger and J. L. Smith. "A one-microsecond adder using one-megacycle circuitry," IRE Trans. on Electronic Computers, vol. EC-5, pp. 65–73; June, 1956.

[80] M. V. Wilkes, "Automatic Digital Computers," J. Wiley & Sons, Inc., New York, N. Y., 305 pp.; 1956.

1957

[81] M. M. Astrahan, B. Housman, J. F. Jacobs, R. P. Mayer and W. H. Thomas, "Logical Design of the Digital Computer for the SAGE System," *IBM J. Res. & Dev.*, vol. 1, pp. 76–83; January, 1957.

[82] N. Chapin, "An Introduction to Automatic Computers," D. Van Nostrand Co., Inc., Princeton, N. J., 525 pp.; 1957.

[83] S. P. Frankel, "Logical design of a simple general-purpose computer," IRE Trans. on Electronic Computers, vol. EC-6, pp. 5–14; March, 1957.

[84] J. M. Frankovich and H. P. Peterson, "A functional description of the Lincoln TX-2 Computer," *Proc. Western Joint Computer Conf.*, Los Angeles, Calif., pp. 146–155; February 26–28, 1957.
[85] D. A. Huffman, "The design and use of hazard-free switching networks," *J. Association for Computing Machinery*, vol. 4, pp. 47–62; January, 1957.
[86] J. A. Rajchman, "Ferrite apertured plate for random-access memory," PROC. IRE, vol. 45, pp. 325–334; March, 1957.
[87] R. K. Richards, "Digital Computer Components and Circuits," D. Van Nostrand Co., Inc., Princeton, N. J., 523 pp.; 1957.
[88] E. L. Younker, "A transistor-driven magnetic-core memory," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-6, pp. 14–20; March, 1957.

1958

[89] F. L. Alt, "Electronic Digital Computers," Academic Press, New York, N. Y., 336 pp.; 1958.
[90] A. D. Booth and K. H. V. Booth, "Automatic Digital Calculators," Academic Press, Inc., New York, N. Y., 261 pp.; 1958.
[91] S. H. Caldwell, "Switching Circuits and Logical Design," John Wiley & Sons, Inc., New York, N. Y., 686 pp.; 1958.
[92] P. Dreyfus, "System Design of the Gamma-60," *Proc. Western Joint Computer Conf.*, Los Angeles, Calif., pp. 130–132; May 6–8, 1958.
[93] W. S. Humphrey, Jr., "Switching Circuits with Computer Applications," McGraw-Hill Book Co., Inc., New York, N. Y., 264 pp.; 1958.
[94] M. A. Maclean and D. Aspinall, "A decimal adder using a stored addition table," *Proc. IEE*, vol. 105B, pp. 129–135, 144–146; March, 1958.
[95] N. Metropolis and R. L. Ashenhurst, "Significant digit computer arithmetic," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 265–267; December, 1958.
[96] D. B. Netherwood, "Logical machine design: A selected bibliography," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 155–178; June, 1958.
[97] M. Phister, Jr., "Logical Design of Digital Computers," John Wiley & Sons, Inc., New York, N. Y., 408 pp.; 1958.
[98] G. E. Poorte and A. S. Kranzley, "The RCA 501 Electronic Data Processing System," *Proc. Western Joint Computer Conf.*, Los Angeles, Calif., pp. 66–70; May, 1958.
[99] M. Shimshoni, "An improved technique for fast multiplication on serial digital computers," *Elec. Engrg.*, vol. 30, pp. 504–505; August, 1958.
[100] A. K. Susskind, "Digital information processing for machine-tool control," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-7, pp. 136–140; June, 1958.

1959

[101] E. E. Bittman, "Thin-film memories," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-8, pp. 92–97; June, 1959.
[102] J. P. Eckert, J. C. Chu, A. B. Tonil and W. F. Schmitt, "Design of UNIVAC-LARC System I," *Proc. Eastern Joint Computer Conf.*, Boston, Mass., pp. 59–65; December 1–3, 1959.
[103] A. L. Leiner, W. A. Notz, J. L. Smith and A. Weinberger, "PILOT—A new multiple computer system," *J. Association for Computing Machinery*, vol. 6, pp. 313–335; July, 1959.
[104] "Microwave Techniques for Computing Systems," IRE TRANS. ON ELECTRONIC COMPUTERS," vol. EC-8, pp. 262–307; September, 1959.
[105] "Proceedings of an International Symposium on the Theory of Switching," Annals of the Computation Laboratory of Harvard University, pt. 1, vol. 29 and pt. 2, vol. 30, Harvard University Press, Cambridge, Mass., 305 and 345 pp.; 1959.
[106] J. A. Rajchman, "Magnetics for computers—a survey of the state of the art," *RCA Rev.*, vol. 20, pp. 92–135; March, 1959.

1960

[107] S. N. Alexander, P. Armer, M. M. Astrahan, L. Bers, H. H. Goode, H. D. Huskey and M. Rubinoff, "Soviet computer technology, 1959," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-9, pp. 72–120; March, 1960.
[108] I. Flores, "Computer Logic," Prentice-Hall, Inc., Englewood Cliffs, N. J., 458 pp.; 1960.
[109] T. W. Kampe, "The design of a general-purpose microprogram-controlled computer with elementary structure," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-9, pp. 208–213; June, 1960.
[110] R. S. Ledley, "Digital Computer and Control Engineering," McGraw-Hill Book Co., Inc., New York, N. Y., 835 pp.; 1960.
[111] A. T. Ling and K. Kozarsky, "The RCA 601 System Design" *Proc. Eastern Joint Computer Conf.*, New York, N. Y., pp. 173–177; December 13–15, 1960.
[112] G. W. Reitwiesner, "The determination of carry propagation length for binary arithmetic," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-9, pp. 35–38; March, 1960.

1961

[113] I. Abeyta, F. Borgini and D. R. Crosby, "A computer subsystem using kilomegacycle subharmonic oscillators," PROC. IRE, vol. 49, pp. 128–135; January, 1961.
[114] F. L. Alt, Ed., "Advances in Computers," Academic Press, New York, N. Y., vol. 1, 316 pp., 1960; vol. 2, 434 pp., 1961.
[115] I. L. Auerbach, "European electronic data processing—A report on the industry and the state of the art," PROC. IRE, vol. 49, pp. 330–348; January, 1961.
[116] A. Avizienis, "Signed-digit number representations for fast parallel arithmetic," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-10, pp. 389–400; September, 1961.
[117] T. C. Bartee, "Computer design of multiple-output logical networks," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-10, pp. 21–30; March, 1961.
[118] F. S. Beckman, F. P. Brooks, Jr., and W. J. Lawless, Jr., "Developments in the logical organization of computer arithmetic and control units," PROC. IRE, vol. 49 pp. 53–66; January, 1961.
[119] W. G. Brown, J. Tierney and R. Wasserman, "Improvement of Electronic Computer Reliability," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-10, pp. 407–416; September, 1961.
[120] H. A. Curtis, "A generalized tree circuit," *J. Association for Computing Machinery*, vol. 8, pp. 484–496; October, 1961.
[121] V. P. Czapla, "The inventor of the first desk calculator," *Computers and Automation*, vol. 10, p. 6; September, 1961.
[122] C. V. Freiman, "Statistical analysis of certain binary division algorithms," PROC. IRE, vol. 49, pp. 91–103; January, 1961.
[123] M. K. Haynes, "Transient analysis of cryotron networks by computer simulation," PROC. IRE, vol. 49, pp. 245–257; January, 1961.
[124] A. S. Hoagland and G. C. Bacon, "High-density digital magnetic recording techniques," PROC. IRE, vol. 49, pp. 258–267; January, 1961.
[125] H. R. Kaupp and D. R. Crosby, "Calculated waveforms for tunnel diode locked pair," PROC. IRE, vol. 49, pp. 146–154; January, 1961.
[126] F. Leary, "Computers today," *Electronics*, vol. 34, pp. 64–94; April 28, 1961.
[127] T. B. Lewis, "Digital computer equipment for an advanced bombing, navigation and missile guidance subsystem for the B-70 air vehicle," PROC. IRE, vol. 49, pp. 313–318; January, 1961.
[128] A. W. Lo, "Some thoughts on digital components and circuit techniques," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-10, pp. 416–425; September, 1961.
[129] R. T. Loewe, R. L. Sisson and P. Horowitz, "Computer generated displays," PROC. IRE, vol. 49, pp. 185–195; January, 1961.
[130] P. R. Low and G. A. Maley, "Flow table logic," PROC. IRE, vol. 49, pp. 221–228; January, 1961.
[131] O. L. MacSorley, "High-speed arithmetic in binary computers," PROC. IRE, vol. 49, pp. 67–91; January, 1961.
[132] F. J. Murray, "Digital computers," in "Mathematical Machines," Columbia Univ. Press, N. Y., vol. 1, 300 pp.; 1961.
[133] E. Nussbaum, E. A. Irland and C. E. Young, "Statistical analysis of logic circuit performance in digital systems," PROC. IRE, vol. 49, pp. 236–244; January, 1961.
[134] G. Ott and N. H. Feinstein, "Design of sequential machines from their regular expressions," *J. Association for Computing Machinery*, vol. 8, pp. 585–600; October, 1961.
[135] R. T. Pearson, "The development of the flexible-disk magnetic recorder," PROC. IRE, vol. 49, pp. 164–174; January, 1961.
[136] W. W. Peterson and D. T. Brown, "Cyclic codes for error detection," PROC. IRE, vol. 49, pp. 228–235; January, 1961.
[137] N. S. Prywes, H. Lukoff and J. Schwarz, "UNIVAC-LARC high-speed circuitry: case history in circuit optimization," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-10, pp. 426–438; September, 1961.
[138] J. I. Raffel, T. S. Crowther, A. H. Anderson and T. O. Herndon, "Magnetic film memory design," PROC. IRE, vol. 49, pp. 155–164; January, 1961.
[139] J. A. Rajchman, "Computer memories: a survey of the state of the art," PROC. IRE, vol. 49, pp. 104–127; January, 1961.
[140] R. C. Sims, E. R. Beck, Jr., and V. C. Kamm, "A survey of tunnel-diode digital techniques," PROC. IRE, vol. 49, pp. 136–146; January, 1961.
[141] O. B. Stram, "Arbitrary Boolean functions of N variables realizable in terms of threshold devices," PROC. IRE, vol. 49, pp. 210–220; January, 1961.
[142] M. H. Weik, "A Third Survey of Domestic Electronic Digital Computing Systems," Ballistic Research Lab., Aberdeen Proving Ground, Md., Rept. No. 1115, 1131 pp.; 1961.

The Evolution of Concepts and Languages of Computing*

R. D. ELBOURN†, MEMBER, IRE, AND W. H. WARE‡, FELLOW, IRE

Summary—**Digital computers are opening exciting new possibilities for progress in language translation, information retrieval, psychological modeling, problem solving and theorem proving. These have resulted not because of the microsecond arithmetic speed but because of the ability to manipulate symbols: to read, write, store, compare, and replace symbols and to follow different courses of action according to differences between symbols. Thus, language in a general sense is a common aspect of these new applications.**

Programmers have been extending the usefulness of computers through the evolutionary development of most artfully conceived languages. Recently, mathematicians and logicians have been proving theorems about formal languages, while linguists have been discovering laws that humans instinctively observe whenever they use natural language. Fruitful cross pollination among such endeavors now promises greatly accelerated progress in determining whether symbol manipulation is for information processing applications what numerical analysis is for arithmetic applications. This paper first reviews the evolution of programming languages from the early days when all programming was done in machine languages, through symbolic coding systems, interpreters, assemblers, generators, and compilers, to the recently developed list processing languages. Then the applications of these languages to game playing, problem solving, theorem proving, and behavior and biological modeling have been described briefly. Finally, in anticipation of extending the capability of computers to accept, use and generate natural languages, the paper concludes with an introduction to some of the contemporary work on formal language theory, including a discussion of six families of abstract languages and their practical implementation.

INTRODUCTION

CLASSICAL numerical analysis, as developed during the 19th century and the first half of the 20th century, replaces the infinite processes of the calculus by a finite, but very large, number of additions, subtractions, multiplications and divisions. It was for the purpose of making such large numbers of arithmetic operations practical that electronic computers were originally conceived. When the first stored-program computers were put into operation in about 1950, the evolutionary use of computers to mechanize quite different kinds of procedures began.

The drudgery of programming and the manifold opportunities for clerical blunders stimulated the development of automatic coding, which is a scheme to make the computer itself take over the drudgery. The successive development of symbolic assembly routines, interpreters, assemblers, generators, and compilers clearly demonstrated that one is not limited to the basic machine language. Suitable translating routines permit a variety of input and output languages to be used.

Experimenters who undertook to program a computer to play a game, to use heuristic procedures to solve problems, or to simulate human behavior found that such problems required specialized languages. Hence, the list-processing languages—IPL-V, LISP, FLPL and COMIT—were invented. For quite general purposes these free the programmer from machine language. More important, such list languages show the automatic computer in the role of a syntactic processor which, in addition to performing arithmetic, manipulates strings of symbols according to finite sets of rules.

Is there a science of symbol manipulation as numerical analysis is the science of arithmetic operations? We see the start of it in descriptive linguistics which is the science that searches for the laws which humans observe naturally as they use natural language. A few descriptive linguists, who also have a background in mathematical or symbolic logic, have formulated models of linguistic structure as abstract mathematical objects and have proved theorems about them. We find that much of natural language and also many artificial languages can be defined precisely by such formal grammars. We conjecture that the newly developing theory of abstract language models will be fundamental for both the applications and the logical design of computers. Therefore, we devote Section I of this paper to a brief review of the evolution of programming languages with an eye toward their description in abstract terms; in Section II we give a glimpse of the abstract theory of formal languages in the hope that readers will pursue further information in the literature.

I. EVOLUTION AND USES OF PROGRAMMING LANGUAGES

In the early days of the computer field, all programming was done in machine language.[1,2] This implies not only that the programmer was confined to only those kinds of instructions which the designers of the machine had built into its comprehension, but also that he had to work in the numeric symbology which the machine accepted. For example, the symbol "Add" was not an acceptable one to the machine. The machine would respond to only such numeric symbols as "010," the machine language equivalent of "Add". Furthermore, the store of the machine was regarded by the pro-

* Received by the IRE, August 14, 1961; revised manuscript received December 22, 1961.
† National Bureau of Standards, Washington, D. C.
‡ The RAND Corporation, Santa Monica, Calif.

[1] W. Orchard-Hays, "The evolution of programming systems," PROC. IRE, vol. 49, pp. 283–295; January, 1961.
[2] J. W. Carr, III, "Programming and coding," in "Handbook of Automation, Computation and Control," vol. 2, ch. 2, John Wiley and Sons, Inc., New York, N. Y.; 1959.

grammer as a collection of individual cells. He was completely responsible for assigning information to particular storage locations, and for being continuously cognizant of what information occupied each cell. The only way to find information in storage was to know the address of the cell which contained it. As an additional burden, the programmer used the same machine addresses in his instructions as the machine would later use to locate information for itself in the store. There was no means, for instance, to specify the location of one piece of information relative to the location of another piece.

The first relief from these problems came with symbolic coding. In such a programming language, the machine accepted such symbols as "Add" or "Multiply", or mnemonic abbreviations. The internal machine structure, however, did not change. With a simple table of correspondence, the machine could be used to transform one symbol into another ("Add" into "010"). Furthermore, the programmer no longer was bound to machine addresses. He could regard the store as divided into regions, each of which had its own set of addresses. If desired, positions within each region could be specified relative to the first location of the region. The first position in each region could also be specified in a symbolic form, rather than as an absolute machine address. A machine routine stated in such a pseudo-language must be translated into machine language prior to execution of the routine by the machine.

The routine which accepts a set of symbolically-coded instructions, translates them into machine language, and at the same time also assigns either symbolic or regional addresses to absolute machine addresses is generally called an *assembly routine*. Usually, the instructions with which an assembler deals are actual machine instructions although expressed in symbolic form. An assembler may allow for macro-instructions, which are kinds of instructions which the machine cannot directly execute but which must be expanded into a longer set of basic machine instructions. There may also be pseudo-instructions which are used to communicate information to the assembler about the routine on which it is operating. Thus, the assembler helps with the symbology problem confronting the programmer, expands the number of apparent operations available from the machine, and eases the chore of assigning storage locations.

An extension of the assembly technique is the *compiling routine*. Although terminology is not very precise, generally a compiler permits more complex macro-instructions than an assembler, and it often excludes machine instructions (even in symbolic form) from the language which it can accept. While the assembler generally deals with each instruction independently of all others, the compiler attempts to capitalize on the information which is contained in the structure or logic of the problem; the context in which each instruction is imbedded is important. Commonly, a compiler is language- or problem-oriented in the sense that it accepts as input the language and operations of a particular class of problems. A compiler usually eases the chore of arranging information in the store.

In the language which the compiler accepts, the programmer can refer to information by its name, rather than having to specify its location in storage. For example, the FORTRAN compiler accepts the symbology and the operations of ordinary algebra; the compiler can accept such symbols as ·, +, ( ), a superscript, a subscript, and letters designating algebraic variables. Sometimes it is possible to describe collectively a large class of routines and to embody this description in another routine called a *generator routine*. For example, a *sort generator* essentially contains a description of the process of sorting. Given information about the items to be sorted, the particular information on which the items are to be sorted, etc., the generator can construct a machine-language routine which will complete the sort.

Automatic Programming

The artificial languages represented by the assembler, compiler, and generator are really automatic coding languages, although they are commonly called automatic programming languages.[3] Building the translating routine for such a language is simply the problem of constructing a routine which makes a given physical sequential machine imitate the behavior of a nonexistent, but still sequential, machine whose language is the particular automatic coding language. The ideas from the theory of sequential machines and the associated theories of formal languages are becoming increasingly important in the construction of coding languages.

There are many coding languages (with their compilers) in existence. For example, FORTRAN[4] is a language for problems which can be stated in algebraic form. The symbols of the language are the symbols of algebra; the syntax of the language is the set of algebraic postulates and rules which leads to properly-written algebraic expressions. A more sophisticated language is ALGOL,[5] the international algorithmic language. COBOL[6] is a language for business problems; its symbols are the symbols of business transactions, and its syntax the set of rules which lead to well-formed expressions which describe business transactions. Other compilers, some of them related more or less to the

[3] The distinction between programming and coding is not always made. Often programming is considered to be that part of problem-solving which starts with formulation of the problem and exhibits the solution of a problem as an information flow diagram. Coding transforms the information flow diagram into a routine for the machine. For such problems as algebraic ones, the mathematical statement of the problem essentially describes the information flow process, and virtually no programming phase exists.

[4] D. D. McCracken, "A Guide to FORTRAN Programming," John Wiley and Sons, Inc., New York, N. Y.; 1961.

[5] P. Naur, *et al.*, "Report on the algorithmic language ALGOL-60," *Commun. ACM*, vol. 3, pp. 299–314; May, 1960.

[6] "COBOL, Initial Specifications," Dept. of Defense, U. S. Government Printing Office, Washington, D. C.

ALGOL language, carry such names as CLIP, JOVIAL, BIOR, MATHEMATIC, MYSTIC, SHADOW III, MADCAP, ALGO, FACT, and FLOWMATIC.[7]

Often the transition from the coding language to the machine language is made in two steps: first, an expansion of the original language to a sequence of machine operations but still in symbolic form; and second, the translation of the symbolic instructions into machine language. The intermediate language could, in principle be formulated in such a way that it would be machine-independent; only the translation from the intermediate language to machine language would depend on details of a particular machine. One proposal for a universal computer-oriented intermediate language is UNCOL.[8]

As the development of coding languages progresses, we can look forward to a large assortment of problem-oriented languages. A problem stated in any one of them will first be expanded into a corresponding statement in the intermediate language, from which a translation can be made to an appropriate machine language. Exchange of routines among users of the same kind of machine has long been practiced; we are beginning to see the exchange of routines (via the problem-oriented languages) among users of different kinds of machines.

List Languages[9,10]

In some very complex problems the allocation of storage space, even with sophisticated automatic coding languages, can be a major difficulty. Generally, this occurs because it is very difficult or impossible to foresee how much information will be produced or have to be stored at each phase of the routine. The concept of *list stores* and *list processing languages* was developed to surmount such difficulties.

"List" is used in the conventional sense to designate a linear sequence of pieces of information which, for some reason, are to be associated. One way to visualize a list store is to think of the individual cells of a conventional store as chained together; they need not be adjacent cells. For example, a given cell might contain not only an item of information, but also information which identifies the location of the next member on the list. Although no large list stores have been built to date, conventional stores can be programmed to behave as such. In a list store the location of information is in terms of the name of the list which contains it, but its position on the list may not be stated. The length of such a list is not necessarily fixed, which implies that there must exist a capability for inserting or deleting an entry anywhere along the list. Such a list is frequently called *push-down*, if the only entry point is at the top. An entry on a list may be the name of a *sublist*, the sublist in turn may reference another sublist, etc. Such a collection of lists and sublists is called a *list-structure*, and its name may itself be an entry on some different list.

List-processing languages permit such operations as the following: 1) insert an entry on a list, 2) delete an entry from a list, 3) create a list, 4) destroy a list, 5) coalesce two lists, and 6) search a list for a given symbol. The problems for which list-processing languages are powerful tend to involve the manipulation of symbols rather than calculation with arithmetic processes. Hence, list-processing languages are sometimes called *symbol-manipulating languages* and obviously must contain such operations as identify a symbol, compare two symbols, define a symbol.

List-processing languages include LISP,[11] FLPL,[12] and the family of Information Processing Languages (IPL-V being the best known).[13] Such languages facilitate storage allocation and give complete freedom for a subroutine to use itself one or more times in executing itself (i.e., recursively).

Artificial Intelligence[10,14]

A rapidly enlarging field of research is that of modeling the higher mental capabilities of man. One group of models, by means of a computer and an appropriate routine, successfully exhibit certain problem-solving abilities[15] which a man appears to use in such situations as proving mathematical theorems[16,17] or playing certain games (*e.g.*, chess).[18] Other models describe various aspects of human behavior and are useful for psychological research.

Routines which exhibit artificial intelligence or simulate natural intelligence are extremely complex; list languages have proved very useful in their construction. Generally speaking, these routines deal only with symbols or strings of symbols (expressions); they must recognize differences between symbols or between expressions, and usually contain a collection of methods for reducing such differences. For example, to prove a

[7] *Commun. ACM*, (complete issue) vol. 4; January, 1961.

[8] T. B. Steel, "A first version of UNCOL," *Proc. WJCC*, vol. 19, pp. 371–378; May, 1961.

[9] B. F. Green, "Computer languages for symbol manipulation," IRE Trans. on Human Factors in Electronics, vol. HFE-2, pp. 3–8; March, 1961.

[10] M. Minsky, "Steps toward artificial intelligence," Proc. IRE, vol. 49, pp. 8–30; January, 1961.

[11] J. McCarthy, "Recursive functions of symbolic expressions and their computations by machine," *Commun. ACM*, vol. 3, pp. 184–195; April, 1960.

[12] H. Gelernter, *et al.*, "A FORTRAN-compiled list-processing language," *J. ACM*, vol. 7, pp. 87–101; April, 1960.

[13] A. Newell and F. Tonge, "An introduction to information processing language V," *Commun. ACM*, vol. 3, pp. 205–211; April, 1960.

[14] M. Minsky, "A selected descriptor-indexed bibliography to the literature on artificial intelligence," IRE Trans. on Human Factors in Electronics, vol. HFE-2, pp. 39–55; March, 1961.

[15] W. R. Reitman, "Programming intelligent problem solvers," IRE Trans. on Human Factors in Electronics, vol. HFE-2, pp. 26–33; March, 1961.

[16] A. Newell and H. A. Simon, "The logic theory machine," IRE Trans. on Information Teory, vol. IT-2, pp. 61–79; September, 1956.

[17] H. L. Gelernter, "Realization of a geometry-proving machine," *Proc. Internatl. Conf. on Information Processing*, UNESCO House, Paris, France, pp. 273–282; 1959.

[18] A. Newell, J. C. Shaw and H. A. Simon, "Chess playing programs and the problem of complexity," *IBM J. Res. & Dev.*, vol. 2, pp. 320–335; October, 1958.

theorem the routine assesses the difference between the beginning point and some intermediate or final step. By means of substitutions or by the application of already proved theorems, the routine attempts to reduce the difference. Problem-solving routines generally contain a collection of techniques for planning the attack on the problem and criteria for determining success. They often contain a form of learning-from-experience so that on future problems the routine will be more effective.

Other models simulate some aspect of the internal behavior of living systems.[19] For example, several models have been built which attempt to imitate the self-organizing ability of networks of neurons.[20] The research in various phases of modeling of intelligent behavior is perhaps the most exciting, challenging and significant work in the field of information-processing routines and machine development. This research may well suggest completely new kinds of machine organization and structure.

Natural Languages

There are many kinds of information-processing problems which are most appropriately stated in a natural language. For example, in a system whose purpose is to receive questions and attempt to answer them from a large file of information (the information-retrieval or library problem), it is most convenient for the user to frame the question in the natural language to which he is accustomed. Whether or not the internal file of information is structured in a natural language, the machine routine will have to deal with the letters, words and syntax of the language used to frame the question. One artificial language especially designed for processing a natural language is COMIT.[21]

Research in many kinds of machine processing of spoken or written language is in progress. The goals of such work include machine recognition of spoken[22] or written[23,24] statements, machine generation of properly formed spoken or written statements, machine translations of one natural language into another,[25] machine generation of abstracts,[26] and machine indexing[27] of files of information. Eventually, in conjunction with appropriate problem-solving routines, it is to be expected that machines will be able to fill the role of intelligent technicians.

II. Abstract Language Models[28]

Language, interpreted broadly to include artificial and formal languages, as well as natural languages and all their problems of definition, generation, recognition and translation, is the best framework we have found for understanding what an automatic computer is functionally and what it can do. In support of this thesis, the evolution of computer programming has been reviewed, beginning with the original basic machine languages and continuing through symbolic assembly systems; interpreters; assemblers; generators; compilers of algebraic, algorithmic, and business oriented languages such as FORTRAN, ALGOL, and COBOL; and the recently-developed list-processing languages LISP, FLPL, IPL-V and COMIT.

Next, the excitingly promising applications of these languages to game playing, problem solving, theorem proving, modeling of behavior for psychological research, and biological modeling have been surveyed. In practically all cases the function of the computer is fundamentally syntactic rather than arithmetic; it is manipulating lists or strings of symbols according to rules of structure. Its basic operations are reading, writing, storing, comparing and replacing symbols, and following different courses of action according to differences between symbols. Perhaps the ultimate achievement in this chain of development will be in providing the computer with the ability to accept, use, and generate natural language.

On this premise, therefore, a most important area of research for advancing the uses of computers is the study of language in the abstract, the inquiry into the forms that a language may take and the properties of formal models for the definition of a language, or the generation, recognition and translation of sentences of a language. Such a theory is now being constructed. It has foundation in the efforts of descriptive linguists to determine mechanically the syntactic structure within natural languages[29] and of mathematical logicians to create artificial languages for mathematics.[30,31] In the latter case, it borrows somewhat from recursive function theory.[32,33] Already the theorists are finding significant

[19] J. K. Hawkins, "Self-organizing systems—a review and commentary," Proc. IRE, vol. 49, pp. 31–48; January, 1961.

[20] F. Rosenblatt, "Perceptron simulation experiments," Proc. IRE, vol. 48, pp. 301–309; March, 1960.

[21] V. Yngve, "A programming language for mechanical translation," *Mech. Translation*, vol. 5, pp. 25–41; July, 1958.

[22] Thomas Marill, "Automatic recognition of speech," IRE Trans. on Human Factors in Electronics, vol. HFE-2, pp. 34–38; March, 1961.

[23] O. Selfridge and U. Neisser, "Pattern recognition by machine," *Sci. Am.*, vol. 203, pp. 60–68; August, 1960.

[24] E. E. David, Jr., and O. G. Selfridge, "Eyes and ears for computers," this issue, pp. 1093–1101.

[25] Y. Bar-Hillel, "A present status of automatic translation of languages," in "Advances in Computers," F. Alt, Ed., Academic Press, Inc., New York, N. Y.; 1960.

[26] H. P. Luhn, "Automatic creation of literature abstracts," *IBM J. Res. & Dev.*, vol. 2, pp. 159–165; April, 1958.

[27] H. P. Luhn, "The automatic derivation of information retrieval encodements from machine-readable tests," in "Information Retrieval and Machine Translation," A. Kent, Ed., pt. 2, pp. 1021–1028, Interscience Publishing Co., New York, N. Y.; 1961.

[28] The assistance of Miss M. Fox, R. A. Kirsch and J. H. Wegstein in reviewing this section is gratefully acknowledged.

[29] Y. Bar-Hillel, C. Gaifman and E. Shamir, "On categorical and phrase structure grammars," *Bull. Res. Council Israel*, vol. 9F, pp. 1–16; June, 1960.

[30] S. C. Kleene, "Introduction to Metamathematics," D. Van Nostrand Co., Inc., New York, N. Y.; 1952.

[31] P. C. Rosenbloom, "The Elements of Mathematical Logic," Dover Publications, Inc., New York, N. Y.; 1958.

[32] M. Davis, "Computability and Unsolvability," McGraw-Hill Book Co., Inc., New York, N. Y.; 1958.

[33] R. M. Smullyan, "Theory of formal systems," "Annals of Mathematics Studies," no. 47, Princeton University Press, Princeton, N. J.; 1961.

problems within programming languages,[34] and programmers are borrowing techniques from the formal theories of languages and automata.[35] Because we confidently expect many more useful results to come out of this research, we have made the final section of this paper an introduction to some of the contemporary work on formal language theory.

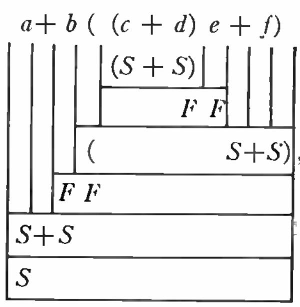
Simple Phrase-Structure Languages

The family called *simple phrase-structure languages* includes most systems of mathematical notation, most of the artificial languages used with computers, and quite possibly natural languages, so that it is not surprising that it has been the most extensively studied of the various abstract models of language.

Ordinary algebraic notation is a simple phrase-structure language. The algebraic expression

$$a + b((c + d)e + f)$$

can be analyzed into phrases as follows:



where F and S denote two kinds of phrases, "factor" and "sentence". Working from the bottom upward, each phrase can be regarded as produced by replacing a symbol F or S with other symbols according to one of six rewriting rules, called *productions:*

$$S \rightarrow S + S \qquad F \rightarrow (S + S)$$
$$S \rightarrow FF \qquad F \rightarrow FF$$
$$S \rightarrow x \qquad F \rightarrow x$$

where x denotes any variable $a, b, c \cdots$. It will be noted that beginning with the symbol S one can produce via these productions any correctly formed expression of this type, *i.e.*, any algebraic expression in which the operations of addition and multiplication are applied any number of times to any number of variables, and parentheses are inserted only where necessary for correct interpretation. Moreover, one cannot produce in this way any improperly formed expression such as $(a(b+)$. One can have the complete expression enclosed in explicit boundary markers $\# \cdots \#$, say for machine purposes, by starting with the production $B \rightarrow \#S\#$.

[34] S. Ginsburg and H. G. Rice, "Two Families of Languages Related to ALGOL," System Development Corp., Santa Monica, Calif., Tech. Memo. (TM Ser.) TM-578; January, 1961.

[35] E. T. Irons, "A syntax directed compiler for ALGOL 60," *Commun. ACM*, vol. 4, pp. 51–55; January, 1961.

These algebraic expressions are an example that satisfies the definition: a (formal) *language* is a distinguished (usually infinite) subset of all the possible finite *strings* (sequences of occurrences) of symbols from a set called the *alphabet.* If a string belongs to a language, we may call it a *sentence* of that language. A *grammar* of a language is a finite set of rules by which all and only the sentences of that language can be generated. *Simple phrase-structure grammars,* like the example given above, consist of the rules of the form $A \rightarrow \omega$, meaning rewrite the symbol A as the string ω.

The most general type of phrase-structure grammar may contain rules by which a symbol can be rewritten only if it occurs in a specified context, *i.e.*, $\zeta A \psi \rightarrow \zeta \omega \psi$ where ζ, ψ, and ω are specified strings. A symbol that can be rewritten under some rule is a *nonterminal* symbol; the other symbols in the alphabet are *terminal. A simple-phrase-structure language* is the set of all strings of terminal symbols that can be produced by starting with one specified nonterminal symbol and continuing to apply any applicable rule of a simple phrase-structure grammar to any occurrence of a nonterminal symbol until only terminal symbols remain. A *general phrase-structure language* is defined similarly using a phrase-structure grammar.

A sentence of a simple phrase-structure language can be produced in the normal left-to-right sequence as follows:

Output produced	Symbols in storage
	S
	$S + S$
$a +$	S
$a +$	$F\ F$
$a + b$	F
$a + b\ ($	$S + S\ )$
$a + b\ ($	$F\ F + S\ )$
$a + b\ (\ ($	$S + S\ )\ F + S\ )$,
⋮	
$a + b\ (\ (\ c + d\ )\ e + f\ )$	

where the convention is that at each step the left-most nonterminal symbol is rewritten.

Notice the properties of the storage used for nonterminal symbols. It behaves like the spring-supported stack of dishes in a cafeteria counter, which pushes down as dishes are added and pops up again when a dish is removed, so that the top dish is always accessible at the level of the counter top. A symbol storage with such properties is called a *push-down store.*

All of the list-processing languages employ push-down stores to hold the return address and the parameters used by a subroutine. If during the execution of a first subroutine a second subroutine is entered, the return address and parameters of the second are simply pushed down on top of those of the first. The latter are thus preserved and pop up again upon completion of the second routine and return to the first. If the push-down store has sufficient capacity, subroutines can be nested to any depth and one routine can even use itself as a subroutine any number of times. This fact corresponds exactly with the use of a push-down store in a simple phrase-structure language where phrases may be expanded within phrases to any depth.

In an application to natural language the various types of phrases denoted by the nonterminal symbols might be sentence, independent clause, noun phrase, verb phrase, singular particle, and so on, through a list of syntactic constituents of a sentence. The terminal symbols can be words or punctuation marks.

The idea that any nonterminal symbol can be considered the name of the set of strings (phrases) producible by starting with that symbol suggests an alternative way of defining simple phrase-structure languages. One can start with certain sets of symbols or strings that are finite sets or are otherwise effectively definable (which shall subsequently be defined precisely and called "decidable") and construct new sets according to the following two definitions in which A and B are any sets of strings:

Union: $A \cup B$, every string in A or in B.[36]

Concatenation: AB, every string $\zeta\psi$ where ζ is in A and ψ is in B. This alternative method would replace the six productions for the algebraic expressions by two equations:

$$S = x \cup FF \cup S + S$$
$$F = x \cup FF \cup (S + S).$$

The first equation means that a string of the set S is a single symbol of the set $a, b, c, \cdots$, or the concatenation of two strings from F, or a string from S followed by "+" followed by another string from S. This definition is constructive in that one initially assumes only the members of the alphabet x are in S and F; then one adds strings to S and to F by repeated use of the implicit defining equations.

The family of the sets that can be defined in this constructive sense by a system of simultaneous implicit equations in unions of concatenations has been shown to be exactly the simple phrase-structure languages.[29,34] ALGOL is thus a simple phrase-structure language, because its official definition[5] uses equations of just this kind; only the notation is slightly different.

A system of defining equations is rather more convenient to use (say, in a machine implementation) if it has the property that for each n the nth equation contains none but the first n variables. Such systems yield a proper subfamily called the *sequentially definable languages:*

$$A, B, C, \cdots \text{ are finite sets}$$
$$X = F_1(X, A, B, C \cdots)$$
$$Y = F_2(X, Y, A, B, C, \cdots)$$
$$Z = F_3(X, Y, Z, A, B, C, \cdots).$$

The identifiers and labels of ALGOL are sequentially definable, while the arithmetic expressions are not.

We will now introduce two more compositions of sets:

Power: $A^n = AA^{n-1}$, where $A^1 = A$ and A^0 is the set that contains only the string having no symbols.

Note that A^n is just A concatenated with itself n times. Power is introduced here only as a tool for the following definition:

Closure:

$$A^* = \bigcup_{n=0}^{\infty} A^n = A^0 \cup AA^*.$$

Closure is thus the first composition of finite sets that produces an infinite set. The two forms of definitions are equivalent, but the second exhibits its construction by iteration.

The structures of the identifiers I, unsigned integers N, and labels A in ALGOL are shown most clearly by the expressions

$$I = L(L \cup D)^*$$
$$N = DD^*$$
$$A = I \cup N = L(L \cup D)^* \cup DD^*$$

where L and D are finite sets of letters and digits, respectively. The meaning of the first equation is that an identifier may be any letter followed by any succession, including none at all, of letters and digits. The expressions that can be written using symbols denoting finite sets and indicating finite numbers of operations of concatenation, union, and closure are called *regular expressions:*

$$D^* = D^0 \cup DD^*$$
$$X^* = X^0 \cup (L \cup D)X^*$$
$$A = LX^* \cup DD^*.$$

The sets denoted by regular expressions are a subfamily of the sequentially definable sets, since each occurrence of the closure operation can be made the symbol for a set, say A^*, which is defined by an equation of the form $A^* = A^0 \cup AA^*$ as in the definition of closure.

Automata

Often an effective procedure for 1) generating the sentences of a language, 2) translating the sentences of one language into corresponding sentences in another, or 3) deciding whether or not any given string is a sentence of a given language can be defined by describing an automaton that will carry out the procedure.

[36] Other notations used are $A \vee B$ and $A + B$.

That is to say, an automaton may be a convenient form of either a *grammar*, a *translator*, or a *recognition grammar* for a language, as described in the following discussion.

Such an automaton can receive a sequence of symbols as input and responds by producing another sequence of symbols as output. This concept of an automaton thus includes all digital computers and switching circuits. The response of an automaton to a particular input symbol may be conditioned by the string of input symbols that preceded it. A state (of conditioning) of an automaton may be identified with a set of input strings that condition the automaton equivalently, which means that, no matter which string of the equivalent set is used to condition the automaton, its response to any particular subsequent string will not be changed. Each successive input symbol carries the automaton into the (next) state whose equivalence set contains the string formed by this symbol following the preceding string.

Abstractly, a finite automaton is a 5-tuple $(S, I, O, \Sigma, \Omega)$, where S (the states), I (the input alphabet), and O (the output alphabet) are finite sets, Σ is a (next-state) function, and Ω is an (output) function that, respectively, associate a member of S and a member of O with each member of the (concatenation) set SI.

A *finite-state language* has both generation and recognition grammar in the form of finite automata. The latter means there is a finite automaton with a designated (starting) state q and a designated subset F of (final) states such that a string is a sentence if and only if the effect of the whole string, applied as an input sequence, is to change the state of the automaton from q into a member of F. Kleene[37] proved that the finite-state languages are just the sets of strings denotable by regular expressions; thus the identifiers, unsigned integers and labels in ALGOL are finite-state languages. If all the productions of a simple phrase-structure grammar G are of the form $A \rightarrow Bx$, $A \rightarrow x$, or $A \rightarrow A^0$ (or $A \rightarrow xB$, $A \rightarrow x$, or $A \rightarrow A^0$) where A and B are nonterminal and x is terminal, then G is a (generative) grammar of a finite-state language.[38–40]

The Polish logician Jan Lukasiewicz discovered a simple way to eliminate the need for parentheses in expressions containing operators that have one, two, or any number of operands. Every operator symbol is simply written as a prefix (or as a suffix) to the sequence of its operands. Correspondence between the usual algebraic notation and the Polish suffix notation can be seen by verifying the equivalence of examples such as

$$a + b \text{ and } ab+$$

$$ab \text{ and } ab\cdot$$

$$ab + c \text{ and } ab\cdot c+$$

$$a(b + c) \text{ and } abc+\cdot.$$

In the last example, we see that a multiplies all the rest of the expression, so the multiplication sign implied appears at the end in Polish notation.

We shall now describe a simple automaton that can test any string to determine whether it is a sentence in this parenthesis language and, if it is, translate it into the equivalent sentence in Polish notation. This automaton cannot be finite state because for any fixed finite memory capacity there exist strings with parentheses nested so deeply that this capacity would be exceeded. Instead it is made of a four-state finite automaton combined with an (at least potentially) infinite push-down store used to remember previous states. The states, denoted 0, 1, 2 and 3, keep track of progress within a parenthetic expression as illustrated by

$$({}_0a_1b_1c_1 +_2 d_3e_3 +_2 f_3),$$

where the subscripts denote the state of the finite automaton after receipt of the subscripted symbol. Whenever the automaton receives a letter as input, it prints the letter as output and, if it was in states 1 or 3, then prints a "·". When the automaton receives a "(" it pushes down its former state into the push-down store, and when it receives a ")", which is proper only in state 3, it first prints "+" and then pops up the most recently stored state and prints "·" if that state is 1 or 3; otherwise, it changes state from 0 to 1 or 2 to 3.

The process of a translation is shown below:

Time step	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Input symbol	#	*a*	+	*b*	(	(	*c*	+	*d*	)	*e*	+	*f*	)	#
Output symbol	#	*a*		*b*			*c*		*d*	+	*e*·		*f*	+·	+#
Next state	0	1	2	3	0	0	1	2	3	1	1	2	3	3	0
Contents of Push-down					3	0 3	0 3	0 3	0 3	3	3	3	3		

A nonsentence will present an improper symbol in states 0, 1, or 2, or overdraw or underdraw from the push-down. This example is a very simplified illustration of the basic actions of the input scanner in a compiler for an algebraic language such as ALGOL, and it shows why a push-down store is usually programmed into such a routine.

[37] S. C. Kleene, "Representation of events in nerve nets and finite automata," in "Automata Studies," C. E. Shannon and J. McCarthy, Eds., Princeton University Press, Princeton, N. J., pp. 3–41; 1956.

[38] N. Chomsky and G. A. Miller, "Finite state languages," *Information and Control*, vol. 1, pp. 91–112; May, 1958.

[39] N. Chomsky, "On certain formal properties of grammars," *Information and Control*, vol. 2, pp. 137–167; June, 1959.

[40] Y. Bar-Hillel and E. Shamir, "Finite state languages: formal representations and adequacy problems," *Bull. Res. Council Israel*, vol. 8F, pp. 155–166; February, 1960.

Two recently designed machines, the Burroughs B-5000 and the English Electric KDF-9, have push-downs built into their hardware. The B-5000 also uses Polish suffix notation for its machine language and is said thereby to save one pass in algebraic compiling.

Recursively Enumerable Sets

Thus far we have discussed a series of families of languages beginning with the simple phrase-structure languages and progressing to smaller and smaller families, each contained in its predecessor. Now we shall investigate more general families.

Language theory can be related to recursive function theory through the device of numbering (assigning an integer to) each sentence of a language. A grammar then becomes a constructive procedure for generating a list of all the members of (*recursively enumerating*) a set of integers, and a recognition grammar is a procedure for deciding whether or not a given integer belongs to a given set. At least six quite different mathematical constructions have been defined[30,33,41–44] in efforts to make precise the notion of effective procedure in its widest possible sense. The striking result of these investigations is that each construction was found able to enumerate exactly the same family of sets, which are called the *recursively enumerable* (re) sets. The remarkable coincidence of capabilities of very different constructions lends great credence to Church's thesis[45] that no constructive procedure can ever be found to enumerate a set that is not re.

A basic result in recursive function theory is that there is a constructive procedure for numbering the re sets; indeed, they can even be placed in one-to-one correspondence with the integers. Define K to be the set of the integers that are numbers of re sets that contain their own numbers. K is re, but consider K', the set of all integers that are not members of K. Every re set contained in K' has a number that is also in K' but is not in the re set; hence K' is not re, which is to say that no constructive procedure can list exactly the members of K'. If the sentences of a language correspond to the members of K', that language can have no grammar; on the other hand, if they correspond to the numbers of K, then the language has a grammar but can have no recognition grammar. A language that has a recognition grammar is *decidable* and one that has not is *undecidable*.

The essential feature that introduces undecidability is the self-reference in the definition K. A language gains great power when it is made able to refer to itself, but it also acquires the power to state paradoxes and undecidable propositions. Most undecidable questions are in one way or another equivalent to questions about membership in K'.

Turing showed that any re set can be enumerated by a suitable (actually quite simple) finite automaton if it is augmented by a (potentially) infinite tape which it can move back and forth and on which it can read, erase, and write symbols. This combination is called a *Turing machine*. Certain Turing machines are universal in the sense that they can imitate any Turing machine if a finite portion of their tape is suitably pre-encoded; thus a universal Turing machine can enumerate any re set.

We have now described six families of languages. Each family is larger than and includes all after itself in the following sequence:

Turing machine (recursively enumerable)
Decidable (recursive)
Phrase structure
Simple phrase structure
Sequentially definable
Finite state (regular expressions).

Because a finite automaton alone suffices to generate or recognize the last family, while a finite automaton plus a tape can generate the first, it might appear that there is really little to choose between families in this hierarchy, but the theoretical power and probably also the complexity of practical implementation vary quite markedly according to the model chosen. Unfortunately, very little is known yet of general principles to guide one's choice. The essential distinction between them is in their memory requirements: *e.g.*, at the bottom, finite automaton; in the middle, finite automaton plus bottomless push-down; and at the top, finite automaton plus endless tape.

[41] A. Church, "The calculi of lambda-conversion," in "Annals of Mathematics Studies," no. 6, Princeton University Press, Princeton, N. J.; 1941, 1951.

[42] A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem," *Proc. London Math. Soc.*, ser. 2, vol. 42, pp. 230–265; 1936–1937.

[43] E. L. Post, "Recursively enumerable sets of positive integers and their decision problems," *Bull. Am. Math. Soc.*, vol. 50, pp. 284–316; May, 1944.

[44] A. A. Markov, "Theory of algorithms," Am. Math. Soc. Transl. Ser. 2, vol. 15, 1960; from *Trud. Mat. Inst. im. V. A. Steklov*, vol. 38, pp. 176–189; 1951.

[45] A. Church, "An unsolvable problem of elementary number theory," *Am. J. Math.*, vol. 58, pp. 345–363; April, 1936.

Development in High-Speed Switching Elements*

ARTHUR W. LO†, SENIOR MEMBER, IRE

Summary—**Recent developments in high-speed switching elements are outlined to indicate the trend of recent work and thoughts on physical implementation of high-speed digital data processing systems. Four outstanding switching elements, thin film cryotron, high-speed transistor, microwave parametric phase-locked oscillator, and tunnel diode, as well as some related circuit and fabrication techniques, are briefly described and commented on.**

The transistor remains as the predominant switching element for these systems, largely because of the vast experience with it. The other technologies are presenting an ever increasing challenge to the transistor in this area.

Introduction

THE phenomenal growth of the data processing science and industry in the past decade was both a cause and a result of the extensive research and development efforts on switching elements particularly suitable for high-speed digital operation. Major developments in this field centered on devices utilizing various solid-state phenomena, as well as the circuit and operation techniques associated with these devices. This paper is an attempt to outline the significant new developments which indicate the trend of effort in this area. Detailed description of specific devices and techniques will not be included, but the major technical approaches will be identified and commented upon. The coverage is limited to the basic, high-speed switching elements useful in handling digital information and excludes devices used for gating analog signals or controlling power in digital form.

Modern digital systems contain a great many (from a few hundred to tens of thousands) switching elements interconnected in complex ways. Such size and complexity, as well as the economics of design, manufacture and maintenance, demand a "building-block" approach, *i.e.*, the use of nominally identical components (of only a few types) and standardized signal waveforms throughout a system. As a basic requirement, switching elements must be able to function as designated in circuits with complex interconnections and under all forseeable combinatorial and sequential signal operating conditions and, at the same time, take into account the fundamental and practical aspects of component tolerance. Such considerations are reflected in the pertinent characteristics required in a basic switching element:

1) The element responds discriminatively to the discrete signals, and can restore deteriorated signals to their respective standard forms. The element must provide unidirectional information flow so that no ambiguity in control can occur through any spurious interaction. These requirements imply that the switching element has well-defined thresholds, discriminative gain and limiting characteristics, and adequate unilateral and isolation properties.
2) The element has high switching speed (considering both the delay time and the transition time in switching from one state to the other), and can operate at a high repetition rate.
3) The physical structure of the element lends itself to easy fabrication in large quantity within specified tolerance, and allows easy assembly and interconnection with associated circuit elements.

A switching element need not fulfill all these requirements to be useful since some of the deficiencies can be remedied, often, by some operation and circuit techniques and/or by incorporating some other components. Nevertheless, these basic requirements serve well as a yardstick for evaluating the various switching elements and help clarify the significance of new developments.

Switching Elements

Many electronic devices have been considered as basic switching elements for the handling of digital information. The various devices utilize widely different physical phenomena and operate in widely different manner. Viewing these devices as circuit elements, however, one can group them into three general categories. This cataloguing permits a systematic discussion of the operational characteristics of the devices and the implications of new developments.

Four-Terminal Devices

This group of devices has the general configuration of a two-terminal-pair network *not* sharing a common electrical node between the input and the output. Notable devices in this group are electromechanical relays, optoelectric devices and cryoelectric devices. The absence of a direct electrical connection between the input and output terminals in these devices allows flexible interconnecting configurations between elements, thus permitting the use of "relay logic" which is highly desirable in many applications. These devices have excellent unilateral action, apparently because their operation involves physical energy of more than one form, each of which can be easily discriminated from the other. [For example, in the Electroluminescent-Photoconductor (EL-PC) pair, the EL material produces light from electrical energy, but it is insensitive to light; while the PC material is electrically sensitive to light but does not

* Received by the IRE, July 5, 1961; revised manuscript received, January 2, 1962.

† Components Division, IBM Corp., Poughkeepsie, N. Y.

produce light.] Although many transducer-type devices have been suggested as possible switching elements, few have gain, especially gain at high operating speeds. Present optoelectric devices, using electroluminescent material or neon bulb as light source and photoconducting material or semiconductor devices as light detector, operate in the millisecond range. Significant improvement in speed awaits, it seems, a better understanding of optoelectronic mechanisms in solids. Recent developments in cryotron techniques, on the other hand, have improved the speed of the device so significantly that they merit consideration for practical applications.

Thin-Film Cryotron: The significant improvement in cryotron speed can be attributed to the evolution of constructing planar, thin-film cryotrons over an insulated superconductive ground plane. The operation of a basic thin-film cryotron circuit is illustrated in Fig. 1, where a constant current I flows, quiescently, in one of the two branches of the loop. An input current applied through control A causes the gate film under the control to switch to its normal (resistive) state. This operation is equivalent to inserting a resistance R into the superconductive loop of total inductance L to cause the steady current to switch to the other branch, in the form of a loop current: $i=I(1-\epsilon^{-R/L^t})$. At $t=2L/R$ approximately 90 per cent of the current I is switched from the left-hand branch to the right-hand branch; and this new division of current remains after the control current ceases, keeping the output gates a' and a resistive and superconductive, respectively. Improving the operating speed of a cryotron circuit is a matter of reducing the L/R time constant. The superconductive ground plane, only a fraction of a micron distant from the film cryotron, supplies an image current to any current in the cryotron. This action reduces the inductance of the cryotron several orders of magnitude below what it would be if the ground plane were not present. The ground plane has the additional effect of localizing the magnetic field produced by a current in its immediate proximity, thus allowing high-density packaging. The resistance R is determined by film thickness and penetration depth, as well as by the resistivity of the gate material. Some alloy gate materials have higher resistivity than the commonly used tin and indium and higher operable temperature, and they are the subject of extensive developmental effort at present. No energy is dissipated in a superconductive circuit in its steady state. For each switching, however, an amount of energy $\frac{1}{2}LI^2$ is dissipated in the form of heat. Because cryotron operation is inherently tied to the critical temperature of the gate material, a narrow operating temperature range must be maintained. Heat dissipation (in high-density packaging) is actually a more serious consideration than the L/R time constant for high-speed operation of cryotrons. Without the significant reduction of L through the use of the superconductive ground plane, the operation of a large-capacity, high-speed cryoton system would be impossible.

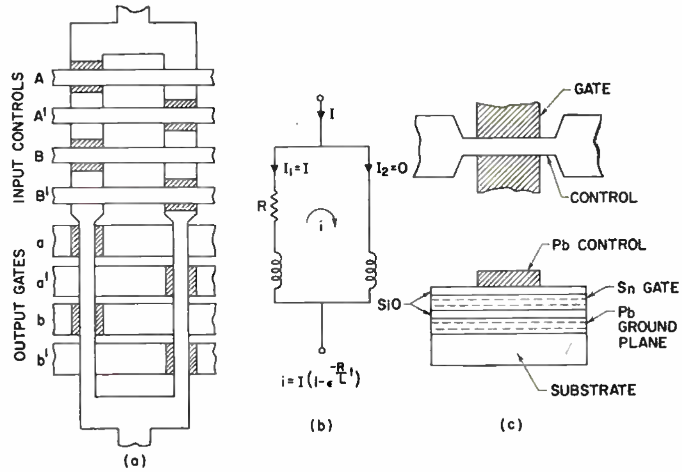


Fig. 1—Thin-film cryotron. (a) Basic cryotron logic circuit. (b) Switching time in terms of R/L time constant. (c) Cross-film cryotron structure.

Two forms of thin-film cryotron configurations have received extensive attention. In the cross-film configuration, the width of the control film is several times narrower than that of the gate film so that a small current in the control can switch the gate film to its normal state while in the absence of a control current the gate can carry a greater current and remain superconductive. Ideally, a current gain greater than unity is all that is required for one cryotron to drive others, but in practice a considerably greater current gain is desired to insure reliable operation against nonuniformity in cryotron units. For cross-film cryotrons of equal control and gate film width (a possible geometry for maximum packing density) and for in-line cryotrons (where the control film and the gate film run parallel in length to obtain a longer and thus higher resistance gate), the current gain is below unity. Operational current gain, however, can be realized by using a constant bias current applied through leads to the control film or through an extra superconductive film over the gate. With the bias arrangement, the amount of current switched is considerably smaller, effectively reducing the heating problem. The biased operation, however, imposes more severe requirements in the control of film dimensions and registrations, as well as of operating temperature. Controlling the thickness of the insulation layer has been a major consideration in the batch fabrication of thin-film cryotrons for high-speed operation.

Switching time of a few nanoseconds has been observed in experiments with a simple, biased in-line cryotron loop. Continuous operation of several interconnected cryotrons at frequencies up to 30 Mc has been reported, and the frequency limit was attributed to instrumentation difficulties in detecting the very weak voltage through the liquid helium dewar. Several hundreds of larger cryotrons (with tin gate film 500-microns wide and 0.5-micron thick, and lead control film 125-microns wide and 1-micron thick) have been batch fabricated on a single glass substrate. These elements have a uniformity of critical current within 5 per cent and operate at 4 Mc. While all the problems of producing a high-speed, high-capacity cryotron system, particularly in heat dissipation, are not solved, the technology is far enough advanced to justify serious consideration. The prospect of manufacturing superconducting systems (switching elements and all interconnections) with homogeneous, bulk-effect materials in an automated batch-fabrication process certainly makes the cryotron more attractive than many other switching elements for very large digital systems in the future.

Three-Terminal Devices

Vacuum tubes and transistors, which are the "work horses" in present-day digital systems, belong to this group. In these devices, there is adequate isolation between the input and output circuits, although they do share a common electrical node. Such an element functions well as a logical gate (conditional regeneration of signal), but it is not readily applicable, as are the four-terminal devices, as a transmission gate (conditional transmission of signal). The most significant progress in this area centers on the fabrication techniques for producing nanosecond bipolar transistors.

High-Speed Transistor: The high performance of recently developed transistors can be attributed to the precise structural geometry evolved from improved fabrication techniques. Early alloy junction transistors suffered in switching speed mainly because of the long transit time and minority-carrier storage time in the base region, since the alloying process was difficult to control in producing a very thin and well-defined base region. With a built-in potential gradient in the base region, the drift transistor offers some reduction in transit time; and at the expense of some circuit complexity, the storage time can be insignificant if the device is operated in the nonsaturating mode. The micro-alloy transistor on the other hand, is constructed by jet etching and surface alloying that produces a very thin base region for high-speed operation, but elaborate process control is required in manufacturing.

Recent development centers on diffusion techniques. Transistor fabrication is essentially a matter of transporting the right kind of impurity atoms over a well defined area on a high-purity semiconductor substrate, to the depth and concentration that form a prescribed impurity-distribution profile. The precise fabrication process is accomplished by a series of masking, photoengraving, and diffusion steps. In the present state of the art, base width in the neighborhood of one micron can be readily reproduced. With such small base width, transit and storage time in the base region is no longer a predominant factor in switching time. High-speed operation is essentially limited by the time required to charge and discharge transistor capacitances (the base diffusion capacitance, the emitter transition capacitance, and the collector transition capacitance). With the presently available thin base width, the diffusion capacitance is in the same order of magnitude as the emitter and collector capacitances. Base resistance is in the order of a hundred ohms because of the required geometry. Emitter and collector junction areas from 50 to 100 microns in diameter are presently feasible, and emitter or collector capacitance of a few picofarads is common. Using such units as an AND-INVERTER or OR-INVERTER basic logical gate with total fan-in and fan-out of five, a total switching time (delay time plus transition time) of 5 nsec is not unusual.

The carrier storage in the collector region has been a significant factor in operation speed. The recently developed epitaxial-growth technique offers a significant improvement. Over a low-resistivity single crystal wafer is grown epitaxially a thin high-resistivity layer, on which the collector junction is formed (Fig. 2). This composite collector structure retains the high collector-breakdown voltage, but effectively reduces storage, saturation voltage, and collector capacitance.

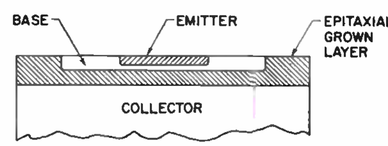


Fig. 2—A high-speed transistor structure (dimensions not drawn to scale).

Over the years hermetic sealing with can and header structure has been identified as the standard procedure in transistor manufacturing. Surface passivation methods, mostly with oxide and glass techniques, are now being developed for diodes and transistors. While thermal-compression bonding is the standard practice for attaching leads to semiconductor devices, various forms of metal-deposition process are being investigated for this purpose. Noteworthy is the growing interest in silicon in spite of its lower carrier mobility. The obvious advantage of higher operating temperature is often a secondary consideration. A significant property of silicon is that an oxide layer can be readily formed over its surface, and present photoengraving and diffusion processes can be readily applied. This is an example where fabrication techniques dictate the choice of a functioning material.

Microelectronics: With digital systems growing in speed and capacity, the task of assembling and interconnecting circuit components in a conventional manner, and meeting cost and performance (particularly speed) requirements, is becoming prohibitive. Developments centered on the idea of treating the assembly, interconnection, and packaging problems at the level of device fabrication have flourished in the past few years. These developments have been publicized under the general term *microelectronics*, although miniaturization is often a by-product of secondary consideration. Many approaches have been reported. The "printed-circuit" board is probably the first attempt. The various "micromodule" versions consist essentially of fabricating and attaching miniature components on suitable substrates with terminal contacts for easy interconnection. The "integrated circuit" or "semiconductor network" approaches are aimed at fabricating components and interconnections in successive steps on a single substrate to form a functional block. Since the commonly used components, such as transistors, diodes, resistors, and capacitors, can all be formulated by selectively controlled inhomogeneity in a semiconductor, the fabrication of a circuit unit consisting of several interconnected components is a task not unlike transistor fabrication, except in degree of complexity. Simple function blocks (such as multi-input logical gates and flip-flops) made in this manner are already commercially available. There is not sufficient evidence that some of the present structures show significant advantage in performance (particularly in speed and reliability) and cost, but the feasibility of this approach is certainly demonstrated. More ambitious approaches of batch fabricating an entire subsystem by physically transporting materials and chemically transforming materials on a molecular scale are currently being investigated in a number of laboratories.

Heat dissipation is a paramount problem in microelectronics. It is unfortunate that some microminiaturization proposals of packing 1,000 to 10,000 components in a cubic inch have not given adequate thought to the heat dissipation problem. Present high-speed transistor switching circuits consume tens to hundreds of milliwatts of power per stage. Noteworthy is the fact that although a transistor can be operated to perform switching at a power level of a fraction of a milliwatt, the high power level in practical circuits is a requirement dictated by component tolerance and operation speed. Component tolerance imposes an uncertainty in switching threshold, and thus a large excursion of signal voltage and/or current is required to insure reliable switching under adverse conditions (including noise). The presence of reactance (in components and interconnections) makes power dissipation a function of operation speed. While reactances themselves do not consume power, the switching of voltage or current level involves storing energy in and withdrawing energy from reactances through dissipative elements. Energy dissipation in this manner increases with switching repetition rate and even more so with switching speed. These considerations emphasize the need of low-level, minimum-reactance components that can be mass-produced to narrow tolerance. It seems apparent at this time that the high packing density of the order desired in microelectronics cannot be achieved simply by size reduction of neoconventional components and circuits.

Two-Terminal Devices

This category includes all the two-terminal devices, as well as devices which have more than two physical terminals, but where there is little or no isolation between the input and output circuits. In order to use such devices for information handling, some special circuit or operating techniques must be employed to provide directionality of information flow. Work in this area in the past few years has resulted in a number of fresh and interesting developments.

Nonlinear inductors and capacitors have long been investigated and used for amplification of relatively low-frequency signals, with a relatively high-frequency energizing source. A derivation of this type of operation, the ferroresonant flip-flop, has shown promise for digital operation at medium speed. It is, however, in the parametric phase-locked oscillator that we witness a fresh and inspiring approach for handling digital information at very high speed.

Parametric Phase-Locked Oscillator (PLO): While digital information is represented by discrete *amplitudes* of some physical quantity in almost all digital systems, the PLO system uses a *phase* script for representing a "0" or "1" signal by a sinusoid of constant amplitude and zero degrees or 180 degrees phase, respectively. When energized by a pump of frequency $2f$, a PLO tank circuit builds up a sustained oscillation of frequency f in either zero degrees or 180 degrees phase with respect to a reference maintained by the pump. The sustained oscillation is locked in phase, and is immune to input signals for the duration of pump energization. Control is executed by applying an input signal of roughly zero degrees or 180 degrees phase prior to or at the beginning of pump energization. This signal steers the oscillation buildup to the respective phase. On account of its perfect symmetry (with respect to the two discrete phases) the circuit has very high gain which is limited only by noise and the consideration of switching time. The pump supplies energy for building up oscillation to a constant amplitude, and simultaneously standardizes the signal by "pulling" the oscillation to one of the two standard phases.

The PLO by itself provides no directivity because the input and output signals share the same terminal. Unidirectional information flow is achieved through separation of input and output signals in *time*, rather than separation in *space* as in three- or four-terminal devices. A sequential pump-energizing scheme (illustrated in Fig. 3) is used to dictate the direction of information flow. It is to be noted that the three-pump scheme alone does not insure unilateral signal flow; the other factor is that the PLO is immune to input signals except for a short duration at the beginning of pump energization.

Logical operation is carried out by *majority logic*, which is essentially a linear summation of inputs (sinusoids of constant amplitude and zero degrees or 180 degrees phase) that allow the majority votes to determine the logical state. With one of the inputs a reference bias, AND and OR logical operation is straightforward. INVERSION is achieved simply by phase inversion through a transformer. An irony in PLO operation, however, is that while information is not represented by

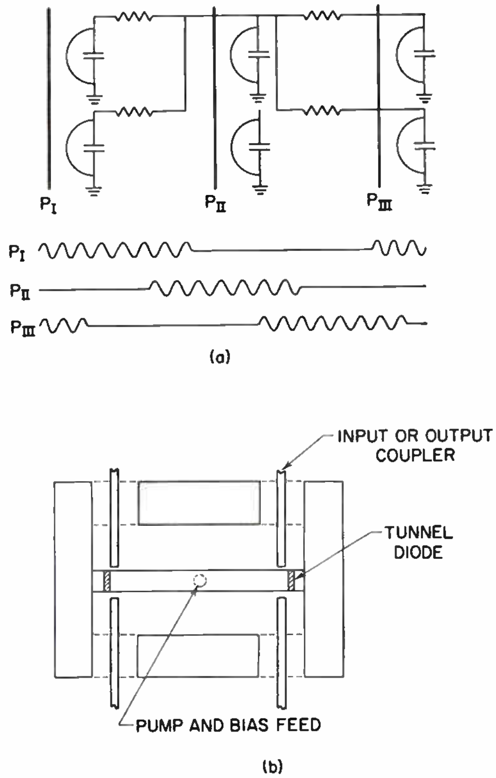


Fig. 3—Parametric phase-locked oscillator. (a) Sequential pumping system. (b) A microwave PLO structure.

amplitude, it is a necessity (and often a major problem in practical systems) to maintain a constant signal amplitude in order to perform majority logic.

The first PLO device was the Parametron using ferrite cores as nonlinear inductance. Practical systems have been built with this device for operation at an information rate of tens of kilocycles with a pump fre-frequency of several megacycles. This speed limit is a result of the heating of the ferrite at higher repetition rates; and attempts at using metallic magnetics have been reported.

Very high-speed operation of the PLO is currently achieved through the use of a semiconductor diode. The diode behaves as a high-Q nonlinear capacitor when slightly reverse biased. Diodes with very small junction area and low bulk resistance, encapsulated in a special, low loss, low capacitance package, offer a cutoff frequency of several hundred kilomegacycles. Using such units, a well-designed PLO circuit can reach a signal information rate of about 1 kMc with a pump frequency of around 30 kMc. Practical design of very high-speed PLO circuits hinges on microwave techniques. Microwave stripline and cavity structures have been explored extensively. An example is shown in Fig. 3, illustrating a two-diode balanced structure with pump and bias connection and input and output couplers. Simple PLO subsystems using symmetrical stripline structures interconnected with microwave transmission lines have been test-built to operate at information rates of several hundred megacycles; they use a pump frequency of 10 kMc.

The practicality of producing a microwave PLO system of usable size is yet to be established. Distribution of kilomegacycle pump power to a large number of circuit elements in proper amplitude and phase relation is a major problem. In microwave structures, geometrical configuration and dimensions dictate circuit behavior; whether this is a blessing or a handicap for building a practical system remains to be seen. These reservations aside, one cannot help but see in PLO an interesting and refreshing approach unlike conventional digital techniques.

Another group of two-terminal devices assume the role of switching elements through their storage and threshold properties. Ferroelectric and ferromagnetic materials characterized by a rectangular hysteresis loop have both been investigated. The ferrite core, in addition to its tremendous success in random access memory, has been used extensively in some logical systems. Although a core can have multiple windings, and turns-ratio can be used to advantage, any flux change in the core links both the input and output circuits. Because of the core's storage property, the output signal can be delivered at a later time than the application of the input signal. Nevertheless, some unidirectional coupling elements, usually diodes or transistors, are necessary in order to provide directivity of information flow. In recent years sophistication in device operation and circuit techniques has been incorporated in magnetic-core logical circuits that do not use unidirectional coupling elements. These developments include the use of multi-aperture cores (where flux transfer between various paths in the core allows the coupling between input and output circuits to be postponed in time), fast-and-slow switching (allowing higher or lower energy dissipation in coupling resistance) and partial switching (to take advantage of minor loop characteristics). Simple systems employing only ferrite cores and copper interconnecting wires have been demonstrated. The switching repetition rate of present ferrite cores is limited to around a megacycle; and metallic magnetics are currently being investigated for higher speed operation.

The conventional diode (which offers discriminative attenuation but no gain) has long been a major supporting switching element and has been used in large quantity in digital systems. Recent units, with shunt capacitance less than 1 pf and storage time measured in 1 nsec are indeed high-speed elements. Several types of negative-resistance diodes, utilizing a minority-carrier regeneration effect, suffer from a long recovery time in switching operation. The development of the tunnel diode now offers a device of very simple structure and extremely high switching speed, as well as well-defined thresholds and both voltage and current gain.

Tunnel Diode: The usefulness of this simple two-terminal device as a practical switching element depends largely on circuit techniques. Two basic modes of operation have been developed. The first mode consists of either bistable or monostable operation with a dc power supply (Fig. 4). In such operation, gain is realized by having the diode quiescently biased close to one of its thresholds so that a small input signal is sufficient to initiate switching in the negative-resistance region. Such operation imposes a severe demand on device uniformity (especially in the peak current of the diode) and the tolerances of other circuit components. These stringent requirements make practical design difficult. Unidirectional coupling elements also are required.

The second mode of operation employs a matched pair of diodes with pulse energization (Fig. 4). In many respects this operation is directly analogous to that of the PLO, permitting the same type of circuit and logic organization that is used in the PLO system. There are, however, some inherent differences between the two cases. The PLO has a natural, perfect symmetry immune to component variation; but the tunnel-diode pair is balanced artificially and requires a minimum input to offset any circuit asymmetry. (Matching diodes by pairs is, in practice, an easier task than that of demanding all diodes in a system to be uniform as in the first mode of operation.) On the other hand, while the information rate in the PLO system is only a fraction of the pump frequency, the tunnel diode switches at the rate of the energization pulse.

Since tunneling is a majority-carrier effect, the switching speed of a tunnel diode is primarily affected by the junction capacitance. As illustrated in Fig. 4, the voltage across the diode can switch only as fast as the condenser is being charged up. At the beginning of switching the condenser charging current is small, thus, the output waveform displays a longer delay time compared with the transition time. In order to reduce the long delay time, current overdrive of five to ten per cent over the peak current is required. Recently developed tunnel diodes of 5-ma peak current have capacitance in the order of 5 pfs and series resistance around 1 ohm. Individual logical gates of total fan-in and fan-out of two or three have been observed to switch in 1 nsec, and simple logic subsystems have been breadboarded to operate at 300 Mc. Although an all-tunnel-diode digital system is feasible, the advantages offered by the diode, including its speed capability, could be offset by the additional complexity in circuits and power supplies needed to operate this two-terminal device. The feasibility of three-terminal tunneling devices are being explored in several laboratories. At present, a promising approach is the transistor-tunnel diode combination, in which the transistor is used for isolation and current gain and the tunnel diode is used for its fast switching in discrete voltage level. Basic emitter-follow type gates have been built, showing a delay of 1 to 2 nsec per stage with fan-in of five and fan-out of four.

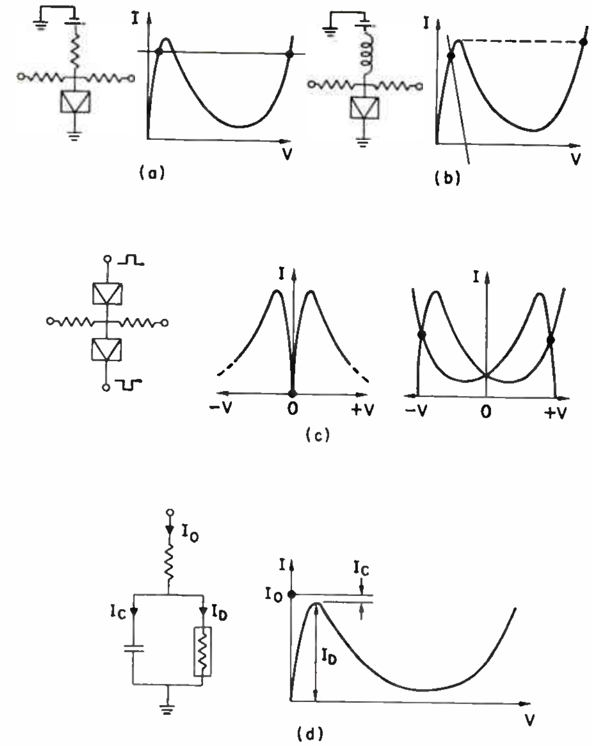


Fig. 4—Tunnel diode operation. (a) Bistable operation. (b) Monostable operation. (c) Balanced-pair operation. (d) Effect of shunt capacitance on switching delay time.

Conclusion

Recent developments in solid-state devices and circuit technology have inaugurated a number of high-speed switching elements, some approaching kilomegacycle rate, suitable for large-capacity digital data processing systems. The transistor remains the predominant switching element for these systems at present and probably will continue to be so for some years to come, largely because of the accumulated knowledge of how to build and use the element for performance and cost advantages. Undoubtedly, the progress of the other technologies cited above is posing an increasing challenge to the transistor as we know it today. When one or more of these competing technologies matures to economic practicability, we can expect more rapid advancement in data processing capability in speed, versatility, and reliability.

Acknowledgment

Recognition of the many individual contributions to the new developments covered in this article would require a bibliography several times the length of the text. It is, therefore, necessary to give a general acknowledgment to the contributors in this field. The author is also indebted to his colleagues at IBM for discussions and suggestions in preparing this paper.

New Concepts in Computing System Design*

GENE M. AMDAHL†, MEMBER, IRE

Summary—New system concepts are discussed within the framework of market requirements and the design goals which are to meet them. These goals are:

1) Programming and compiling aids for increased operating efficiency and flexibility,
2) Higher machine speeds for increased throughput and reduced job cost,
3) Multiprogramming (time sharing) and multiprocessing (load sharing) for greater concurrent utilization of system hardware,
4) System exchanges for greater flexibility in system hardware complement and allocation.

Introduction

THE SELECTION of items to be discussed in this abbreviated article depends upon the degree of association with areas recently having the greatest published activity in computing system design.

These design activities all appear to be focussed on rather general requirements in the market place. Market needs seem to be growing more rapidly in the direction of variety, within computing installations as well as among them. On any one machine it is very likely that: problem running times will be distributed over a range from several seconds to several hours; problem characteristics may vary from scientific computing to data processing to logical manipulation; compiling and debugging time may exceed production time; and real-time processing must be interspersed with scheduled processing. Also, as machine speeds increase, one finds that the communication between man and machine becomes more remote. In part this problem is alleviated by corresponding advances in software and hardware techniques which reduce requirements for "on-line" exchanges. However, not all of this communication can be replaced without serious loss in both man and machine effectiveness.

For purposes of exposition, principal design activities may roughly be categorized by four of the goals sought by the computing system designer. These four goals are:

1) Programming and compiling aids for increased operating efficiency and flexibility,
2) Higher machine speeds for increased throughput and reduced job costs,
3) Multiprogramming (time sharing) and multiprocessing (load sharing) for greater concurrent utilization of system hardware,
4) System exchanges for greater flexibility in system hardware complement and allocation.

Programming and Compiling Aids

Statistics obtained from a large number of computing installations indicate that programming costs have been running approximately 50 per cent of the total cost of the installation. Training programmers to prepare problems for computer solution is in itself an enormous task, and, with the growth of computer usage, the training programs have not been able to keep pace. Because of this, an increasingly larger percentage of problems are being written in a language more closely related to the problem than to the computer instruction set. These are then compiled by the computer into its own instruction set. Statistics relating to computer usage indicate that approximately 30 per cent of useful computer time is being used for compiling.

Because of these cost and manpower problems, more and more attention is being given to the area of reducing programming and compiling tasks for computer systems. In the past the most common way of attempting to reduce programming costs has been to design more sophisticated and larger sets of instruction codes in the computer. This approach may well reduce programming difficulty for writing a problem directly in machine language. However, the difficulty encountered by a compiling program in translating from a source language to a complex machine instruction set is much greater than that encountered in translating to a simple order structure.

Historically, the major emphasis in determining the speed of a computer was placed on raw arithmetic speed. In future systems the ability of a computer to perform the functions required by compiling, debugging and supervisory systems must be taken into account. Computers must be able to operate efficiently with variable sized fields, to scan data records, to search tables, to construct and use lists, to edit information for output, etc.

One approach to these functions is the use of an instruction set that is primitive in a computational sense. A primitive code is defined here as a set of operations in which the entire function of each is required whenever employed. The primitives should be selected for simplicity and generality so that required functional sequences can be efficiently constructed. It is of interest to note that the best programming systems use simple rather than complex coding techniques. A primitive instruction set has been employed in the Thompson

* Received by the IRE, September 19, 1961; revised manuscript received, November 21, 1961.
† IBM Corporation, Data Systems Division, Poughkeepsie, N. Y.

Ramo Wooldridge AN/UYK-1. This set concentrates on simplifying execution of macro and interpretative codes rather than reducing compiling difficulty.

An attempt to reduce compiling tasks by use of Polish notation during problem execution is encountered in the B-5000. In order to permit the use of Polish notation, the B-5000 employs a "pushdown" store. A "pushdown" store, in effect, retains operands, either from previous results or from loads, and automatically returns them to the arithmetic unit upon encountering an operation to be performed. The earliest hardware implementation of a "pushdown" store appears in the Aeronutronic Logic Evaluator, delivered in early 1961 to RADC. In this application Polish notation is not employed, but rather automatic translation from parenthetical algebraic notation to parenthesis-free Polish notation is performed during execution of the algebraic logic statement.

Another approach is to automate the task of storage allocation. This technique involves using a much larger memory for programming purposes than is actually available in high-speed random-access memory. This technique has been employed in the Ferranti Atlas. Backing up the large programmers' memory is a large drum memory. During program execution, requests for memory references go through an associative memory. The contents of the associative memory indicate whether the requested information is actually in high-speed memory and, if so, where. If the requested data is not in the high-speed memory, the supervisory program interrupts and calls for transfer of this data from the drum. In principle, the one for one relationship of large programming memory to backup drum storage is not required, for the supervisory system could determine the location of data in any of a number of peripheral media.

Related to the principle employed in the Atlas is the more generalized approach taken by the B-5000. For programming purposes one has, in effect, an almost unlimited pseudo-memory array addressed entirely symbolically. After compiling, the symbolic addresses have been replaced by entries to tables of arrays of actual addresses within the storage media. At execution time all memory references are executed indirectly through these tables set up by the compiling program.

In both the Atlas and B-5000 pseudo-memory approaches an indirect reference is required for all memory requests. This adds an extra memory cycle to each request. It is clear that these indirect references need not necessarily be performed every time, but only on condition that a reference to a previously unidentified block occurs.

Higher Speed

The principal reason for use of a higher-speed computer is that it generally provides more processing capability per dollar than a lower-speed computer. There is an empirical law—Grosch's law—which states that the speed of computers has increased generally as the square of the cost. We need not be concerned with the accuracy of this law, but only with the fact that the relationship between cost and speed does yield a lower cost per job the faster the computer.

A number of approaches to increasing the speed of computers have been employed or proposed in recent years.

Approaches of interest in the area of speeding up the arithmetic and control unit relative to memory-cycle time are *overlap* and *look-ahead.* These require a number of instructions to be in the process of execution at any given time within semi-independent functional units. In overlap the execution of a single instruction can normally be considered to take place in a fixed sequence progressing through these units. The sequencing is momentarily halted whenever a situation arises in which a given unit is busy, and this unit is required by a succeeding instruction. A machine which has developed overlap to a fairly high degree is the Remington Rand LARC. Look-ahead extends this philosophy in that a queue for the use of the functional unit is initiated, and other functions to be performed in the execution of the instruction continue if logically feasible. The IBM Stretch and the Bull Gamma 60 are two computers that make use of look-ahead.

To reduce the effective access time for data and instructions to less than the full memory-cycle time, a number of computing systems have employed multiple memory boxes operating in parallel. In some configurations, *e.g.*, Remington Rand LARC, the memories are available for accessing at separate phases of a clock which cycles in sequence through the core memories and makes a complete revolution every memory cycle. In other computing systems, *e.g.*, IBM Stretch, a clock is used purely for time sharing the memory buses. With this system any memory box may be referred to at any phase provided that the memory has completed its cycle from earlier references.

Another approach to speeding up computations involves reducing the number of references to large memory through the use of two types of high-speed stores. These stores may employ very-high-speed core or active registers. The two types of stores are FIFO (First In-First Out) and LIFO (Last In-First Out) (commonly called list structures) organizational structures. The FIFO store technique has been primarily considered for instruction storage. Instructions in a sequence being executed by the machine are simultaneously recorded in the FIFO store. When a transfer of control is encountered which transfers control back in the sequence a distance less than the capacity of the FIFO store, the instructions can be picked up from FIFO store for execution rather than requiring a lengthier access to the large memory. This technique was first used in the RCA BIZMAC I and was contemplated for, but not included in, the IBM Stretch. The FIFO store is in the Illinois Oddity and is planned for the CDC 6600.

The LIFO store (referred to earlier as "pushdown" store) is more commonly associated with the arithmetic unit and permits efficient automatic storage and return of intermediate results in a local portion of a computation. This technique has also been considered in program-control-sequence use for holding subroutine linkages. The LIFO store technique has been employed in the Aeronutronic Logic Evaluator, Burroughs B-5000, English Electric KDF-9, and Ferranti Atlas.

To take advantage of the higher performance attainable through the concurrent use of functional units and the reduced memory accesses with FIFO and LIFO stores, it is necessary that the arithmetic units in the computing system have a correspondingly reduced execution time.

Due to space limitations, the reader is referred to an excellent discussion of some high-speed arithmetic-unit techniques by MacSorley.[1]

Multiprogramming and Multiprocessing

In this paper *multiprogramming* will be defined to mean time sharing of a single CPU (Central Processing Unit) for a variety of tasks, while *multiprocessing* will refer to load sharing on one or more problems by more than one CPU. Primarily multiprogramming and/or multiprocessing have been concerned with the simultaneous handling of input-output buffering and computing. More recently, editing has also been executed simultaneously with internal computational tasks, as in the Remington Rand LARC.

Unless multiprogramming is restricted to jobs specifically planned to run concurrently and to programs that are "debugged," a minimum of five requirements must be provided for in hardware and/or in the programming system.

1) Memory protection—If all or portions of several programs are to be used concurrently in the high-speed memory, a given program cannot be permitted access to sections of the memory allocated to other programs. Further protection may be required if subroutines within the supervisory area are to be available upon request to all programs. The latter case may require that these areas can be read from, but not written into, by a program.
2) Program and data relocatability—Since the memory area available to a program and its data cannot be determined at compile time, a minimum need is for relocatability at load or reload (when a problem is dumped due to a high-priority job) time with a relocating load routine. Indirect addressing of some nature is required if the program and data are to be relocated independently.
3) Supervisory program—Since a problem program does not have full information about the status of the computing system, there must be a supervisory routine which handles storage and Input/Output allocation, controls I/O operations and processes interrupt signals.
4) Interrupt system—While the current trend is to provide a multiplicity of interrupt signals, a minimum of one—a timed return to the supervisory program—must be available. The timed return will permit the supervisory program to keep control of the computing system even if a problem program is in an endless loop or has executed a program halt.
5) Symbolic addressing of I/O—I/O units are similar to memory areas in that they cannot be allocated at compile time. The designation of a unit by the problem program must be replaced by an assigned number at loading or execution time, or else must transfer control to the supervisory routine for I/O execution.

This list is certainly not adequate for efficient multiprogramming and is not necessarily minimum if a multiprocessing system is considered. A number of things have been tried to improve the properties of a system for multiprogramming. One approach is the multiple program counters in the Minneapolis Honeywell 800 where one instruction of each of 8 programs is executed in a cyclical basis. Most of the concurrent programs for this system are related to I/O. Another approach is the MIT TX2 in which high-speed core registers hold the equivalent of a multiplicity of arithmetic-unit register sets. On an interrupt signal a given "arithmetic unit" is effectively brought into play until the next interrupt appears or a release is executed. Means are provided for returning to the previous "arithmetic unit" employed upon release by the current unit.

Approaches to multiprocessing include the Bull Gamma 60 which has a variety of independently-controlled functional units, some of which may be duplicated, such as a printer which executes its own program sequences (including limited editing functions), an arithmetic unit, a comparator unit, etc. Another approach is the TRW-400 where multiple computers may be switched between I/O units and communication channels via a real-time exchange (described under System Exchanges). The Remington Rand LARC has an I/O control and editing computer running simultaneously with a large-scale scientific computer. There are some examples of IBM 7090 and 1401 units linked together through tape units or 1301 disc files to compute and to perform I/O editing, respectively.

A special subgroup of requirements of multioperation includes memory protection, relocatability and symbolic addressing of peripheral equipment. In the B-5000 and Atlas computer these functions are performed by a

[1] O. L. MacSorley, "High-speed arithmetic in binary computers," Proc. IRE, vol. 49, pp. 67–91; January, 1961.

supervisory system implemented by combined program and hardware means. All memory references by the CPU are indirect through tables maintained by the supervisory system. These lists make available to a program only those areas of core storage belonging to that program. They also permit supervisory intervention to make available in high-speed storage that data which has been placed on backup storage due to internal space limitations. Since I/O is handled by the supervisory-system protection, relocation and symbolic addressing of I/O is achieved by programming. It is of interest to note that symbolic addressing of I/O is achieved in hardware in the TRW 400. A memory is provided which transforms the number used in the program to the proper I/O unit address.

Another approach to memory protection is the boundary registers in the IBM Stretch which cause an interrupt when memory access is attempted beyond these two limits. There are several versions of memory protection available on the IBM 7090, and one of the most interesting is the division of memory into blocks. Corresponding to each block is a bit in a memory-mask register. When access to memory is attempted, the bit corresponding to the block which contains the desired address can be interrogated to determine if this memory block has been assigned to the program controlling the CPU. In conjunction with this type of memory protection, a relocation register is provided to modify addresses generated in a program to the area assigned to that program and its associated data.

There is some question concerning the efficiency of multiple computers, for one cannot ignore the difficulties of making several computers appear like a single computer. If cooperating on a single task, coordination is necessary to ensure that computers do not get too far ahead or behind others in the system. If performing separate tasks, assignment, arbitration and monitoring functions for distributing work loads and peripheral equipment must be provided. There is also the problem of the amount of information that may have to be sent between the various units in the system. In general, two computers are less than twice as powerful as one computer, while doubling the equipment in a single computer usually results in much more than doubling the power of the resultant computer.

From a system-reliability viewpoint, there are advantages to a multicomputer system, provided that there are multiple units of each kind and that it is possible for one of these units to be disconnected from an active role in the system. A disconnected unit must make its internal register contents available to the system for recovery purposes. In such a system it may be possible to obtain graceful degradation in event of failure of a given unit. It would appear that the requirements for graceful degradation are probably greater than the requirements for multiprogramming or multiprocessing alone. Historically, the approach has been to have standby facilities in the event of failure of the operating facility. This method guarantees the minimum performance of the system without taking advantage of maximum computing capability available when all units are functioning properly.

System Exchanges

The computer system exchanges are considered to be the interconnection system between CPU's, memory boxes, peripheral storage and input/output units within an installation.

At present the most flexible computer exchange is in the RW-400 system. This exchange is fully electronic, being designed in both transfluxor and diode switching technologies. It is effectively a 16- by 64-channel rectangular cross-bar switch. CPU's and buffers, which are active controlling elements, are connected on the 16-channel side. On the 64-channel side the buffers again appear along with the array of peripheral storage and I/O devices. In addition, on one of these channels there is a small core memory which holds actual peripheral storage and I/O unit numbers assigned to each computer, addressed by symbolic unit numbers employed in each computing unit's operating program.

Of growing interest to the system designer is the handling of a large number of low-speed lines, where the communication network is a large or major portion of the system. These networks operate primarily under the control of devices other than the CPU, although in cases of traffic loads exceeding the capacity of the network the CPU can normally inhibit operation of any input unit. To bring such a large number of channels into a CPU economically, multiplexed (generally sequential scanning) I/O channels are used. In order to service requests for such a large number of I/O units, interrupt facilities must be provided which contain identification of unit and communication line. These systems require large transaction queue-buffer storage capacity, with essentially random access characteristics. Magnetic drums or disc files satisfy this requirement, but magnetic tapes must be precluded unless the system operation can be modified. In order to utilize magnetic tapes, the queues must be sorted before processing or retransmission rather than handling by input sequence alone. In addition, further backup random access storage is generally required to contain the large number of service routines which are likely to be needed for the variety of transactions. An early example of this is the SAGE system with its large number of telephone links with radar sites plus its large number of CRT displays. In the commercial area there is the SABRE system, in which up to 1000 agents' sets are connected into 7090's, 7080's or 7070's and where flight seat reservations are stored in disc files. The SABRE system requires large random access storage for file data but not too many routines for processing the data.

A third application is one in which a computer with large peripheral storage is used exclusively as a buffered communication exchange. Such a system has been designed, using the RCA 601, which handles 50 teletype or Kineplex input channels and 50 teletype or Kineplex output channels. These incoming messages contain priority and routing data along with the message. Messages are accepted by the central processor and stored in sequence in queues which are ordered by output line and priority. Provision is made to accommodate extremely long queues in the event of unusual communication loads on a very small number of output channels.

Similar systems are under consideration by some scientific computer users to simulate the operation of a multiplicity of small private computers. In this application terminal sets would be placed wherever convenient for the users of the system, and the terminal sets tied to a large CPU with its associated memory. The user would request service from his terminal set and, after the problem has been run, the output returned to the terminal set. Such a system is now undergoing limited tests at M.I.T. If its effectiveness meets expectations, the system may have considerable impact on the computing system designs of the future.

Conclusion

The new system-design concepts are aimed at making the computing system a more powerful, efficient, effective and flexible tool, in spite of the compounding complexity of hardware assemblies and the variety of applications. An adequate discussion cannot be presented in an article of reasonable length, or even by a single author. Furthermore, many of the new concepts are as yet unpublished. Even at this time so many of the computing machines and systems referred to in this article do not have nearly adequate descriptions published; therefore, the author feels compelled to omit all bibliographical material.

The Impact of Hybrid Analog-Digital Techniques on the Analog-Computer Art*

GRANINO A. KORN†

***Summary*—In a sense, many present-day analog computers exist in a gap left by the cost-speed characteristics of available digital computers. But even as digital computers improve, the all-parallel "live mathematical model" afforded by analog-type simulation remains indispensable to the system designer. As a result, new hybrid computing elements and systems combine analog and digital features, and analog computers are becoming faster. The techniques discussed include combination of conventional analog and digital computers, digital programming and checkout of analog computers, operational digital computing elements, new high-speed analog computers with memory and digital switching, true hybrid computing elements, and some hybrid computer systems.**

Introduction

MODERN electronic analog computers permit real-time and fast-time solution of the possibly formidable differential equations describing complex dynamic systems, such as space vehicles and chemical plants [1]. In such applications, analog computers have not merely filled a cost-speed gap left by existing digital computers; they serve as "live mathematical models" facilitating direct intercourse between system designer and simulated system, as well as the inclusion of actual hardware in real-time simulation. A system designer will tend to use the analog computer himself, while he may want to delegate digital-computer work to numerical analysts and computer programmers [2].

This situation is not static. The speed advantage of real-time analog computers is crowded not just by the increasing speeds of digital computers, but by their ability to work automatically-programmed night shifts. New compiler routines permit digital-computer setups directly from analog-computer-type block diagrams [3], [4]. At the same time, improved "operational" digital integrators, multipliers, and resolvers can replace analog components in some exacting applications and still preserve the direct-model all-parallel nature of the computation.

In the face of these developments, analog-computer designers have emphasized ever-higher computing speeds together with new devices and systems combining desirable features of both analog and digital computation. It is likely that the more expensive "analog" computers of the future will contain as many digital components as operational amplifiers.

* Received by the IRE, June 28, 1961; revised manuscript received, November 20, 1961.
† Electrical Engineering Dept., University of Arizona, Tucson, Ariz.

Combined Analog-Digital Simulation

Real-time simulation problems requiring high computing accuracy (*e.g.*, trajectory computations with three-dimensional coordinate transformations) and simulation of digital system components have led naturally to combination of existing analog and digital computers in many of the larger simulation laboratories.

In the examples of Fig. 1 [5] a multipurpose digital computer is linked to a real-time analog computer through one of the conversion units now commercially available from several manufacturers [6]. The digital computer performs that part of the real-time computation for which it is most suited, namely, accurate coordinate conversions, trajectory computations, and the simulation of digital-control equipment. The analog computer simulates dynamic control functions requiring greater bandwidth and less accuracy. Another striking application of combined multipurpose computers is the postflight analysis of tape-recorded aircraft-maneuver flight-load histories [7].

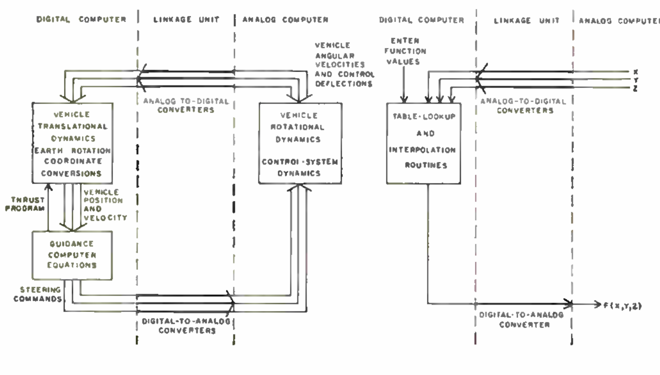


Fig. 1—Typical applications of combined general-purpose analog and digital computers [5]. (a) Combined real-time space-vehicle simulation. (b) Digital function generation.

The combined simulation technique was pioneered in the aerospace industry [5]–[7]. Painful experience seems to indicate that the provision of small multipurpose digital computers is greatly preferable to part-time utilization of large digital computers. In combined simulations requiring greater digital-computer speed, operational digital computing devices such as digital differential analyzers can supplement or replace the conventional digital computer.

Operational Digital Computation

In the author's opinion, the "operational" block-diagram structure of the analog computer, which closely parallels that of the system being simulated, will remain indispensable for system design. It is possible, however, to replace analog computing elements by digital computing elements capable of real-time operation. *Operational digital computers* represent variables basically by binary numbers, but unlike most general-purpose digital computers, they employ separate digital adders, multipliers, integrators, etc. in the manner of analog-computer setups. Such all-parallel operation permits relatively high computing speeds with a minimum of intermediate storage. The nuisance of transferring n-binary digits on parallel lines can often be avoided by *incremental computation* which transfers only changes ΔX of variables X between certain of the computing elements; during each clock period, ΔX is either 1, 0, or -1, so that, at most, two lines are required to transfer ΔX and its sign. An incremental computing element, then, represents the rate of change of the variable X by the repetition rate $\Delta X/\Delta t$ of the ΔX pulse train, and simple binary counters can generate the variable X itself wherever it is needed. Fig. 2 (next page) shows examples of simple operational digital computing elements employing binary and incremental representations [8], [9].

The best known operational digital computers are *all-parallel digital differential analyzers* which employ incremental integration and multiplication, together with digital function generators based on table-look-up and interpolation routines [8], [10]. A general-purpose computer of this type (Packard-Bell TRICE) can be patched just like an analog computer and uses 3-Mc digital circuits to produce 100,000 binary increments per second. Note that a given increment rate can buy either greater accuracy for slowly changing variables, or greater bandwidth with less accuracy [10]. TRICE permits all-digital real-time simulation of aircraft, but requires several hundred semiconductors for a single integrator. The resulting cost seems staggering when compared to that of dc analog integration. The author believes, however, that operational digital computation is still a relatively neglected field. In the future, costs are bound to be reduced through new arithmetic schemes, higher-order integration algorithms, new circuits and devices, and also through competition, mass-produced microcircuits, and amortization of development expenses.

Faster Analog Computers with Digital-Computer Features

In recent years the analog-computer industry has expanded its market to smaller design and research groups with small and medium-cost electronic analog computers which are likely to remain quite safe from digital competition. In view of the digital threat to a portion of their large-computer market, analog-computer manufacturers have attempted to increase the computing-speed advantage of their machines and to include automatic features which greatly simplify computer-setup and check-out procedures. At the same time, analog computers are acquiring a measure of the memory, decision-making ability and automatic programming abilities of digital computers.

Practically all large analog computers permit setup of coefficient potentiometers and function generators by digital keyboard control and punched-tape storage of coefficients and functions [11]. Such machines also per-

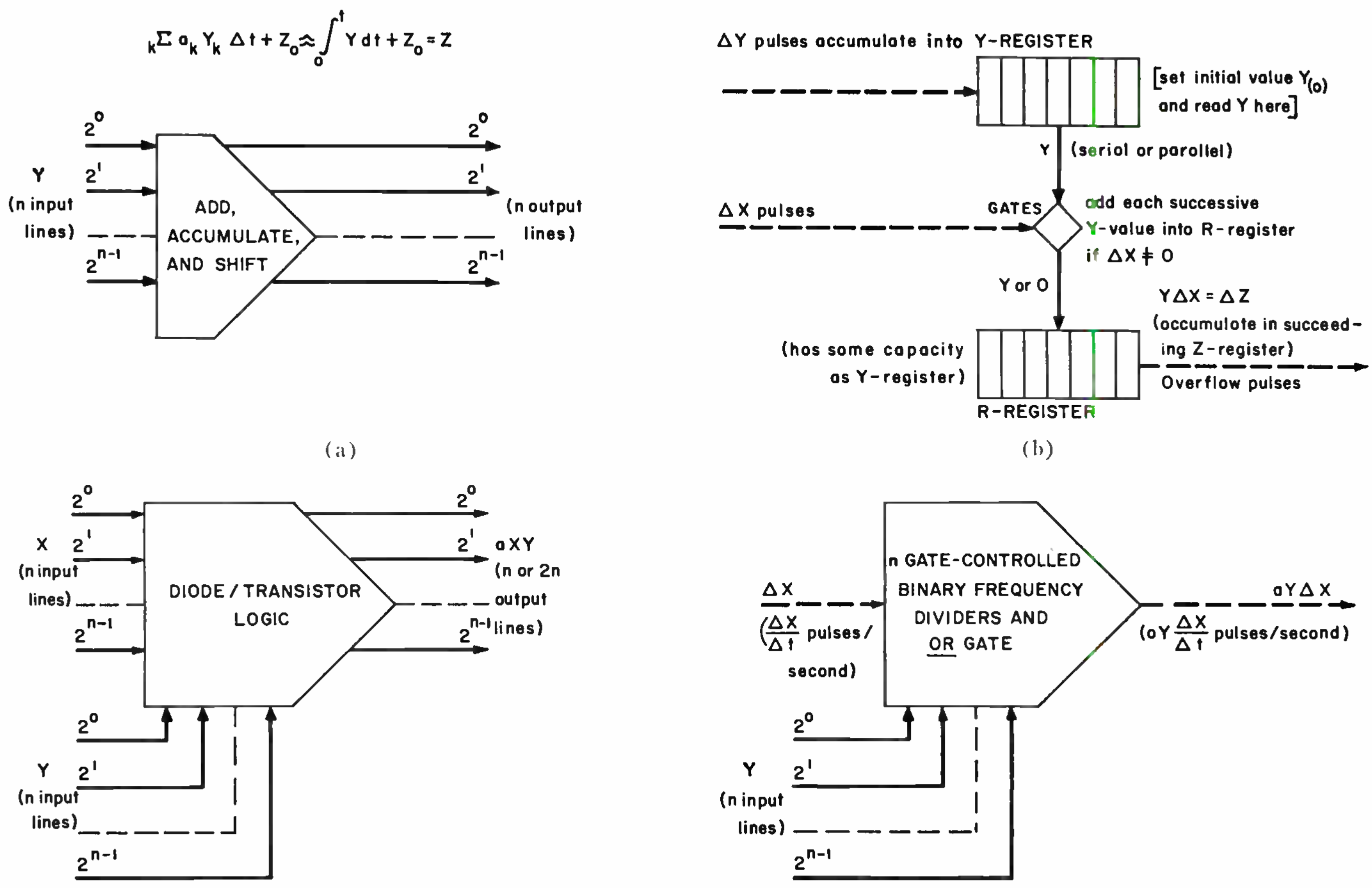


Fig. 2—Examples of simple operational digital computing elements capable of real-time operation. To simplify the representation, only operations on positive quantities are illustrated; in practice, a sign bit would be transmitted with, or coded into each variable. In the "incremental" computing elements *b*, *d*, the variables ΔX, ΔY, ΔZ represent pulses occurring at rates proportional to $\Delta X/\Delta t$, $\Delta Y/\Delta t$, $\Delta Z/\Delta t$. (a) Simple all-parallel integration in a digital computer. (b) Basic incremental integration in a digital differential analyzer. (c) All-parallel multiplier for multiplication of two binary numbers. (d) Multiplication of a pulse rate of a binary number.

mit digital printout of all initial conditions, thus speeding static setup checks [11].

The next step towards completely automatic programming of large electronic analog computers is to let a digital computer compile scale factors, potentiometer settings, and analog interconnection diagrams (setup sheets) directly from the given problem equations. Programming automation will be complete when the digital computer can actually accomplish analog-computer interconnections, say by means of a super-crossbar switch array, and proceeds to check out the analog-computer solution. Since the scaling-programming-checkout procedure for large analog computations is tedious and often consumes much more time than the computation itself, there is considerable interest in such complete programming automation [6], [37]. The very real difficulties inherent in switching a large number of analog-signal connections is simplified in some applications by various combinations of prepatching and automatic switching.

Wherever the inclusion of actual system hardware does not necessitate real-time operation, there is a universal trend towards *faster analog computation*. All-electronic multipliers, function generators, and resolvers are now more accurate as well as faster than computer servomechanisms. New diode function generators with a claimed static accuracy of 0.02 volts are available as sine-cosine generators [12] and for quarter-square multiplication [12]; they may well replace modulation-type multipliers if operating experience proves their reliability. The highest computing rates are afforded by repetitive and iterative analog computers which can implement vast parameter-optimization studies and statistical analyses at speeds impossible to duplicate by other means within the foreseeable future.

New Repetitive and Iterative Analog Computers

In conventional electronic differential analyzers, the integrator control relays cause each operational-amplifier integrator to follow an initial-condition (IC) input (RESET or TRACK mode), to integrate an input (COMPUTE mode), or to hold the last output value (HOLD mode) [11]. The use of high-speed relays or (preferably) *electronic integrator switching* permits accurate *repetitive operation* at between 5 and 1000 computer runs per second, so that the effects of parameter changes are readily seen on an oscilloscope presentation [2].

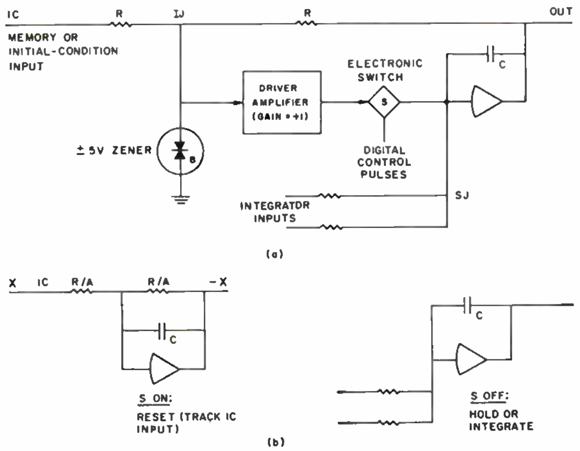


Fig. 3—The electronically-switched integrator is the key element of the modern analog computer and serves as an analog memory as well as a repetitive-computer integrator. With the electronic switch on (TRACK or RESET mode), the output will track the IC input. With the switch off (HOLD or COMPUTE mode), the output holds its last value or integrates an input sum. A typical unit can track a 20-v peak-to-peak sine wave at 50-kc and will hold within 0.1 v for 0.1 second with a 0.005-μf capacitor [14].

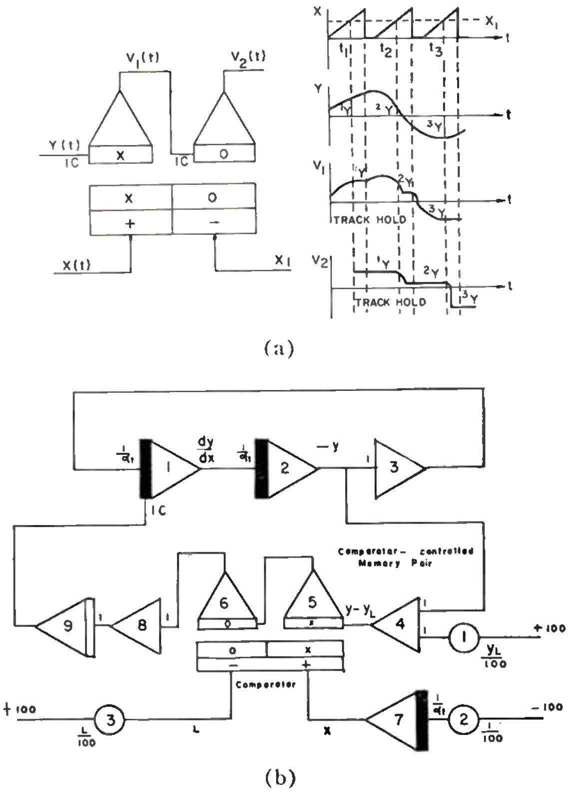


Fig. 4—(a) A comparator-controlled memory pair. Two electronically switched integrators alternate between TRACK and HOLD to permit solution of a wide variety of problems by finite-difference techniques. The comparator input voltage x can be any repetitively reset computer variable [16]. (b) A simple example of the iterative-differential-analyzer technique: iterative solution of the split-boundary-value problem

$$\frac{d^2y}{dx^2} = -y \quad \text{with} \quad \frac{dy}{dx} = 0 \quad \text{for} \quad x = 0,$$

$$y = y_L \quad \text{for} \quad x = L.$$

Darkened integrators are under rep-op control. x is represented by the computer time (real time) $\tau = a_t x$, so that $d/dx = a_t d/d\tau$. Amplifiers 1, 2, 3 constitute a differential-analyzer setup which solves the given differential equation repetitively about 10 times per second with the given initial condition $dy/dx = 0$ and the trial initial-condition change ${}^n\Delta y(0) = [{}^{n-1}y(L) - y_L]\Delta t$ for the nth computer run, where $y(0) = 0$. This trial initial-condition change is supplied by the comparator-controlled memory pair. Amplifier 5 tracks the output $y(x) - y_L$ of amplifier 4 from the start of each computer run until x reaches the value $x = L$ and trips the comparator. Amplifier 6 then acquires ${}^{n-1}y(L) - y_L$ and holds this value until x again reaches $x = L$ during the next cycle. The process converges rapidly to the desired solution.

Recent order-of-magnitude improvements in the speed and offset characteristics of electronic integrator switches [13]–[15] have not only improved repetitive computation but establish the electronically-switched integrator as a novel analog computing element which can serve as an *analog memory cell* as well as an integrator or summer. In the TRACK mode the new integrator can sample values of a variable applied to the IC terminal; in HOLD the integrator holds the last sample value and can introduce it into subsequent computer runs. Entirely new *sequential and iterative analog computing techniques* are thus made possible.

Fig. 3 shows an electronically-switched integrator and illustrates its modes of operation. Two such integrators are combined into a *comparator-controlled memory pair* [Fig. 4 (a)] which can sample fast computer solutions at sampling times determined either by a timing system or by any voltage in the computer. The step-wise output of the memory pair can be introduced iteratively into successive computer runs and may be said to simulate a digital-computer output. Fig. 4 (b) illustrates the iterative solution of a simple two-point boundary-value problem.

The repetitive-iterative differential analyzer with memory permits the introduction of "subroutines" obtained by fast repetitive analog computation into "slow" analog-computer setups [15], [19]. Examples of such applications include repetitive computation of the inner integral in a double integral [16], approximate solution of partial differential equations by difference-differential equations [16], and sequential solution of heat- and mass-balance equations at different points of a chemical reactor or distillation column [15]. Another significant application is automatic parameter determination or optimization by steepest-descent techniques [20]. The sequential optimization of, say, 40 parameters of a complex nonlinear system practically calls for high-speed repetitive operation; this need is even more acute when a *sample* of random-input computer runs must be taken for each combination of parameter values.

The repetitive-iterative analog-computation technique combines the operational model of the analog computer with the decision-making ability and memory of the digital computer. Because of the partially continuous and partially discrete nature of the computation, some of the new iterative techniques are not directly comparable to iterative digital-computer techniques. This is a fertile field for theoretical investigation.

Fast Analog Computers Incorporating Digital Logic

The availability of fast electronic switches, comparators, and sample-hold devices in the iterative differential analyzer has increased the utilization of these elements for direct simulation of control systems incorporating digital logic, sampled-data operations, and analog-to-digital conversion. Again, the switching elements have been operated at a sufficiently rapid rate to synthesize electronic time-division multipliers and electronic resolvers capable of operation with the slow portion of the analog computer, so that, in a sense, the "slow" portion of the iterative analog computer requires only linear computing elements, comparators and switched integrators. In the author's opinion, this type of operation is somewhat wasteful, since it simulates inexpensive digital flip-flops and digital switches by expensive chopper-stabilized operational amplifiers and accurate analog switches. It is suggested, therefore, that fast analog computers be equipped with a set of *digital modules*, such as Schmitt triggers, flip-flops, and diode gates, which can be operated together with switched integrators and analog switches to permit a new variety of simulation work as well as convenient synthesis of new hybrid analog-digital computing elements [21].

True Hybrid Computing Elements

The inclusion of digital computing elements in analog computers, or vice versa, brings us to the design of *true hybrid analog-digital computing elements*. It should be clear that the really basic transducers relating analog (continuous-variable) and digital (binary) operation are not complex analog-digital-analog converters but *comparators* (output changes state when the input sum goes through zero) and *analog switches* (analog voltage ON or OFF as dictated by a binary input). The development of accurate and fast components of this type, not an easy task, must be the basis of all hybrid-computer work [21]. Fig. 5 shows some currently used comparators and analog switches [13].

Repetitive operation and hybrid-computer techniques require *operational amplifiers* orders of magnitude faster than slow conventional analog computers. Fig. 6 shows a feedforward design yielding a chopper-stabilized amplifier capable of operation in the megacycle range.

Analog switches and fast comparators are combined in the well-known *binary-to-analog and analog-to-binary* converters shown in Fig. 7. Since the reference voltage in each of these circuits can be a variable, these circuits also permit *multiplication or division of analog and digital voltages*. The combination of both circuits has served as a hybrid multiplier-divider in both analog and digital systems [2], [24].

It is an intriguing fact that a pair of analog voltages each accurate within only 1 per cent can, in principle, represent a numerical variable to within 0.01 per cent. Although various vernier devices and two-speed servo-

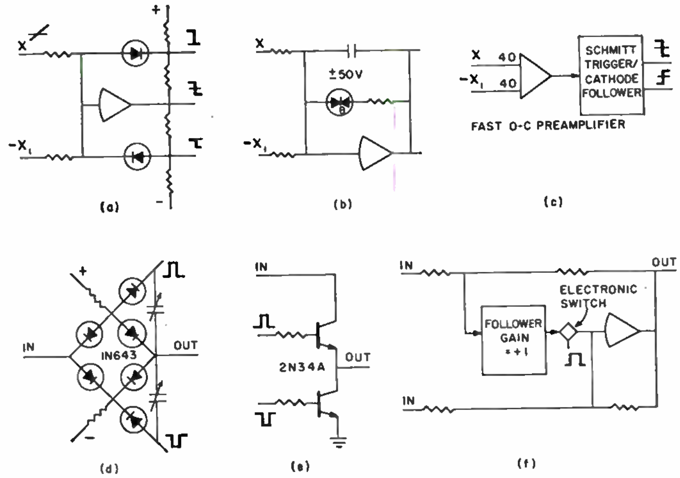


Fig. 5—Comparators [13] (a)–(c) and analog switches [13] (d)–(f) are the really basic analog-digital-analog conversion links. The basic comparators (a) and (b) yield 0.1-v accuracy at low frequencies (to 2 cps) and about 0.3-v accuracy at 100 cps. (c) shows a Schmitt-trigger comparator using a chopper-stabilized preamplifier to increase the static accuracy. This circuit is useful at 5–10 kc. The switching circuit (f) [15] can control ±100-v analog voltages with ±8-v pulses from transistorized digital modules.

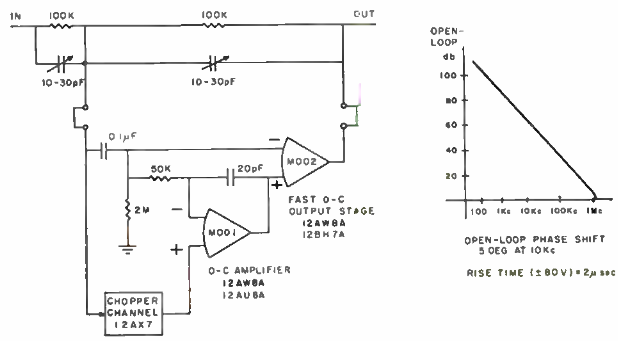


Fig. 6—Repetitive operation and hybrid techniques require operational amplifiers faster by orders of magnitude than "slow" conventional analog computers. The feedforward technique used in this ASTRAC amplifier yields 5° open-loop phase shift at 10 kc; the phase inverter shown is still useful at 1 Mc [21].

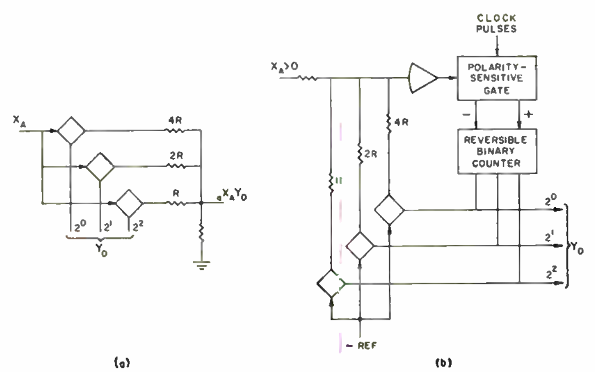


Fig. 7—Simple digital-to-analog converter (a) and analog-to-digital converter (b). In hybrid computers these circuits do not only serve for conversion but are used as relatively simple analog-digital multipliers, dividers, and scale changers (see also Fig. 9).

mechanisms and also perturbation methods for the analog solution of differential equations [25] can be considered as two-digit analog devices with two amplitude scales, this technique has not been tried to any extent.

On the other hand, a 3-digit binary number X_D (0 to 8) combined with an analog voltage X_A accurate to within 0.1 per cent can represent a numerical variable X to within 1/80 per cent. It should be possible, then, to combine simple 3-bit operational digital computing elements with analog devices so that the continuously-variable voltage X_A interpolates the less significant digits of X between the accurately computed coarse digital values X_D. Each hybrid computing element based on this principle will require interaction between its analog and digital components, including generation of *carry pulses.* Whenever the output voltage of an analog computing element exceeds a threshold corresponding to one digital unit, a *carry pulse* must be generated so as to add one unit to the binary output and subtract the threshold value from the analog output.

Fig. 8 shows a number of analog computing elements capable of generating a carry pulse whenever their analog outputs exceed set limits [26]–[28]. The carry pulse is then used to advance the digital output by one binary unit and to subtract a corresponding value from the analog output; the carry pulse may also serve as an increment pulse in an incremental computer. In each of the novel computing elements of Fig. 8 a transistor or bridge limiter senses when the threshold is exceeded and

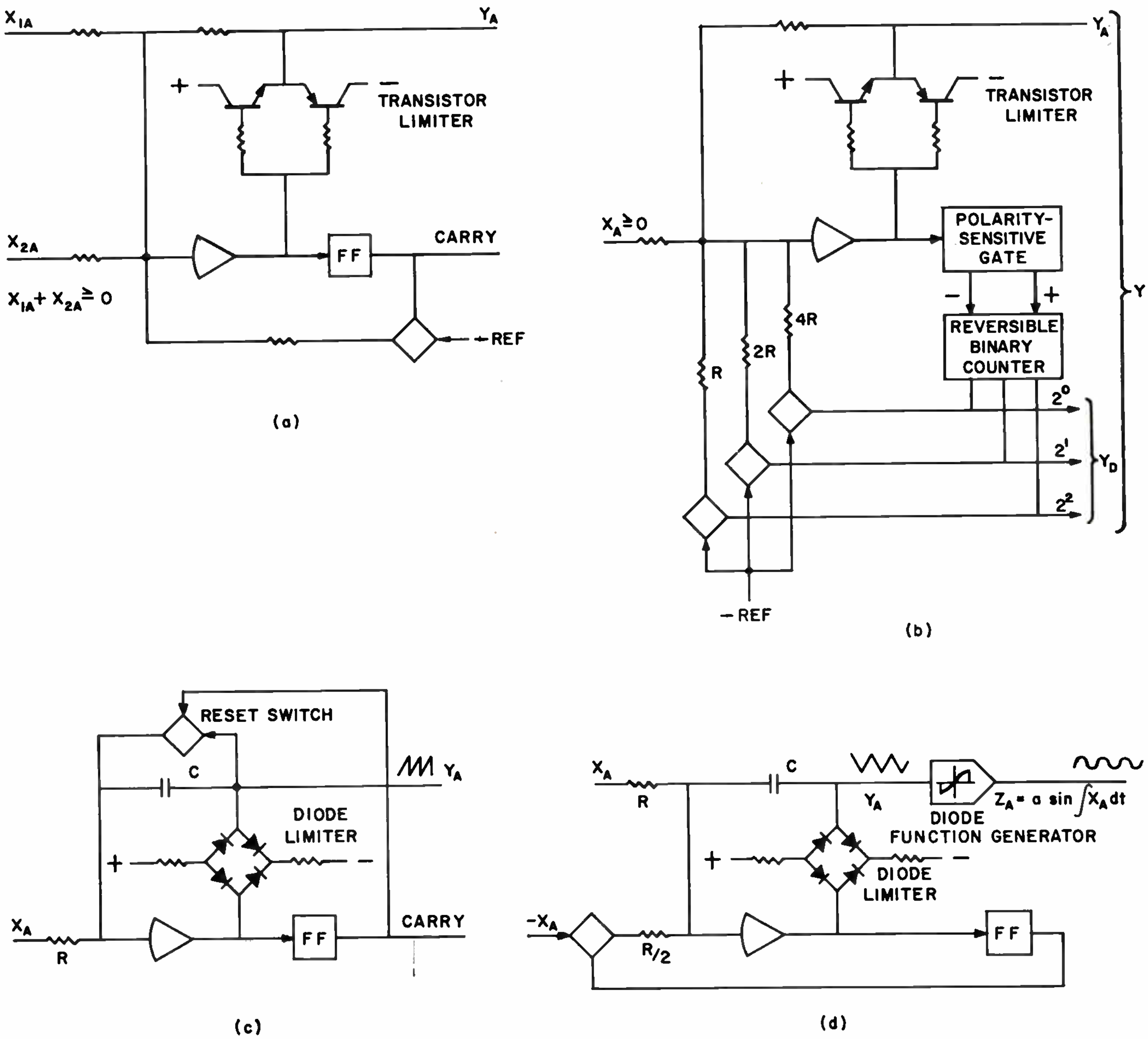
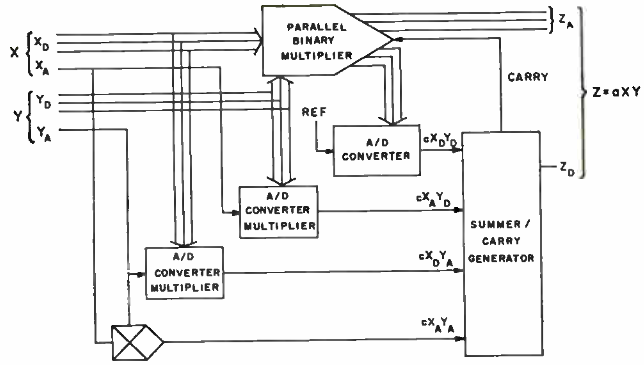
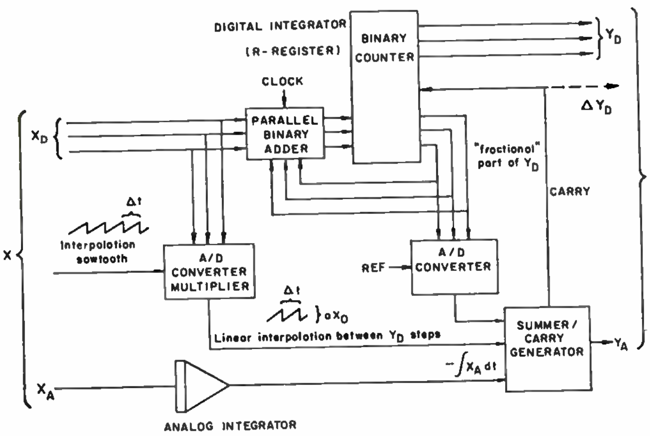


Fig. 8—Development of hybrid computing elements from operational amplifiers with built-in comparators: each circuit resets itself and generates a digital carry pulse whenever the analog output exceeds set limits. The circuits shown are (a) a summer-carry generator [28] (b) a partial analog-digital converter [26] (c) a self-limiting integrator-carry generator [27] and (d) a self-reversing integrator used in a scale-factor-free rate resolver [27].



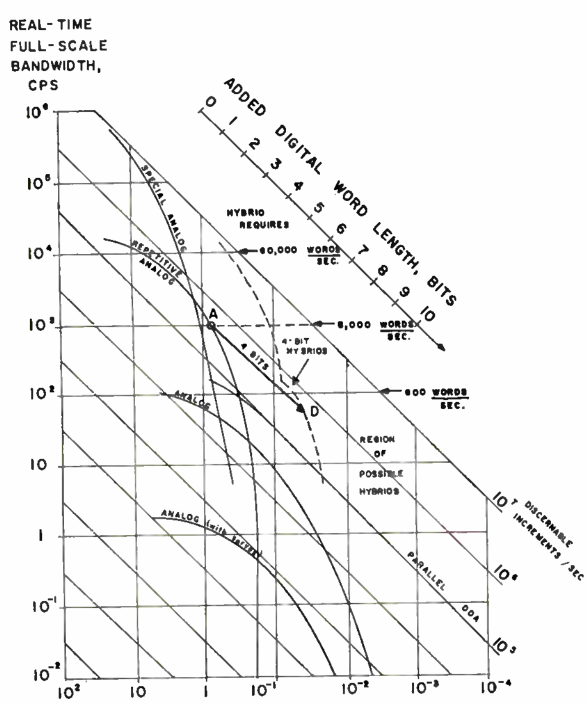
(a)

This true hybrid multiplier [29] combines circuits of the types illustrated in Figs. 2, 7, and 8. For simplicity only positive quantities are considered here. The product $XY=(X_D+X_A)(Y_D+Y_A) = X_DY_D+X_DY_A+X_AY_D+X_AY_A$ is generated by remarkably simple digital-digital, analog-digital, and analog-analog multipliers together with a summer-carry generator. Note that in this example the accuracy required of the analog-analog multiplier is only 1 per cent.



(b)

A true hybrid integrator. A simple 3-bit binary accumulator combines the digital integral with a carry generated when the analog part of the circuit reaches the value corresponding to 1 digital unit. The summer-carry generator analog inputs are 1) a sawtooth of amplitude X_D to interpolate between successive Y_D values; 2) a voltage corresponding to intermediate fractional values of the digital integral; and 3) the analog integral. Again only positive quantities are considered. The carry may be used to generate an incremental integrator output, if desired.



(c)

Gross accuracy-bandwidth relations for analog, digital, and NBS-type hybrid differential analyzers. The resolution-bandwidth product (discernible increments per second) for the DDA is constant, except for effects of increased truncation errors at low accuracies. Analog computers are limited by static accuracy and high-frequency phase shift. The vector AD exemplifies a practical hybrid using, for simplicity, only positive analog voltages (one-half analog full scale—simply add one extra digital bit for machines using negative as well as positive analog voltages). In this example, combination of a 1 kc, 1.5 per cent FS analog computer (point A) and a 4-bit 6000 words/second digital computer yields an over-all accuracy of 0.25 per cent FS with an over-all full-scale bandwidth of 60 cps (point D). Note that the bandwidths obtainable with NBS-type hybrids are limited by the state of the analog-computer art; and note the boundary of practical 4-bit hybrids.

Fig. 9.

interrupts an operational-amplifier feedback loop to produce the desired carry pulse.

The *complete hybrid computing elements* of Fig. 8 [28], [29] are developed from the simpler devices shown in Figs. 5–7. Each variable X is represented by an analog voltage X_A together with a binary number X_D expressing the most significant digits. In the *hybrid multiplier* of Fig. 9 (a) the digital components X_D, Y_D are multiplied in a parallel binary multiplier. Analog carries are introduced into the output of this multiplier to form the correct digital component of the product XY. An analog summer-carry generator accepts the three products $X_A Y_A$, $X_D Y_A$, $X_A Y_D$, which are generated, respectively, by a coarse analog multiplier and two digital-analog converter/multipliers [see also Fig. 7 (a)].

The *hybrid integrator* shown in Fig. 9(b) is the basic element of the National Bureau of Standards and Link hybrid differential analyzers [28], [29]. A digital integrator (binary adder-counter) integrates the digital input X_D with suitable modification through analog carries. A digital-analog converter-multiplier multiplies the output of a sawtooth generator common to all integrators of the machine by the input rate X_D of the output Y_D to interpolate between digital output values. The corresponding analog voltage is combined with an analog voltage corresponding to fractional values of the digital integrator output and with the integrated analog voltage in a summer-carry generator. The digital output Y_D can be transferred directly, or the carry pulses may be treated as increment pulses in the manner of a digital differential analyzer.

Hybrid Analog-Digital Computer Systems

Apart from their use in computing systems as such, hybrid analog-digital computing elements may have just the combination of accuracy and speed needed to save much digital equipment in many data-processing systems. This is especially true in instrumentation and control systems required to handle continuous-variable inputs, for it may be possible to build computing operations into the analog-digital conversion equipment.

Existing complete computer systems utilizing true hybrid computing elements are all of an experimental nature and will serve to indicate the possibilities of such systems rather than to indicate a definite trend.

The *National Bureau of Standards hybrid differential analyzer* [28] employs computing elements of the type illustrated in Fig. 8 with 5-digit digital components. Interest in such hybrid systems stems from the following features:

1) Digital components used need not be excessively fast and require a relatively low number of digits, so that all-parallel logic can be used at moderate cost.
2) Computing speed is determined by the relatively fast analog components.

Note, however, that in *every* computing system required to represent real-time phenomena over a given frequency spectrum *any increase in accuracy requires a proportionate bandwidth increase.* This is true simply because any increase in scale factors requires analog or digital data to change at greater rates.

Hybrid differential analyzers may bridge a speed-accuracy gap between analog and digital computers especially in fast trajectory computations (space-vehicle reentry, range-safety impact prediction). Such computers could conceivably be designed to serve as accessories to increase the speed of existing digital computers. Relatively little development work has been done on hybrid computing elements, and substantial improvements and simplifications are quite likely.

The *Arizona Statistical Repetitive Analog Computer* (ASTRAC) developed at the University of Arizona [21], [30], [31] is an all-electronic repetitive-iterative analog computer intended mainly for Monte-Carlo-type studies of control systems and communications systems [32], [33].

Random-input solutions of system differential equations are generated at rates between 10 and 100 cps; a flexible digital-control unit samples these solutions at push-button-controlled times for a preset number of 1000 to 10,000 computer runs. The sample values are read into a hybrid analog-digital statistics computer which computes probabilities [34] and sample averages of x, x^2, $x(t)\ y(t-\tau)$, etc., for nonstationary as well as stationary random processes. The repetitive computer contains digital multivibrator and diode-logic modules as well as analog memory units, switches and conventional analog computing elements. Random inputs as well as randomly-timed events are obtained from noise generators which generate random telegraph-wave or Gaussian outputs. ASTRAC memory-integrators can be comparator-controlled in the manner of Fig. 4; the all-digital control unit permits convenient control of iterative subroutines. Fig. 10 shows a block diagram of the ASTRAC system, and Fig. 11 illustrates the operation of a hybrid averaging unit in the statistics computer [21].

The possibility of *program storage* on magnetic tape or drums is an attractive feature of hybrid as well as digital computers. A hybrid differential analyzer proposed at M.I.T. as a real-time flight simulator for the U. S. Navy [13], [35] would represent each variable by a positive analog voltage, together with a digital word expressing sign and scale factor (floating-decimal-point position). The analog arithmetic unit comprises parallel analog integrators, but only one analog summer, one multiplier, and one divider, which are time-shared with the aid of electronically-switched sample-hold circuits. A relatively simple digital arithmetic unit performs the corresponding operations on the digital sign and scale words. Analog comparator circuits sense the magnitude of summer, multiplier, and divider output voltages and

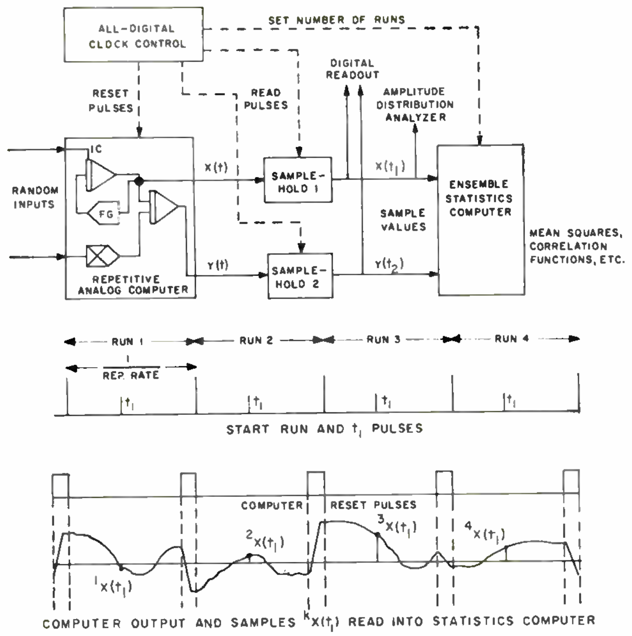


Fig. 10—ASTRAC implements Monte-Carlo techniques by simulating a random-input dynamic system 10 to 100 times per second. Digitally-controlled sample-hold units read out selected samples of dynamic variables $x(t)$, $y(t)$ at push-button selected times t_1, t_2 after the start of each computer run. The statistics computer averages functions of successive sample values to produce estimates of mean-square values, correlation functions, probabilities, etc., for nonstationary as well as stationary processes.

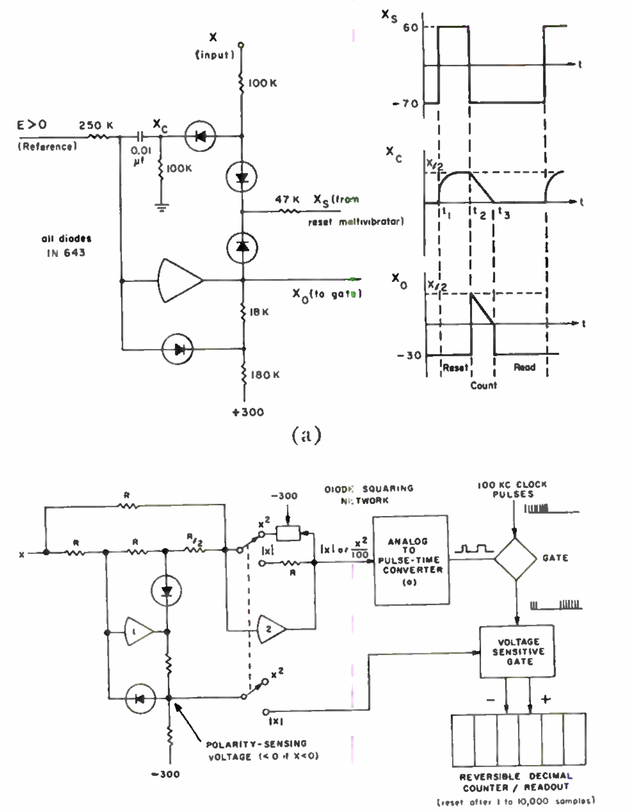


Fig. 11—(a) ASTRAC analog-to-pulse-time converter (b) Hybrid sample-averaging circuit. A precision absolute-value circuit (amplifiers 1 and 2) applies positive voltages $|X|$ or $X^2/100$ to the converter and, if X is to be averaged, reverses the counter input when X changes sign. The decimal readout counter accumulates readings for between 1 and 10,000 samples of X or $X^2/100$. Note that the average of a product $XY/100$ may be read out through alternate sampling of $(X+Y)^2/4$ and $-(X-Y)^2/4$.

automatically change the analog-circuit gain, together with the floating decimal point of the digital system, so that favorable scale factors are selected automatically. The digital arithmetic communicates with the analog arithmetic by gating scale-setting resistors into operational amplifiers.

A hybrid function generator employs digital circuitry, together with analog addition, multiplication, and division. Discrete values of functions of one, two, or more variables as well as values of constant coefficients are entered in digital form on a magnetic-memory drum. These digital data are decoded and entered into the analog-digital arithmetic unit which can perform suitable interpolations. The drum also controls the machine program by a sequence of command pulses which control the successive operations of the arithmetic elements through a switching matrix and a multiplicity of analog and digital gates. No information is entered on the drum during a computer run [13].

The final version of the M.I.T. hybrid computer omitted the floating-point hybrid arithmetic, since it was believed that digital bookkeeping on the exponents associated with each variable would be either too slow or too complicated [38]. The resulting system is, essentially, a very fast multiplexed analog arithmetic unit combined with analog integrators and a digital-control unit. The idea of automatically scaled hybrid representation is, nevertheless, interesting and might deserve more investigation, especially in connection with special-purpose computers.

Analog Variables Other Than Amplitude

Although analog voltages are easy to add and integrate, they are not easily stored on magnetic tape or drums. They are similarly difficult to transmit over long distances, and analog switching, which is required for all automatic programming schemes, requires relatively critical drift-free circuitry. These three disadvantages are readily overcome through analog representations of variables by *frequency modulation*, *pulse-width modulation*, or *pulse-position modulation* of pulse trains. These representations require merely the transmission of timing pulses, which are easily stored, transmitted and gated. Such modulation schemes also permit particularly simple conversion to digital representations with the aid of gated counters. The intriguing possibilities inherent in such a representation were demonstrated in a new and different flight simulator designed at the University of British Columbia, which uses pulse-time representation of data [36]. Although a complete set of new computing elements had to be designed for the simple operations of addition, multiplication, and integration, the simple gating and drum storage made possible by the new representation gives the machine unique automatic-programming capabilities. Methods of this type, possibly in conjunction with operational digital techniques, deserve continued attention in the future and may play a significant part in the continued impact of analog-digital techniques on the analog-computer art.

Conclusion

To summarize, many present-day analog computers exist, in a sense, in a gap left by the cost-speed characteristics of available digital computers. But even as digital computers improve, the all-parallel "live mathematical model" afforded by analog-type simulation remains indispensable to the system designer. As a result, new hybrid computing elements and systems combine analog and digital features, and analog computers are becoming faster. The trends discussed include combination of conventional analog and digital computers, digital programming and checkout of analog computers, operational digital computing elements, new high-speed analog computers with memory and digital switching, true hybrid computing elements, and some hybrid computer systems.

References

[1] A. S. Jackson, "Analog Computation," McGraw-Hill Book Co., Inc., New York, N. Y.; 1960.
[2] G. A. Korn and T. M. Korn, "Electronic Analog Computers," McGraw-Hill Book Co., Inc., New York, N. Y., 2nd ed.; 1956.
[3] J. R. Hurley, "DEPI (Differential Equations Pseudo-code Interpreter), An Analog-computer Simulator," Allis-Chalmers Mfg. Co., Memo dated; January, 1960.
[4] L. A. Ohlinger, "ANATRAM—First step in breeding the DIANALOG," *Proc. WJCC*, vol. 17, pp. 315–329; May, 1960.
[5] J. L. Greenstein and R. M. Leger, "Computer Handbook," H. D. Huskey and G. A. Korn, Eds., McGraw-Hill Book Co., Inc., New York, N. Y.; 1962.
[6] *Proc. Combined Analog-Digital Computer Systems Symp.*, Philadelphia, Pa., December 16–17, 1960; Simulation Councils, Inc., La Jolla, Calif.
[7] H. Levy "New NASA/AF Navy Joint Computing Facility Combines Analog and Digital Techniques," *Datanation*, Beckman/Berkeley Div., Richmond, Calif., 1960.
[8] E. M. Grabbe, *et al.*, "Handbook of Automation, Computation, and Control," John Wiley and Sons, Inc., New York, N. Y., vol. 2; 1958.
[9] M. Palevsky, "Digital differential analyzers," in "Computer Handbook," [5], pp. 19-30 to 19-60.
[10] R. M. Beck and J. M. Mitchell, "DAFT, a digital/analog function table," *Proc. WJCC*, vol. 17, pp. 109–119; May, 1960.
[11] B. D. Loveman, "Semiautomatic programming," in "Computer Handbook," [5], pp. 4-12 to 4-18.
[12] C. D. Morrill, "Electronic multipliers and related topics," in "Computer Handbook," [5], pp. 3-40 to 3-61.
[13] G. A. Korn, "Analog-storage techniques for electronic analog computers," in "Computer Handbook," [5], pp. 6-26 to 6-34.
[14] T. Brubaker, "Electronic Analog Memory," Elec. Engrg. Dept., Univ. of Arizona, Tucson, ACL Memo No. 22R; March, 1961. *Electronics*, pp. 141–143; September 29, 1961.
[15] Eastern Simulation Council Presentations, "DYSTAC Applications," Computer Systems, Inc., Monmouth Junction, N. J.; 1960.
[16] J. M. Andrews, "Mathematical applications of the dynamic storage analog computer," *Proc. WJCC*, vol. 17, pp. 199–233; May, 1960.
[17] L. B. Wadel, "An electronic differential analyzer as a differential analyzer," *J. ACM;* July, 1954. (Note that Wadel's pioneer work precedes the introduction of commercial iterative analog computers.)
[18] L. B. Wadel, "Automatic iteration on an electronic analog computer," *Proc. WESCON Computer Sessions*, pp. 13–19; August, 1954.
[19] M. C. Gilliland, "Iterative differential analyzer function and control," *Instruments and Control Systems*, April, 1961.
[20] W. Brunner, "An iteration procedure for parametric model building and boundary-value problems," *Proc. WJCC*, vol. 19, pp. 519–535; May, 1961.
[21] G. A. Korn, "New high-speed analog and analog-digital computing techniques: the ASTRAC system," *Proc. 3rd Internatl. Conf. on Analog Computation*, Opatija, Yugoslavia, September, 1956; Presses Académiques Européennes, Brussels, Belgium (in print).
[22] C. S. Deering, "A wide-band direct-coupled operational amplifier," *Proc. Natl. Simulation Conf.*, Dallas, Tex., 1956; Simulation Councils, Inc., La Jolla, Calif.; 1956.
[23] H. Koerner, "How to extend operational-amplifier response," *Electronics*, vol. 33, pp. 90–91; November, 1960.
[24] H. Schmid, "Transistor analog computing elements," in "Computer Handbook," [5], pp. 7-2 to 7-29.
[25] L. Bush and P. Oralando, "A perturbation technique for analog computers," IRE Trans. on Electronic Computers, vol. EC-8, pp. 218–222; June, 1959.
[26] H. Schmid, "A hybrid function generator," *Proc. Combined Analog-digital Computer Systems*, [6].
[27] H. R. Eckes, "Self-Resetting Integrator," Elec. Engrg. Dept., Univ. of Arizona, Tuscon, ACL Memo No. 37; 1961.
[28] H. K. Skramstad," A combined analog-digital differential analyzer," *Proc. EJCC*, vol. 16, pp. 94–101; December, 1959.
[29] H. Schmid, "Combined analog-digital computing elements," *Proc. WJCC*, vol. 19, pp. 229–315; May, 1961.
[30] G. A. Korn, "Repetitive analog computers at the university of Arizona," *Instruments and Control Systems*, September, 1960.
[31] T. Brubaker and H. R. Eckes, "Digital control unit for a repetitive analog computer," *Proc. WJCC*, vol. 19, pp. 353–361; May, 1961.
[32] W. E. Vander Velde, "Make statistical studies on analog simulations," *Control Engrg.*, June, 1960.
[33] E. L. Berger and R. M. Taylor, "Optimization of radar in its environment by GEESE techniques," *Proc. WJCC*, vol. 19, pp. 490–507; May, 1961.
[34] T. Brubaker and G. A. Korn, "Accurate amplitude-distribution analyzer combines analog and digital logic," *Rev. Sci. Instr.*, vol. 32, pp. 317–322; March, 1961.
[35] A. W. Herzog, "Pulsed analog computer for simulation of aircraft," Proc. IRE, vol. 47, pp. 847–851; May, 1959.
[36] E. V. Bohn, "A pulse-position-modulation analog computer," IRE Trans. on Electronic Computers, vol. EC-9, pp. 256–261; June, 1960.
[37] "EURATOM/ISPRA, le code APACHE (analog programming and checkout)," *Proc. 3rd Internatl. Conf. on Analog Computation*, Opatija, Yugoslavia; September, 1961; Presses Académiques Européennes, Brussels, Belgium (in print).
[38] M. E. Connelly, "Analog-digital Computers for Real-time Simulation," Electronic System Lab., Mass. Inst. Tech., Cambridge, Rept. No. ESL-FR-110; 1961. (Note bibliography of project reports.)

Mass Storage*

A. S. HOAGLAND†, SENIOR MEMBER, IRE

Summary—One key to expanding the applications potential for information processing systems is the economic availability of on-line memory of extremely large capacity to permit the automatic servicing of record files. Further, mounting emphasis on "management" decision making will accelerate the emergence of true file-centered systems. This paper reviews the evolution and future trends in the implicit "mass storage."

The term mass storage indicates a unit capacity ranging upward from one million characters. Such memory cannot currently be realized or projected with solid-state devices (*e.g.*, fabricated matrices of discrete, bit elements such as core arrays). Thus, continuous (and generally interchangeable) surfaces with associated coupling transducers are exploited.

The access time variability to memory locations, arising from the requisite mechanical motion, makes the structuring and organization of mass memory (software) of major significance to effective systems integration.

Magnetic recording, with its unlimited surface reusability, is the firmly entrenched mass storage technology. The last decade has seen almost two orders of magnitude improvement in storage density, with in excess of 25,000 bits/sq in already used in announced products. The major technical innovation to date has been in the air-floated head which, by permitting both high density and high surface speed, brought about the introduction of mass random access memory (disk array-1956).

Areas covered include systems perspectives and memory organization, memory modularity and mass storage structure, reliability and economic considerations, magnetic recording technology, and nonmagnetic media (including image storage).

For the forseeable future, magnetic recording will continue to be intensively exploited as a major factor influencing the growth of information processing with further major advances in storage density and mass storage performance capabilities developing rapidly.

I. Introduction

THE EXPANDING commercial applications for information processing systems are increasingly dependent upon the economic availability of on-line memory of extremely large capacity for the automatic servicing of sets of records (or files). Utility billing and airline reservations are two examples of file maintenance requiring "mass storage." The growing need in business operations to have masses of organized information readily available for effective management decision making will accelerate the emergence of true file-centered systems.

In the scientific field of data processing, the temporary storage of intermediate results dictates a mass store which can rapidly interchange data with main memory. In addition, the trend towards multiprogramming makes a large back-up store for high-speed reloads and dumps of internal memory (program shuffling) advantageous.

"Mass storage" is a term used here to imply a unit capacity in excess of one million alphanumeric characters (of the order of 10^7 bits). Certain existing mass storage units store more than a 100 million characters of information, and developments now under way will materially extend this figure.

Such memory is far beyond the present or projected economic feasibility of electronic devices. Solid-state memory is generally based on a matrix-like assembly of individual bit storage elements, (*e.g.*, magnetic core arrays) which are directly addressed and selected by a switching net [11]. This type of storage provides high-speed (microsecond) access to any memory location needed for accomplishing data and instruction manipulation within a central processor. In contrast, for mass memory, continuous (and generally interchangeable) storage surfaces with associated coupling transducers are exploited. The recording density (bits per square inch) is principally a function of the three-dimensional mechanical registration tolerances that can be realized between the storage film and the coupling transducer(s). Relative motion is required for access to memory locations. While very large capacities are practical, access times are in milliseconds to seconds, since mass storage involves mechanical movement. The access time variability to memory locations, arising from the requisite mechanical motion, makes the structuring and organization (software) of mass storage a key factor to effective systems utilization.

Modularity is an extremely important concept in mass storage. A module is an independently addressable block of memory, and a given memory system will, in general, consist of a number of such modules. In contrast to a core-type memory, a mass storage module can have a multiplicity of simultaneously addressed and operated access means. The most common attributes used to characterize a mass storage module are: capacity, access time (*i.e.*, the average access time for randomly located information), the instantaneous data rate on sequential information transfer, availability of storage member interchangeability, and hardware-imposed address location formats.

Systems-derived performance factors such as effective character rate and useful throughput (record processing rate) depend not only on the mass storage specifications but also on memory allocation (file layouts), per cent activity of a given file (ratio of records actually processed to total), number and availability of the data transfer channels servicing the memory (the maximum would be one per access mechanism), provisions for queuing and ordering of access requests, checking procedures, etc. The associated control logic is an integral facet of any mass storage appraisal.

* Received by the IRE, June 19, 1961; revised manuscript received, November 22, 1961.

† IBM Advanced Systems Development Division, San Jose, Calif.

The need for operational reliability of mass memory units, which involve the integration of complex mechanical and electronic assemblies, has been and still remains one of the major, even though less exciting, challenges faced in the further development of mass storage systems. The trend to communications-oriented multi-terminal systems with their inquiry requirements and emphasis on real-time response greatly magnifies the importance of "on-line" reliability.

To date, magnetic recording is the technology primarily used for the mass storage of digital information. Although it is only in recent years that magnetic recording has come into wide general use, its invention by the Danish engineer, Valdemar Poulsen, dates back to 1898. The paramount functional advantage of magnetic recording surfaces is their unlimited reusability. This property permits the direct modification of stored information (for digital information storage, binary coding with saturation recording is used). Additional advantages of magnetic recording for mass storage of data over other potential storage film media are: the simplicity of the recording transducer (a magnetic head); the flexibility in mechanical structure possible (and hence, choice of performance specifications) due to the ability to place the storage film on almost any supporting surface in conjunction with ease of mounting a magnetic head; the high bit storage densities and read-write transfer rate obtainable with magnetic recording; and great ruggedness with respect to handling and environmental conditions. Consequent from the above characteristics is extreme memory economy. Typical figures place on-line magnetic recording storage at a cost per bit of 1 per cent or less than that for magnetic core memory. The replaceability and off-line shelf storage potential (*e.g.*, tape reels) makes inactive mass memory less costly by even more orders of magnitude.

Magnetic recording represents the integration of several basic engineering fields and has been generally characterized by rapid progress achieved by evolutionary advances rather than dramatic innovation. The one "breakthrough" that can be identified with the computer field is the air-floated head. Otherwise, advances in the magnetic recording art have largely emanated from increasingly higher precision and quality in components. No technology that might challenge the position of magnetic recording now appears in sight. Thus, an increasing effort will continue to be placed on magnetic recording and associated mechanisms in spite of their already long period of exploitation.

Certain exploratory work has been carried out on photographic films for "read only" types of applications (for example, language translation and telephone exchange routing). In such uses, one can capitalize on an ability to carry out high-speed local searching with the effectively inertia-less optical beam transducer. However, for a large capacity requiring a great amount of surface area, mechanical positioning is required, and that is where the basic access limitations inherently arise in a mass store.

There is developing interest in image storage within the growing sphere of information processing. While concepts and techniques have yet to be generally established, it is in the storage of fixed images that films of photographic and thermoplastic materials have drawn most attention. Such materials permit optical or area-type transducers and hence, with simplicity, the ability to project, minify, magnify, etc. Thermoplastic materials do permit rewriting, and this feature distinguishes investigations into their potential role.

The body of the paper is divided into three major sections: historical evolution of mass storage, present state of the art, and future trends, with emphasis given to key problem areas faced in achieving further progress

II. Historical Growth

The work which ushered in mass data storage was firmly initiated by 1947. This activity was associated and concurrent with the explosive "take-off" of the digital computer field at that time. Since then, progress has paralleled the rapidly expanding growth of data processing. Early work was oriented to the scientific computer market. The initial mass memory device was the magnetic tape unit, to provide both auxiliary "back-up" storage to main memory and input-output (electromechanical terminals) buffering in large-scale systems. The later emergence of commercial data processing brought with it a wider variety of functional usages and mass storage hardware.

Commercial or business data processing, as it was evolving as a main facet of activity in the computer field in the early 1950's, gave a tremendous impetus to mass storage development and had a major impact on its direction. File storage for records maintenance was the central requirement. The volume of business records certainly justified the introduction of the descriptor "mass."

Magnetic tape was immediately exploited for mass records storage and to the present, almost all large-scale commercial data processing systems are tape based. Tape transports are inherently suited for sequential access to information, and conventional batch processing procedures (punched card methods) were therefore instituted. However, a major percentage of available machine time is then needed for the tape sorting of records preparatory to processing against a master tape file. For low file activity, tape devices are very inefficient, due to the constraint of sequential access. Further, effective file inquiry operations are not possible. To keep head and surface wear down to acceptable rates, tape speed is quite limited, representing a severe restriction on bit rate and access time.

The character of business data processing thus indicated the need for an entirely different type of mass storage. The desirability of storing large volumes of information with any record available rapidly gave stimulus to the development of a mass random access memory. A short access time is then necessary and a high-speed rotating surface provides this feature. The extremely high surface velocity then arising is only feasible by providing a slight separation or spacing between the magnetic head(s) and moving surface. This arrangement is categorized as noncontact magnetic recording.

The air bearing supported head (using an air cushion to control head-to-surface spacing) was the innovation which, combined with the above memory concept, brought about this entirely new hierarchy of mass storage, beginning in 1956. By this spacing technique, it was possible to develop a high-capacity rotating disk array since a head could closely follow the appreciable runout of large disks (the disk array permits a high recording surface area to volume ratio).

The first version (the IBM RAMAC 350) could store five million characters with an access time to any record of less than a second. Secondary technical features of note were the use of self-clocking and a wide erase narrow read-write head unit. These design approaches, combined with the use of an air-supported head, provided techniques that compensate for the head-to-track registeration tolerances of such a gross mechanical structure, and thus permitted the high track density and high bit density necessary for large capacity.

Since the introduction of mass random access memory there has been great activity in this type of mass storage and a number of units of rather widely varying configurations are now available. One other, uniquely different in form, is the NCR CRAM (1961), a randomly selectable magnetic card unit. This random access memory also features replaceable storage, formerly only available with magnetic tape. In terms of performance, a disk array has been announced (1960) with a capacity of approximately 90 million characters and an access time of less than 0.2 second.

As one perspective on the progress in digital magnetic recording over the last decade, consider storage density:

RR Uniservo I (tape), 1950—1400 bits/sq in
IBM RAMAC 350 (disk), 1956—2000 bits/sq in
Potter 906 (tape), 1960—24,000 bits/sq in
RR Randex drum file, 1960—30,000 bits/sq in
IBM 1301 (disk), 1961—25,000 bits/sq in.

These specifications are indicative of both the progress and emphasis placed on such storage in the last several years. The cost per bit of on-line mass memory has steadily dropped until now it is in the range of 0.1 to 0.01 cent per bit, a major factor for its increasing acceptance.

III. Present Status

There is such a wide diversity in mechanical configurations for mass storage that attention will be focused primarily on systems perspectives and technology. With respect to magnetic recording, four areas of technical development can be distinguished: the magnetic head, the magnetic recording surface, head-surface registration mechanics, and read-write electronics. Each of these will be briefly discussed in turn; then the present status of storage density, which involves the integrated performance from all these components, will be summarized.

A. Systems Features

Today, mass storage can essentially be distinguished as primarily sequential or random in character of access and, hence, mode of operation. For sequential record transfer, the effective character rate is less than the instantaneous rate due to hardware constraints on address location format. In tape, "interrecord gaps" define record groups for block transfer to main memory and are necessary to accommodate the consequent start-stop tape motion. Interrecord gaps, although much shorter, are also necessary on high-density disk or drum tracks to make feasible "single record" modification.

Magnetic tape has been the predominant means for mass storage, and its unique systems features result from the fact that a tape transport is a sequential access device. The only practical way to use tape is to address by record content. In updating a file, for example, the master and transaction tape reels are serially read (information is arranged and maintained in ordered sequence) and a "new" tape is created, on another transport, with the unmodified as well as the altered records being transferred. While this procedure is inefficient with low activity files, it is easy to make insertions and deletions as well as to handle variable length records, as physical sections of tape have no specific indentity. Further, a tape reel is relatively cheap and therefore low-cost, off-line, archival storage is feasible.

Random access memory involves addressing by physical location to a single record or a particular block of records (one track), which is then scanned. This addressing problem is acute since the "set of keys" (record identifiers) of a file will, in general, bear no direct relation to the closed set of machine addresses. Various randomizing techniques (key transformations) are used to convert scattered keys covering an extensive range to a dense and relatively uniform distribution of numbers to obtain automatic addressing capabilities. Another unique technique derived for random access memory is "chaining" or the linking of address locations (or records) by recording at a given address location the address of the desired succeeding location, etc. Chaining is a systems technique to overcome hardware constraints on "bucket size" but will generally incur the price of cascaded (multiple) access times. Certain hardware allows "selective" record length since intratrack sector addressing divisions are not mechanically imposed.

A disk array (noncontact recording) can easily read and write information at high instantaneous character rates. If the processor "compute" time is closely matched to this record transmission speed, there can arise a serious imbalance with the much longer mass memory access time for purely random access jobs. For this reason, continual cost reduction pressures exist to make more modularity of mass storage feasible, thereby effectively multiplying the number of possible channels that may concurrently service the total store and better mask the mechanical motion times inherent to mass storage accessing.

The high word rates with noncontact recording and multiple-head groups on a rotating unit have made such devices attractive for short-term high-speed main memory dumps and reloads even though this application involves a sequential transfer by its nature.

B. Magnetic Recording Technology

1) Magnetic Heads: Magnetic head design skills have advanced to the point where the principal concerns are to evolve fabrication techniques that will provide an improved and smaller precision assembly. Currently, head gaps as small as 20 microinches can be realized. When thin laminations (from ½ to 2 mils) or ferrite pole pieces are used, magnetic heads can yield considerably in excess of 5000 bits per inch *resolution* with frequency bandwidths extending considerably beyond one megacycle.

2) Magnetic Recording Surfaces: Higher quality, thinner recording surfaces is the mainstream effort in magnetic surface work rather than a search for improved magnetic properties. There is a broad range for the magnetic properties, within which performance is not noticeably affected.

Magnetic films in use today are oxide coatings formed from a dispersion of either Fe_3O_4 or $\gamma-Fe_2O_3$ in an organic binder, and Co-Ni platings. Oxide surfaces today can be coated to somewhat less than 100 microinches in thickness with a 2–5 microinch surface finish. The high quality achieved in magnetic coatings is extraordinary. Disk surfaces are made without any bit defects at 25,000 bits/sq in on a 24-inch diameter disk storing in excess of 5 million bits. The quality of high-performance magnetic tape is equal or superior to this.

Co-Ni platings require extremely good substrates, since for equivalent linear bit densities it is necessary to go to thinner layers than required with oxide surfaces. Co-Ni films have been made as thin as 20 microinches which have proven satisfactory. The unique advantages of metal films are their superior wear characteristics and their inherent suitability to the production of a thin layer.

3) Head-Surface Registration Mechanics: Spacing is the key bit density parameter. By using the boundary layer of air carried by the moving surface for a bearing, it has been possible to operate reliably at spacings down to 100 microinches while following several mils of surface runout. The development of the air bearing head, with its surface-following characteristic, represents the most significant technical breakthrough to date in mass storage.

4) Read-Write Electronics: Read-write electronics represents a development area where considerable future progress may be unpredictably realized since it is the one aspect affecting density least defined in terms of theoretical factors. From early work based on distinct output pulses and amplitude detection, we now see peak sensing and phase modulation techniques being exploited with a consequent capability to operate at higher densities where considerable pulse crowding occurs.

In digital recording, *each* bit must be read correctly (although simple parity checking can detect an error and cause a reread). For this reason, digital magnetic recording is the most demanding application of the art. Some digital recording standards call for less than one error per billion bits read, and the bit size may be less than one thousandth by twenty thousandths of an inch! Transient variations in head-to-surface spacing (*e.g.*, from dust particles) at high densities can cause large signal fluctuations and thereby multiple bit errors. Burst error correction codes and other sophisticated coding techniques now are attracting attention for such high-density digital magnetic recording.

5) Storage Density: Track densities up to 50 tracks per inch are now found in practice. Two criteria establish the limit on this figure: 1) signal-to-noise ratio (head output is proportional to track width); 2) mechanical ability to keep the head and track aligned. The IBM 1301 disk file stores data at 500 bpi[1] and 50tpi[1] while the Potter 906II tape unit records at 1500 bpi and 16 tpi (the low ratio of tpi to bpi on tape being typical).

C. Nonmagnetic Media

A photoscopic disk memory has been developed to serve as a dictionary for language translation. By its very nature, this information is "permanent." The model stores 30 million bits (2000 bpi, 700 tpi) with an access time of about 50 milliseconds. Optical sensing permits "servo tracking" (through an electron beam) with extreme sensitivity, allowing a much closer balance to be obtained between bit and track densities than in magnetic recording. A dictionary is an ordered file organized for rapid lookup. A search operation can be executed very efficiently with this device because of the speed of track-to-track repositioning with an electron beam readout mechanism. This application represents an excellent case where the advantages of a photographic material for a file are capitalized upon. Another such example is Minicard. The Minicard system is a document retrieval system in which photographic microimages (60:1 reduction) are placed on a card along with digital index information.

Thermoplastic tape as a medium for storage is a relatively recent innovation. It has limited reusability, but to date the mechanism must be kept in a vacuum, which limits design flexibility. An electron beam serves as the transducer. Compared to magnetic recording, writing is

[1] bpi = bits/in; tpi = tracks/in.

very slow and reading requires a relatively complex optical system, although its inherent compatibility with electron-optical transducers again leads to a high storage density potential.

IV. Future Trends

A. Systems Organization

Key systems problems, extending from applictions methodology and procedures to machine organization concepts, accompany the growing role of mass storage. Software will be as significant as hardware activities. We are at the beginning of the era of *file-oriented* data processing systems. "Real time" will become more meaningful in commercial information processing systems as there is more and more emphasis on management decision making. The evolution of low-cost random-access mass storage will bring about a file-oriented systems approach to all levels of business. Modularity of mass memory permitting continuous growth in capacity will be increasingly stressed.

The trend in the future will be on random plus sequential capabilities for mass storage. Consider a typical random access disk memory with an average access time of 100 msec, a character rate of 50 kc, and containing a file consisting of 100 character records. The effective record rate for sequential flow is nearly 500 per second, while for random access transfers, this figure would drop to 10. This wide differential highlights the desirability of planning around a random plus sequential (or batch) philosophy.

The control by programming of recorded format indicators will become standard, permitting dynamic masking and layout of memory space. The assignment of storage space for key-address combinations (index files) and their manipulation rather than record movement will be emphasized. Table lookup and cross-reference techniques must be improved by introducing special "rapid scan" features to optimize "built-in memory logic." Facilities will be introduced for address and record queues within the store with priority logic capabilities to maximize channel throughput. Sophisticated techniques and hardware combinations will be developed to provide for ready report generation without having to preplan for a desired correlation order. Communications-oriented processing systems will necessitate mass storage "buffer exchange" capabilities to handle record assembly and distribution traffic as multiterminal systems become commonplace.

A major problem in file-centered systems is establishing the proper balance in mass storage configurations. Hierarchies of memory modules (of widely varied performance parameter mixes) integrated to meet total systems requirements can best optimize over-all systems performance, recognizing that cost tends to be inversely related to access time for a given capacity. Mass storage control logic will eventually extend to the functional ability to readily rearrange records within the hierarchies of memory according to their activity, significantly improving access utilization. The concept of several processors working against a mass store will become established, radically changing the perspective of what is an "attachment" and what is the "mainframe."

For the immediate future, off-line storage must be used where possible for economy, but in time it should be possible to transfer increasingly inactive information successively to very cheap (and consequently long access) sections of storage until it truly becomes archival. In time, cost reductions will also permit a relatively high percentage of unassigned memory space. This fact will greatly simplify memory organization.

B. Mass Storage Structure

There will be continuous efforts at cost reduction to make larger mass storage with more modularity economic. On-line mass storage cost is a major barrier today to numerous applications susceptible to data processing, some proposed systems requiring the on-line storage of billions of characters. Manufacturing production techniques are rapidly improving the cost picture, indicating mass storage systems will also soon impact the conventional electromechanical accounting equipment area. Overlapped (read-compute-write), multichannel file operations will be conventional, giving an effective balance between access time, record transfer time, and process time to maximize throughput. Present systems configurations tend to be distorted in being dominantly either file limited or processor limited. The storage of interlaced (or multiplexed) data and/or the balancing of mechanical motions between surface and transducer movement will be introduced as a technique to control character rate independently while permitting the upgrading of capacity and access speed as technology progresses.

Looking to the future, there appear to be two unique approaches that will eventually characterize the principal on-line mass storage systems component. One is the disk array and the other is the magnetic strip file. The former offers a more economic access modularity, although this advantage is achieved through acceptance of a common disk drive (single shaft), and a low-cost "turntable" potential for smaller capacity storage modules. The tape strip memory basically uses a tape strip cartridge in the manner of a tape reel, emphasizing simplicity of interchangeability as well as high volumetric efficiency for extremely large capacity. For logging of journal-type data, tape still seems without a peer.

C. Reliability Considerations

A major problem which must be approached from a systems (as well as a technical) viewpoint is that of high reliability. Achieving extreme reliability in complex electromechanical assemblies, which are at the same time intimately integrated with high-speed, low signal level electronics, represents a challenging task. Increasing performance demands on technology coupled with increasing requirements for on-line "real time" file availability compound this problem.

Currently, in many random access memories, it is standard practice to provide automatically an address compare by recording the address of the physical location in the first section of the record located there. It is highly advantageous to make a readback check after writing to provide confirmation of proper data recording before the newly created or received information is eradicated from the transfer buffer area. While some tape transports provide immediate checking by means of a special dual element head unit, a separate read cycle has been required after writing in head positioning devices. Special head structures to provide this checking feature automatically as well as address compare without an access time penalty must evolve for all forms of mass storage. This capability will also bring about automatic data reallocation without system interruption when defective regions are detected.

Systems organization must provide for the requirements of operational reliability by suitably accommodating occasional hardware failures and maintenance needs. Replaceable storage concepts will be exploited since the storage member can then be placed on another mechanism to continue file processing during "down time."

D. Technology

Investigations made of a magnetic recording unit regarded as a communications channel will yield bit density gains by the acceptance of more sophisticated recording electronics. To date, progress has primarily been based on scaling down "geometry"; however, with no new storage technology warranting a redirection of activity, signal handling techniques must eventually draw major attention.

There has been and will be continuing exploratory research on novel transducers. However, no physical phenomena for a magnetic recording transducer appear competitive to the simplicity and flexibility of a magnetic head. Rather, the next major stage of progress in the transducer area will be automated head fabrication using thin film technology (rather than watchmaker-like assembly), resulting in cost reductions which will have an impact upon the hardware composition of mass storage mechanisms.

Reduced head-to-surface spacings, approaching theoretical limitations for air bearing heads of about 25 microinches, will bring about a large increase in noncontact bit density. The recording frequency bandwidth is inversely proportional to spacing (until this parameter is reduced to the order of the head gap and recording surface thickness). Track seeking and following servo access techniques, which can circumvent the track density limitations imposed by the buildup of cascaded mechanical dimension tolerances, will be applied to effect significant increases in this parameter. Within the next few years, one million bits/sq in (*e.g.*, approximately 5000 bpi, 200 tpi) will become the "state of the art" in digital magnetic recording.

There is continuing work looking toward the development of other, and hopefully better, film media for mass storage. There appears to be no evidence to suggest that a replacement can be anticipated for magnetic recording surfaces. Further, the upper bit density limit of magnetic surfaces is far in excess of present practice and laboratory studies, which are now set by achievable mechanical registration tolerances.

Prophesying the very long-range future course of mass storage, it is intuitively felt that eventually all such memory will not only be digital in form but also electronic in character. It seems evident that this latter "revolution" will not come about from further progress in the conventional matrix store as presently conceived and pursued. However, the studies gradually being initiated in, a) the field of neuron net simulation and the memory function, and b) microminiaturization, lead to the conjecture that one day these fields may merge and permit entirely new means and concepts to be applied to the problem of realizing mass memory of extended capability. For the foreseeable future, however, it is clear that mass storage based on magnetic recording has a vital and increasingly important role in information processing.

Bibliography[2]

[1] W. S. Buslik, "IBM magnetic tape reader and recorder," *Proc. Eastern Joint Computer Conf.*, New York, N. Y., December 10–12, 1952; pp. 86–90.

[2] G. W. King, G. W. Brown, and L. N. Ridenour, "Photographic techniques for information storage," PROC. IRE, vol. 41, pp. 1421–1428; October, 1953.

[3] G. E. Hanen, "Air floating, a new principle in magnetic recording of information," *Computers and Automation*, vol. 2, pp. 23–25; November, 1953.

[4] T. Noyes and W. E. Dickinson, "Engineering design of a magnetic-disk random-access memory," *Proc. Western Joint Computer Conf.*, San Francisco, Calif., February 7–9, 1956; pp. 42–44.

[5] D. N. MacDonald, "Datafile—A new tool for extensive file storage," *Proc. Eastern Joint Computer Conf.*, New York, N. Y., December 10–12, 1956; pp. 124–127.

[6] M. L. Lesser and J. W. Haanstra, "The random-access memory accounting machine, system organization," *IBM J. Res. and Dev.*, vol. 1, pp. 62–71; January, 1957.

[7] W. W. Peterson, "Addressing for random-access storage," *IBM J. Res. and Dev.*, vol. 1, pp. 130–146; April, 1957.

[8] M. M. Astrahan, "Role of large memories in scientific communications," *IBM J. Res. and Dev.*, vol. 2, pp. 310–313; October, 1958.

[9] D. D. McCracken, *et al.*, "Programming Business Computers," John Wiley and Sons, Inc., New York, N. Y.; 1959.

[10] W. E. Glenn, "Thermoplastic recording," *J. Appl. Phys.*, vol. 30, pp. 1870–1873; December, 1959.

[11] J. A. Rajchman, "Computer memories: A survey of the state-of-the-art," PROC. IRE, vol. 49, pp. 104–127; January, 1961.

[12] A. S. Hoagland and G. C. Bacon, "High density digital magnetic recording techniques," PROC. IRE, vol. 49, pp. 258–268; January, 1961.

[13] *Proc. Symp. on Large Capacity Memory Techniques for Computing Systems*, Information Systems Branch, Office of Naval Res., Washington, D. C., May 23–25, 1961.

[14] A. S. Hoagland, "A high track density servo access system for magnetic recording disk storage," *IBM J. Res. and Dev.*, vol. 5, pp. 287–296; October, 1961.

[2] A selected bibliography of pioneer work or generalized treatments.

Eyes and Ears for Computers*

E. E. DAVID, JR.†, SENIOR MEMBER, IRE, AND O. G. SELFRIDGE‡

Summary—Attempts to mechanize character reading and speech recognition have greatly accelerated in the past decade. This increased interest was prompted by the promise of computer inputs more flexible in format than punched cards or magnetic tape. Research has shown that automatic sensing can be done reliably if the task is suitably delimited. Cleverly designed marks on standard forms can be both machine and man readable. A single type font or a few fixed ones are tractable if the print quality is controlled. Handprinting can be handled for careful writers, as can meticulous handwriting. Isolated spoken words taken from a small number of talkers and a limited vocabulary can be automatically recognized. Typical error rates for these machine-sensings run between 0.5 and 25 per cent. These results imply that reading unrestricted typestyles, handwritten scrawl, or recognizing conversational speech is beyond the reach of present methods.

From the engineering viewpoint, questions of values enter. Might it not be wiser to punch cards or tape while making copy rather than depend upon complex character recognition hardware? Is it useful to have voice input to a computer when a finger and typewriter are available? Answers to such questions will depend upon the specific application. Certainly, the utility of automatic sensing will depend upon what is to be done with the material after it enters the computer as well as the internal organization of the machine itself. Perhaps all would agree that we should have automatic inputs before the Russians, but it is not so clear that we need them soon as practical computer inputs.

As MAN RUSHES to build his replacements, he notices an interim requirement for man-machine communication. In the meantime at least, computers must be able to, but cannot, understand the writing and talking of men. We are protected from technological unemployment so long as we are buffered by punched cards, magnetic tapes, and on-line or off-line printers. But the day will come!

Current computer inputs are peculiarly inflexible in format, and inconvenient in use. To communicate with the machine by these means we almost invariably must employ an extra man or man-plus-machine, who transcribes our words into shaped holes exactly positioned or magnetic pulses precisely timed.

Men profess to be ill-suited to such menial labor. Though women seem more easily subjugated, they too tend to be expensive, slow, and inaccurate. Therefore, we are increasingly calling on computers to handle those highly variable input signals whose relevant features are not simply related to their physical properties.

In the early days of computers we all learned a congeries of theorems by Turing and von Neumann which told us (or so we thought) that a computer could do anything we told it. We would merely (!) have to specify sufficiently accurately just what it was we wanted the machine to do. And it is true that some highly variable input signals can be categorized by elaborate, exhaustive programs, but it is just not feasible thus to program recognition of printing, speech, handwriting, radar and sonar signals, and objects in photographs (clouds in satellite weather pictures, for example).

Such material can be converted to a standard format, of course, by a suitable scanner, microphone or other transducer whose output is amplitude-quantized samples. We are not much better off, because we then have the same excess of data in two forms, neither related to the classificatory representation we want in the end. Rather, for instance, from speech waveforms we ask a phonetic or literal transcription, from printing and handwriting a literal ("alphanumeric") transcription and not a TV raster, from radar a target inventory, and from weather pictures a cloud map. Such representations imply abstraction of raw data into more meaningful perceptual coordinates. Computers able to convert inputs to such forms will be much more flexibly, powerfully, and economically coupled to the real world.

Though such abstraction is difficult, we already have given some of our machines limited ability to read printing in certain type faces [1], [2]. But reading scratch-pad handwriting or transcribing conversational speech by machine is far beyond our ken. Also, it seems clear that, in any case, the computer needs much extra, expensive equipment to handle the patterned inputs we are discussing. Yet as computers get faster, and as we can program more sophisticatedly, input flexibility will become more crucial than it already is.

Computer engineers, in a spasm of interdisciplinary enthusiasm, are rediscovering that perception is hard to understand and even harder to simulate. Armed with the tautological knowledge that computers *can* do what we tell them to do, we tell them to perceive and then ask the psychologists to fill the gaps in our instructions. They don't know how to do it either, but *they've* known it longer. Pattern recognition, which refers to the actual categorization involved, the actual decision that some sound or sight is "mother," is learned by children at an early age. Computers are very young too and must learn pattern recognition; we do not say that pattern recognition must come full-blown from the machine unaided and untaught. Some men, at least, can construct and write pattern recognition programs. That is, men can do the learning and generalizing, and then

* Received by the IRE, August 11, 1961; revised manuscript received, January 8, 1962.
† Bell Telephone Laboratories, Murray Hill, N. J.
‡ Lincoln Laboratory, M.I.T., Lexington, Mass., operated with support from the U. S. Army, Navy and Air Force.

instruct the machine. But it does seem to us that in the long run the machine will have the ability to carry some of this burden itself and will, in fact, need to do so as it tackles subtler problems.

METHODS [3]

Automatic input (that is, without a man) is the classification of input signals (images or sounds) into categories. Basically, the objects or signals are presented one at a time to the input device which makes certain measurements on each. These measurements are inputs to a decision "logic" which assigns the input to one of the available categories [4]; that is, the machine measures and then decides.

What distinguishes the techniques we are discussing here from customary ones is that they concern recognition of complex variable patterns of many elements rather than detection of a few fixed elements. For instance, alphanumeric text from a typewriter comprises some 80 complex shapes. For machine recognition of a particular font, however, templates of these shapes can be prepared and input letters matched against these, reducing the recognition process to one of identity. Of course, the input letters must be (or must be transformed to be) the same size as the templates and must be properly oriented. *Template matching* then, involves samples or pieces of samples of actual or typical or ideal inputs in their original representations. The method gives accurate results for any finite alphabet of shapes each of whose examples are identical within the resolution of the processor.

Unfortunately handwritten or handprinted letters, spoken words and even some printed text typically show wide variations from any set of templates which can be prepared in advance. We might provide a metric to measure the degree of fit of the unknown to every template and also supply transformations to normalize the unknown input to reduce the variability. These measures do not rescue the situation. The greater the variability, the less successful template matching becomes. By human standards, very little variability may reduce performance to unacceptable levels.

We conclude from history that template matching in its pure form, or modified with a metric and normalization, is adequate perhaps only for printing and typewriting. It was evident (after some years) that in many cases people do not work with templates, but prefer to abstract features. Parallel lines, cusps, or totally-enclosed spaces, for example, are features that may aid in separating alphanumerics.

Characterizing shapes by their distinctive features or properties may ease the problem of variability. Such properties tend to be invariant over a far wider range of either hand- or machine-printed letters than can be accommodated by a like number of templates. A list of the properties of known samples can provide a basis for classifying unknown ones. This *property-list matching*, however, does not eliminate the variability problem but transfers it to another level. The "invariant" properties must be recognized regardless of their size and orientation, despite the variability inherent in human motor performance, despite our individual personal habits, and despite the "noise" in reproduction devices. Yet property or feature extraction is a step in the right direction and has proved to be less fallible than a "grand" transformation intended to reduce all inputs to one of a finite set of precise templates.

If the property list is appropriately designed, recognition need not hinge strongly on any one feature, and may incorporate correlated or dependent properties for correcting errors in feature extraction. That is, the property-list technique may work well enough to be useful in situations where template matching would not be useful.

In either case, recognition performance can be improved by using context. Most obviously, letters and sounds make words, but even knowing digram letter or syllable frequencies can improve decisions about letters or phonemes. We may be more sophisticated and use words, perhaps incorporating their syntactical and grammatical restrictions; we might even use semantics. We guess that using more context will mean using much more storage—the number of English words being vastly greater than the number of letters and the number of semantic possibilities overwhelmingly greater.

Again, both template and property-list methods can be supplemented by limited "learning" or "self-organizing" or "adapting." In effect, properties in a list or elements of a template are examined statistically for their relevance to particular categories by noting their influence in determining correct and incorrect classifications. Those leading to correct decisions are reinforced; that is, their influence is increased, and vice versa. The "adapting" can be programmed into the recognition system itself if so desired [5]. Thereby, the machine may vary its parameters to take advantage of "good" features—"good" features being those which make for correct classification. Further the machine may be able to make large-scale adjustment in its parameters adaptively. For example, guessing that a piece of text is in English rather than in French would call up different *a priori* letter digram and word frequencies. The flexibility of computer inputs can be increased in this way, but the method has been explored only scantily.

The ability to cope with new input patterns hinges upon the relevance of the features incorporated in the early stages of the recognizer, or upon the relevance of templates which can be formed from the available elements. Yet to be explored at all is any method, other than random combinations of available elements, for formulating new features to be tested for relevance.

Thus the crucial, often the least mentioned, aspect of recognition is the selection of suitable features. In many studies, particularly involving alphanumerics, experimenters have been guided by their intuition. Others have taken the exhaustive (and exhausting) approach. The accomplishments of these methods ride on *ad hoc* inspiration and dogged persistence. Still other researchers have consulted generating mechanisms for speech and handwriting, searching for constraints to confine the range of choices. Thus in speech recognition, studies of vocal tract acoustics and speech synthesis have established the vocal resonances as central features. Also at the contextual level, language and grammar reveal how phonemes and words are contrived to make phrases, clauses, and sentences. At its best, successful *synthesis* reveals a vocabulary of "atomic" features. These then provide the acid to attack verbal or literal compounds. At its worst, synthesis misleads the unwary to accept oversimplified, overspecialized elements which have no counterparts in the league of everyday usage.

Yet to the extent that it reveals the structure of words, letters, or sounds, synthesis is not to be ignored, and it provides a rationale beyond the helter-skelter of randomness. Studies of human perception and sensory mechanisms too, hold a promise of "features-to-come," but the meat is yet to be put on the table even though the aroma whets the appetite. Here also, bona fide knowledge of behavior and physiology can aid in avoiding naïveté.

Character-Recognition Devices

The goal here is to introduce alphanumeric, text into the computer, including *printed* text, handwriting, and handprinting. In the first the question of segmentation is usually not bothersome, since separate characters are completely isolated [6]. The characters are usually arrayed in horizontal lines, but subscripts and superscripts (as in reference notes and mathematical formulas) must be handled by any technique claiming to be complete. Fig. 1 shows a range of relatively easy to relatively challenging examples.

The first technical choice arises here: what range of fonts shall be admitted? There are many useful cases where just one font does suffice: an instructive example is that of the American Banking Association. Realizing that fonts intelligible to people are hard to make intelligible to computers, they settled on one that was intelligible to computers and only barely intelligible to people. (This example is also interesting because it is not a visual one for the machine. The font is printed in magnetic ink and scanned by magnetic sensors.)

If many fonts are to be recognized, we may suppose that they have been previously specified, and that successful recognition of the characters to some extent requires recognition of the font they come from. This case then must be distinguished from the most general one, when all fonts must be recognized, even those not met before—though perhaps we may exclude those like *The New York Times*, Fig. 1.

The next big choice is the extent to which variation

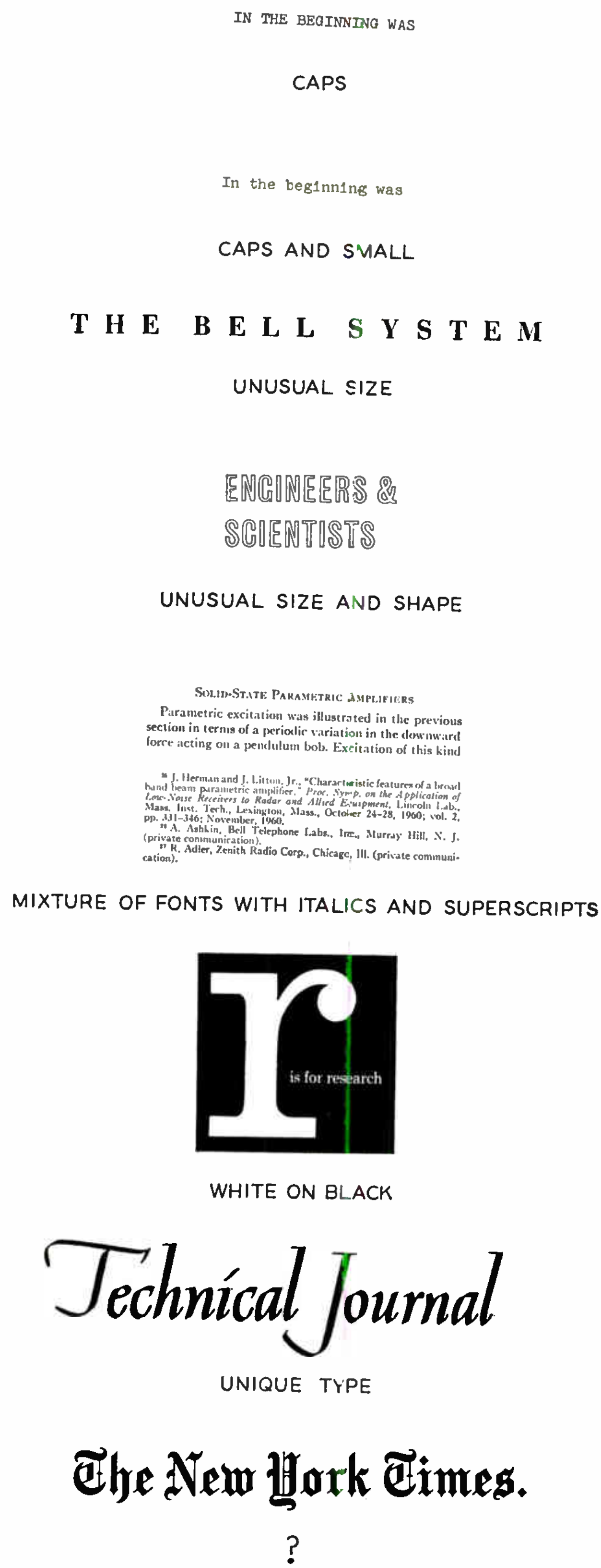


Fig. 1—Variations of type in common use complicate the automatic reading problem.

can be tolerated. We must expect a certain amount of additive or subtractive (or even multiplicative) noise, which should not (but usually does) seriously degrade performance. The noise is usually not independent from resolution cell to resolution cell, as has often been assumed for analytical purposes. In fact, noise comes in clusters like infections. There may be errors of registration, especially from typewriters, and occasionally errors of rotation. Probably the reasonable standard here is that errors from noise or misregistration should be no more frequent than typographical errors.

Now we reach the choice of techniques, as we described them above: template matching is well adapted to known fonts, especially when one can rely on good registration. Many commercial programs use what is essentially template matching. It cannot be said that these are entirely satisfactory, though some are interesting and ingenious.

Finally, there is a choice of mechanism. Though, of course, the critical element is some photosensitive material, it may be preceded by much or little optical processing, which can, in effect, do a large amount of parallel processing cheaply and fast. For example, in Fig. 2, two photosensitive elements can detect at once the degree of template matching of the image with a known character. Many people, however, have preferred scanning as a desirable compromise: here a vertical row of (typically 5 to 20) photocells sweeps horizontally along the line of type and the matching is done electronically. There are some advantages: misregistration in the optical example of Fig. 2 can scarcely be handled except by actual relative motion of the template and image, but electronically there are several plausible ways, for example, treating the lowest active channel as a base-line. Scanning does not restrict us to template matching but can extend fairly directly to property-list matching. The property list will naturally be biased by the fact of horizontal scanning, but should nevertheless be adequate. For either "matching," the stored patterns can be incorporated as logical circuitry. This acts as a "convoluted slot" through which the input pattern must pass. Further, extension to different fonts need not require physically new templates and one can, at least in principle, extend to completely new fonts.

There have been many character recognition experiments which illustrate these techniques and choices. Results from representative ones are listed in Table I [7]-[9].

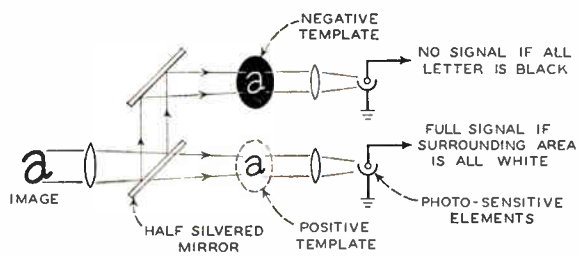


Fig. 2—Template matching is a useful technique where the type face can be specified in advance, but misregistration can reduce its effectiveness.

TABLE I

Investigator	Input Representation	Vocabulary	Number of Writers Tested	Approx. Error Rate	Additional Facts and Comment
Baran, Estrin	X-Y matrix	Machineprinted numerals	One machine over ribbon life	9 per cent	Maximize *a posteriori* probability (Bayes' Eqn.)
Highleyman	X-Y matrix	Machineprinted numerals	One machine, 50 heads	0.6 per cent	Cross-correlation against probability matrices
Bledsoe, Browning	X-Y matrix	Handprinted alphanumerics	1 person	22 per cent	Matching 2-tuples of matrix elements against table
Lewis	13 features	Machine and hand-printed letters	15 mixed people and machines	18 per cent	
Bomba	Local features	Handprinted alphanumerics	2 persons	0 per cent	Decision tree
Doyle	Local features	10 Handprinted letters	≈20 persons	12 per cent	Maximize *a posteriori* probability (Bayes' Eqn.)
Kamentsky	Presence or absence of intersections of character with scanning lines	Handprinted numerals	≈20 persons	1.7 per cent	Correlations of code lists which describe scan interactions
Sherman	Nodes describing line endings and junctions	9 Handprinted letters	≈20 persons	5 per cent	Connection matrix describing graph which joins nodes; tested against master matrices
Marill, Green	Distance of character from field edge along selected line segments	3 Handwritten letters	10 persons	3 per cent	Likelihood function assuming independent normal distribution of measures
Highleyman	X-Y matrix	Handprinted numerals	50 persons	6 per cent	Linear decision functions

Notwithstanding these possibilities, commercial ventures now are concentrating on optical and electronic template matching and hope to achieve speeds startlingly small compared with the easily attained 50,000 characters per second from modern tape units. It does not seem, though, that this limitation is inherent in the techniques themselves because photosensitive elements can certainly have better than a microsecond response time. But 50,000 characters per second is on the order of 18,000 words or some 30 pages per second, which is dangerously rapid for paper.

An interesting exception to the previous paragraph is the IBM 1418 [10]. A vertical line of 17 cells moves through 12 positions to blanket a character. The cells give binary responses (setting the thresholds is tricky) which are stepped into a 12×17 shift register. The character is thus normalized since at some stage it is centered; errors of rotation are not allowed for. The machine uses a large number of empirically derived features as inputs to a binary decision tree. It can recognize some, though not all, of the numerical printing of IBM printers, with a rejection rate as low as one in 10^5 or 10^6, but more typically one in 10^3 or 10^4. Extension to full alphanumeric coverage should be merely a matter of time and demand. The speed is 480 letters per second or about four Kennedys [7], [11]. (As many as 7 pieces of paper per second can be handled, though not necessarily processed.) The system's speed is thus about 20-db below available tape mechanisms, which can handle a full length novel every 6 seconds or so.

It is proper to ask who or what would want to read very many full length novels at 6 seconds each. The point is that to tie up a large digital computer at much slower speeds is expensive; thus, we conclude that probably for some time the print-reading device will include a buffer and will run essentially off-line.

This conclusion will limit for some time effective application of adaptation and learning. These have a depth of structure difficult to handle without a full general-purpose computer, which, as we have seen, we are reluctant to tie up while pages are slowly turned.

The acquisition of more general devices will, therefore, depend greatly on the need. Leaving aside the question of very high speed, who needs all-font capabilities and is willing to pay for it in computer time or extra equipment? The intelligence services might be expected to need something to read *Izvestia* and *Aviation Week*, but so long as the computer is acting as a mere transcriber or storer, the effective bottleneck is the man who must read the material to understand it. (And we are a long way, indeed, from a computer's understanding any but the *most primitive* English.)

If the reader, working with the electronic scanner, could show and identify samples of every letter for a piece of text, the scanner could then proceed without him. In reading *Izvestia*, at least, this could relieve the man's boredom by occasionally replacing one level of tedium with another, and would not necessitate complicated learning programs. But even this much would be a very large and complex piece of electronic equipment indeed, and even so, it would incorporate little of the sophisticated learning techniques many of us espouse as intellectual goals.

A hopeful sign is the increased emphasis on multiprogramming. If we can find ways of interleaving large programs, we may be able to use the computer to supply the sophistication for character recognition while it is performing unconnected computations most of the time. Were this possible, it would be more feasible to take advantage of context to correct character misrecognition or typographical errors [12].

One branch of character recognition where context is clearly essential is handwriting. The segmentation problem for characters is very difficult and that for words is hard enough. Handprinting has received some fair amount of attention (chiefly because it avoids the segmentation problem) though the applications are obviously limited. Representative results are listed in Table I [13]–[18]. An easy solution here seems tantalizingly out of reach: an IBM exhibit at the Western Joint Computer Conference in Los Angeles, in 1961, had a *small* machine to recognize handprinting of numerals only, but they had to satisfy a set of rules about which the illuminating comment was often heard, "Well, I could live with those rules, but I doubt if most people would."

Pure handwriting recognition is being studied at Bell Telephone Laboratories by Harmon and Frishkopf [19], and nowhere else that we know of. Actually, they treat not just the completed writing, but the dynamic formation of the strokes, and their technique is not easily extendible to previously written material. An interesting program by Eden and Halle [20] of the Research Laboratory of Electronics at Massachusetts Institute of Technology generates handwriting, but the constraints inherent in it have not as yet been applied to handwriting recognition.

Automatic Speech Input

Reduction of speech to machine-readable form is a problem of long standing. We do not know the origin of the idea and it is probably lost in antiquity. One early instrumentation of this notion was revealed by Paget [21], whose work in the 1920's on the basic nature of speech set the stage for much that followed. His contribution, among other important ones, demonstrated a correspondence between the individual phonetic symbols that represent a spoken language and the acoustic phenomena associated with the human voice. Thereby the voice-operated toy, "Radio Rex," became feasible. It consisted of a celluloid dog with an iron base held within its house by an electromagnet against the force of a spring. Current energizing the magnet flowed through a metal bar which was arranged to form a bridge with two supporting members. This bridge was sensitive to

500-cps acoustic energy which vibrated it, interrupting the current and releasing the dog. The energy around 500 cps contained in the vowel of the word "Rex" was sufficient to trigger the device when the dog's name was called. Modern speech recognizers utilize this same principle, elaborated and differently instrumented to be sure, but based upon similar phonetic-acoustic relations.

Such relations arise from the properties of the human vocal tract. In speaking, different positions of the tongue, lips, and jaw give the vocal tract diverse shapes. The shapes and gestures correspond to the various phonetic symbols. Each shape gives rise to a distinct frequency spectrum and each gesture to a spectral transition [22]. Acoustically, then, speech can be considered as a succession of spectral steady states and transitions. These features are in close correspondence with the phonetic content and we may see them clearly in a sound spectrogram which displays a "running short-time" speech spectrum, Fig. 3 [23]. Note that over most of the time scale, the spectrum contains three or four concentrations of energy, which arise from the natural modes of the vocal tract when it is excited by short pulses of air from the vocal cords. These energy concentrations are known as formants.

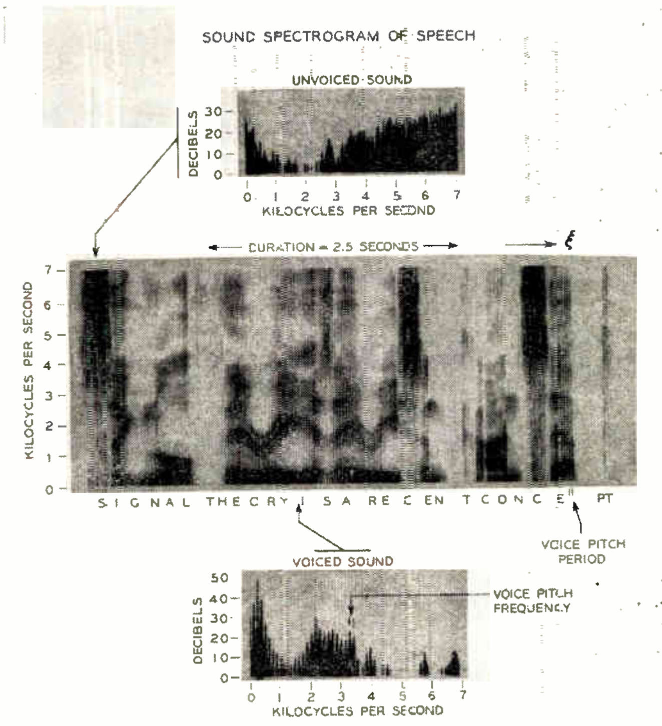


Fig. 3—Sound spectrograms of speech show spectral steady states and transitions which are intimately related to the phonetic content of the utterance.

At other times, the spectrogram contains no clearly-defined structure, but is noise-like. This characteristic arises when the vocal tract is excited with a random disturbance created by air-flow turbulence at a constriction in the vocal tract. An example is the final sound in the word *this*. Sounds with vocal cord excitation are known as *voiced;* those with turbulent excitation, as *unvoiced;* those with no excitation, as *silent.*

Speech spectrograms may be described quantitatively by dividing the time-frequency plane into rectangles $\Delta f \times \Delta t$ in size, and specifying numerically the energy in each. Intelligible speech can be recovered from such a matrix of numbers, and so in that sense it is a "complete" template representation [24]. Also complete, but more succinct, are abstract descriptions based upon time-dependent parameters, that is, properties such as formant locations, over-all intensity and voiced-unvoiced status. Sufficient resolution for the time-frequency matrix representation requires, *inter alia*, about 1000 degrees of freedom ($\Delta f \times \Delta t$ rectangles) per second of speech, while the parametric version has succeeded with as few as 300 per second. In terms of adequately quantized values, these numbers become 3000 and 1000 binary digits. Such representations achieve a great economy over a speech-time waveform which typically requires 10,000 degrees of freedom and 40,000 bits per second. In storing speech data for recognition, this feature is clearly of great importance, not only for efficient storage, but also because the spectrographic and parametric forms of speech-acoustic data are basic to phonetic interpretation. Most experimental speech recognition systems rest upon these fundamentals.

Typical recognition paradigms use either spectral templates or parametric property lists. Utterances to be identified are analyzed and represented in the same form, and then compared to the stored data. The closest fit, according to some metric (correlation or rms error, for instance) fixes the identity.

Schemes of this kind work less than perfectly. The same phonetic sound spoken by various talkers can give rise to spectrograms whose general features are qualitatively similar, but whose frequency-time matrices differ in detail. This fact is all too clear in Fig. 4. These differences reflect variations in vocal tract dimensions (which determine mean formant frequencies), the range of vocal pitch found among the population, the different gestures of articulation from person to person, the diverse timing and intensity patterns typical of the individual, and so on. Some of these factors are discounted in a parametric representation, but here too, any simple description taken over a number of talkers is far from invariant.

Another difficulty arises from the different dialects in common use. Here the same *words or phrases* spoken by different talkers will have different phonetic content. Thus transliteration from a sequence of phonetic elements to English words may involve complex linguistic structure.

Still another difficulty is that acoustically, speech is continuous in time, whereas its phonetic or literal representation consists of discrete symbols or words. How is a recognizer to reconcile or match a continuous acoustic

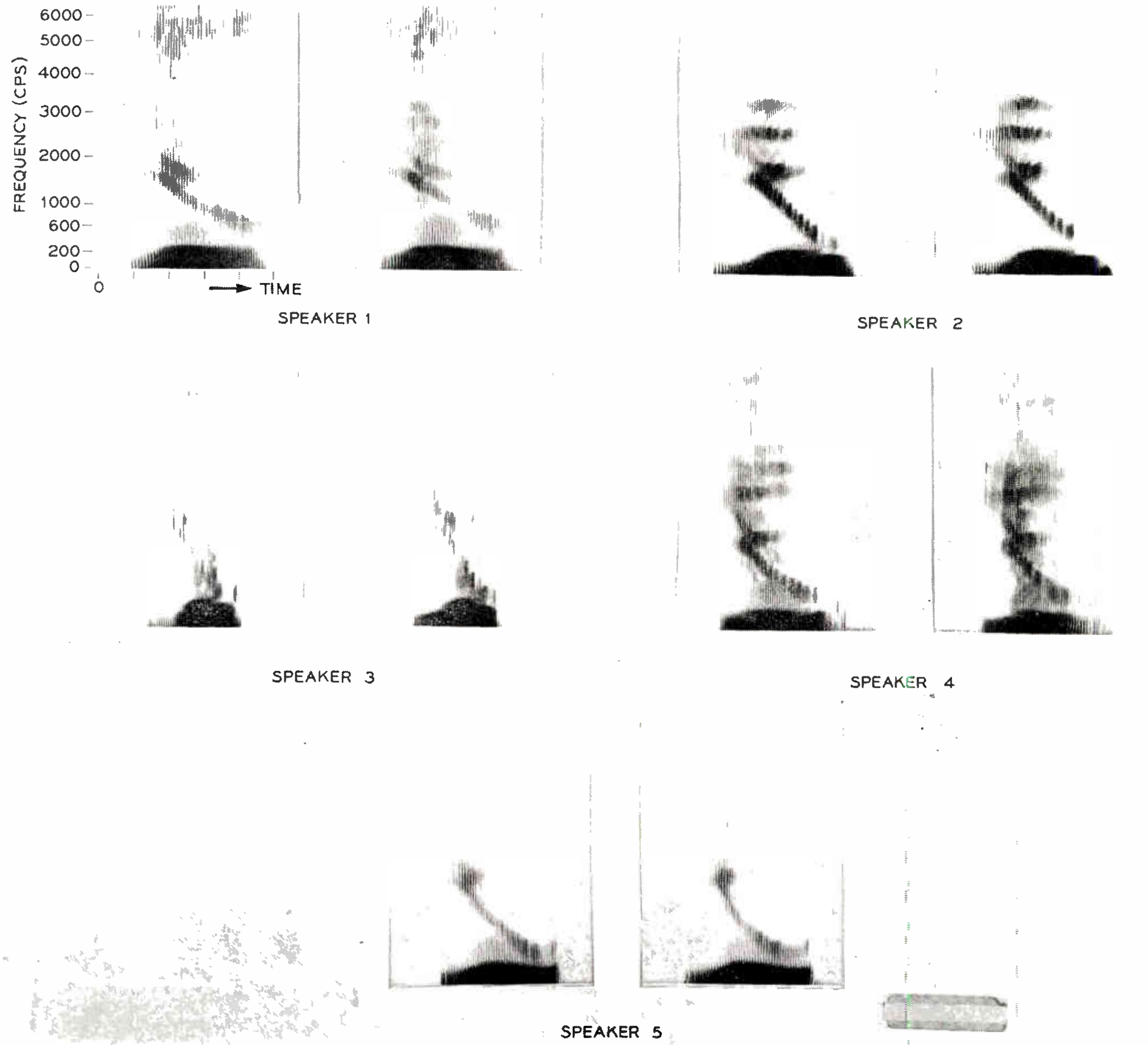


Fig. 4—Sound spectrograms of five talkers each pronouncing the word *you* twice, show that variations in spectral pattern from talker-to-talker are much greater than between successive utterances from the same talker.

flow at its input with the discrete acoustic segments in its memory?

These thorns have been avoided, rather than plucked out, in most speech recognition devices by imposing the following constraints:

1) The stored spectral-temporal patterns and the samples to be recognized are taken from the same talker, or a small number of talkers.
2) The recognition vocabulary is limited to a few words or sounds such as the 10 decimal digits.
3) Each word or sound must be spoken in acoustic isolation.

These crutches aid greatly in retrieving respectable performance from a difficult situation. Typical recognition results are listed in Table II [25]–[37]. Examination of this table shows that the rate of correct identification can approach 100 per cent when the input is protected from the rough and tumble of everyday talk.

As far as we know, no one has yet used such a scheme for practical computer input. Yet there seems no technical reason why speech recognition could not serve under constrained conditions like those where character readers have proved useful [38]. One obstacle lies in the amount and complexity of equipment required. Spectrographic analysis calls for an extensive bank of contiguous band-pass filters or their equivalents [39]. Their rectified and smoothed outputs may require further processing, normalization with respect to intensity and time for example, before being routed to the comparator.

TABLE II

Investigator	Speech Representation	Vocabulary	Number of Talkers Tested	Approx. Error Rate	Additional Facts and Comment
Kersta	Selected entries from Δf-Δt matrix (200 cps ×67 ms)	10 digits	9 men, 5 women	0.2 per cent	Spectrograms quantized into 2 levels
Davis, Biddulph, and Balashek	Formants 1 and 2 as a function of time	10 digits	1 talker	2.0 per cent	Correlation metric
Fry and Denes	Selected entries from Δf-Δt matrix	14 speech sounds in 139 words	1 talker	28.0 per cent (Sounds) 56.0 per cent (Words)	Phoneme digram frequencies used to supplement primitive recognition from acoustic data
Olson and Belar	Δf-Δt matrix	10 words or syllables	1 talker	2.0 per cent	
Dudley and Balashek	Δf-Δt matrix	10 digits	2 men	5 per cent	Temporal sequence disregarded
Mathews and Denes	Δf-Δt matrix	10 digits	6 men	6 per cent	Spectral pattern time and amplitude normalized
Hughes	Spectral features	11 sound categories in 100 words	4 men, 3 women	30 per cent	Feature selection based on linguistic analysis
Shultz	Spectral features	10 digits	25 men, 25 women	3 per cent	
Petrick and Willett	Δf-Δt matrix	10 digits	1 talker	<1.0 per cent	Spectral patterns time normalized
Forgie and Forgie	Spectral features	10 vowels	11 men, 10 women	7 per cent	
Keith-Smith and Klem	Δf-Δt matrix	10 vowels	11 men, 10 women	6 per cent	Statistical decision procedure used to select relevant spectral features
Sebestyen	Δf-Δt matrix	10 digits	10 speakers	<1 per cent	
Suzuki and Nakata	Formants 1 and 2	5 vowels in consonant contexts	5 speakers	≈20 per cent	Additional experiments on vowels in bisyllable words and short sentences yield higher error rates

For a parametric representation, this processing may include complex logic aimed at selecting significant spectral peaks, separating voiced from unvoiced speech and both from silence, etc. Time multiplexing may be convenient at some point in this chain to avoid multiple processors and comparators. Similar multiplexing must be adapted to read-out from the storage. If the stored data are to be adapted to different talkers or groups, then the equipment must operate in a "store" or "learning" mode in addition to the recognition mode. Suffice it to say, the engineering design problem for practical speech recognition is not simple.

Another unavoidable thorn is reliability. How is the performance implied in Table II affected by degradation of the input speech? One should not forget in practice that there are extraneous noises, like airplanes or other talkers or typewriters, and that people talk loudly or softly, holding the microphone close or far away, hoarsely or musically, and so on. We have barely started evaluating such difficulties; neither have we paid much attention to the permissible tolerances.

All these difficulties may well be eased by further research on speech and perception. More versatile speech recognizers may distinguish not only particular spectral features but particular dynamic transitions as well. Linguistic information concerning speaker dialects, intonation, stress, and timing may all aid in reconciling acoustic patterns with their literal counterparts. Surely the complexity of recognition paradigms will not decrease.

We should reiterate that mere extensions of the techniques in the works cited are almost certainly inadequate for handling conversational speech, which we believe to be much harder than handling isolated words.

We cannot easily predict the utility of speech recognition as a computer input. At present the amount of necessary hardware seems forbidding and the competition is simplicity itself, a keyboard and a finger, but the need for speed, for rapid feedback, for technological bravado surely dictate the availability, at least, of a speech typewriter in the next ten years. It will probably turn out that its utility will hinge on what is to be done with human utterances after they enter the machine.

Conclusion

We believe that progress towards useful and effective eyes and ears for computers has been slow, possibly slower than necessary (especially for print readers); partly because we seriously underestimated the difficulties; and there is an inherent question of values. What is the utility of a speech typewriter? Telephones now have dials instead of operators, and one big customer has vanished, alas! Why can't printing devices produce machine-readable tapes while they are printing? Present-day character readers provide a ponderous way of reading carefully-controlled print. All agree we should get these exotic sensors before the Russians, but we are not sure that we obviously need them soon as practical, useful, economical computer inputs. For most of us, contact with the computer is, as it were, by 24-hour mail;

telegraphy or teletype would be a bountiful improvement and telephones altogether too confusingly fast.

But in another vein, as these techniques appear, they will uncover their own applications just as computers themselves have. The commercial interest nowadays is largely a competitive one—successive advertising disclosures arouse successive new bursts of research activity. At some critical point the techniques will draw support from performance rather than promise, though we can at present only ill-describe such future performance.

Costs will be large. But remember that one reading input device can (potentially) replace hundreds of people reading onto keypunches. A real-time speech typewriter can replace 5 people if it works 24 hours a day; it will be a large piece of equipment that can be amortized on $20,000 a year.

The real difficulty then will be the input-input problem: given words thundering in, what do we do besides store them? The state-of-the-art of automatic indexing and abstracting is even more inadequate and as we said before, it will be a long time before we handle English in all its full meaningful glory. But we are a step ahead. The problems we discuss today are being solved. Here at least we can feel fairly sure that we have identified one of tomorrow's problems.

References

[1] References in this paper will attempt to be representative, rather than exhaustive. The interested reader is referred to M. L. Minsky, "A selected descriptor-indexed bibliography to the literature on artificial intelligence," IRE Trans. on Human Factors, vol. HFE-2, pp. 39–55, March, 1961, for a comprehensive listing.

[2] D. H. Shepard, P. F. Bargh, and C. C. Heasley, Jr., "A reliable character sensing system for documents prepared on conventional business devices," 1957 IRE WESCON Convention Record, pt 4, pp. 111–120.

[3] M. L. Minsky, "Steps toward artificial intelligence," Proc. IRE, vol. 49, pp. 8–30; January, 1961.

[4] The term "logic" refers to the traditional techniques that are closely akin to truth-table-like manipulation of binary signals. Today the more meaningless term "process" is appropriate.

[5] O. G. Selfridge, "Pandemonium: a paradigm for learning," *Proc. Symp. on Mechanization of Thought Processes*, H.M. Stationery Office, London, England; 1959.

[6] Note that a typewriter leaves a blank rectangular annulus around each character, but that typically other printing techniques do not, even combining several characters into one character; *e.g.*, fi, ffl, &c.

[7] P. Baran and G. Estrin, "An adaptive character reader," 1960 IRE WESCON Convention Record, pt. 4, pp. 29–41.

[8] W. H. Highleyman, "An analog method for character recognition," IRE Trans. on Electronic Computers, vol. EC-10, pp. 502–512; September, 1961.

[9] P. M. Lewis, "The characteristic selection problem in recognition systems," IRE Trans. on Information Theory, Special Issue on Sensory Information Processing, vol. IT-8, pp. 171–178; February, 1962.

[10] J. J. Leimer, "Design factors in the development of an optical character recognition machine," IRE Trans. on Information Theory, Special Issue on Sensory Information Processing, vol. IT-8, pp. 167–171; February, 1962.

[11] Our President is reputed to set the reading pace of Washington, D. C., with a reading speed of 1200 words per minute.

[12] See, *e.g.*, G. A. Miller "Language and Communication," McGraw-Hill Book Co., Inc., New York, N. Y.; 1951.

[13] W. W. Bledsoe and I. Browning, "Pattern recognition and reading by machine," *Proc. Eastern Joint Computer Conf.*, Boston, Mass., December 1–3, 1959, pp. 225–232.

[14] J. S. Bomba, "Alpha-numeric character recognition using local operations," *Proc. Eastern Joint Computer Conf.*, Boston, Mass., December 1–3, 1959, pp. 218–224.

[15] W. Doyle, "Recognition of sloppy, hand-printed characters," *Proc. Western Joint Computer Conf.*, San Francisco, Calif., May 3–5, 1960, pp. 133–142.

[16] L. A. Kamentsky, "The simulation of three machines which read rows of handwritten Arabic numbers," IRE Trans. on Electronic Computers, vol. EC-10, pp. 489–512; September, 1961.

[17] H. Sherman, "A quasi-topological method for the recognition of line patterns," *Proc. Internatl. Conf. on Information Processing*, Paris, June 15–20, 1959, pp. 232–238; UNESCO, 1959.

[18] T. Marill and D. M. Green, "Statistical recognition functions and the design of pattern recognizer," IRE Trans. on Electronic Computers, vol. EC-9, pp. 472–477; December, 1960.

[19] L. S. Frishkopf and L. D. Harmon, "Machine reading of cursive script," *Proc. 4th London Symp. on Information Theory*, Butterworths, London, August 29–September 2, 1960, pp. 300–316; 1961.

[20] M. Eden and M. Halle, "The characterization of cursive writing," *Proc. 4th London Symp. on Information Theory*, Butterworths, London, August 29–September 2, 1960, pp. 287–299; 1961.

[21] Sir Richard Paget, "Human Speech," Harcourt Brace and Co., pp. 79–80; 1930.

[22] C. G. M. Fant, "Acoustic Theory of Speech Production," Mouton and Co., The Hague, The Netherlands; 1960.

[23] This spectrogram was made by a 300-cps bandwidth scanning spectrum analyzer whose output was recorded on electrically sensitive paper; see E. E. David, Jr., "Signal theory in speech transmission," IRE Trans. on Circuit Theory, vol. CT-3, pp. 232–244; December, 1956.

[24] E. E. David, Jr., "Naturalness and distortion in speech processing devices," *J. Acoust. Soc. Am.*, vol. 28, pp. 586–589; July, 1956.

[25] E. E. David, Jr., "Artificial auditory recognition in telephony," *IBM J. Res. & Dev.*, vol. 2, pp. 294–309; October, 1959.

[26] K. H. Davis, R. Biddulph, and S. Balashek, "Automatic recognition of spoken digits," *J. Acoust. Soc. Am.*, vol. 24, pp. 637–642; November, 1952.

[27] P. B. Denes, "The design and operation of the mechanical speech recognizer at University College, London," *J. Brit. IRE*, vol. 19, pp. 219–229; April, 1959.

[28] H. F. Olson and H. Belar, "Phonetic typewriter," *J. Acous. Soc. Am.*, vol. 28, pp. 1072–1081; November, 1956.

[29] H. W. Dudley and S. Balashek, "Automatic recognition of phonetic patterns of speech," *J. Acoust. Soc. Am.*, vol. 30, pp. 721–732; August, 1958.

[30] P. B. Denes and M. V. Mathews, "Spoken digit recognition using time-frequency pattern matching," *J. Acoust. Soc. Am.*, vol. 32, pp. 1450–1455; November, 1960.

[31] G. W. Hughes, "On the recognition of speech by machine," Sc.D. dissertation, Dept. of Elec. Engrg., MIT, Cambridge, Mass., 1959.

[32] G. L. Shultz, "Investigation procedures for speech recognition," *Proc. Seminar on Speech Compression and Processing*, AF Cambridge Res. Ctr., Bedford, Mass., September 28–30, 1959, Tech. Rept. No. 198; 1959.

[33] S. R. Petrick and H. M. Willett, "A method of voice communication with a digital computer," *Proc. Eastern Joint Computer Conf.*, New York, N. Y., December 13–15, 1960, pp. 11–24.

[34] J. W. Forgie and C. D. Forgie, "Results obtained from a vowel recognition computer program," *J. Acoust. Soc. Am.*, vol. 31, pp. 1480–1489; November, 1959.

[35] J. E. Keith-Smith and L. Klem, "Vowel recognition using a multiple discriminant function," *J. Acoust. Soc. Am.*, vol. 33, p. 358; March, 1961.

[36] G. S. Sebestyen, "Recognition of membership in classes," IRE Trans. on Information Theory, vol. IT-7, pp. 44–50; January, 1961.

[37] J. Suzuki and K. Nakata, "Recognition of Japanese Vowels," *J. Rad. Res. Lab.*, vol. 8, pp. 193–212; May, 1961.

[38] Probably the real reason is economic: the marginal utility of a speech input adapted to just one talker, when he does not suffer from grippe or laryngitis, is just too small.

[39] There are only three computer installations that have gone as far as a real-time microphone-with-filter-bank input, we believe.

Section 13
ENGINEERING MANAGEMENT

Organized with the assistance of the IRE Professional Group on Engineering Management

The Art of Engineering Management*

HECTOR R. SKIFTER†, FELLOW, IRE

Summary—**Challenged by larger and more complex tasks, the engineering manager must place greater emphasis on the motivation and intellectual development of his people if he is to obtain the maximum creative effort from individual engineers.**

Management of the environment rather than the individual is suggested as the key to motivation—and to successful engineering management. Problems facing us in electronics will be easier to solve if managers build an environment characterized by sound organization, stimulating goals, economic rewards and, most important of all, freedom of initiative for the individual with maximum emphasis on self-direction and self-control.

Introduction

The need for excellence in engineering management has never been more urgent than it is today. With the exponential growth of science during the past few decades has come a vast change in the role and responsibility of those who manage technical personnel.

Management was not so difficult a generation ago when small groups of engineers could design and build most of our equipment. But today, in electronics, we are faced with tasks of unprecedented magnitude and complexity. The proposed lunar program, involving manned exploration of the moon before the end of the decade, will exceed in complexity and scope any of our previous technological enterprises. The skills of tens of thousands of engineers in government and industry will be required to achieve this step into space.

Today's manager faces vastly more difficult administrative and technical problems than his predecessor. He must not only coordinate the activities of large groups of engineers but also direct the application of highly advanced technology.

All this requires new management techniques and skills, and it has lent urgency to improvement in the art of engineering management. In fact, there are many today who believe that our technology has outstripped our management—that the biggest job we face is in learning how to motivate and lead effectively.

Today's manager has serious problems. One of these is the shortage of skilled technical manpower. The Engineering Manpower Commission has estimated that the present critical shortage of engineers in the United States will become even worse. By 1965 it is forecast that we will have only 32,000 new engineering graduates each year, in contrast to an average annual demand for more than double this figure. It will be more difficult than ever to build technically superior engineering groups. Even the oft-used, and usually unsuccessful, approach of subsituting "bodies for talent" will offer no solution, because the numbers simply will not be available.

Another problem is that of preventing the technical obsolescence of the engineering staff. In these days of scientific breakthrough and fast-moving, proliferating technology, an intensive and continuing effort is required by engineers at all levels, including the manager, to keep technically up to date.

Also, the manager is faced with balancing group needs against individual needs. On large, complex projects it is essential to maintain tight over-all direction and control in order to live within contract commitments. Yet, experience has shown that creative engineering activity cannot be programmed or closely supervised. The manager has responsibility not only for the group but also for the individual engineer. He must provide an environment in which the individual can reach his maximum potential.

Thus, it is appropriate that we give serious thought

* Received by the IRE, January 3, 1962.
† Airborne Instruments Laboratory, Deer Park, N. Y.

to the responsibilities of the engineering manager—and particularly to those functions which require emphasis at this time. To a large degree, the progress of our society will depend upon the skill with which we manage our scientific and technical resources.

This discussion will deal with engineering management in both large and small organizations. The ideas presented are undoubtedly influenced by the problems posed by extensive government-sponsored research and development programs. The type of engineering organization we discuss is engaged neither in pure research nor in routine design. It is, rather, typical of the groups involved today in the application of advanced technology to military and industrial problems. Its stock in trade is not only technical competence but imaginative leadership.

With this background, I would like to review the evolution of engineering management, the responsibilities of the manager, some approaches for the motivation and intellectual development of technical personnel, and some key problems of the future.

Evolution of Engineering Management

Engineering management, as we know it today, is relatively new. About the turn of the century several pioneers in management, such as Taylor, Gilbreth, and Gantt, made a profound contribution to modern production methods. In France, Fayol pointed the way to new techniques in administration.

During these early years, management techniques were developed and improved—primarily in manufacturing organizations. A body of experience and knowledge grew up during this period which made possible the establishment of large business enterprises. Management did not evolve into a science measured by the standards of modern technology, but it did become a well-established discipline. The problems of applying these production management principles to engineering management were first experienced by the managers of the small specialized engineering groups which grew up during World War I and characterized the early years of the radio and electrical industry.

World War II brought explosive growth to the research and development activity in electronics, nuclear physics, and allied fields. Organizations ranged in size from a single scientist and his staff to large groups such as Radiation Laboratory at M.I.T., Radio Research Laboratory at Harvard, and Airborne Instruments Laboratory at Columbia, operating under the Office of Scientific Research & Development. The Manhattan Project, of course, was the largest of them all. During the same period a number of commercial firms, such as American Telephone & Telegraph, General Electric, and Radio Corporation of America, organized large engineering staffs.

These groups were built with a minimum of formal organization. Work was organized around project engineers and department heads reporting to a director. Service and planning functions were accomplished on an *ad hoc* basis with a minimum of scheduling and paper work. Efficiency often suffered, because it was necessary to employ all available manpower, and many engineers were added without training or supervision. Still, these groups had able people and a variety of talent. No limits were placed upon the creative engineer, who maintained close liaison with the user.

Wartime pressures provided a motivating external environment that made these research groups highly effective, but this type of informal organization and management was not adequate for the long pull in peacetime. To obtain consistent results from an engineering group, it is essential to have planned objectives, sound organization, and carefully defined work assignments.

Several factors tended to bring about improvement in engineering management practices in the early postwar years. A number of groups "spun off" from the wartime research laboratories to form independent organizations. Many new companies were formed, and the larger commercial firms expanded their activities in defense electronics. As new enterprises proliferated and old ones expanded in a peacetime environment, management was forced to re-evaluate organization and engineering procedures. Managers found themselves answerable to employees, stockholders, and boards of directors. At the same time, their customers placed greater emphasis upon reliable systems and equipment, on-schedule delivery, and effective cost control.

Engineering management matured considerably during the first postwar decade. Engineers became more familiar with techniques that had been evolved over the years in industrial management. They participated in trade association seminars and management meetings. Most important, they learned on the firing line by successfully handling tasks of increased complexity.

During the period after the Korean War, important new lessons were learned. Many companies that had concentrated on research and development established production operations and experienced the problems of moving systems and equipment from research and development into production.

In recent years the emphasis in defense electronics has been on research and development. We have concentrated on the development and fabrication of a limited number of complex systems and equipment. We have had to anticipate the requirements of, and work with, newly created procurement organizations responsible for weapon systems. Within our companies we have had to create nearly autonomous project task forces; new organizations to handle schedule and cost control, reliability, and quality control; new groups to handle liaison between subcontractors and team members; new lines of communication within the organization and to the customer. To those of us who have participated in these electronic systems and space programs, this latter period has offered perhaps the greatest challenge to the imagination and capability of our engineering managers.

Responsibilities of the Engineering Manager

Today we are faced with a dilemma. On the one hand we must turn out vastly more complex systems and equipment within rigid cost and time schedules; this implies much greater "front office" direction and control. On the other hand, we must maintain the most favorable possible environment for creative engineering effort; this implies the maximum freedom of initiative for the individual engineer.

What does this mean to the engineering manager? His basic responsibilities remain the same. He is still charged with establishing goals, organizing, motivating and communicating, measuring performance, and developing the personnel in his organization. But he must shift emphasis to the most important elements of his job to meet changing conditions.

Who is this engineering manager? The president of the company? In many cases, yes. The working engineer? Certainly yes! The management principles, tools, and objectives outlined here apply throughout the organizational ladder. As Peter Drucker has so well stated, we are concerned with the management of managers. Each echelon of the management chain has a responsibility for the application of the appropriate techniques to the direction of the managers reporting to it. The objective can well be a team approach by all levels in the most effective management of the resources and capabilities of the organization toward clearly stated goals.

Obviously the engineering manager must be concerned with goals. They must be broad objectives which excite the imagination and give people something they want to work for, something they don't yet know how to do, something they can be proud of when they achieve it.

The engineering manager must also give careful thought to organization structure. As our engineering groups grow, it becomes increasingly difficult for the manager to pay detailed attention to the individual engineer's work accomplishments. The engineering activity must be subdivided into small enough groups for the individual engineer to achieve recognition and identity. Groups and section leaders must be selected who combine project engineering experience with leadership ability. And supporting engineering services must be provided to relieve the technical workforce of administrative and housekeeping details.

Performance evaluation is another of the manager's important responsibilities. In an engineering organization, intelligent measurement of performance requires that the manager demonstrate a personal interest in and technical understanding of the work being performed. The results of such an evaluation obviously are used for administrative purposes—for salary administration, promotion, and the transfer of employees. But, properly applied, performance evaluation results can provide the recognition which a group or individual deserves when a goal has been met or, conversely, the stimulant required when results are less than desired.

These functions—goal-setting, organizing, and performance measurement—are all important. But if today's engineering manager is going to face up to his really fundamental and most demanding tasks, he must place increasing emphasis on how he motivates and develops people. Effective motivation and development of technical personnel are so vital to successful engineering management, so related to working out a compromise between rigid management control and complete freedom of initiative, that this discussion will be devoted largely to these aspects of engineering management.

Motivation of Engineering Personnel

Managing the environment rather than the individual appears to be the most successful approach to the motivation of engineering personnel.

Experience has shown that the creative intellectual effort of the scientist or engineer cannot be programmed. The engineer's intellectual contribution cannot be obtained by giving orders or by the type of close supervision used in manufacturing operations. If the engineering manager places pressure on his people, indicating in detail what they can and cannot do, initiative and creativity are weakened or destroyed.

Those who study management believe that a revolution is taking place in management that is particularly applicable to the engineering organization. A long-standing concept of managerial control has held that it is necessary to impose direction and limitations on the individual in order to get him to perform the work for which he is hired. The trend is away from this toward recognition of the capacity of individuals to exercise self-direction and self-control in the service of objectives to which they have committed themselves. Management by objective and self-control will inevitably replace management by authority and externally imposed control.

This suggests that the engineering manager must adopt a new approach to establishing goals. He must continue to communicate and interpret the broad objectives of management. But he must give increasing personal attention to aiding, encouraging, and counselling his engineering groups—and individual engineers—to establish their own performance goals. These include not only time and cost objectives for the typical research, development, or production project but any number of goals that the individual engineering group can work toward in furtherance of the over-all corporate objectives—new technical areas that a group will investigate, new proprietary developments that the group will begin, new competence that the group will acquire through self-development or the addition of qualified personnel.

The important thing is that, with the encouragement and assistance of engineering management, these goals are self-determined. And the individual engineer, committed to objectives which he has helped to establish, develops an inner motivation and drive which he cannot possibly have if these goals are imposed upon him from above.

Before he can create a favorable environment for engineering work, the engineering manager must understand the values, interests, and expectations of technically trained personnel.

The typical scientist or engineer is keenly interested in exercising full use of his talents and training on the job. And he seeks recognition and the opportunity for development in his professional career. Thus, the engineering manager must keep alert to the nature of work assigned to individual engineers, the challenge involved in particular jobs, and the freedom which engineers have to carry out their responsibilities. The engineering manager must provide for recognition of status—both outside the company in the professional world and within the company. He must encourage publication of technical results and participation in the affairs of professional organizations. Within the company, the manager must provide the engineer with rewards and recognition for excellence; he must provide the engineer with the opportunity to advance up the "technical ladder" as well as up the administrative or managerial ladder.

There is growing recognition of the problem of providing opportunity for the talented engineer to advance. Even in a rapidly expanding organization, not everyone can become a vice president or department head. Many very competent engineers do not desire to switch from technical to administrative responsibilities. We need positions of prestige and challenge for these valuable individuals. Staff positions (such as that of consultant) have, in many cases, been filled or created for this purpose. Carefully chosen personnel can be very effective in such positions. If, on the other hand, these posts become a refuge for ineffective or incompetent personnel, great harm can result.

Financial incentives are a vital part of the environment which provides motivation. Though salary is not the primary incentive to peak performance, it must be competitive with what the engineer can obtain elsewhere; within the company, salaries must be fairly administered. The temptation to hire at salaries above those of comparable or superior persons already on the payroll must be controlled. The salary structure in the engineering organization must be set up so that the money an individual receives is fair relative to the market, economic conditions, importance of the job, and the individual's contribution. The company must also provide a fringe benefit program consisting of group insurance, medical insurance, and retirement benefits which are competitive and offer adequate protection to the employee.

Beyond this, it is essential to consider programs that provide increments of money, stock, or deferred income to provide incentive for and compensate different increments of risk taking and effort. In the intensely competitive electronics industry, incentive rewards must be provided for technically trained personnel who are making the basic risk-taking decisions that determine the future of the business. These rewards usually take the form of stock options, or, in a closely held enterprise, the opportunity for stock participation. The stock option is usually considered an incentive for future performance, not a reward for past efforts. It gives the valued employee a chance to build capital, to participate in ownership, and to identify personal with corporate goals. Many of us believe that a personal stake in the business is one of the best possible guarantees for extra effort on the part of key technical personnel.

Bonus plans are used to provide incentive reward to middle engineering management not directly involved in risk-taking decisions. The bonus offers a means of recognizing the work of individual performers throughout the organization who have made unusual contributions to over-all results.

Good communication is another essential element in an environment favorable to creative engineering. The establishment of objectives and financial rewards is meaningless unless both vertical and horizontal communication is effective throughout the engineering organization.

One of the most difficult problems facing the engineering manager is that of maintaining adequate communication as the organization grows. The typical engineering development group that has doubled or tripled in size over the past few years is no longer a happy little family. Whereas the department head previously knew all members of his staff personally, he now has two or three levels of command interposed between himself and his engineers. He must act through section leaders, group heads, and project engineers. He finds himself faced with the problem of managing managers rather than managing engineers. And he realizes that he must organize, delegate, and communicate in order to accomplish tasks that he previously handled directly with the responsible individual. This problem is acute because highly trained engineering personnel want a voice in policy—want to understand the total picture of what is going on.

The techniques of communication, and methods for improvement in such channels, are too involved to cover in this discussion. However, it should be emphasized that there is no substitute for informal face-to-face contact between the manager and his people. Meetings must take place often enough and last long enough for the manager to learn and, in turn, demonstrate his knowledge of and interest in the projects which are underway.

Some supervisors develop the facility of building strong personal friendships within their group. While avoiding any semblance of paternalism, improved communication and stability can be developed in a group in which personal friendships and common recreational interests help create mutual respect and understanding.

The engineering manager should be alert to a condition of unstable equilibrium that may develop in any organization. An interesting parallel with a collegiate football team can be pointed out. After a notably successful season, the cream of the crop of new players is attracted, resulting within a year or two in an unbeatable team. Subsequently a perturbation (a crushing defeat or change of coach) starts a decline that is just as spectacular as the earlier rise to fame. This is a thought to chill the heart of the engineering manager: the talented engineering organization is a potential victim of the same instability. The pinnacle of success is a mighty slippery perch. Intellectual challenge and strong motivation help provide the stability so important to the environment in which the engineer works.

Managing the environment by encouraging self-determination of goals and by providing adequate economic incentives and good communication thus appears to be the best way to direct engineering groups involved in tasks of increasing complexity. This does not mean that management shifts responsibility for over-all planning, organization, and direction to lower levels in the organization; this kind of responsibility cannot be delegated. But it does mean obtaining the maximum participation from highly trained technical personnel in every phase of the operation of an engineering enterprise.

Intellectual Development

In a period of rapidly advancing technology, such as now exists, technical obsolescence of the individual engineer—or of the entire organization, for that matter—is a real threat to the growth and effectiveness of the company. Maintaining an environment favorable to intellectual development is one of the engineering manager's most difficult and challenging tasks.

The task is difficult because there are so many intangibles among the factors which stimulate intellectual development on the part of an engineer. We know engineers who have kept up to date through an intense curiosity about their own and other disciplines. And we know others who have fallen behind and are no longer equipped to make technical decisions in their own fields, let alone embark upon new areas. The reasons for this difference are hard to understand. Intellectual capacity and drive certainly account for a good deal; yet, there are other elements. In industry there is growing recognition that we must do more to create an environment favorable to self-development.

Several points of view have been expressed about advanced training for scientific and technical personnel. Some consider such educational programs part of fringe benefits—something done to attract and hold young engineers and to remain competitive with other companies. Others claim that if a company is fortunate enough to have talented personnel, it has an obligation to society and to itself to encourage their self-development.

I would not argue with these concepts. But I believe there are other urgent reasons for attempting to develop our technical personnel. First is the recognition that technical specialists require scientific training beyond that received in the normal college program. The creative engineer must be able to apply the most advanced techniques to the problems that confront him. His perspective must be broad so that he can work in new fields required by the demands of dynamic technology.

Second is the problem of sheer technical obsolescence faced by engineers who have been out of school for a number of years. Much too large a proportion of our senior personnel have not kept current with fast-moving and complex developments in fields such as space technology, communications, computers and data processing, materials technology, energy sources, and industrial automation. Many lack the technical base and flexibility to meet the challenges that lie ahead.

The problem of technical obsolescence is related to the responsibility of engineering management to protect the assets of a business. Toward this end, we maintain physical equipment and set up depreciation reserves on plant and machinery. In much the same way, we have the obligation to protect and enhance our investment in people, the most valuable asset in the technically oriented business.

Fortunately, there is growing concern in the electronics industry over this problem. A number of companies have recognized the need to provide advanced training for newly hired engineers. One company has organized a program of graduate study in communications engineering and related sciences, conducted during working hours, which leads to a master's degree at the end of the second year. A number of companies have two- or three-year programs for newly hired engineers which emphasize advanced study in technical fields related to the engineer's future work. Other companies make it possible for newly hired personnel to take advanced work at nearby educational institutions. These are indications that progress is being made on the problem of advanced training for newly hired personnel.

The problem of advanced training for senior engineers appears to be more difficult to solve. Because of demands on the time of these responsible engineers, it is difficult to set aside time for advanced work during working hours; nevertheless, one large company recently established a program under which more than 100 engineers are enrolled each year at colleges and universities for resident work leading to advanced degrees. This company found early in the program that participation had become identified with status, that many participants were more interested in personal advancement than in advanced training. They discovered that a number of men were not adequately prepared in languages and mathematics to embark upon doctoral work and that refresher courses would be essential prior to sending such men for advanced work. With modifications, the program is being continued and expanded.

Despite cost and the impact upon the organization of the temporary loss of key personnel, other companies will no doubt establish programs involving resident college work for senior engineering personnel. Some will move ahead with out-of-hours educational classes, lectures, seminars, and other methods for providing advanced training. However, until such formal steps are taken, the engineering manager has an obligation—and an opportunity—to train both newly hired and senior engineers. Within his group, he must create an atmosphere favorable to self-development. By example, support, and counsel, he must influence his people toward placing effort upon self-development in areas vital to the individual and the company. He must encourage use of reference material already available within the company; many companies have discovered the desirability of providing extensive library services and facilities for engineering personnel. Beyond this, he must provide opportunities for the cross-fertilization of ideas; must permit wide dissemination of information on work in progress in different areas; must provide recognition to those who have taken advanced training, with or without company assistance; and must cultivate a spirit of competitiveness with respect to intellectual achievement.

Steps such as these should at least accelerate self-improvement on the part of engineering personnel and contribute toward greater capability and better utilization of existing scientific and technical talent. There is room for considerable imagination and experimentation in this area.

In the long term, I believe a change of philosophy will be required to accomplish all we need to do. Engineering management must recognize the benefits from educational programs and be willing to underwrite their cost. At the same time, the new generation of engineers must develop a realization that their careers will be divided between work and advanced study and that only through exposure to new disciplines and new technology can they hope to obtain advancement and the full realization of their potential.

Key Problems Facing Us

Of the many problems that face engineering management today, it appears to me that three deserve particular attention: planning our technical future, staffing and organizing to handle the increasingly complex systems jobs, and reducing the cost of engineering work.

Careful planning as to the technical areas in which an enterprise will concentrate is more essential now than ever before because military and industrial requirements are constantly changing. Emphasis today is on advanced equipment, larger and more complex systems. Even the largest companies cannot keep pace in all areas of technology. Hence, it becomes vital to be selective and to develop outstanding capability in specific technical areas.

The spadework necessary to make sound decisions on corporate technical direction is an important task for engineering management working in cooperation with finance, marketing, production, and other elements of the organization.

Many approaches to providing technical direction have been adopted. Some companies have established a planning organization within the engineering or marketing groups with clearly defined responsibilities for long-range planning. Others have established planning groups at a top corporate staff level whose job it is to initiate and coordinate forward-looking thinking throughout the organization. Several of the large companies have set up "think" groups of highly trained scientists and engineers who are completely separated from the day-to-day operations of the company. Each of these methods has advantages and disadvantages.

The planning groups within engineering may become too engrossed with technical factors, with the result that the needs of customers receive second priority. Those within marketing may place too great emphasis upon customer needs and market expectations, not enough emphasis upon technical aspects. Both types of groups, however, can make progress, if they have well-defined responsibilities.

The corporate-level planning staff may be too far removed from current technical and marketing operations to provide sound guidance. This is particularly true of the planning group whose members do not maintain sufficient personal contact with key staff members and potential customers. On the other hand, this type of planning organization enjoys the advantage of an over-all picture of the company's operations. Properly run, it can make a valuable contribution.

The most controversial of the planning mechanisms is the so-called "think" group of highly trained technical personnel, who often are completely separated from the rest of the organization, organizationally and geographically. Such groups may tend to concentrate to such an extent on long-term factors that they lose touch with intermediate-term factors. Some planning organizations of this type have been known to disassociate themselves from realistic corporate objectives. Others have lost sight of their original charter and have turned their efforts from planning to proposal effort and the handling of certain types of government contracts. On the positive side, such a group can be highly creative and look well into the future without being distracted by day-to-day considerations.

What is important is not where the planning function is located within the corporate structure, but that somewhere all the elements of effective planning are performed. This means that the engineering manager, regardless of the level at which he works, must accept planning as one of his crucial responsibilities. He must set up the mechanism for looking ahead regardless of whether the group for which he is responsible is a division, a department, or a section. He must block out time to participate in, provide liaison with, and feed information to the planning group, which may be operating elsewhere in the company. And the manager must create an atmosphere in his group that encourages looking ahead.

Staffing and organizing to handle the increasingly complex systems job may turn out to be the most difficult problem of all for the engineering manager. During the past decade we have had the challenge of gearing up for large, complex tasks. But this is only the beginning. In space technology and other areas we will be faced with tasks beyond anything we have accomplished to date.

The trend toward larger, more complex systems has several implications for the engineering manager. Internally it means that all of a company's related capabilities must be brought to bear on proposal effort that may lead to participation in systems work—and in the performance of these complex jobs after the proposal has been accepted. No longer can we afford the luxury of individual groups who draw aside from the mainstream of company effort and are not interested in joining the commitment for large tasks. These artificial barriers must be broken down; the specialized groups who see their particular challenge in highly creative engineering and advancing the state of the art must be made part of the total effort that individual companies bring to bear in obtaining and handling systems work.

Furthermore, we must place greater emphasis on joint bidding, which implies closer prime-subcontractor relationships, team efforts, and joint ventures with other companies. No one company, not even the largest contractors, can hope to accumulate the array of talent and resources necessary to handle the very large systems. Industry must take the initiative in evolving effective techniques so that the best qualified organizations in specialized technical areas can pool their capability to handle the very large systems jobs. Companies must join forces earlier and communicate more effectively during the precontractual stage. Once contracts have been awarded for major systems, steps must be taken to improve the working relationships between the companies involved, to improve cost and schedule control, and to speed the integration of subsystems into the over-all system.

Within individual companies, engineering management is faced with the problem of how to organize in order to handle participation, whether as a prime contractor or a subcontractor, in these large tasks. We have learned that supervision of these large projects is vastly different from the engineering work we have accomplished in the past. Program management groups are needed to coordinate and assume responsibility for all aspects of the job—cost, schedule, manpower, material, and other requirements. They subcontract to other elements within the company and establish relationships with team members or subcontractors outside. They visualize the task in its entirety, shift to meet changing technical, schedule, and budget requirements, and focus the collective energies of groups inside and outside the company on the accomplishment of the task.

Engineering management must face up to certain realities with respect to systems work. First, the management of these projects at all levels requires a unique combination of administrative and technical skill. Personnel selected for these positions must be people who are challenged by the rewards of creative management as well as those of creative engineering. Second, the maximum possible degree of permanence must be built into program management personnel assignments. The insecurity which has characterized the lot of engineers on many large systems jobs can be eliminated if the company builds program management groups permanently into the organization structure. This, of course, implies a commitment by the company to make systems work a permanent part of its activity. Beyond this, it means establishing the type of program management or task force type of operation that can be self-perpetuating. Such groups must not only keep on top of the work at hand, but must participate in planning and proposal effort that will find new fields to conquer and new applications for techniques, systems, and equipment already developed. When a major task has been completed, key personnel in the group are shifted to the business-getting function.

Finally, more effective tools must be developed for the direction and control of our systems work. At frequent intervals today we measure progress and arrive at estimates of time, cost, manpower, material, and other requirements needed to complete the task. On major systems this is a difficult and time-consuming undertaking. Engineering management must develop new review techniques and make greater use of computers and data-processing equipment, to speed and increase the accuracy of such controls. Engineering management needs more accurate and timely information during the progress of the work which can be used to remove bottlenecks, correct mistakes, adjust to changes—in short, to maximize performance on the job.

Reducing the cost of engineering work is another of the significant problems facing engineering management. By and large the electronics industry has done a creditable job of developing and producing complex systems and equipment during the postwar years. But from the controller's standpoint some of our efforts have been extravagant. We hear it said by members of Congress and special committees that defense research and engineering has become the greatest luxury the country has ever known. Our customers among the armed services tell us that greater emphasis must be placed upon financial management and cost control.

We in electronics realize that cost control will be essential if the major government programs ahead are to be carried out within the limits of funds available. It is already apparent that only the able and efficient contractors will survive in the long run.

For the engineering manager this means greater emphasis on accurate cost estimating, particularly on cost-plus-fixed-fee development programs. And it will require intensified and continuous effort to live up to our obligation to deliver complex systems within the original estimates of time and cost.

In carrying out his functions—planning, organizing, motivating, evaluating performance, and developing people—the engineering manager must be keenly alert to cost factors. He must develop within his organization an awareness of cost and the methods necessary for effective cost control. Still he must recognize that, in evaluating the performance of a research and development group, efficiency is not a simple ratio of output per unit of input. Often a creative engineering approach can substantially reduce the cost of a system or equipment under development.

It is the engineering manager's task to encourage this kind of creative approach, to make certain that costs are under control, and to obtain the cooperation of all project personnel in working toward cost reduction. If these steps are accomplished, the engineering manager will have gone a long way toward ensuring the success of the complex and difficult projects that lie ahead.

Conclusion

Technical personnel engaged in programs vital to government and industry must be strongly motivated and working at maximum potential if our society is to progress.

To achieve this result, today's engineering manager must place increasing emphasis on how he motivates and develops people. By effective motivation and intellectual development, the manager can arrive at a sound compromise between too rigid control and a "laissez-faire" approach to engineering management.

Faced with projects of greatly increased complexity, the engineering manager might be tempted to increase top-level direction and control. But we have learned that much more effective leadership is provided by managing the environment rather than by managing the engineer. We must create an environment favorable to creative engineering with all that this implies—good organization and communication, stimulating goals, economic rewards and, most important of all, freedom of initiative for the individual with maximum emphasis on self-control and self-direction.

The difficult problems facing us today—such as reducing costs, planning the technical future, and organizing to handle increasingly complex systems—become easier to solve if engineering managers create the kind of environment that encourages participation by technical personnel in every phase of the operation. Such an environment will also stimulate and accelerate intellectual development, which has become essential because of the danger of obsolescence of our knowledge in the midst of rapidly advancing technology.

In the long run, I am confident that emphasis on individual creativity, initiative, and self-expression will prove more effective in producing results than the closely directed and controlled systems of management that have sometimes been used. Engineering organizations that impose direction solely from above will never match the achievement of those who make self-direction and inner motivation driving forces. To create the kind of engineering groups in which free men cooperate and achieve the fullness of personal development is the challenge ahead for engineering management.

Section 14
ENGINEERING WRITING AND SPEECH

Organized with the assistance of the IRE Professional Group on Engineering Writing and Speech

Growth and Importance of Engineering Publication*

KEITH HENNEY†, FELLOW, IRE

Summary—**The pages of past issues of the PROCEEDINGS provide an interesting and accurate picture of how the electronics art and science has developed over the years. A few examples of early writings are given which reflect prior stages of the art, and some observations are made on the place of technical publishing in the electronics industry.**

BY ITS VERY Charter and Constitution the Institute of Radio Engineers has always had as one of its objects "the publication of papers." Among the evidences of distinction recognized by the Constitution in its qualifications for Senior Member is the "publication of important original engineering or scientific papers, books or inventions" and at least one Fellow had this honor bestowed on him because the Board of Directors felt that his editorial work had "served to advance the progress substantially" in the fields in which the Institute serves.

In all scientific, and most technical disciplines, it is well recognized that the publication of papers is a vital part of one's job. Radio engineers have not been backward in this respect. Even a casual glance through the pages of early issues of the PROCEEDINGS shows not only who was doing what and when but gives an excellent picture of how the electronic art and science has developed.

A casual glance, however, is impossible for any reader old enough to have had some contact with the early days. The casual glance is apt to consume hours because these early days constitute a fascinating panorama of what is now so well known (or forgotten) but what was then a complete puzzle or undreamed of. It is certain that a young engineer who never heard of Austin or Cohen or Morecroft or Arnold or of a quenched gap or a carborundum detector will come away from a reading of the PROCEEDINGS of the 1920's with a feeling of awe that engineers of those days accomplished so much with so little knowledge. It is also fairly certain that young engineers of today will discover that what seems so new now may really be rather ancient.

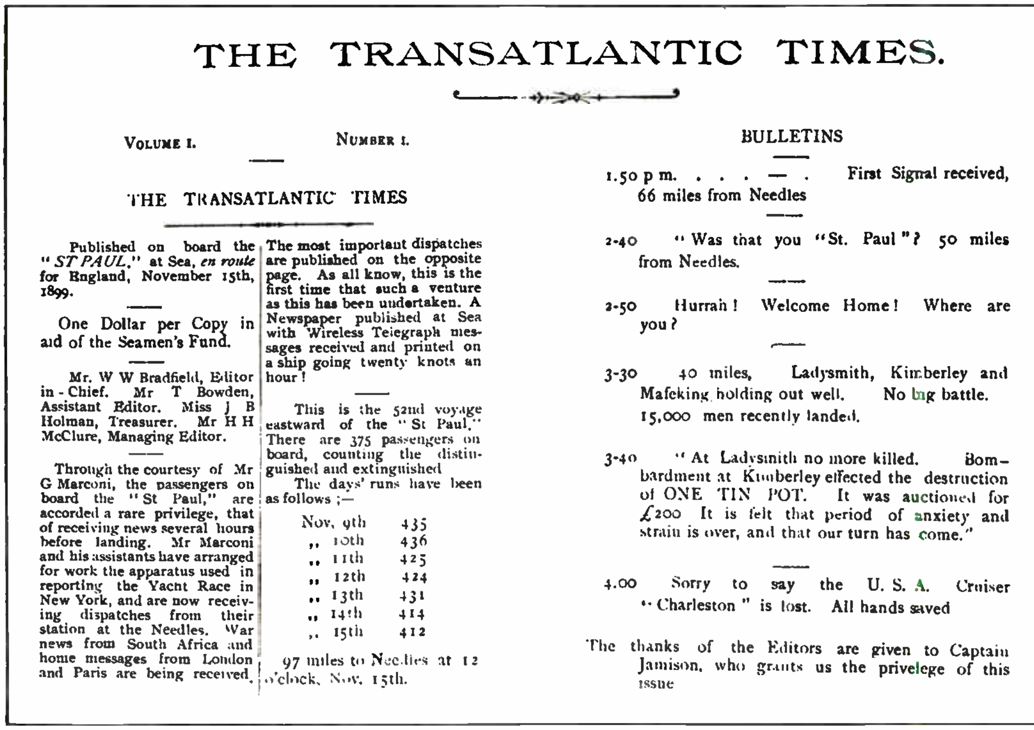
Thus Marconi (Fellow, 1916, Medal of Honor, 1920) writing in 1922 says,[1] "it should be possible to design apparatus by means of which a ship could radiate or project a divergent beam . . . in any desired direction, which rays, if come across a metallic object . . . would be reflected back to a receiver . . . on the sending ship, and thereby . . . reveal the presence and bearing of the (object) in fog or thick wheather."

Or, Moorhead and Lange in 1921[2] who lament that "in spite of wide applications of the tube, in spite of its being in daily use all over the world, no satisfactory solution of a mathematical theory has been given." Or, Alexanderson (Fellow, 1915, Medal of Honor, 1920) speculating[3] on how it is that radio will carry over such extraordinary distances on a quiet night or, in fact, how radio goes anywhere at all.

Between 1899 when Marconi made possible the first newspaper published at sea (Fig. 1) and 1962, the whole of radio and electronics has unfolded, at first slowly

* Received by the IRE, June 5, 1961.
† Snowville, New Hampshire.

[1] G. Marconi, "Radio telegraphy," PROC. IRE, vol. 10, pp. 215–238; August, 1922.
[2] O. B. Moorhead and F. C. Lange, "The specifications and characteristics of Moorhead vacuum tubes," PROC. IRE, vol. 9, pp. 95–129; April, 1921.
[3] E. F. W. Alexanderson, "Central stations for radio communication," PROC. IRE, vol. 9, pp. 83–94; April, 1921.



THE TRANSATLANTIC TIMES.

VOLUME I. NUMBER I.

THE TRANSATLANTIC TIMES

Published on board the "*ST PAUL*," at Sea, *en route* for England, November 15th, 1899.

One Dollar per Copy in aid of the Seamen's Fund.

Mr. W W Bradfield, Editor in-Chief. Mr T Bowden, Assistant Editor. Miss J B Holman, Treasurer. Mr H H McClure, Managing Editor.

Through the courtesy of Mr G Marconi, the passengers on board the "St Paul," are accorded a rare privilege, that of receiving news several hours before landing. Mr Marconi and his assistants have arranged for work the apparatus used in reporting the Yacht Race in New York, and are now receiving dispatches from their station at the Needles. War news from South Africa and home messages from London and Paris are being received.

The most important dispatches are published on the opposite page. As all know, this is the first time that such a venture as this has been undertaken. A Newspaper published at Sea with Wireless Telegraph messages received and printed on a ship going twenty knots an hour!

This is the 52nd voyage eastward of the "St Paul." There are 375 passengers on board, counting the distinguished and extinguished

The days' runs have been as follows ;—

Nov. 9th	435
,, 10th	436
,, 11th	425
,, 12th	424
,, 13th	431
,, 14th	414
,, 15th	412

97 miles to Needles at 12 o'clock, Nov. 15th.

BULLETINS

1.50 p m. . . . . — . First Signal received, 66 miles from Needles

2-40 "Was that you "St. Paul"? 50 miles from Needles.

2-50 Hurrah! Welcome Home! Where are you?

3-30 40 miles, Ladysmith, Kimberley and Mafeking holding out well. No big battle. 15,000 men recently landed.

3-40 "At Ladysmith no more killed. Bombardment at Kimberley effected the destruction of ONE TIN POT. It was auctioned for £200 It is felt that period of anxiety and strain is over, and that our turn has come."

4.00 Sorry to say the U. S. A. Cruiser "Charleston" is lost. All hands saved

The thanks of the Editors are given to Captain Jamison, who grants us the privelege of this issue

Fig. 1—"The Transatlantic Times," courtesy of the Harvard College Library.

and then with great rapidity, but all the steps are to be found described in the pages of the PROCEEDINGS and in the pages of other radio publications of the times. The whole history is there, the editorial pages giving the concepts, the theory, the experimental methods and results, the advertising pages giving the components, devices and systems which resulted. For later-day patent lawyers, the advertising sections of the early publications have proved to be gold mines of information on the dates when certain matters entered the art.

Whatever the purpose of technical publishing today, there is no doubt that early engineers wrote papers to be read.

Any engineer of that time or this can easily get past the first page and more without giving up, and the proof of high readership lies in the extended discussions of individual papers which appeared in later issues. No one can read Lee deForest's (Fellow, 1918, Medal of Honor, 1922)[4] discussion of Roy Weagant's (Fellow, 1915, Morris Liebmann Memorial Prize, 1920)[5] antistatic inventions without visualizing the scene, the researcher-inventor setting forth what he has accomplished, the critics stuggling to understand or to deny.

Not all of the papers of the time were actually published—a situation which differs not at all from the present day. The writer remembers with great vividness an evening in Massachusetts spent with G. W. Pickard (Fellow, 1915, Medal of Honor, 1926) and Henry S. Shaw (Fellow, 1935). Pickard was then in the midst of his famous nightly recording of the strengths of signals from WENR in Chicago and the writer was sent to gather material for a report on this research. He broke into a rather hilarious session consisting in the reading of a paper by George Lewis (Fellow, 1951) on a proposed method of rating signal strength. This was a matter that was not very well standardized at the time. Signals were to be rated, according to the Lewis method, somewhat as follows:

1) Audible with the headphones tightly clamped to the head.
2) Audible with the headphones on the table.
3) Audible with the headphones in the next room.
4) Audible with the phones in the next room and the door shut.
5) Very loud signals.
6) Signals loud as hell.

For the record, and for the benefit of modern engineers, it is worth noting that Pickard's homemade comparison oscillator sat on the floor in the middle of a copper wash boiler (where can one buy such an instrument today?) and that this original research did not require a lush government contract to be revealing and effective. As the needle of Pickard's microammeter gave the instantaneous vagaries of WENR's signals, this young editor began to glimpse what the Appleton (Fellow, 1931, Vice-President, 1932, Medal of Honor, 1962) and Kennelly (Fellow, 1928, Medal of Honor, 1932) layers in the ionosphere were all about.

[4] L. de Forest, "Further discussion on reception through static and interference," PROC. IRE, vol. 7, pp. 543–553; October, 1919.

[5] R. A. Weagant, "Reception through static and interference," PROC. IRE, vol. 7, pp. 207–244; June, 1919.

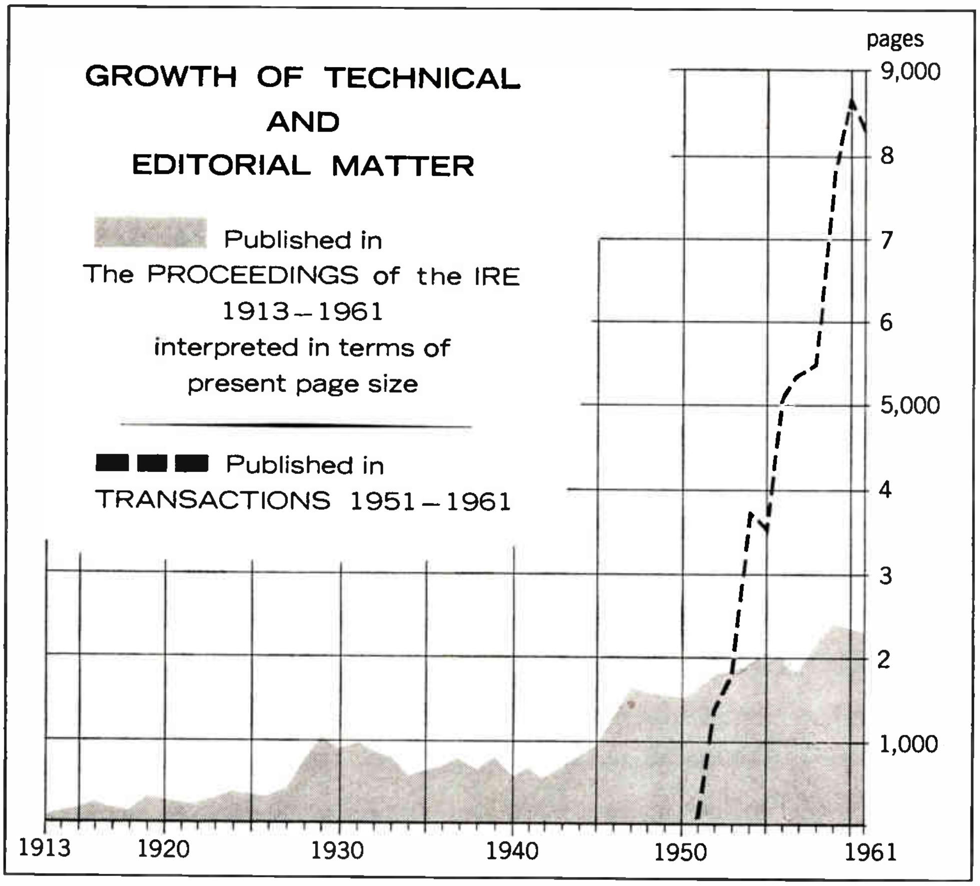


Fig. 2—Volume of technical and editorial matter published in the PROCEEDINGS and the TRANSACTIONS. The figures do not include advertising pages.

Many of the early papers were descriptive and not highly analytical. They dealt with important but broad subjects; they interested and concerned the whole field. All were readable.

Among other matters discussed in the PROCEEDINGS was a report by WJZ covering the period May 15, 1923 to May 1, 1924. In that time WJZ, then on top of the Aeolian Building, New York, and operating on 455 meters, broadcast 4325 events. Among them were the following:

Alto solos	3
Baritone solos	165
Bass solos	32
Contralto solos	104
Soprano solos	505
Tenor solos	159.

It is easy to see why engineers soon invented soprano killers which led, quite naturally, to the modern blab-off which makes television pleasant for many and possible for some.

It must not be supposed that all matters discussed in the PROCEEDINGS culminated in great advances; or that subjects once covered disappeared. Crystal detectors were in disfavor only to come into the limelight again after many years.

Whatever the eventual result, the work had to be done; something had to be studied, dissected, analyzed, reported, accepted or rejected. Even at this date, who is bold enough to say that the articles were not worth the paper to print them on?

The extraordinary growth of the Institute has, of course, occasioned or paralleled an equally great growth in the publication activities of the industry. While the nation as a whole was doubling in population (1910 to 1960) the Institute increased its membership by a factor of 40. For the first ten years IRE and its 2000-odd members produced approximately 300 pages per year. When the real and sustained rise in membership began in 1940, the 5000 members had nearly 600 pages per year to digest. Prior to this time, the pages of the PROCEEDINGS were small, constituting approximately 40 per cent of the word content of the present standard format.

The great leap in membership took place within the last ten years, attaining nearly 90,000 by the end of 1960. In the same period the publishing burden assumed by the Institute rose correspondingly (Fig. 2). In 1959 some 18,000 pages of text and advertising were written, edited, typeset, proofread and published. Over half of the editorial pages appeared in the TRANSACTIONS of the various 28 Professional Groups. Established in 1953, the STUDENT QUARTERLY contributed over 200 pages per year of readable and forward-looking articles.

From its very earliest days, the PROCEEDINGS under the editorial guidance of Alfred Goldsmith (IRE Founder, 1912, Fellow, 1915, Medal of Honor, 1941) educated its membership in the very forefront of the advancing art. Its standards were high; many papers probably written more for the benefit of the author than for the edification of the reader never saw the printed page. As the fledgling radio-electronics field grew from a purely empirical science to a highly sophisticated and analytical one, other journals more "practical" in nature found a niche and not only told their audiences what to do but how to do it. Few still exist, *Wireless Age*, *Radio Broadcast*, *Popular Radio*, *Radio Engineering* and others enjoyed editorially successful careers even if they were not too fortunate financially. Of the ancients, *Radio News*, *Wireless World* and *QST* continue, the latter being the *vade mecum* of the amateur and the engineer alike for many years and constituting the repository of numerous highly important articles, the results of experiments which proved more than once that all the technical brains are not bottled up in the great laboratories.

Although the audiences of the early journals were enthusiastic and often large, the publishing economics were not highly favorable. One journal, *Electronics*, begun in 1930 swam in tepid water for years and its editors, O. H. Caldwell (Fellow, 1944) and K. Henney (Fellow, 1943) were called on the carpet annually to explain why the magazine should not be thrown out the window. Yet they both saw the day (1956) when this journal published more pages of advertising than any other monthly magazine in the world.

The place of technical publishing in any industry as fluid, technically, as electronics is important and secure as a way of recording current events. What the historian of 50 years hence will think as he reads the pages of today lies completely in the realm of speculation; only one endowed with an excess of temerity would say today that much of what is printed is not worth the paper it is impressed on. In scientific circles, time has its own way of teaching the future-tellers to be careful.

Communication: A Responsibility and A Challenge*

ELEANOR M. McELWEE†, SENIOR MEMBER, IRE

Summary—**Throughout his career, an engineer has a responsibility to communicate the results of his work to four major audiences: his management, his colleagues, students or trainees, and the lay public. He must, therefore, be proficient in various forms of oral and written communication, ranging from simple conversations to formal reports and articles. To transmit information effectively, an engineer must have a complete understanding of his audience, his subject and his language. The Professional Group on Engineering Writing and Speech sponsors publications, meetings and workshops to help IRE members meet this communication challenge successfully.**

THE PROFESSIONAL responsibility of an engineer goes far beyond the mere obligation to keep himself informed of current developments in the industry which may affect his work. As his competence in his particular field increases, so too does his responsibility to make his own contribution to the body of knowledge which makes it possible for others to be informed. Thus, the successful engineer finds himself spending an increasing amount of time communicating what he has learned to the many audiences who have a right to expect to hear from him. Furthermore, he finds that his ability to express himself effectively—orally or in written form—becomes ever more important to his continued professional growth.

This paper, which is sponsored by the Professional Group on Engineering Writing and Speech (PGEWS), discusses the four major audiences with whom the engineer must communicate: management, his colleagues, students or trainees, and the lay public. It also describes the various methods or forms of communication which may be used for each audience, and outlines some of the activities of PGEWS which are designed to help engineers write and speak more effectively.

Reports to Management

An engineer's best and most interested audience is usually his own management. In today's rapidly expanding electronics industry, management has a vital

* Received by the IRE, January 5, 1962.
† Radio Corporation of America, Electron Tube Division, Harrison, N. J.

need to know and understand the results of research and development projects in many diversified areas. Because of the growing size and complexity of most companies, a manager generally must depend on his engineers to supply him with whatever information he needs. The reporting of such information is not only a fundamental responsibility of the engineer, but also a valuable opportunity to demonstrate the advancing status of his work, his understanding of its place in the company's over-all activities, and his ability to state things simply, clearly, and concisely.

Reports to management take many forms, ranging from simple conversation at one extreme, through conferences, memos, and letters, to formal engineering reports at the other extreme. Much of the reporting is oral, and most of it is rather informal. However, the formal engineering report deserves special mention here, if for no other reason than that it is used so much less than it should be.

The company that pays an engineer's salary certainly has the right to a written record of whatever results his work produces. All too often, however, valuable know-how is passed on from one generation of engineers to another by word-of-mouth only, with inevitable distortion and losses. The written report, when properly prepared, preserves engineering information in a useful, permanent reference source, and eliminates wasteful duplication of effort.

Professional Papers

The communications in which the engineer himself is usually most interested are the ones intended for his colleagues—for other equally experienced engineers working in his own or allied fields. In general, this category of communications is also the one of major concern to IRE because it encompasses the technical papers presented at the various national and local meetings and the articles published in the Proceedings, the Professional Group Transactions, and many of the Section publications.

Such professional papers enhance the reputation and prestige of the companies which employ engineers, as well as the status of the engineers themselves. In fact, in contributing such a paper an engineer discharges many responsibilities at once: to his company, to his professional society, to his colleagues, and to himself.

Education or Training

One of the most important communication responsibilities of the engineer is toward the engineering trainee or student. The electronics industry is advancing so rapidly in so many areas that information on the multitude of new developments can be made generally available only with the help of engineers closely associated with the actual work.

The engineer has a wide range of opportunities to discharge this obligation to trainees and students. The most immediate is, of course, the opportunity to train less experienced engineers working under his own direction. In many cases, the management of his company also makes it possible for him to participate in more formal training programs for engineers working in other activities. In addition, professional societies have a continuing need for qualified engineers to take part in lecture series, to talk to student chapters, and/or to write tutorial articles. Colleges, too, are in need of guest lecturers and of instructors for special courses.

Engineers can make a permanent contribution to the educational field by writing books. The "technical" departments of many of the large book publishing companies have long lists of subjects which they feel have not yet been adequately covered in book form. In most cases, they are more than willing to discuss prospective books with engineer-authors and work out a mutually satisfactory outline. This area of communication is perhaps the most rewarding to the engineer—personally, professionally, and financially.

Speaking to the Public

The most difficult, and consequently the most ignored, field of engineering communication is that addressed to the lay public, *i.e.*, the nontechnical audience. Too often this field is left entirely to "public relations" specialists whose technical knowledge may be inadequate to fill the need for complete, reliable, accurate, and easy-to-understand information. Newspaper reporters and magazine writers make a valiant attempt to serve as "middle men" between the engineer and the public, but their numbers are few and their information is necessarily limited by their separation from the actual work.

Engineers have a professional responsibility to promote understanding, and to prevent misunderstanding, of both their individual work and the profession itself. Particularly in the present "electronic age," it is increasingly important for the engineer to do his part to extend public knowledge of new developments and of engineering in general.

Opportunities for this type of communication may not be immediately apparent, but usually just the faintest indication of interest and availability brings forth invitations from various local civic, church, and business groups. Organizations such as the P.T.A., the Lions Club, the League of Women Voters, the Ladies Aid Society, and the Merchants Association are constantly looking for interesting speakers for their meetings. The engineer who takes advantage of such opportunities to communicate with his nontechnical neighbors performs an invaluable service to both his community and his profession.

Forms or Methods of Communication

The engineer who wishes to communicate with all these various audiences will, of course, use many different methods. In some cases, the communication will be oral, consisting of simple conversation with one or two others, conferences, panel discussions, informal talks to small groups, or formal lectures to large audiences. In other cases, the communications will be written, informally in letters or memos, or formally in reports, articles, or books.

Whether oral or written, however, the main purpose of engineering communications is to inform, not to impress. Some engineers mistakenly believe that formal papers call for long and unusual words, complex sentences and paragraphs, and a large quota of mathematical equations. Perhaps an engineer's simplest guide to what a good paper should be is his own reaction to those he hears or reads. An effective paper may well impress the audience, but the effect results from the engineer's achievements, not his vocabulary.

In all his communications, the engineer faces the challenge of transmitting knowledge from his own mind to those of his audience with minimum loss and distortion. To answer this challenge successfully, he must have a complete understanding of three important factors in the communication process: his audience, his subject, and his language.

An effective speaker or writer is always aware of his particular audience, and speaks to them on the basis of *their* knowledge and experience, as well as his own. He lets *their* interests and needs govern the length of his discussion, the amount of theory or detail he presents, and the language he uses.

To communicate effectively to any audience, the engineer must also know his subject thoroughly. In many cases, the actual preparation of a talk or article forces an engineer to take a broad, over-all view of his work—to place it in the proper perspective with regard to other developments, to organize his thoughts logically, and to do whatever additional work may be necessary to fill in gaps in his story.

An effective communicator is also thoroughly familiar with his language. His vocabulary is adequate to express even the most complex concepts simply, directly, concisely, and clearly. He uses language so skillfully that his audience is never aware of his individual words or sentences, but only of the information that they convey.

Well-prepared communications reflect the thought processes that are basic to both good engineering and objective reporting. When an engineer's thought processes are logical and disciplined, his comments are organized, accurate, factual, forceful, and clear. His communications are complete, but concise; thorough, but coherent; detailed, but understandable.

PGEWS

IRE took special notice of the engineer's responsibility to communicate just five years ago this month, when it approved the formation of a Professional Group on Engineering Writing and Speech (PGEWS). This Group was organized in the hope of improving the effectiveness, that is, the readability and the "listenability," of IRE publications and convention sessions. To date it has attempted to attain this goal by three types of activities: publications, meetings, and workshops.

Publications of the Group include a *Newsletter* which serves as an internal communication medium, TRANSACTIONS which include articles and book reviews, and *Symposium Proceedings* which contain all papers presented at national PGEWS meetings. The Los Angeles Chapter of the Group has also prepared a small eight-page booklet entitled "Techniques for Better Talks" which is distributed by IRE to all scheduled speakers for the Annual International Convention. Publication plans for the future include at least one new area of effort, the preparation of detailed instructional monographs which will cover various phases of the writing and speaking processes, as well as special aspects of various specific jobs.

PGEWS also sponsors its own national symposia, sessions at IRE national meetings, and local Chapter meetings. These meetings cover a wide variety of topics, ranging from basic principles of writing, editing, or speaking, to theoretical discussions on such subjects as information retrieval and automatic abstracting. Various formats are also used, including lectures, panel discussions, and formal debates.

Perhaps the most directly useful activity of PGEWS is the workshop type of program conducted by local Chapters. Because writing and speaking are skills to be mastered, rather than sciences to be learned from a book, both practice and criticism are essential parts of any training or improvement program. Local workship programs give all members of PGEWS an opportunity to increase their proficiency in oral, written, and graphic expression.

CONCLUSION

In the future, as in the past, the engineers who leave their mark on the profession will be those who recognize and accept their responsibility to share their accumulated knowledge with others; those who, having something to say, are willing to devote the necessary time and effort to saying it well to the right audience.

Section 15

HUMAN FACTORS

Organized with the assistance of the IRE Professional Group on Human Factors in Electronics

Human Factors in Electronics—Historical Sketch*

HENRY P. BIRMINGHAM†, SENIOR MEMBER, IRE

***Summary*—The growth of the new science dealing with man-machine relationships is traced from the first organized efforts during World War II to the present. The current rapid growth is underlined by the organization of professional societies, the largest of which is the IRE Professional Group on Human Factors in Electronics. The discussion includes reference to the several government, university and industrial laboratories which have pioneered in this field.**

ENGINEERS HAVE BEEN designing machines for human use since the start of civilization, and in doing this they have had to look closely at the man. In some instances only his physical structure and kinetic properties had to be considered, as was the case in the design of the screwdriver, the axe and the hand-saw. But often, as for example in the design of the automobile and the airplane, it was necessary to take into account more subtle properties. In these cases the engineer was required to consider not only the amount of force a man could apply, but also how quickly he could sense changes in energy and the way in which he responded to them.

However, during World War II, as the development of highly complex man-machine systems became commonplace, the need emerged for specialists in man-machine relationships. This has culminated in the start of an interdisciplinary science called variously biomechanics, engineering psychology, human engineering or human factors engineering. "Human factors engineering" has been popularly shortened to "human factors" and is the term generally applied to the science which is concerned with the characteristics of the human which are important in the design of systems.

* Received by the IRE, November 6, 1961.
† U. S. Naval Research Laboratory, Washington, D. C.

The new field has developed rapidly in the seventeen or eighteen years of its existence, and its importance has been recognized by the establishment of human factors groups within several of the older and larger professional societies including the American Society of Mechanical Engineers and the American Psychological Association. In 1956, a separate society, devoted to the new science, and called the Human Factors Society, was established.

In 1958, the Professional Group on Human Factors in Electronics was organized within the IRE with an administrative committee consisting of engineers who have been leaders in the new discipline. Assuming the continuance of existing growth curves, PGHFE, at the time of this publication, is the largest of these human factors groups. It is currently publishing Transactions on a semi-annual basis and has sponsored three annual human factors conventions.

The "human factors" approach to man-machine relationships emphasizes modification of the equipment to acheive man-machine compatibility, in contrast to the "selection and training" approach which attempts to fit the man to the machine. It was during World War II that such modifications of machine design was first applied on a large scale by people who became specialists in these relationships. Of particular significance during this early period was the work conducted under the National Defense Research Committee, exemplified by research on problems in anti-aircraft fire control at NDRC—project N-111 under Dr. William Kappauf. At the same time, in Great Britian and in Germany work was beginning on similar problems. It is of interest to observe that whereas in the United States and Great Britain this work was carried on primarily by engineers and psychologists, in Germany it developed under the Medical profession.

In the period immediately following World War II, growth of the young science continued in the Military and Industrial areas and in the Universities. In 1945, a human engineering group was established at the U. S. Naval Research Laboratory in Washington, D. C. by Dr. Robert M. Page. The group was headed by the late Dr. Franklin V. Taylor. At about the same time, a group at the U. S. Air Force Aeromedical Laboratory in Dayton, Ohio, also began work in research and design under Dr. Paul M. Fitts. It was not until a few years later that groups dealing in the new science were established in industrial laboratories. Early industrial groups formed include one under Dr. John Karlin at the Bell Telephone Laboratories, Murray Hill, New Jersey, and another under Dr. Stanley Roscoe at the Hughes Aircraft Company, Culver City, California.

Today, there is hardly a large industrial concern which does not boast at least one human factors group. Some, such as General Electric, have several. There are many in the various laboratories of the Department of Defense, and there is little reason to believe that the new scientific area has reached the end of its growth.

For the near future, a greater increase is seen in the interest of engineers, physicists and mathematicians in human factors engineering. Increasing numbers of young graduates in these fields are taking positions in human factors organizations with the intention of making this field their career.

It is expected that problems and approaches in the future will be largely in the systems area, and will call for knowledge of engineering, physics, mathematics and psychology on the part of all its workers, regardless of their primary field of education.

The Man-Machine System Concept*

D. T. McRUER†, SENIOR MEMBER, IRE, AND E. S. KRENDEL‡, SENIOR MEMBER, IRE

Summary—**The man-machine concept as developed in this paper refers to a closed-loop control system comprising a machine and an actively participating human operator. The human component attributes are described with both general and specific engineering models. Particular emphasis is placed on the human adaptive characteristics which make for an unusual control system.**

Predictive behavior is discussed in terms of actual-physical and conceptual-effective display organizations resulting in compensatory, pursuit, and precognitive system organizations. Adaptive-operator transfer characteristics are presented, and applications of the adaptive model for random-input compensatory system design are listed. As an example, the model is used to estimate operator and system characteristics for a specific system, and the estimates are compared with experimental results derived from associated empirical data. In a discussion of performance assessment, the basic attributes which any criterion for system performance should possess are reviewed, and a *modus operandi* for manual control system performance assessment is proposed.

Introduction

IN THE BROADEST sense, the man-machine system concept encompasses all the interactions between human and mechanistic elements when these elements are united into a system. However, as treated here, the concept refers to a system comprising a machine and an actively participating individual man. Central to the man-machine system concept in this context is the specification of human skills under circumstances where man's mechanical outputs serve as machine inputs, and machine outputs are among the inputs to the human. Historically, descriptions of human skills evolved by engineers were reflections of the then current state of technology. Early man-machine systems used man primarily as a power source, and descriptions of man were couched in such terms as "engine efficiency" and "duty cycle." With the rising importance of control devices, the language and techniques of control engineering have become increasingly appropriate for describing human skills. As the coupling of men with data-processing devices becomes increasingly critical for effective man-computer symbioses, perhaps even more effective communications and computer engineering descriptions of man will evolve.

Those human contributions to a control system which make manned control systems most desirable are unfortunately the very contributions most difficult to describe in conventional engineering terms. Such human component attributes as judgment, multimode capability, and adaptability make for a remarkable control system. Of these, human adaptability is the attribute which can most nearly be described in engineering terms. Within the confines of the physical constraints

* Received by the IRE, July 5, 1961; revised manuscript received, November 24, 1961.

† Systems Technology, Inc., Inglewood, Calif.
‡ The Franklin Institute Labs. for Research and Development, Philadelphia, Pa.

present in a given application, human adaptability includes:

1) Organization of the system structure.
   a) Selection and use, as feedback quantities, of those sensible machine outputs best suited for control purposes.
   b) Setup of an internal organization (equivalent to the construction of several signal-processing paths within the human) which effects highly efficient use of any coherence in the presented stimuli (system forcing function and feedbacks).
2) Adjustment, within the system structure as organized, of transfer characteristics to forms appropriate for control.

The attributes of judgment, multimode behavior, and adaptability in a control system component allow the design of systems of great reliability and versatility. By using judgment, the man can evolve and modify performance criteria, select relevant inputs, decide between programmed and impromptu responses, and so forth. A multimode capacity enables the man to implement different modes of behavior and thus to effect the diverse requirements which judgment may dictate. To design such control systems, engineering descriptions of the critical component, man, are needed.

As descriptions of man improve, at least three desirable effects will result: the synthesis of better man-machine systems will be possible; fruitful new insights into the mechanistic inner-working of man will evolve; and, since man is the archtype adaptive controller, potent analogical insights for new automatic adaptive controllers will emerge.

Implementation of the Man-Machine System Concept

Implementing the man-machine system concept implies that the techniques and data necessary to synthesize a desirable system are available in a convenient form. The generalized model of Fig. 1, which illustrates the variety of human adaptability listed above, defies description in terms of present-day analytic techniques. Perhaps the only means for describing the contents of Fig. 1 in a quasi-analytic sense would be to generate an elaborate catalog of cause-and-effect relationships. Science and engineering make progress, however, by simplifying their object of study so that a succinct catalog of cause-and-effect relationships is feasible, and then expanding this kernel of knowledge by successive elaborations. Thus, if the system is restricted by design to include only the path indicated by the heaviest line in Fig. 1 (redrawn as Fig. 2), the result is a rudimentary system which is fundamental to many man-machine applications, convenient for measuring a large amount of human dynamics data, and suitable for subsequent elaboration. In this compensatory system, the forcing function and controlled-element are fixed. The human's internal organizing ability is suppressed by providing Gaussian or otherwise random-appearing inputs, thus making the organization of both the physical and the effective system structures identical. Within this fixed system, the human is characterized by a describing function and remnant[1] achieved by measuring a smoothed-over view of the control process via cross- and power-spectral densities.[2,3] Generalization beyond the fixed parameters is accomplished by empirically determining analytical models and adjustment rules for Y_p and Φ_{nn} as functions of the constraints imposed by the other system elements.

As an alternative to the describing function technique, the experimenter can pit his skill as a computer specialist against the problem.[4–6] This results in the selection of useful analog computer mechanizations to simulate the human operator in a constrained control task. The criterion for goodness of simulation can be visual, for a point-to-point correspondence; best in a mean-square sense; or best according to other tests for goodness of fit. The analog model is more difficult to generalize for other system situations than analytical approximations to the describing function plus remnant data. However, human time-dependent effects and such nonlinearities as can be generated on an analog computer are more easily measured directly using simulation methods, although these effects are implicitly contained in describing function and remnant measurements.

As in most nonlinear system studies, an intermix of describing function and computer simulation models is appropriate. Their application is somewhat unconventional, because any description of human control dynamics requires that the measurements generated include the gross variability inherent in sampling human behavior. Such variability imposed by the nature of the human subject can be minimized by the nature of the task, as well as by stratifying the subjects according to levels of skill, experience, or other criteria.

Engineering Models of Human Behavior

The most complete engineering model for human behavior has been developed for the compensatory, random-appearing forcing function situation described by Fig. 2. When averaged data generated by different experimenters are brought together and generalized, a

[1] A. Tustin, "The nature of the operator's response in manual control and its implications for controller design," *J. IEE*, vol. 94, pt. IIA, pp. 190–202; May, 1947.

[2] W. H. Huggins, "Memo on the Experimental Determination of Transfer Functions of Human Operators and Machines," Cambridge Field Station, AMC, Cambridge, Mass., Memo E-4070: 1949.

[3] E. S. Krendel and G. H. Barnes, "Interim Report on Human Frequency Response Studies," Flight Control Lab., Wright-Patter-AFB, Dayton, Ohio, WADC TR 54–370; June, 1954.

[4] "Final Report: Human Dynamic Study," Goodyear Aircraft Corp., Akron, Ohio, Rept. No. GER-4750; April, 1952.

[5] N. D. Diamentides, "A pilot analog for airplane pitch control," *J. Aeronautical Sciences*, vol. 25, pp. 361–370; June, 1958.

[6] J. R. Ward, "The Dynamics of a Human Operator in a Control System—a Study Based on the Hypothesis of Intermittency," Ph.D. dissertation, University of Sydney, Australia; May, 1958.

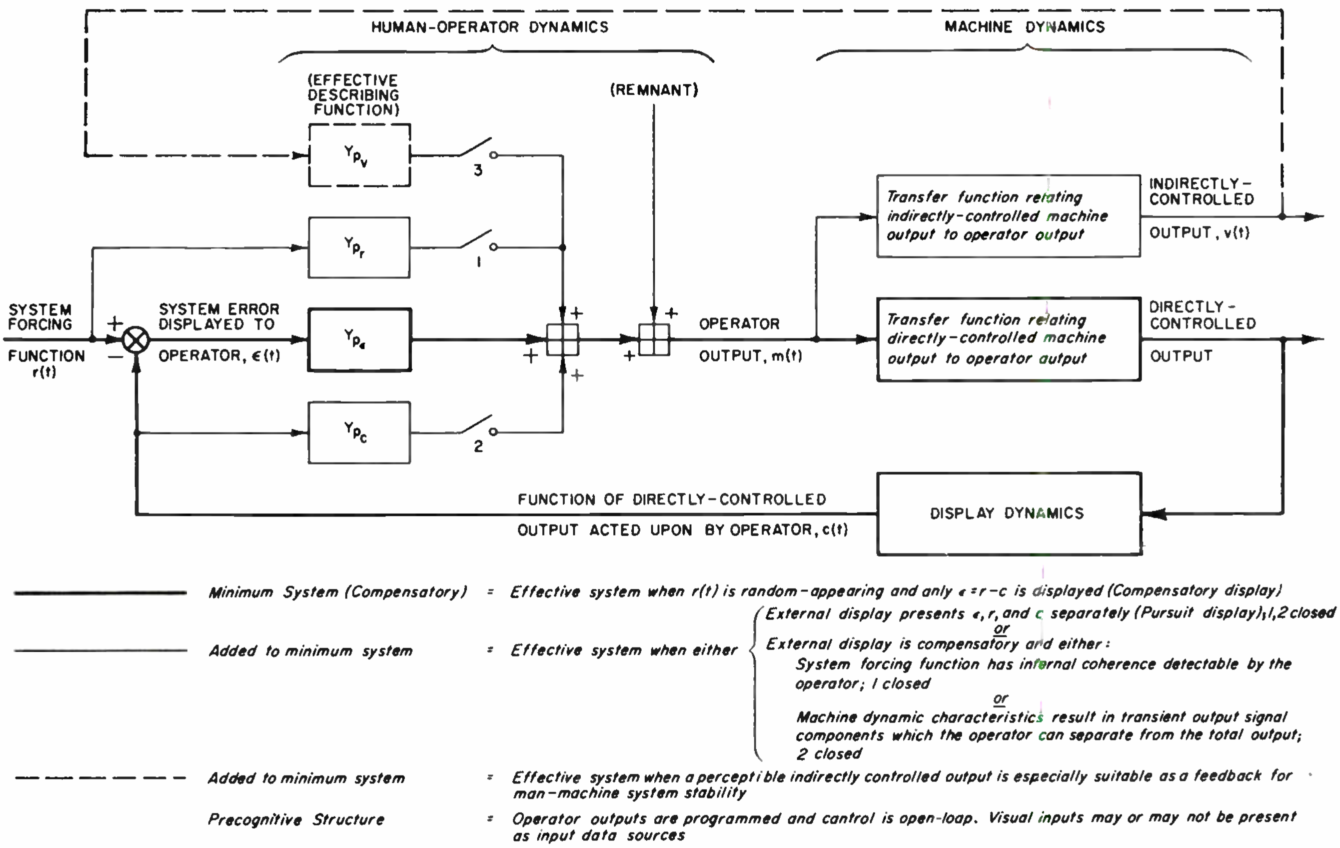
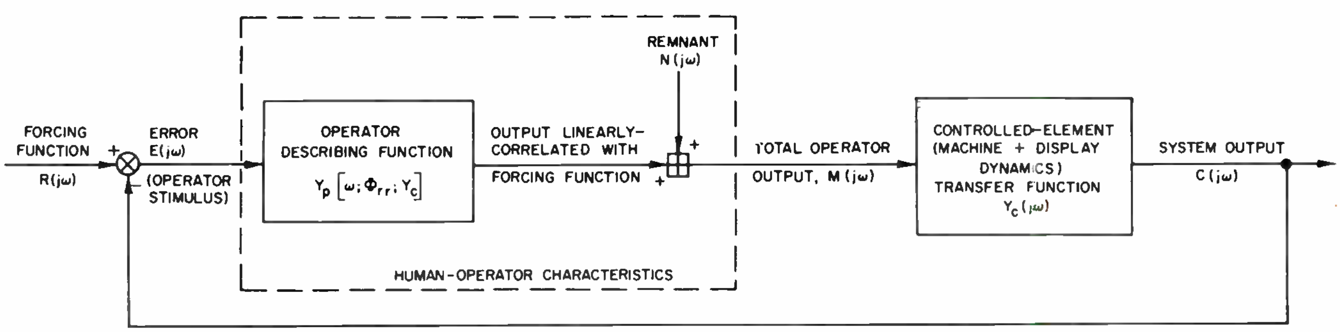


Fig. 1—Possible man-machine effective system structures.



MEASURABLE QUANTITIES

$R(j\omega)$, $E(j\omega)$, $M(j\omega)$, $C(j\omega)$ — FOURIER TRANSFORMS OF OBSERVABLE TIME SIGNALS

$\Phi_{rr}(\omega)$, $\Phi_{ee}(\omega)$, $\Phi_{mm}(\omega)$, $\Phi_{cc}(\omega)$ — POWER-SPECTRAL DENSITIES OF OBSERVABLE SIGNALS

$\Phi_{re}(j\omega)$, $\Phi_{rm}(j\omega)$ — CROSS-SPECTRAL DENSITIES BETWEEN FORCING FUNCTION AND OPERATOR INPUT AND OUTPUT

DERIVED OPERATOR CHARACTERISTICS

DESCRIBING FUNCTION, $Y_p(j\omega) = \dfrac{\Phi_{rm}(j\omega)}{\Phi_{re}(j\omega)}$; REMNANT POWER-SPECTRAL DENSITY, $\Phi_{nn}(\omega) = |1 + Y_c Y_p|^2 \Phi_{mm} - |Y_p|^2 \Phi_{rr}$

$[\Phi_{rn}(j\omega) \equiv 0$ BY DEFINITION; $\Phi_{rr}(\omega)$ AND $Y_c(j\omega)$ HELD CONSTANT$]$

Fig. 2—Human response measurements for compensatory systems.

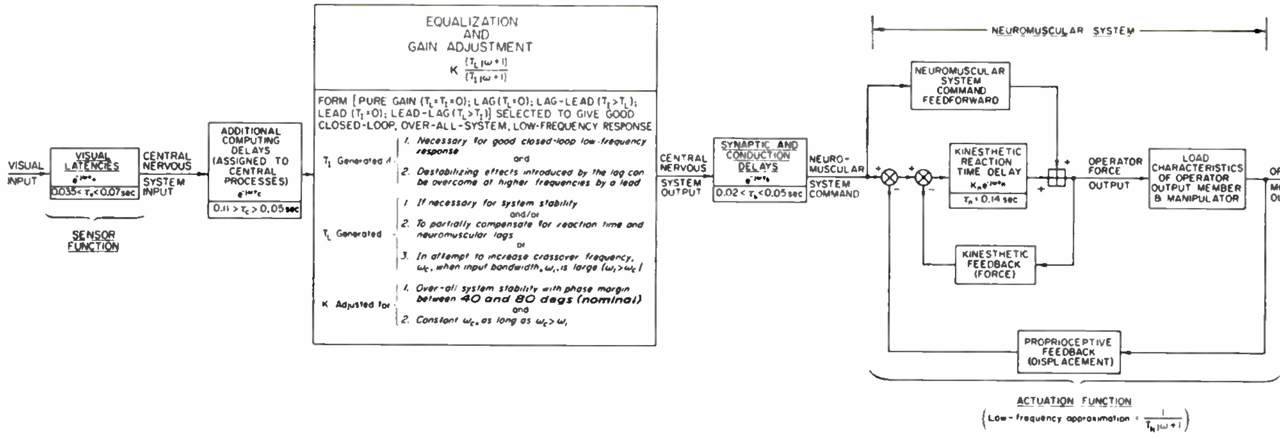


Fig. 3—Fully articulated hypothetical model of human operator.

model for human dynamics consisting of two essential elements emerges:[7–9]

1) A generalized describing function form.
2) A series of "adjustment rules" which specify how to "set" the parameters in the generalized describing function to form an approximate model of pilot behavior for the particular situation considered.

Fig. 3 illustrates a detailed model of this generalized describing function; some nonlinear and internal feedback features are not included, and no effort has been made to describe neural conduction in physiological detail. This detailed model is based not only upon spectral measurements of the form noted in Fig. 2, but it also reflects data from psychological and physiological experiments on various human "subsystems." The block diagram components do not necessarily correspond to physical elements within the human, but the structure shown is the minimum required to exhibit equivalent behavior on even a quasi-linear basis over the frequency range from zero to 5 cps.

This model applies only when the operator cannot effectively restructure the physical system. Consequently, while providing simplicity of analysis, the measurement situation of Fig. 2 also obscures the extent to which the human can take advantage of the coherence and predictability which may be present in the forcing function. For predictable aspects of the forcing function to be detected, the forcing function must be readily distinguishable from the control responses which the operator makes. Furthermore, for predictive behavior to be most effective, the human must have a preview of the forcing function so that he may prepare his response in advance. As an analogy, consider the steering of an automobile over a winding road. The speed of the vehicle and the bends in the road define a system forcing function. When visibility is clear and the road is flat, it is possible to predict the steering actions well in advance, and to proceed for intervals of time in an essentially open-loop fashion. Were the road hilly and the visibility poor, the driver would be able to predict the curves for only a short distance, and he would have to maintain closer control of his vehicle. Finally, if a dense fog closed in on the driver, so limiting his visibility that he could only see his hood ornament and the white center line of the road a few feet ahead of his vehicle, prediction would be extremely difficult, and the driver would be tracking an essentially compensatory display (as presented in Fig. 2). These three conditions, which merge at the boundaries, can be idealized in the precognitive, pursuit, and compensatory organizations subsumed by Fig. 1.

The compensatory, pursuit, and precognitive "displays" given physical meaning above can also be thought of as effective system organizations which result when the human learns to distinguish and to act upon predictable elements in the forcing function.[10] It is this ability to construct a succession of conceptual displays, which are effectively equivalent to more elaborate and often unfeasible physical displays, which constitutes adaptation to the input. These successive organizations of perception in terms of effective displays provide a model for the learning of motor skills and give the design engineer an upper limit for possible human performance in a tracking task.

[7] A complete summary, interpretation, and correlation of all data available to 1957 appears in D. T. McRuer and E. S. Krendel, "Dynamic Response of Human Operators," Flight Control and Aeromedical Labs., Wright-Patterson AFB, Dayton, Ohio, WADC TR 56-524; October, 1957.

A condensed version is given by these authors in "The human operator as a servo system element," *J. Franklin Inst.*, vol. 267, pp. 381–403, May, and pp. 511–536, June, 1959.

[8] Data generated after 1957 are presented in the same context as above in D. T. McRuer, I. L. Ashkenas and C. L. Guerre, "A Systems Analysis View of Longitudinal Flying Qualities," Flight Control Lab., Wright-Patterson AFB, Dayton, Ohio, WADD TR 60-43; January, 1960.

[9] A summary of quasi-linear operator models developed prior to 1956, with emphasis on psychological aspects, is J. C. R. Licklider, "Quasi-linear operator models in the study of manual tracking," in "Developments in Mathematical Psychology," R. D. Luce, Ed., Free Press of Glencoe, Ill., pp. 169–279; 1960.

[10] E. S. Krendel and D. T. McRuer, "A servomechanisms approach to skill development," *J. Franklin Inst.*, vol. 269, pp. 24–42; January, 1960.

Human adaptation to a control task may also occur by the selection of the most effective feedback paths (Fig. 1). Although this form of adaptation has not been submitted to as extensive a measurement program as has adaptation to controlled-element or adaptation to the input, it nonetheless holds great promise for future research in that it may define a clear-cut engineering basis for the design of complex displays.

Performance Assessment

The choice of criterion is the essential problem in the assessment of any control system. However, in a man-machine system the human component adds a new dimension to the evaluation of system performance, since the human is both an adaptive-operational device and a measuring device.

The basic attributes which any criterion for system performance should possess are reliability, ready applicability, and selectivity; and for control systems with deterministic inputs, criterion forms which possess each of these characteristics to a sufficient extent can readily be selected.[11] However, when Gaussian random inputs are used, "optimum" linear systems based upon minimization of the mean-squared-error are substantially the same as "optimum" systems derived by using non-mean-squared-error criteria.[12] Time-weighting and other devices to improve selectivity are usually not applicable, and such criteria as time-on-target are difficult to relate to system and forcing function characteristics. For these and other reasons, the performance measures ordinarily used for assessment of continuously operating man-machine systems are $\overline{\epsilon^2}$ or the closely related $\overline{|\epsilon|}$.

Unfortunately, for constant forcing function characteristics, human adaptability renders the mean-squared-error particularly unselective as a system performance measure. The human tends to maintain a fairly constant mean-squared-error by modifying his dynamics to compensate for changes in the controlled-element dynamics. Ideally, man-machine system assessments would combine two features: system performance, as measured by an average such as $\overline{\epsilon^2}$, and insight into how this performance is achieved by the system. When detailed measurements of describing functions are possible, they can serve the second purpose. An alternate procedure is the use of expert evaluations, as in pilot opinion ratings, to give a single, over-all measure. When tightly constrained situations are imposed upon the human controller, and several competing systems are to be compared, such expert ratings exhibit far less variability and unreliability than might be expected from common experience with routine opinion polls. Indeed, recent work has shown that the basically subjective expert ratings can give surprisingly objective indications of both over-all system performance and the operator describing function generated; this work has also shown that the man can rank, in a highly reproducible and valid manner, the system he controls in terms of these objective characteristics.[8] Thus, either human response measurements or expert operator differential ratings which are highly selective, are useful adjuncts to the more conventional, but less selective, average performance measures.

Applications of Human Behavior Models

All the ramifications of the general model in Fig. 1 have been used for design purposes; the simplest to illustrate in analytical application is the specialized model (Fig. 3) which applies only to compensatory situations. For most system studies, a simpler form (Fig. 4) is adequate to describe the system low-frequency (0–1.5 cps) characteristics. Here the adaptive features—the human's prime virtue as a system component—are separated from the relatively unalterable characteristics. Prominent among the latter is the reaction-time delay τ, which is ordinarily the dominant human-con-

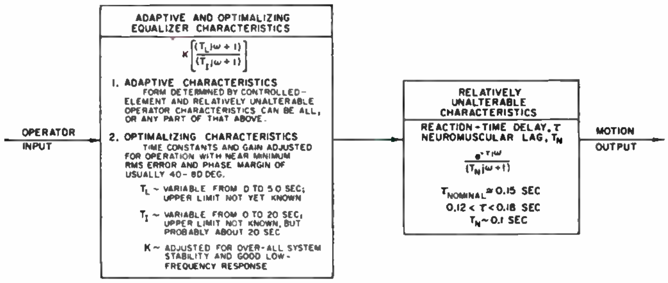


Fig. 4—Condensed version of human operator model for compensatory operation.

tributed liability in a manual control system. To estimate operator behavior in a given task, the adaptive equalizer characteristics are adjusted to values similar to those a servoengineer would select if he had the equalization $K(T_L j\omega+1)/(T_I j\omega+1)$ available to compensate for and close the loop about process dynamics comprising the unalterable operator characteristics in series with the controlled-element transfer function. When the resulting gain crossover frequency ω_c is greater than the effective system forcing function bandwidth ω_i, and when any dominant closed-loop system quadratic mode has a damping ratio ζ_{CL} greater than 0.35, these estimated operator transfer characteristics will be a reasonable approximation to experimentally obtained values. If either $\omega_i > \omega_c$ or $\zeta_{CL} < 0.35$ result, the actual operator describing function form will differ from the model characteristics; the operator will generate higher-order leads (in a generally unsuccessful attempt to make $\omega_c > \omega_i$), and will exhibit erratic nonlinear behavior.

The adaptive model for random-input compensatory systems described above is useful for many different an-

[11] D. Graham and R. C. Lathrop, "The synthesis of 'optimum' transient response: Criteria and standard forms," *Trans. AIEE*, vol. 72, pt. 2, pp. 278–288; 1953.

[12] S. Sherman, "Non-mean square-error criteria," IRE Trans. on Information Theory, vol. IT-4, pp. 125–126; September, 1958.

alytical purposes in manual control design. For example, it has been used to:

1) Estimate human and over-all man-machine system response and performance when the system forcing function and controlled-element dynamics are known.[7,8,13]
2) Determine the barely uncontrollable controlled-element dynamics and controllability boundaries.[8,13,14]
3) Find the maximum forcing function bandwidth compatible with reasonable control action on the part of the human.
4) Indicate the type of additional system equalization (to be achieved via the display, in the control system, or by controlled-element modifications) desirable to achieve best manual control in the sense of minimum operator-equalization requirements.[15,16]
5) Determine desirable machine outputs (for multi-degree-of-freedom machines) for feedback to the display.
6) Delineate those features of the machine dynamics which are most important from a vehicle-handling-quality standpoint.[8,13,14]

The analysis of compensatory manual control systems can be approached in much the same way as that for ordinary automatic-feedback systems. As a simple example, typical system performance predictions based upon the adaptive-operator model will be illustrated for the control system shown in Fig. 5. The operator-adapted describing function for control of the pure gain controlled-element will be

$$Y_p \approx \frac{K_p e^{-j\omega\tau}(T_L j\omega + 1)}{(T_I j\omega + 1)(T_N j\omega + 1)} \approx \frac{K_p e^{-0.15 j\omega}}{(T_I j\omega + 1)}. \quad (1)$$

In (1), the lag time constant T_I will be much greater than the lead time constant T_L to achieve good low-frequency closed-loop system response; and T_L will be adjusted to approximately cancel the neuromuscular lag T_N. With the simplified form of (1), the phase margin ϕ_M is

$$\phi_M = \pi - \tan^{-1} T_I\omega_c - \tau\omega_c. \quad (2)$$

[13] I. L. Ashkenas and D. T. McRuer, "The Determination of Lateral Handling Quality Requirements from Airframe Human-Pilot Systems Studies," Flight Control Lab., Wright-Patterson AFB, Dayton, Ohio, WADC TR 59-135; June, 1959.

[14] H. R. Jex and C. H. Cromwell, "Theoretical and Experimental Investigation of Some New Longitudinal Handling Quality Parameters," Flight Control Lab., Wright-Patterson AFB, Dayton, Ohio, ASD TR 61-26; March, 1961.

[15] "The Artificial Feel System," Northrop Aircraft, Inc., Hawthorne, Calif., BuAer Rept. AE-61-4V; May, 1953.

[16] H. P. Birmingham and F. V. Taylor, "A design philosophy for man-machine control systems," Proc. IRE, vol. 42, pp. 1748–1758; December, 1954.

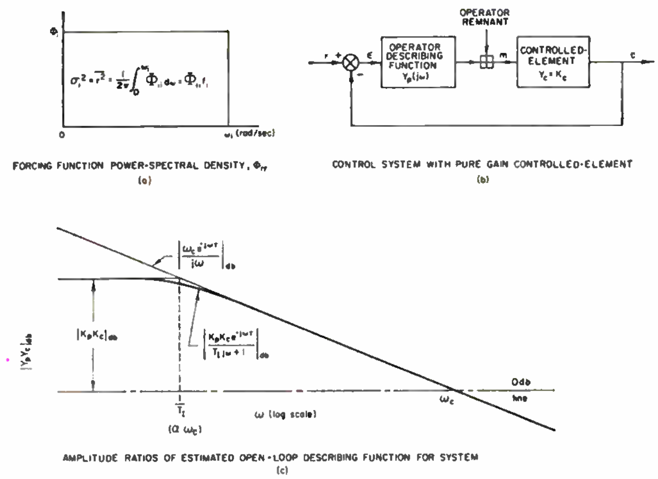


Fig. 5—Given and estimated characteristics for manual control system with $Y_c = K_c$.

Since $\omega_c \gg 1/T_I$ for good closed-loop system response, (2) becomes, approximately,

$$\phi_M \approx \pi - \left(\frac{\pi}{2} - \frac{1}{T_I\omega_c}\right) - \tau\omega_c.$$

Letting $\alpha = 1/T_I\omega_c$,

$$\omega_c \approx \frac{1}{\tau}\left[\frac{\pi}{2} + \alpha - \phi_M\right] \quad (3)$$

This relationship is shown in Fig. 6 for a phase margin of 40 degrees. Three values of τ are used to bound this parameter, or alternatively, to take into account slight phase margin shifts, for a nominal τ value, from the 40 degrees assumed. Any value of α less than 0.1 is compatible with good closed-loop system response, so an $\alpha = 0.1$-line provides the final boundary of a probable crossover frequency region. Using the middle of the region, the estimated crossover frequency will be approximately 6 rad/sec. This estimated crossover will remain approximately the same for all forcing function bandwidths less than 6 rad/sec. When the forcing function bandwidth exceeds 6 rad/sec, the describing function form will be modified drastically from that of (1) to one exhibiting far more lead, and the system low-frequency closed-loop response will be poor. Thus, the estimated crossover frequency is not only an important system response parameter for $\omega_i < \omega_c$, but it also gives the maximum forcing function bandwidth compatible with reasonably good system response.

As long as $\omega_c \gg 1/T_I$, an open-loop system describing function given by

$$Y_pY_c \approx \frac{\omega_c e^{-j\omega\tau}}{j\omega} \quad (4)$$

is suitable to estimate closed-loop system characteristics with little error. If operator remnant is ignored, the

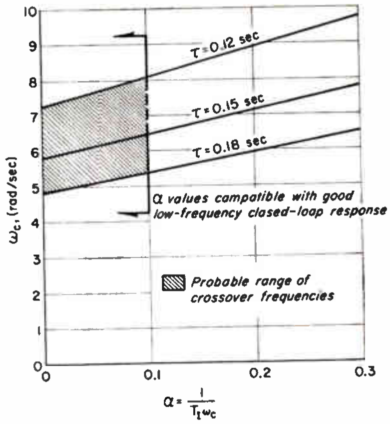


Fig. 6—Crossover frequencies for 40 degrees phase margin.

mean-squared-error, using this simplified form, will be

$$\overline{\epsilon^2} = \frac{1}{2\pi}\int_0^{\omega_i} \Phi_{\epsilon\epsilon}(\omega)d\omega$$

$$= \frac{\Phi_{ii}}{2\pi}\int_0^{\omega_i} \frac{\omega^2}{\omega^2 - 2\omega\omega_c \sin \omega\tau + \omega_c^2} d\omega. \quad (5)$$

Since $\sin \omega\tau \approx \omega\tau$ for $\omega < \omega_i$, the relative mean-squared-error becomes

$$\frac{\overline{\epsilon^2}}{\sigma_i^2} \approx \frac{1}{(1-2\omega_c\tau)}\left[1 - \frac{\omega_c/\omega_i}{\sqrt{1-2\omega_c\tau}} \tan^{-1}\frac{\omega_i}{\omega_c}\sqrt{1-2\omega_c\tau}\right] \quad (6)$$

or

$$\frac{\overline{\epsilon^2}}{\sigma_i^2} \approx \frac{1}{3}\left(\frac{\omega_i}{\omega_c}\right)^2 \quad (7)$$

if the arctangent is carried only to the third term in its expansion.

The various operator and system characteristics estimated above are compared in Fig. 7 and Table I with experimental results derived from data originally obtained by Elkind.[17] The major difference between the analytical and experimental operator models is the experimentally observed variation of T_I with ω_i. This is significant for fine-grained operator descriptions (for example, the observed T_I, ω_i variation gives credence to a more refined version of the model in Fig. 3 containing nonlinear elements enclosed by an internal feedback structure), but it is not extremely important for preliminary design calculations.

This elementary example illustrates pragmatically the major conclusion of this paper: that man, acting as a control system component, *can be* described in useful engineering terms.

[17] J. I. Elkind, "Characteristics of Simple Manual Control Systems," Lincoln Lab., Mass. Inst. Tech., Cambridge, Rept. 111; April, 1956.

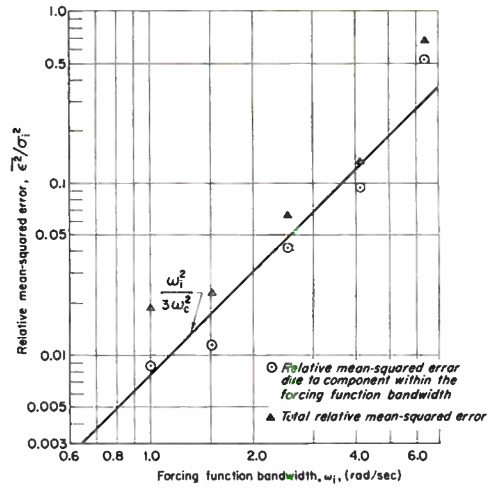


Fig. 7—Comparison of estimated and experimental mean-squared error.

TABLE I

COMPARISON OF ESTIMATED AND EXPERIMENTAL OPERATOR CHARACTERISTICS

Item	Estimated from Analytical Model	Averaged Data from Experiment*
Operator Describing Function Form	$\frac{K_p e^{-j(0.15)\omega}}{T_I j\omega + 1}$	$\frac{K_p e^{-j(0.14)\omega}}{T_I j\omega + 1}$; $K_p/T_I \approx 8.2$; $K_p = 87/\omega_i^2$
Crossover Frequency	6 rad/sec	6–8.2 rad/sec
Critical Forcing Function Bandwidth	About 6 rad/sec	About 6 rad/sec
Phase Margin	40 degrees	42 degrees

* Derived from interpretation and, in some cases, extrapolations of Elkind's[17] R 0.16, R 0.24, R 0.40, R 0.64, R 0.96, R 1.6, and B6 experimental results.

Communication Between Man and Machine*

J. E. KARLIN†, MEMBER, IRE, AND S. N. ALEXANDER‡, FELLOW, IRE

Summary—**The inclusion of computer-based assemblies into man-machine complexes has enhanced the relevance of characteristics of the human link in the system. In particular, the virtues and limitations for transmission of information among the physical subsystems through a human channel is an important consideration in systems design. The current status of the supporting research for this area is briefly summarized. The related but more recondite area covering communication between man and computer is also important to the effectiveness of man-machine complexes. However, supporting research for this area is meager and most of it is contained in engineering psychology studies pertinent to the needs of specific systems. Research facilities more appropriate for the exploratory needs of both areas are becoming available and these presage the broad intensive programs needed to provide research guidance to the systems design of man-machine complexes.**

Introduction

THIS PAPER will consider two basic aspects of communication between men and machines:

Part I: Man as a transmission channel with short-term information storage properties, and

Part II: The relation between man and computer.

We will discuss basic problems and principles rather than summarize design information. The latter is quite voluminous and is readily available.[1-5]

Part I: Properties of the Human Transmission Channel

The general problem, diagrammed in Fig. 1, concerns the transmission of information between two points in a physical system via the human channel. Three aspects of the human channel will be considered: input, output and input-output compatibility. The principal discussion will be of the input.[6]

Input and Output

What are the important properties of man's input (sensory end organs) and output (motor end organs) which determine the design of the machine input (display) and output (control) for optimum information transmission?

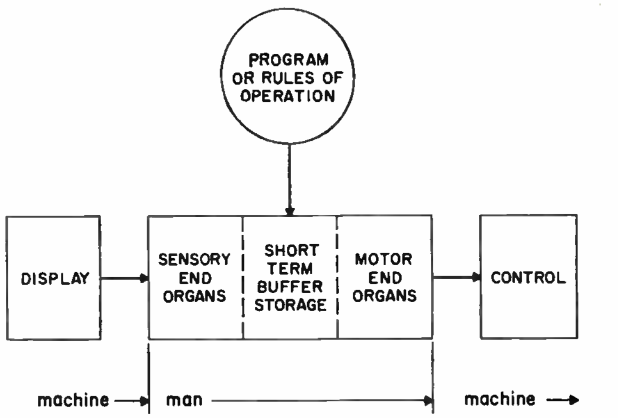


Fig. 1—Man programmed through rules of operation for "one to one" transmission where output and input information should be identical if transmission is perfect.

Input-Output Compatibility

In certain cases optimum matching of devices or circuits to man's input and output characteristics depends on a particular compatibility relationship between the two.

The discussion of these three aspects will be restricted on two scores: 1) to simple "one-to-one" information transmission and processing situations which involve only man's short-term buffer storage properties—long-term storage and information transformation are taken up in the next section on computers, and 2) to open-loop situations where man's output is not fed back to modify signals to his input. The feedback problem is taken up in a companion paper in this issue.[7]

Input Properties

Unidimensional Displays (for a single signal)

In a man-machine system the input to the man is often a single light or a sound, chosen from an ensemble of lights or sounds and invariant over its duration and spaced at relatively long intervals, where there is no premium on his speed. Where the signals in the ensemble differ only in one dimension, like frequency or amplitude, the capacity of the human channel for transmitting information is very limited—of the order of two to three bits. For example, as shown in Fig. 2, untrained

* Received by the IRE: Part I—November 17, 1961; Part II—February 14, 1962.

† Bell Telephone Laboratories, Murray Hill, N. J.

‡ National Bureau of Standards, Washington, D. C.

[1] P. M. Fitts, "Engineering psychology," *Ann. Rev. Psych.* vol. 9, pp. 267–294; 1958.

[2] A. W. Melton and G. E. Briggs, "Engineering psychology," *Ann. Rev. Psych.*, vol. 11, pp. 71–98; 1960.

[3] A. T. Welford, "Ergonomics of Automation," Problems of Progress in Industry, Dept. of Scientific and Industrial Research, London, England, No. 8; 1960.

[4] W. E. Woodson, "Human Engineering Guide for Equipment Designers," University of California Press, Berkeley, 1954.

[5] C. T. Morgan, Ed., "Human Engineering Guide to Equipment Design," McGraw-Hill Book Co., Inc., New York, N. Y., to be published in 1961–1962. (Prepared under direction of Joint Services Steering Committee.)

[6] Additional detailed information is contained in G. H. Mowbray and J. W. Gebhard, "Comparison and Interaction Among Sensory Input Channels," Johns Hopkins University, Applied Physics Lab., Baltimore, Md., TG-264; March, 1956.

[7] D. T. McRuer and E. S. Krendel, "The man-machine system Concept," this issue, pp. 1117–1123.

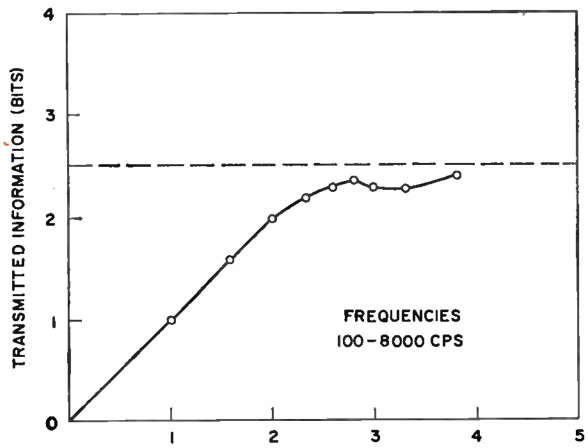


Fig. 2—As the amount of input information per signal is increased by increasing from 2 to 14 the number of different frequencies to be identified, the amount of transmitted information approaches, as its upper limit a channel capacity of about 2.5 bits per judgment.

man can no longer identify a single pure tone accurately if it can assume more than about six values of frequency over the audio range. The same limitation on "absolute" discrimination holds for intensity of tone, visual position of a pointer on a line and color hues.[8] This small number of bits tends to be constant for all of man's sensory inputs for signals presented singly and differing only in a single dimension.

It is evident in everyday situations that the human channel can transmit considerably more than this handful of bits. Three variables have been studied experimentally to provide better understanding of the problem of increasing such transmission of information: more complex single signals (multidimensional displays), sequence bursts of signals and multiple sensory inputs.

Multidimensional Displays (for a single signal)

A two-dimensional display normally results in the human channel transmitting more information than from a one-dimensional display but typically less than the sum of the information in both dimensions. For example, a tone with 2.5 bits of pitch information and 2.3 bits of loudness information results in only 3.1 bits being transmitted through the human channel; similarly a dot in a square gives 4.6 bits rather than twice the 3.5 bits from a dot on a line. In general, as Fig. 3 shows,[8] the proportionate gain in human channel capacity diminishes as the number of display dimensions increases; the upper limit to this kind of single signal transmission still appears to be small, something less than 10 bits. Furthermore, the human being requires more and more training to benefit as the dimensionality is increased.

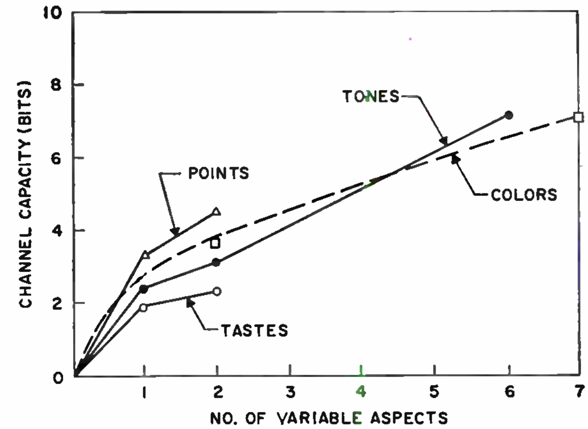


Fig. 3—The general form of the relation between channel capacity and the number of independently variable attributes of the signal.

Sequential Displays

Transmission through the human channel can be increased by the use of a signal burst input. The limit to such an increase depends primarily on the time elements involved and the decay properties of the human buffer storage.

How much information can be transmitted by the human channel in a signal burst? In classical psychology this limit has been linked with the span of immediate memory. As Fig. 4 shows, it turns out that almost regardless of the code or information content of each signal this limit remains in the region of 5 to 9 signals.[8] It can be seen that after looking at or listening to a sequence of letters or digits or words the average man can call back only about seven. Although this number of signals tends to be invariant, the amount of information transmitted through the human channel depends on the code and increases as the information content of each signal increases.

The limitation in information transmission for a burst does not appear to be imposed by the sensory input of the human channel but rather by the decay properties of the buffer storage and possibly the interference set up by the motor output action. It has been demonstrated, for example, that when a random sequence of 18 letters is presented to the human channel almost all 18 (between 70 and 80 bits) enter the channel even though only the 6 or 7 (about 30 bits) permitted by the memory span can be called back.[9] The rest are presumed lost through decay in his buffer storage while he is calling back the letters; this decrement is shown to occur within one second, as shown in Fig. 5. In this technique the man is exposed to the 18 letters (6 groups of 3) for 0.05 sec and then calls out one group of three consecutive letters where the group is identified at random immediately *after* the termination of the signal. The number of letters read out correctly is mul-

[8] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *J. Psych. Rev.*, vol. 63, no. 2; 1956.

[9] E. Averbach and G. Sperling, "Short term storage of information in vision," *Information Theory, Fourth London Symposium*, Butterworth and Co., Ltd., London, England; 1961.

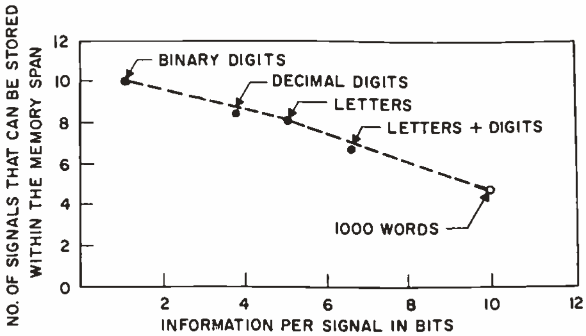


Fig. 4.

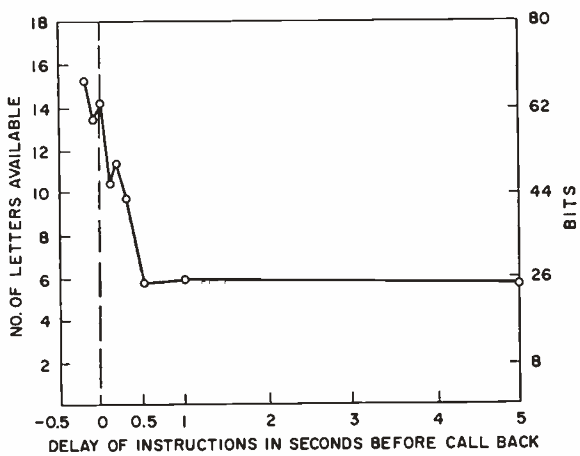


Fig. 5.

tiplied by 6 and this estimate, on the average, represents the information reaching the buffer store.

How fast can information be transmitted by the human channel? This is difficult to measure. In the preceding experiment it would appear that about 70 bits reached the input short-term store as a result of the 0.05-sec signal and that any part of this information is available to the human channel for immediate readout. However, the amount available decreases to an asymptotic value of less than 30 bits within a second.

Output and Input-Output Compatibility Properties

Data on man's output properties and the effect of input-output compatibility[10] are still quite meager. The references cited in this article summarize the information available. The most efficient form of output is probably speech rather than hand movement both because vocal movements are more rapid[11] and accurate and because speech signals are very familiar and contain a large number of bits per signal. However, the availability of machines responsive to voice signals is still very limited.

Multiple Sensory Inputs

Attempts to increase information flow through the human channel by using multiple sensory inputs have not been very fruitful. Optimizing transmission through a single input has shown better results. Perhaps of more importance to systems designers is the situation where the human channel is forced to respond to two message sources in parallel. In such cases some help can be obtained by having the sources feed different sensory inputs.[6]

Over-all Human Channel Capacity

What is the upper bound to human channel capacity? It was pointed out above that 70 bits of visual information presented for 50 msec to the eye are immediately available for readout. This suggests that the storage rate of the input may be considerably above 70 bits per second. However, it was also pointed out that this availability diminishes considerably within one second and, indeed, most measures of the output rate have shown a figure between 10 and 40 bits per second.[12] The highest rates have been in connection with the reading of words.[13] The limitation is apparently in the processing of the input information rather than because of mechanical output speed. For sustained communication beyond a few seconds the practicable rates for most human activities have been closer to 10 bits per second than to 40.[12] What happens when the human channel is overloaded has been studied in a number of experiments.[14]

Finally, the human channel has highly developed properties for resolving differences between signals. Two signals presented within a fraction of a second of each other can be discriminated with 1 per cent or less of the difference required for identification when presented singly under the same conditions.[6] Transmission through the human channel should be improved very considerably if ways could be found to use such "relative" information in displays.

Part II. The Relation Between Man and Computer

The foregoing summary of current information about human channel capacity in the simple communication chain shown in Fig. 1 must be interpreted carefully in shifting from laboratory environments to operating man-machine systems. While the theoretical measure of the information being transmitted during experiments may be held fixed, the way in which the human transforms the format of this information can not be disre-

[10] P. M. Fitts, "Human Information Handling in Speeded Tasks," IBM Corp., Research Center, Yorktown Heights, N. Y., Res. Rept. RC-109; 1959.

[11] For example, M. Braunstein and N. S. Anderson, "A comparison of the speed and accuracy of reading aloud and keypunching digits," IRE Trans. on Human Factors in Electronics, vol. HFE-2, pp. 56–57; March, 1961.

[12] H. Quastler, Ed., "Information Theory in Psychology," The Free Press, Glencoe, Ill.; 1955.

[13] J. R. Pierce and J. E. Karlin, "Reading rates and the information rate of a human channel," *Bell Sys. Tech. J.*, vol. 36, pp. 497–516; March, 1957.

[14] For example, J. G. Miller, "Information input overload and psychopathology," *Am. J. Psychiatry*, vol. 116, no. 8; February, 1960.

garded in an operational environment. His ability to filter out or abstract the designated information from a rich sensory pattern is dependent upon prior conditioning or learning. In terms of the model of Fig. 1, a significant element of "programmed response" is employed in filtering and converting the input sensory pattern into the format for an output response. This provides opportunity for subtle inclusion of prior information derived from his long-term memory which is associated with the "programmed response." This effect may distort the communication experiment, but often it is the source of enhanced operational performance that justifies use of a human link in the communication chain. An example is the way in which subjects cope with occasional informational overload.[14,15]

Man-Machine Communication: Programming

Thus far, only the characteristics of the human link as a combined information filter and format converter have been considered. Other characteristics of this link are brought into play when the man-machine communication involves an interchange of data and systems-status information as part of their joint performance of a unitary task. This broadened "communication" can take the form of a coherent set of instructions for the computer, known as programming. Since this quickly becomes a demanding form of communication, a succession of synthetic languages for improving such communication is evolving and these are the subject of a companion paper.[16]

The way programming is presently being accomplished, the concept of "reaction time" is inapplicable, but it does have meaning for the "real-time" interchanges of information that can occur between man and machine after the programs are in the computer. The scanty literature that is available conveys the unjustified impression that effective real-time interchange has been realized. This overstates the current situation, and the level of accomplishment is better characterized as "mechanically extended man."[17] Current efforts to introduce automation into air traffic control tend to follow this pattern, as did earlier studies for applying radar aids.[2]

There are a few instances in which the intent is to use the man more as a systems monitor and decision maker and less as a conventional systems operator. The first serious effort along these lines appeared in the man-computer communication features of the Air Defense System.[18] As this trend grows, so will the appetite for research results that bear on how to better accomplish man-machine communications. The journal literature presently offers limited resources to meet such demands.[19] To date research information is sparse and it is usually imbedded in informal documents intended to report on the over-all performance of a particular system.

General vs Specific Systems Research

The potential yield from engineering psychology investigations that are independent of particular systems have been handicapped both by the lack of sufficient theory to guide the research and by the inadequacy of tools for conducting such research.[20] With respect to the latter, appropriate research facilities are beginning to make their appearance.[21,22] However, most of the contributions will continue to flow from the experience with specific man-machine complexes in which different distributions of the functions between the man and the machine are attempted. In this connection, Licklider has pointed up some thought-provoking extrapolations for enhancing the power of such man-machine combinations.[17]

An Example: Key Sort and Code Sort Systems

By comparison with the complexities of air defense and traffic control systems, schemes for automating the sorting of letter mail[23,24] show a far simpler pattern of man-machine relationships. Consequently, the human factors considerations can be more readily related to reported research results. Benjamin Franklin did the original human factors studies for the manual sorting system which is only just now being challenged by two approaches called Key Sort and Code Sort.

Key Sort is an example of "mechanically extended man" in that it removes the limitation of the manual system of giving efficient access to only 80 sorting bins and extends it to well over 300 bins. These systems require that the operator memorize the sorting scheme relationships which define the way in which the address on the letter determines the distribution into the bins. However, the operator of a Key Sort System must memorize a larger fraction of the relationships for the total scheme if this mechanical extension is to be effective.

The man-machine communication in Key Sort con-

[15] D. B. Yntema and W. S. Torgerson, "Man-computer cooperation in decisions requiring common sense," IRE Trans. on Human Factors in Electronics, vol. HFE-2, pp. 20–26; March, 1961.

[16] W. H. Ware and R. D. Elbourn, "The evolution of concepts and languages of computing," this issue, pp. 1059–1066.

[17] J. C. R. Licklider, "Man-computer symbiosis," IRE Trans. on Human Factors in Electronics, vol. HFE-1, pp. 4–11; March, 1960.

[18] R. R. Everett, C. A. Zraket and H. D. Bennington, "SAGE—A data processing system for air defense," *Proc. Eastern Joint Computer Conf.*, Washington, D. C., December 9–13, 1957; pp. 148–155.

[19] J. A. Kraft, "A 1961 compilation and brief history of human factors research in business and industry," *Human Factors*, vol. 3, pp. 253–283; December, 1961.

[20] P. M. Fitts, *op. cit.*, see especially p. 272.

[21] H. K. Skramstad, A. A. Ernst and J. P. Nigro, "An analog-digital simulator for the design and improvement of man-machine systems," *Proc. Eastern Joint Computer Conf.*, Washington, D. C., December 9–13, 1957; pp. 90–96.

[22] A. A. Ernst, "Feasibility Study for a Man-Machine Systems Research Facility," WADC, Tech. Rept. 59-51; March, 1959.

[23] I. Rotkin, "The mechanization of letter mail sorting," *Proc. Eastern Joint Computer Conf.*, Washington, D. C., December 9–13, 1957; pp. 54–57.

[24] B. M. Levin, M. C. Stark and P. C. Tosini, "Post office mechanization," *Elec. Engrg.*, vol. 80, pp. 1–6; February, 1961.

forms partly to the model in Fig. 1. It departs from the model in that the operator is required to perform a conversion of the information content perceived in the address on the mechanically displayed letter before his handkeying operation conveys this converted information to the sorting machine. The conversion uses the scheme knowledge stored in the operator's long-term memory, and this "programmed response" serves a purpose similar to that of looking up a number in a telephone directory. Although some relevant psychological investigations have been reported, these now need to be supplemented by more specific investigations.

The Code Sort approach provides the assistance of an electronic directory that relieves the operator of the need to memorize any part of the sorting scheme. He initiates a lookup in the electronic directory for use by the sorting machine through keying in an abbreviated form of the perceived address. His long-term memory is called upon only to supply an appropriate abbreviation rule rather than to recall scheme relationships. Because the time and effort required to think through abbreviation rules are greater than the time and effort required to remember frequently used abbreviations, the operator will tend to memorize frequently occurring ones. Since the operator can always revert to the abbreviation process whenever presented with an unfamiliar address, the proportions of the two modes of action will depend on the extent of the operator's training and the occurrence of familiar addresses among the letters.

Doubtless there are other versions of the automatic electronic directory which might benefit from the contributions to engineering psychology that will result from code-sort experimentation. In line with comments above, this knowledge will initially have to be gleaned from systems reports on the automation of mail sorting.

Conclusions

The study of communication between man and machine is still in its infancy. Available methods and models for information transmission are as yet restricted to very simple communication situations, so that solutions to problems in everyday complicated communications such as radarscope search or air traffic control are still approached through realistic experiments. One principal impediment to model building is the difficulty in determining and quantifying the information content of complicated signals; a bit of information transmitted by man in distinguishing between two dissimilar signals can not meaningfully be compared with a bit for almost identical signals. The problem is, however, becoming increasingly acute and the study effort needs to be strengthened appreciably.

Acknowledgment

Grateful acknowledgment is made to G. A. Miller and the American Psychological Association for permission to use the information in Figs. 2–4, and to E. Averbach and G. Sperling and to Butterworth and Co. for the information in Fig. 5.

Section 16

INDUSTRIAL ELECTRONICS

Organized with the assistance of the IRE Professional Group on Industrial Electronics

Early History of Industrial Electronics*

W. C. WHITE†, FELLOW, IRE

Summary—**Industrial electronics is the phrase usually employed to cover the use of vacuum tubes in fields outside of communication. The term was originally applied to the Tungar Rectifier (about 1915) and the use of phototubes (about 1927).**

By 1930 Industrial Electronics became a growing business. However, its early growth was slow, because there was a general mistrust about the dependability of tubes. Wider use of thyratrons that could handle currents in amperes was an important contribution; the development of ignitron tubes, and the use of steel envelopes for them, were considerable factors; also, such new applications as high frequency for induction and dielectric heating helped in the advancement.

THE TERM INDUSTRIAL ELECTRONICS is generally accepted as the use of electron-tube devices in the control and operation of machines employed in industry. This involves such fundamental applications as: photoelectric relays; Tungar Rectifiers for battery charging; railway signaling; high-voltage rectifiers; motor control; and induction and dielectric heating for the control of spot and seam welders.

Introduction

During the years of World War I and up through the early years of broadcasting, the demand for new types of tubes for applications to radio were urgent. As a result, not too much attention was paid, except in a few cases, to the use of tubes for industrial applications.

There were a few early non-radio uses for tubes, such as Tungar battery-charging bulbs and high-voltage kenotrons for cable testing and dust precipitation. As of January 1, 1930, there were some drastic changes in the picture as regards the future and the business outlook for industrial tubes. Three factors, all occurring at about the same time, contributed to this situation:

1) As of that date RCA took over from General Electric and Westinghouse the manufacture of all radio receiving tubes and some small transmitting tubes. They also had the right to take over the remaining transmitting tubes as soon as they were in a position to do so.
2) The big depression of the 30's had started and the demand for tubes, along with all other products, started to drop alarmingly.
3) It had been decided that development, design, engineering and manufacture of vacuum tubes had now grown to a point where they could be divorced from a laboratory.

It is hard to imagine a more inauspicious combination of circumstances for General Electric and Westinghouse. In order for them to exist it became absolutely necessary to develop new tube applications and new types of tubes to take the place of production transferred to RCA. A great deal was accomplished because of this urge.

The philosophy of tube work in industrial electronics had to be quite different than that in radio. In the latter case, advances in radio communication were largely based upon new and improved tubes and it was not too difficult to "sell" ideas and products. In industrial electronics, however, the situation was just reversed. In the field of power equipment and control, highly satisfactory methods and devices had been pretty fully developed and there was no great urge to substitute tubes. It was soon found that the best approach to this problem was developing applications that could not be done satisfactorily in any other way. Examples of this were high-speed spot and seam welding, motor control (dc motors from an ac supply line) and railway signaling, and elevator control and the electrocardiograph. There

* Received by the IRE, December 27, 1961.
† Research Laboratory, General Electric Company, Schenectady, N. Y.

was a common saying at that time that General Electric was not in competition with any other company but rather with the old way of doing things.

The Tungar Rectifier

It is interesting to review this development for two reasons. Not only is it the first application of a gas-filled, hot-cathode tube, but it is also very typical of the sort of procedure that seems to have been involved in several early, highly successful projects. Also like several other successful tube accomplishments, it was a by-product of lamp work.

During the spring and summer of 1912, Dr. Irving Langmuir of General Electric became interested and active in the development of the gas-filled, tungsten-filament, incandescent lamp. These early laboratory lamps were constructed with heavy-coiled tungsten-spiral filaments that operated at low voltage. The bulb space around the filament was relatively small so that it would operate at high temperature and a few drops of mercury were placed in the bulb. There was usually another portion of the bulb where the mercury vapor condensed and ran back into the filament portion.

These lamps were frequently tested by operation from a 110 v dc line through a series resistance. As one of the objects of these tests was to study the life and characteristics of the lamp at very high temperatures, there were frequent filament burnouts. It was noticed that, when the filament sometimes burned out during operation, an arc formed at the break.

In all of this early work, the primary interest was in the arc as a source of light and apparently no tests on the arc as a rectifier were made at that time. However, one of the men sketched a tube with a hot filament and a separate anode plus liquid mercury.

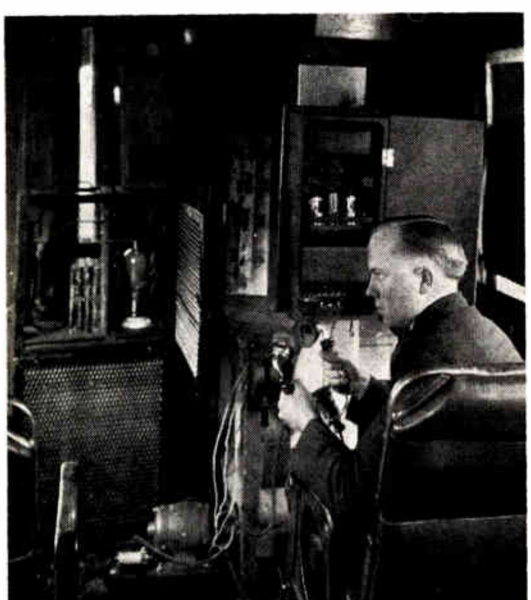


Fig. 1—Early carrier current equipment installed in a trolley car (1927).

In the early part of 1915 the possible need for a garage battery charger was considered and during 1916 it became a going business.

The Story on Early Phototube Work

Photoelectric effects in a voltaic cell were observed over 100 years ago. The phenomenon of photoelectric emission of electrons from the alkali metals (such as potassium and sodium), which forms the basis of modern phototubes, was first noticed over 50 years ago. Such tubes, custom built, could be bought at a high price from foreign scientific apparatus makers abroad. The U. S.-made tubes were really not commercially available until about 1928.

During the early years, work on photoelectric tubes was sporadic, largely because there was no apparent important demand for the use of such tubes. A few were made and used for incandescent lamp photometric work. Although the principle of the photoelectric relay was demonstrated to many persons between 1920 and 1928, no one seemed to envision the industrial uses for such a device.

The factor which really activated the quantity manufacture and use of phototubes was the development of talking motion pictures. Several methods for making such pictures were proposed. One involved the use of a black and white sound track on the film. This idea was the one that finally won out and it was demonstrated in a theatre in September, 1927. It used a phototube that picked up light variations from a film track and transformed them into audio currents for the sound effects. During 1928 a program of phototube development was started to meet the RCA demand for their photophone talking motion picture projectors.

Up to this time all phototubes had been made by distilling the light-sensitive electron-emitting surface of the alkali metal (at that time usually potassium) on the inside of the glass bulb, leaving an opening of clear glass for the light beam to enter. This type proved to be difficult to make in a factory using lamp or vacuum tube making techniques. In 1928 the idea of distilling the light-sensitive layer of alkali metal on a sheet metal electrode mounted on a stem inside the bulb was originated. This idea we accepted at once. In that same year, the American Standards Association adopted the word "phototube" to replace photoelectric tube to designate devices in which the electron emission resulting from exposure to light was the operating principle. A phototube, called the UX-868, was used by RCA Photophone for several years in practically all of their production of projectors for talking motion pictures. About this time cesium came into general use for the light-sensitive cathode surface since its spectral response is greater and more similar to that of the human eye than is potassium. In the fall of 1928 experiments were conducted on transmitting voice and music over a light beam using, of course, a phototube for a pick-up device.

By the new agreement with RCA, effective January 1, 1930, the manufacture of phototubes and their sale for television and motion picture use were transferred to that company. The only field left to General Electric and Westinghouse was that for industrial applications. Work was renewed on such applications during 1930, but, as in other industrial electronic applications, commercialization was hard, slow work. During 1929 a practical photoelectric relay unit was developed and during 1930 commercialization was underway.

Production of phototubes dropped to nearly nothing during the next few years because the industrial field was slow in developing; however, by the middle of 1931, business had picked up to a rate of a few thousand tubes a year. In 1932 added research work resulted in increased phototube sensitivity.

One interesting development was an automatic bean sorter that was put into commercial operation in 1931. In this device each dried bean was picked up and "looked at" by a phototube. If it were discolored, it was automatically rejected. Banks of these machines sorted beans at the rate of many thousands a minute.

Although new types of phototubes were developed and new uses found for them, high manufacturing costs further restricted the small market and, for a number of years following the outbreak of World War II, very little progress was made in this work.

Vacuum Tubes for Railway Signaling

Although business in tubes for train control was a very small percentage of the volume of tube business, it is significant as the first really successful wide scale industrial tube application. For over 25 years an average of several thousand tubes a year went into this service.

In 1916 the Union Switch and Signal Company of Swissvale, Pa., became greatly interested in GE's idea for a method of transferring the green-orange-red block signals from the trackside into the cab of the locomotive in front of the engineer.

If this could be done, not only would the danger of lower visibility due to fog, smoke, rain and snow be eliminated but automatic slowing down or stopping of the train if a signal was "passed" could be accomplished in a practical way.

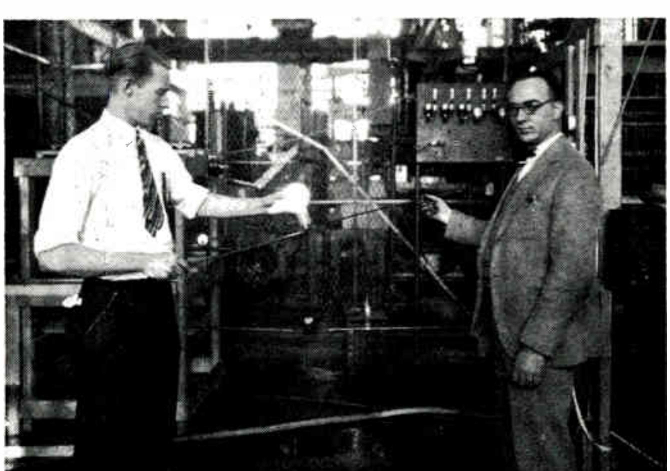


Fig. 2—10-kw 50-Mc oscillator with which artificial fever was first studied (1927).

To gain the broad objective there had to be a method of picking up, on the locomotive, a voltage or current of useful value proportional to the values of current in the rail used to activate the conventional trackside signals. The amplifying property of the vacuum tube provided a means of doing this from coils on the locomotive frame located a few inches above the rails.

Remember that it was 1916, five years before broadcasting, when vacuum tubes were considered expensive, frail, glass bottles used chiefly by a small group of somewhat eccentric individuals known as radio engineers and radio amateurs. The idea of putting them on a locomotive for a vital safety device seemed a little preposterous even to the early enthusiasts in the field of what was later known as electronics.

Progress was meager and slow until 1922 when the start of broadcasting necessitated a real tube manufacturing set-up. Two types of tubes were developed: a high-mu triode for voltage amplification, type TC-3 (later PJ-2), and a power amplifier tube, type TC-4 (later PJ-4).

During 1923 the General Railway Signal Company also became interested in GE's tubes. By June of that year a few locomotives and a short section of a Pennsylvania Railroad branch line near Sudbury, Pa., were equipped by Union Switch and Signal and put into regular scheduled operation. In 1924 installations were underway for the Santa Fe, Nickel Plate and New Haven Railroads and by the end of the year the tube demand was nearly 500 a month.

Much of the credit for the successful development of this electronic system of continuous automatic train control must be given to L. V. Lewis, an engineer of the Union Switch and Signal Company, whose vision, persistence and resourcefulness brought this development to a successful conclusion.

High-Voltage Kenotrons and Their Applications

The Field of Application

For a number of years there have been a few uses for high-voltage (up to millions of volts) dc for which only a few milliamperes or amperes were necessary:

1) For electrostatic precipitation of smokes, dusts and suspended air particles. (Active work in this field began about 1910 using a mechanical type rectifier and this business grew and has continued to the present.)
2) For a voltage supply for X-ray tubes where for certain applications dc is preferable to ac.
3) Testing of power transmission and distribution cables. (Due to the capacity between the conductor and the sheath, a very large kva rating of ac test equipment is required; therefore, dc is preferable.)

4) For test equipment to simulate the lightning of nature.
5) To supply voltages up to 20 kv for the plate supply of high-power radio transmitters. (This use disappeared after the development of the hot-cathode mercury-vapor tube in about 1928.)

High-Voltage Kenotrons

Before the work on the high-vacuum electron tube began in 1913, high voltages could not be employed with tubes having a hot cathode because of gas ionization troubles. However, the GE Research Laboratory developed techniques that made such tubes possible and practical. That same year Dr. Coolidge applied these methods to the making of a greatly improved X-ray tube. Tests using high-vacuum rectifier tubes for smoke precipitation were made in 1917.

Due largely to commercial reasons, progress in using high-vacuum rectifier tubes for smoke precipitation made little or no progress for many years.

Thyratrons

The present-day thyratron represents a combination of several pioneer developments:

1) In 1914 Dr. Langmuir first suggested a method of controlling the arc in a mercury-pool tube by means of a grid. He showed how a grid voltage could be used to control the starting of the main arc in each rectifying cycle. Thus, the average arc current through the tube was controlled when an ac anode voltage was used.

   In 1922 Toulon, a French scientist, improved on this method of control by varying the phase of the grid voltage with respect to the anode voltage rather than to its amplitude. Thus, the arc could be made to start at any point in the anode voltage cycle and this resulted in a very practical and convenient method of controlling the average value of a rectified anode current.

   Here then were the necessary tools for a whole new chapter in the progress of electronics. But they were somewhat ahead of their time and their value and applicability were not at once appreciated.
2) In 1936 Dr. A. W. Hull developed the idea of operating a hot-cathode diode in a low pressure of an inert gas or vapor. As a result, the space-charge effect was eliminated and the voltage drop of the discharge dropped to a low and more or less constant value of 5 to 10 volts. At this low arc drop voltage, the positive-ion bombardment of the cathode is not destructive. There results, therefore, a highly efficient rectifier.
3) Due to the elimination of space charge, the cathode could almost be surrounded by a heat reflecting and heat conserving shield. Only a few small holes are necessary for the conducting ionized vapor. Thus, the current-carrying capacity of the cathode for a given amount of heating power is enormously increased. An electron-emitting filament in space may give an emission of only about 50 ma/w whereas in a heat-shielded cathode in a well-designed gas-filled tube the figure may approach one ampere per watt.
4) In 1931 came the shielded-grid thyratron, the principle of which is incorporated in most modern thyratrons, even if the lead to the added electrode is not brought out to a base pin or other external connection. The function of the "shield" is to minimize the current and power required to activate the thyratron control grid.

The thyratron was announced in May, 1928 and the word first appeared in print in July, 1928. Actually hot-cathode mercury-vapor diodes were made available for use before the thyratron was brought out. There was a ready market for these rectifier tubes in the radio industry, particularly for rectifiers for the high-voltage plate supply of transmitters. By the end of 1928 such tubes were in regular use in several stations. Soon the hot-cathode mercury-vapor type completely superseded the high-vacuum type.

Thyratrons, however, found no early ready market in the radio field and to develop their use in industry took some time and much promotional work. During 1929 a line of thyratrons was developed which included the use of an inert gas as well as mercury vapor. These ranged in current ratings between a fraction of an ampere up to $12\frac{1}{2}$ amperes. Three types of shielded-grid thyratrons were made available early in 1933.

The first real commercial application of thyratons in

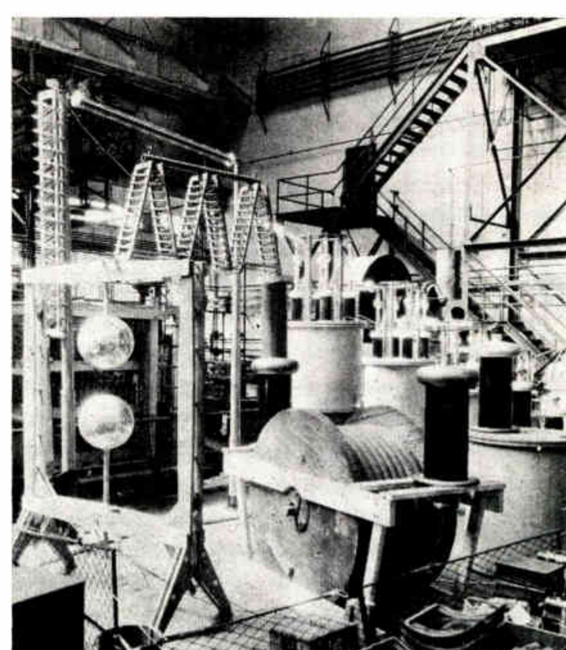


Fig. 3—High voltage, high vacuum tube rectifier for cable testing rated at 4,000,000 volts (1931).

appreciable numbers was made in the illumination control equipment for the then new Chicago Civic Opera House. This included about 250 small thyratrons in combination with saturation reactors to dim the lights in the many circuits of that large theatre. The theatre was officially opened in November, 1929, and the thyratron installation was highly successful. For its time, it was a very ambitious undertaking in the new science of industrial electronic control. A few other relatively large installations of illumination control were made.

In 1930 active engineering started on the use of thyratrons for the control of resistance welding (spot and seam). During the next year or two, a number of installations were made. The first large installation, and the forerunner of the wide use of electronic welding control in the automobile industry, came from the Ford Company in 1932 and numbered 12 equipments. These not only used small control thyratrons in the timing circuits but also $12\frac{1}{2}$ ampere tubes in the main-welding current-power supply circuit. Ignitrons were not known or available for the next few years. As soon as they were available they were substituted in place of thyratrons for the power supply control, but thyratrons continued to be used for the accurate timing and control circuits. Year by year such timing circuits have become more highly developed and complicated. Their accuracy of timing and flexibility of control were unobtainable except by electronics.

As early as 1930, an installation of a generator voltage regulator using thyratrons was made. This worked quite well and a few other installations followed, but it took many years for this field to develop. Before 1933 trial installations of thyratrons for synchronous motor exciters, railway signals, telegraph circuits, X-ray tube circuit control, and motor control were also made.

Fig. 4—Photoelectrically controlled drinking fountain (1931).

During 1931 active work started on a system of fire control for the directing of large guns on U. S. naval vessels. This system employed thyratrons for supplying dc to the motor for moving the gun mountings. It was fully developed and installed on a number of large ships. At the time it represented quite a business for thyratrons.

Building up the thyratron business was slow, hard work during the first ten years. Unlike transmitting tubes, there was no active business corresponding to radio broadcasting awaiting them and engineers in industry had not yet accepted electronic devices. There was little competition between makers of thyratrons; the real and serious competitor was "the old way of doing things."

Most of the initial successes were in applications, like welding control, where results were obtained which just could not be done in any other way and, therefore, expensive equipment was justified.

Sealed Steel Ignitron Tubes

Early Work on Pool Tubes

Work on mercury-arc phenomena was started by Cooper-Hewitt about the turn of the century.

There was a more or less steady increase in the use of sealed glass bulb rectifiers in the U.S.A. during the first quarter of the century. These were used largely for series arc street lighting and battery charging for electric vehicles. The increasing use of incandescent lamps for street lighting, the improvements in gasoline powered automobiles and the development of Tungar battary chargers have largely replaced glass bulb mercury-arc rectifier tubes in the U.S.A.

The Ignitron

In 1933 the Westinghouse Company announced its Ignitron. Its potential value was at once recognized and an active developmental program soon commercialized it extensively. Progress in making and applying Ignitrons was rapid. By the end of 1934 a welding control unit using glass Ignitrons was installed in a customer's shop.

The Sealed Steel Envelope

About the time of the Westinghouse Ignitron announcement, General Electric was well along on its program for making sealed steel envelope vacuum tubes of various kinds and types. This latter development was made possible by two other new techniques: fernico alloy for making heavy current seals with glass, and electronically controlled vacuum-tight resistance seam welding.

In February, 1936, the first successful sample of an all-steel envelope water-cooled ignitron with fernico seals was built. These all-steel ignitrons soon largely superseded the glass tubes.

As regards the power rectifier field, during 1936 the American Gas and Electric ordered for installation at Fostoria, Ohio, a sealed steel tube rectifier of about 150-kw rating for the excitation of a large (30,000 kva) synchronous condenser. By the end of 1936, this was in successful operation. Thus, the start was made in what proved to be an active field of application.

The Early Use of High-Frequency Induction Heating

Early Work

The production of heat by induced currents was recognized as early as 1880. The heating of transformer cores due to eddy currents was, of course, understood at an early date. Probably the real engineering and developmental pioneer in induction heating with frequencies above the power range was Professor E. F. Northrup of Princeton University. In 1918 he built practical "furnaces" for frequencies above 10,000 cycles and powers as high as 60 kw using a spark gap oscillator.

Largely as a result of the work on radio transmitting tubes, circuits and their components during World War I, it was possible in 1919 to produce, in a practical way, outputs of several kilowatts up to about 1-Mc frequency from vacuum tubes.

Not much further progress was made until 1921 when work became active in developing the UV-199 receiving tube for the new but fast growing broadcasting industry. This was before the ac heater type of tube was available and the low filament current of the UV-199 (60 ma) made it a popular type much in demand for battery operated receivers (even at a price of $6.50 each). Due to the very small filament, the outgassing of the anode by electron bombardmentheatingduring manufacture was impractical. High-frequency induction heating of the electrodes proved to be the only practical method. This at once created a demand for high-frequency heating equipment. During the next few years, the use of induction heating increased in the tube manufacturing industry and has continued to this day to be a widely employed technique.

Fig. 5—Photoelectric door opener in a restaurant (1931).

High-Frequency Dielectric Heating and its Application to Therapeutics

Prior Work

During the early years of the century when "wireless" was being developed, the problem of overheating the dielectric material of the transmitting condensers was a common one. For many years such heating was chiefly a nuisance rather than an interesting phenomenon leading to a possible business.

Actually much of the very early experimental work was on living things. In 1925 a Dr. Petroff of the Trudeau Sanitorium at Saranac Lake visited Dr. Coolidge of the GE Research Laboratory with an interesting story. At an earlier date, W. K. Kearsley of the GE Research Laboratory (now retired) had made up and given to Dr. Petroff a 2-meter oscillator of possibly 10 to 15 watts output. Dr. Petroff told Dr. Coolidge that small tadpoles had been killed when placed near the coil of this oscillator.

Except for a successful and lethal test conducted a little later on a cockroach discovered prowling near a short-wave oscillator in a laboratory, little work was done on this subject until about 1926.

In 1926 electrostatic heating tests were run on mice and flies. For flies it was found that, for the field strength available, a frequency above about 75 Mc was necessary for a lethal dose. About that time an egg was also cooked by high-frequency heating.

Later in 1926, Dr. Whitney of General Electric became much interested in this subject and began experimenting on mice in both electrostatic and electromagnetic fields. His tests covered long periods of exposure in fields not strong enough to kill or disable the mouse.

Two of his 1927 test results were of interest. The mouse was enclosed in a large diameter glass cylindrical tube in a horizontal position with the ends screened. The coil or condenser plates for heating were at one end of the tube. For some of the tests, this enclosure was put outdoors on fairly cold days. The mouse would soon learn to adjust its position just near enough to the electrode end of the cylinder to keep comfortably warm. Also after long exposure, during which time the mouse acted normally, its tail dried, withered and finally fell off. However, the mouse lived on. This latter test received considerable publicity.

The next step took place late in 1927 while tests were being run by E. D. McArthur of the GE Research Laboratory on a high-power (5 to 15 kw) radio transmitting tube under development. The men near the circuits while making these tests noticed a peculiar sensation and some complained of headaches. A doctor found an increase in bodily temperature (fever) of these men. Dr. Whitney was much interested as he knew of the promising work of others on certain diseases when the patient was given an artificial fever with drugs.

Equipment was soon constructed to heat parts or most of the human body between two large parallel plates at frequencies around 50 Mc. By the fall of that year, low-power equipments had been sent to Albany Medical College, Mayo Bros. Clinic and Johns Hopkins Medical School.

In the summer of 1929 a 400-Mc oscillator using a small split-anode oscillating magnetron was supplied to the Woods Hole Laboratory for some biological experiments. This was probably one of the earliest uses of the oscillating magnetron that was to become such a useful tool in the radar of World War II.

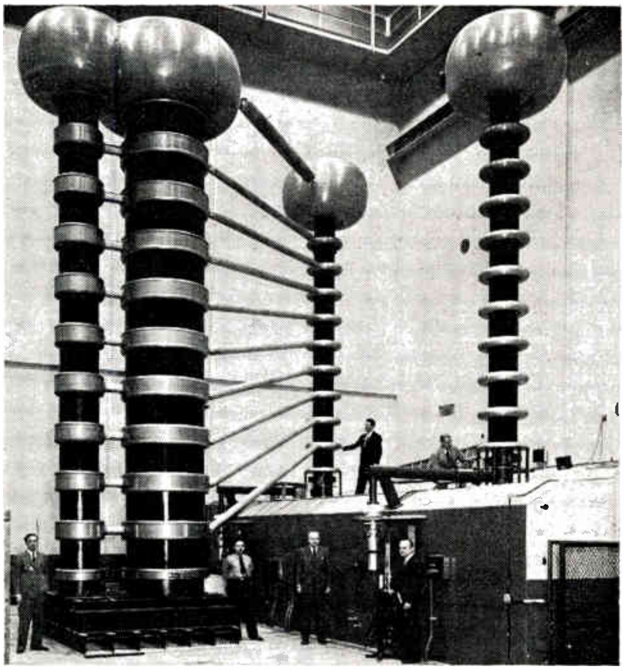


Fig. 6—X-Ray equipment 1.4 million volts (1941).

In 1930 the Cudahy Packing Company became interested in using this phenomenon for sterilizing and precooking meat. Tests were also made on foods, fruits, fruit juices, hot dogs, vegetables, eggs, nuts, etc., as well as on the drying of wet materials, including laundry.

There was publicity on some of these tests, with their frequencies up in the many megacycles. As to be expected, some manufacturers of medical equipment and publicity-seeking doctors saw an opportunity. Early in 1931 a New York City doctor coined the name "Radiothermy" and this helped to encourage commercialization. By the end of 1933 a model was in manufacture. By 1935 a number of manufacturers were active and equipments were in use by doctors, clinics and hospitals (and some by rental in homes) treating various aches and pains.

Except for cooking there was little interest or activity outside the field of therapeutics. Some tests were made on foods, golf-ball cores, gluing laminated wood for tennis racket frames and rayon drying, but, no applications resulted at that time. In fact, commercial industrial use of dielectric heating did not evolve until 1940 when a firm in Richmond, Va., developed the technique for the quick hardening of the bonding cement used in making plywood. Pilot plant operation was carried on in a mill on the West Coast.

Shortly thereafter the use of this method of heating was started in the manufacture of plastic articles for very quickly preheating the raw material pellets to soften them just before they were placed in the mold.

Active work was also started in several places on the rapid thawing and cooking of frozen foods.

In therapeutics the high hopes of some of its proponents that here was a general curative agent have not been realized, but it shows usefulness in alleviating the pain and stiffness accompanying some ailments of joints and muscles.

Industrial Electronic Developments in the Last Two Decades and a Glimpse into the Future*

WALTHER RICHTER†, FELLOW, IRE

Summary—**During the past two decades significant improvements have been made in all types of electron tubes. Small tubes became still smaller, and the output of large tubes has been increased. Greater reliability has been achieved, and frequency limits have been extended upward. Hydrogen thyratrons have been developed, providing high power for pulse work. Ignitron ratings have been increased significantly. The problem of shutting off gas tubes by grid control has received attention and has led to two new types of tubes which at this time have, however, only limited output.**

To the list of photosensitive devices have been added the lead-sulfide and similar cells, providing means of detecting infrared radiation of long wave length.

A host of ingenious transducers has extended the use of electron tubes in the measurement and control field; smaller magnitudes of physical quantities, such as distance, pressure and magnetic field intensity can be observed and recorded. On the other end of the scale, tubes have provided HF power for induction and dielectric heating.

But by far the most outstanding development of the last 15 years has been in the field of solid state or semiconductor devices. Here the transistor, announced in 1948, occupies first place, and improvements in its characteristics are coming at a breath-taking pace. In 1957 the silicon-controlled rectifier, aptly called a solid-state thyratron, joined the march. These two devices are giving industrial electronics a tremendous new drive.

About fifty years ago a witty observer remarked that the dividing line between the two then existing branches of electrical engineering was the decimal point. To the left of the decimal point was the power branch of the profession, with kilowatts, hundreds of volts and amperes, whereas the region to the right, with milliamperes, microwatts, etc., was the domain of the communication branch. Where should the electrical engineer specializing in electronics be placed in this picture? If Lee De Forest's invention, which is without doubt the most important milestone in the history of electronics, is considered as the starting point, then electronics for almost two decades operated essentially on the right side of the decimal point. After all, his "Audion" was an almost exclusive possession of the communication branch. However, mercury-arc rectifiers are decidedly electronic devices, and some of them were in operation as battery-charging equipment on that November day in 1906 when De Forest placed a grid between the cathode and anode of his experimental tube and thus became the "father of radio" and of a large part of electronics in general as well. The industrial electronic engineer of 1962 can definitely not be relegated to one or the other side of the decimal point; he has to jump nimbly from one side to the other (unless he chooses to confine himself to one side or the other). As a matter of fact, the two decades since the beginning of the Second World War have continued a trend already established in the two preceding decades; the number of digits to either side of the decimal point has been increasing for him, a development which should not be too great a surprise.

The list of tools that the industrial electronic engineer had at his disposal when entering the war is well known. The list included a great variety of vacuum tubes of the receiving type, high-power vacuum tubes as used for broadcasting which began to serve industry in applications to be discussed later in this paper, thyratrons and mercury-pool-type tubes, of which the controlled type, the ignitron, assumed a dominant position, and photosensitive devices. That improvements were made in all of these should, of course, be no great a surprise, but space limitations here do not permit the enumeration of more than a few of the changes that have taken place.

In the field of the receiving-type tubes—and these are also the workhorses for the industrial electronic engineer when he operates on the right side of the decimal point—metal tubes have given way to glass tubes again, but they are a good deal smaller than their predecessors. They were called miniature tubes when they appeared commercially, but they are already beginning to look like giants compared to what is just around the corner (more about this later). Triggered by an alarming failure rate when receiving type tubes began to be used in business machines and were operated for long periods under cutoff conditions, considerable work was done on cathodes, with the result that we now have much better cathodes. Reliability, which in industrial applications is, of course, of immeasurably greater importance than for a hi-fi set, has been the subject of intense study and many apparently small changes have contributed to making these tubes much more reliable than they were 20 years ago. Disk-seal-type tubes, a structure which shows minimum lead inductance, have increased the frequency capabilities of tubes. As far as the most recent developments in tubes are concerned, it may be suspected that the appearance of the transistor has brought on, or has at least hastened, the development of tubes of very small size. Thus, there have been developments of

* Received by the IRE, October 11, 1961.
† Consulting Electrical Engineer, 8575 N. Hawthorne Rd., Milwaukee, Wis.

stacked ceramic tubes, using ceramic washers not only as insulators, but also as the walls of the tube. RCA's Nuvistor uses cylindrical parts instead of washers. At the time of this writing the smallest tube is probably General Electric's microminiature heaterless tube. It consists of a coated cathode disk, a perforated grid disk and a solid anode disk, all of them made of titanium, which turned out to possess some properties just made to order in tube design. Two ceramic disks, hermetically sealed to the three metal disks complete the structure. Such a tube is about $\frac{1}{4}$ inch in diameter and less than $\frac{1}{8}$ inch high with a spacing of .01 inch between the electrodes. It gives a transistor of comparable power handling ability quite some competition. For such a tube to operate, the cathode must, of course, be brought up to a temperature sufficient to cause emission from its coating. In this case the tube, together with resistors, capacitors and inductances is placed in a heat insulated enclosure which is heated to 580°C. Naturally, resistors and capacitors must be able to operate under this condition also, but the circuit loss in the resistors can be utilized to keep the temperature of the enclosure at the desired point. The life expectancy of these tubes is about 100 times as great as that of regular tubes. These developments show that tubes are by no means ready to abdicate to transistors.

High-power vacuum tubes might be considered mostly of interest to the communication engineer where radar especially requires very high-power outputs. But high-power vacuum tubes are also serving industry as we shall see in the discussion of applications. Here significant changes in the design and utilization of new materials, especially ceramics, have resulted in much higher power as well as more rugged and reliable construction. Thus, a recently developed "super-power tube" can deliver 500 kw at 75 Mc. The tube operates with a plate voltage of 20 kv with a dc plate input of 1000 kw. Magnetron and klystron tubes have also seen spectacular advances but will not be discussed here.

As the communication engineer can claim the high-power vacuum tube as his own, so can the industrial electronic engineer claim the gas tube, and especially the mercury-pool cathode-type tube as his own. The workhorse for him was, and still is, the ignitron which served him well during the war years. The series of these tubes has been extended upward; the latest ignitron (7151) can switch 4800 kva (two tubes in back-to-back connection) at potentials of 250 to 600 volts.

In contrast to heated cathode type of gas-filled tubes (thyratrons) the ignitron is not controlled by a grid, but the arc is initiated by a relatively large current flow, fortunately of very short duration, through an ignitor rod. A successful attempt to combine the advantages of the mercury-pool-type cathode (its tremendous overload capacity) with a grid control is represented by the "Excitron" developed by the Allis-Chalmers Manufacturing Co. This tube appeared on the maket just about at the beginning of the war. In it a "keep-alive" dc arc provides an emitting spot on the surface of the mercury at all times. In a way we have here a heated-cathode-type tube, but the heating is accomplished by an arc instead of a filament. The tube has been used in similar applications as ignitrons. In the low-power field, the year 1941 saw the introduction of the type 2050 gas-filled thyratron which became one of the most useful tools in control circuits. In this tube an inert gas replaced the previously used mercury vapor, making the control characteristics of the tube independent of temperature.

The de-ionization time of both types of tubes just discussed is in the order of tens or hundreds of microseconds which is a serious drawback when the tube is to be used in pulse work. In 1946 the Sylvania Company placed on the market the first hydrogen thyratron, and this type of tube has become the standard in pulse work involving high power. Recent tubes of this type are capable of controlling discharges of pulse networks of nearly 50 Mw, and development work on tubes of even higher power is in progress.

However, the mercury-pool-type tube has also been the subject of further development work. From the RCA laboratories came the "Empretron" which is capable of operation in the kilocycle range. In this tube the arc is not started by an ignitor dipping in the mercury pool as in the ignitron, but by a spark from a starter electrode. This principle is not new, but in the new tube the starter is a small pot of mercury with an orifice at the bottom; a drop of mercury clinging to the orifice forms the starter electrode. Since the electrode thus renews itself, the problem of electrode consumption is avoided.

As is well known, the control element in a gaseous tube, be it a grid as in the thyratron or an ignitor as in the ignitron, can control only the start of the arc, but cannot shut the current off again. To accomplish this the anode circuit must be interrupted, or the anode must be made negative for a time about equal to the de-ionization time. This is a serious shortcoming of the tube, and has received the attention of research workers. Again from the RCA laboratories came two approaches to the problem; in 1951 the Plasmatron, a hot-cathode helium-filled diode capable of controlling a current of about one ampere continuously at a low voltage (6 volts) and in 1954 the Tacitron, a low-noise thyratron capable of current interruption by grid action. While these tubes are evidently low-power devices, their development shows that the value of a gaseous tube that can be turned off is well recognized.

In the field of light-sensitive devices, the war years added the lead-sulfide cell, soon followed by the cadmium-selenium and tellurium-sulfide cells; these types of cells extend the range deep into the infrared part of the spectrum, and found application in the solution of problems dealing with radiant heat.

Industry also profited from spectacular developments in the field of X-ray tubes. In 1941 industry was offered the first one-million volt completely packaged industrial radiographic unit; by 1946 the Betatron had emerged from the developmental stage and was providing 20-million volt X-rays. Within the last five years the amplification or intensification of X-ray pictures has also received much attention; there are two successful approaches to the problem: one makes use of electroluminescent phosphors to amplify the X-rays directly, the other is an image intensifier operating on a principle similar to a television system, which permits the observation of the X-ray picture at a remote place.

This may also be the place to call attention to a device which can be classified as an electron tube with as much justification as a cathode ray tube; it is the electron microscope. The development of this device into a practical tube for the physicist, metallurgist and others is due to Zworykin and Hillyer of RCA. In 1941 the first commercial model was offered; in 1943 a table top model, only 16 inches long, capable of providing magnification of 100,000 times was made available; and in 1951 a model in which the magnetic coil system was replaced by permanent magnets was developed.

It is claimed that at one third of the expense and complexity of larger electron microscopes, the instrument solves 90 per cent of the problems pertaining to electron microscopy.

So far we have been discussing what new types of tubes, or modifications of old ones, the last two decades have brought forth in the field of industrial electronics. But this is, of course, not the whole story. A brand new family of devices has been added to the tool chest of the industrial electronic engineer, and the magic word is "solid-state" devices. Crystal diodes, a refinement of the venerable cat whisker and galena crystal of the twenties, appeared in 1945. Then history appeared to be repeating itself. Fleming's valve was a diode, which became a control device when Lee De Forest added the grid to it. In 1948 Shockley, Bardeen and Brattain of the Bell Laboratories announced the point-contact transistor; they too had converted the diode into a control element by adding a third electrode. Some two years later the junction transistor was announced, which took its place very rapidly in the communication branch as well as the industrial branch of electronics.

In order to familiarize electrical engineers as quickly as possible with this new tool, it was perhaps natural to compare the transistor with a tube and to stress the similarities between the two devices. Recent books on electronic circuits quite often show such standard building blocks as multi-vibrators, Schmitt trigger circuits, etc., both in the tube and in the transistor version next to each other, thus apparently justifying the thought that a transistor is nothing but a solid-state tube. While this is certainly not the place to engage in a discussion of transistor characteristics, it may serve a purpose to call attention to some rather significant differences between the two devices. In the vacuum tube, control of the anode, or load current, is obtained by a volage applied to the grid, and the chief merit is that no current—and therefore no power—is demanded from the controlling voltage. If the grid does draw current, it is a deviation from the ideal condition. In a transistor operating in the common emitter connection (which is closest to the standard tube connection), on the other hand, control of the collector, or load current, is obtained by a current into (or out of) the base. The first feature to keep in mind is, therefore, that the transistor is a current-controlled, and not a voltage-controlled device.

The second significant point is the voltage drop across the tube or transistor. Both are control elements that are placed in series with a load across a source of voltage, similar to a rheostat, so to speak. A rheostat usually can be reduced to zero, in which case the full source voltage is applied to the load. To obtain such an action with a tube is nearly impossible, even by making the grid quite positive (although a look at the plate characteristics of a pentode show that, with this type of tube, one can come surprisingly close to this "bottoming" condition, as the British call it). A transistor, on the other hand, can be saturated quite easily; voltages in the order of only 10 to 100 mv are lost across the transistor. These are values that compare with the drop across an ordinary switch; it explains why in switching circuits transistors have become so popular. Owing to this low voltage drop under saturation conditions the heat developed in the transistor is, of course, also correspondingly small (a short-circuited rheostat does not become hot, neither does an open circuited one)!

The next important difference is far more radical than the two discussed so far and usually comes as a surprise to many electrical engineers. A vacuum tube can control the current in a load; but it is, at the same time, at all times a rectifier. Current flow cannot reverse; the supply voltage must always be positive. A tube cannot be used as a switch in an alternating current circuit, at least not if we wish to obtain positive as well as negative values. A transistor, on the other hand, with the proper direct current injected into the base will pass current either from collector to emitter, or in the opposite direction, so that it can be used as a switch in an ac circuit.

The fourth difference does not have to do with the characteristics of a given transistor itself, but is the fact that there are two kinds available, the pnp and the npn transistor. The two operate with reversed polarity. Only the man who has struggled with the design of dc amplifiers employing tubes can appreciate what it would have meant to him if he had a vacuum tube at his disposal that would operate with a negative voltage on the anode and in which the current would be decreased by making the grid more positive with respect to the cathode.

The characteristics just described evidently permit the transistor to do things which cannot be done by tubes; but this does not mean that it can in turn do all the things that a tube can do. For instance, where a very high input impedance is essential, tubes can outperform transistors; temperature drifts are also more serious for transistors, and so are the effects of nuclear radiation. It would appear that in the foreseeable future there is a field for both devices. As a matter of fact, it is gratifying to see in the recent literature descriptions of "hybrid" circuits with both tubes and transistors, utilizing the best features of both.

The high state of excitement produced by the advent of the transistor may have somewhat obscured the importance of a development that took place just in plain semi-conductor rectifiers. At first we had copper oxide rectifiers, then we had selenium rectifiers and therefore the appearance of the silicon rectifier in the middle of the last decade did not seem to be something to get very excited about. From the point of view of the circuit engineer that is perhaps justified. But when one compares the current densities and forward voltage drops of the silicon rectifier with those of the older devices, one realizes that nothing short of a revolution has taken place as far as space requirements go. With 18 ampere silicon rectifiers, about $\frac{1}{2}$ inch diameter, and not much more in length, available at quantity prices of about 60 cents, the days of the low-voltage dc generator in our cars seem to be numbered; alternators will take their place. The silicon rectifier is truly a most significant development in its own right.

In the case of the silicon rectifier, history also appears to have repeated itself. Vacuum diodes became control elements by the introduction of the grid; gaseous rectifiers with heated cathodes became controlled rectifiers which can turn on the current, but cannot turn it off. In 1958 General Electric announced the "solid-state thyratron" a name which aptly describes the action of the device, but which was quickly superseded by the three letters SCR, denoting silicon-controlled-rectifier. The SCR, just like its tube counterpart, the thyratron, can only initiate the current flow, but cannot stop it. But it has an advantage over the thyratron in two significant aspects. The voltage drop across it, while in the conducting state, is about one volt compared to the ten to fifteen volts arc drop of the gas tube. The circuit losses and the heat developed in the device are therefore also only a fraction of those encountered with gas tubes. Secondly, the turn-off time—the equivalent of the deionization time of the gas tube, is at least by an order of magnitude less than for gas-filled tubes. This increases the frequency range over which the SCR can be used, and it brings about a reduction of the size of certain circuit components, such as commutating capacitors, when the SCR is used in inverter circuits. At this time the current ratings of SCR's are still below 100 amperes, which is way under the several thousand amperes of the ignitron, but in applications within their ratings they certainly offer a number of advantages over tubes.

Even if this review of recent developments of interest, or possible interest, to the industrial electronic engineer had to be kept very sketchy, one new member should at least be mentioned, and that is the tunnel diode. It is clearly a low-power—indeed a very low-power device, but the frequency range over which it is capable of operating is truly fantastic. What jobs industry will assign to it, only the future will tell.

After having discussed the tools which were available, or were made newly available to the industrial electronic engineer in the past two decades, let us take a look at how he has applied them. Quite evidently, space does not permit a complete list or even a list of a very small fraction of all the applications that electron devices have been put to. Only very few examples can be selected for presentation here, and it can hardly be expected that the readers of this review will agree with the writer on the choice of the examples. If what is in your opinion the most outstanding application of electron tubes in the past 20 years is missing here, apologies are offered herewith. Again, we turn our attention first to the receiving-type tubes. They are relatively low-power devices and serve in industry essentially as measuring or controlling devices. In applications involving the control of a considerable amount of power, they are, of course, used only in the first stages, usually in the role of sensing elements. The story of the use of electron tubes or transistors in the field of industrial measurements and controls is really not the story of the tubes themselves, but the story of the transducers. Industrial processes call for the measurement or sensing of such things as distance force, stress, torque, speed, temperature and many others. But a tube or a transistor is a device which responds to voltage or current, and that is all. It can be of use only after somebody has invented a "transducer," that is, a device converting a physical quantity into another one. The transducers needed by industry in connection with electron devices are those that convert the quantities just mentioned into electrical voltages or currents. Now, if we have a transducer that converts the quantity we are interested in into a voltage or a current, why not measure it with a conventional volt meter or ammeter? Why throw in tubes or transistors? The answer is simple, and it is found in two characteristics possessed by both the tube and the transistor. The first is their ability to amplify, which permits them to accept signals of very low power, and the second one is their high speed of response. Most transducers by their nature can furnish only very minute amounts of electrical energy and could not possibly operate ordinary meters.

Many physical effects have been used in transducers; thus, a force applied to a piezoelectric crystal (quartz for instance) produces a voltage; a coil moving through a magnetic field has a voltage induced in it which is proportional to the speed with which the coil moves. these are examples of so-called "active" transducers. But there are also "passive" types. If a resistance changes its values, it can be made part of a Wheatstone bridge, where such a change will produce an unbalanced voltage, which is then a measure of the variation of the quantity originally changing the resistance value. The more commonly used transducers employ the following effects:

resistive, capacitive, inductive, piezoelectric, thermoelectric, photoelectric and magnetostrictive. They are used in the measurement of displacement, velocity, acceleration, force, stress, flow, humidity, temperature, and others. Many of the quantities can be sensed by more than one kind of transducer. Thus, displacement has been sensed by all of the transducers mentioned, except thermoelectric. However, most of these transducers could not possibly operate without the help of some amplifying means.

The most useful, and the most famous transducer is the well-known resistance strain gauge. Although the change of resistance of a wire when it is stressed, was discovered by that master experimenter, Lord Kelvin, the practical application of the principle had to wait until it could be teamed up with the vacuum tube or transistor. It is hard to assess how much the strain gauge contributed to our winning of the war. But strain gauges were plastered all over our experimental aircraft, and the results they furnished permitted the strengthening of the structures where needed, and the removal of material where the stresses were less than anticipated. In the testing of the Boeing B-52G, 2000 strain gauges were mounted on the wing and fusilage structure, giving the information needed by the designer. It is of interest to note that although strain gauges were no war secret and articles about them could be found in most magazines, the German literature shows that their importance was evidently completely overlooked. Strain gauges have also been combined with telemetering equipment, where their resistance change is used to modulate the frequency of an oscillator, thus broadcasting the desired information by frequency modulation. This permits measuring stresses in inaccessible or dangerous locations, and from moving or rotating parts of a structure.

As far as the measurement of small distances or displacements is concerned, the highest sensitivity is probably obtained today with transducers of the capacitive type. Circuits have been described capable of detecting a capacity change of 0.001 $\mu\mu$f. Two plates, each about $\frac{5}{16}\times\frac{5}{16}$, and spaced 0.005 inch apart change their capacity by such an amount if their distance is changed by approximately one microinch.

Transducers utilizing the inductive principle have been the subject of much study; there are several forms of these transducers which space does not permit to discuss in detail here. One of the most promising is based on observing the effect of the impedance of a coil when a conducting body is brought into its field. Sometimes the method is called the "eddy current method." The pioneering work in this field was done by the Foerster Research Institute in Germany. The method is also quite useful in the detection of flaws in a given test specimen.

Photoelectric transducers have almost limitless applications, which is no great surprise. Two applications in the steel industry may be worth mentioning. In 1960 a detector for pinholes in sheet steel was described capable of detecting pinholes 0.0005 inch in diameter in sheet steel traveling at a speed of 1700 feet per minute (about 20 mph). Another application is a photoelectric width gauge; mounted 15 feet agove the rolling table, it monitors the width of the steel strip which is several feet wide to between $\pm\frac{1}{8}$ inch.

In the early days of industrial electronics the steel industry was quite reluctant to make use of electron tubes, these fragile little bottles, but its attitude has changed considerably. The measurement of the width of a sheet has just been mentioned; measurement of the thickness of it while it is being rolled is, of course, of equal, if not greater, importance. In 1947 and 1948 there were described thickness gauges making use of beta rays and of X-rays.

One more example demonstrating the ability of electronic gear to respond to very small quantities will be mentioned. The magnetic field of the earth is in the order of $\frac{1}{2}$ to 1 Oersted. Electrical engineers are used to dealing with fields thousands of times more intense than this and, consequently, consider the magnetic field of the earth as quite a weak field. Yet, in geophysical prospecting for bodies of ore, the variation of this field from point to point on the surface is often an indication of the existence of iron containing bodies of ore. A magnetic survey of a large area used to be a very tedious and time-consuming project. In 1947 the Bell Laboratories described an air-borne magnetometer which was at the end of a cable trailing from an airplane. To express the variation of the intensity of the field from point to point, the Oersted (that weak field) is much too large a unit, and it is customary to express the variations in a smaller unit, the gamma, where one Oersted is 100,000 gammas. The Bell unit, when operating on its most sensitive range gives full scale deflection on a meter with a variation of 50 gammas, or one gamma per scale division; this is ten micro-oersteds per division! (Who ever heard of a micro-oersted!)

We shall now turn to the applications of high-power vacuum tubes in industry. At this time large amounts of power and high frequencies can be produced only with the aid of vacuum tubes operating as oscillators. Industry makes use of HF power mostly for two applications: induction and dielectric heating, both of which have seen spectacular growth and advances in the last two decades. For induction heating of thin metal or for surface hardening, the frequency is usually below 500 kc. One of the outstanding installations in the war years was furnished in 1943 by Westinghouse for the purpose of flowing tin on steel strip. The strip which had been electroplated with tin previously was traveling at a speed of 500 to 1000 feet/minute through the coils, which were supplied with 1200 kw at 200 kc. This method permitted a speed from 4 to 5 times as high as could have been accomplished with a gas-fired furnace or a hot oil bath. Within the past few years HF power has also been successfully applied to the welding of tubing, including the production of copper tubing from copper strip.

The second field of use of HF power is in dielectric heating. The first installation attracting wide attention was furnished by the Girdler Corporation in 1942 (the principle was demonstrated two years earlier). It was used for the bonding of plywood. Input was 400 kw at two Mc. In 1943 electronic sewing machines for plastics, operating at 60 Mc and RF guns for spot gluing, operating at 200 Mc were introduced. In 1946 Westinghouse furnished a 125 kw unit at 13.6 Mc to the Firestone Co. for curing and drying foam rubber. In 1952 industrial magnetrons for dielectric heating at frequencies from 915 to 2450 Mc were under consideration.

There are other uses of HF power in industry such as the production of ultrasonic power, but the two just mentioned represent by far the largest part.

Let us now examine the applications of gas tubes in industry. It has already been stated that these tubes are really the workhorses of the industrial power engineer, and it is again quite impossible to discuss more than a few applications. Spot, resistance and flash welding have assumed a dominant role in many manufacturing processes in this country and today almost all welding machines are controlled by ignitrons. As far as the number of controlled kilowatts is concerned, no other tube has contributed, or is contributing, as much to industry as the ignitron. But, while welding applications account for the largest number of ignitrons, they have, of course, found other uses, such as motor control in locomotives, in strip mills and in the control of pot lines in the aluminum industry. In 1943 a converter using ignitrons was furnished by General Electric to the Carnegie Illinois Steel Corporation; it served to convert 20,000 kw from 44 kv at 25 cps to 69 kv at 60 cps. The frequencies do not have to be exact.

The use of HF power for induction heating has already been discussed, but for the majority of heating applications such high frequencies are neither needed nor, as a matter of fact, desirable. Even today, there exists a crying need for economically priced conversion to power in the kilocycle range, from thousand to ten thousand cps. Gas tubes can be made to operate as inverters, and in 1945 Allis-Chalmers offered an inverter using excitron tubes to produce power in the 2000 cps range for induction melting furnaces. One very unusual application of ignitrons will serve to close this part. It is the power plant for the Bevatron, the high-energy particle accelerator of the University of California Radiation Laboratory. The magnetic field in the chamber must be built up to its maximum value and reduced to zero ten times per minute. During the buildup of the field, which takes about two seconds, a dc voltage varying between 18,000 volts at the start and dropping to 12,000 volts at the end of the two seconds is applied to the coil and the current rises from zero to 8333 amperes. The energy stored in the magnetic field is then 80 megajoules (80,000 kw seconds)!

The dc power is obtained through ignitrons from two ac generators, each rated at 46,000 kva, and each driven by an induction motor, rated at 3600 hp equipped with a solid steel plate flywheel weighing 67 tons. Most of the energy required during the current buildup comes from the flywheel. To bring the field down to zero, again in about two seconds, the firing control of the ignitrons is instantaneously changed (by back phasing) so that they act as inverters. The dc voltage at the coil changes instantaneously from +12 kv to −6 kv, increasing from this to about −16 kv during the time that the current decreases to zero. During this time the inverter action causes the energy to flow back into the ac generator and the flywheels increase in speed again. The driving motor has to furnish only enough energy to cover the losses during each cycle. The ignitrons are similar in design to those used in 3000 volt railroad service; the high direct voltage requires four of them in series, and the high peak current require four anodes in parallel. This is truly an engineering accomplishment of the first order.

X-rays have already been mentioned several times in this survey. Industry has made use of X-rays for many problems dealing with the inspection of products, as is well known. What is not so well known is that X-ray pictures can be produced with very short exposure times. In 1942 it was already possible to obtain the X-ray picture of a bullet while penetrating a steel plate, and in 1949 X-ray moving pictures were obtained on 9.4 inches wide film taken at a rate of 150 frames per second with an exposure time of 1/100,000 second.

We are now ready to turn to the latest member of the family, the solid-state devices. As far as transistors are concerned, their contribution is found mainly in a tremendous reduction in weight and size, and an increase in reliability in many devices formerly employing tubes. This might not appear to be a significant point, when we look at photoelectric relay, or some device operating with perhaps half a dozen tubes, but when thousands of tubes, as in a computer, are involved, space and weight savings produce really spectacular results. Add to this the fact that transistors can do many things that tubes cannot do (as explained earlier in this discussion) and it becomes understandable why a great number of electric devices are being "transistorized."

The silicon-controlled rectifier is only about four years old, and its cost is still quite a serious obstacle to widespread use. Nevertheless, there are already some applications indicating the trend that developments will take. In March, 1961 General Electric introduced a 50 kva converter, for converting 125 volts dc to 220/440 three-phase power at a frequency variable from 50 to 500 cps. This is the kind of development that industry has been waiting for and will watch with intense interest.

This short discussion of semiconductors should not be closed without at least mentioning a device which may well become one of the most important building blocks for the engineer interested in measurement, control and computing, and that is the Hall generator. The discovery by E. H. Hall in 1879 that a stream of electric charges in a metal (the word "electron" was not known at that time) could be deflected by a magnetic field perpendicular to the flow of charges was a landmark from a theoretical point of view, but the voltage produced by this squeezing of the charges to one side of the conductor was so minute as to be of little practical value. It was reserved for later generations of physicists to discover the relation between the Hall voltage and the carrier density and mobility in the conductor material. Suffice it to say here that the Hall voltage obtained from semiconductors is of practical value (in the order of hundreds of millivolts). Since the Hall voltage is proportional to the magnetic field density and to the current, we now have a device which is capable of generating a voltage that is proportional to the product of two quantities, provided we can convert the two quantities into a current and into a magnetic field density respectively. Simple wattmeters, operated over a wide frequency range, would appear to be possible by utilizing this effect.

Statistics showing how much industrial electronics has contributed to the economy are difficult to obtain, mainly because the dividing line between industrial and other electronics is hard to define.

One set of statistics divides the field in industrial and consumer electronics, the latter including entertainment, broadcast, television, radio, services and distribution. With this division the statistics[1] are as follows (the 1965 figures are projected):

	1940	*1950*	*1960*	*1965*
1) Industrial	4.5 million	680 million	2.8 billion	4.6 billion
2) Consumer		3.1 billion	5.35 billion	6.7 billion

It is seen that the rate of growth in the industrial field exceeds by far that in the consumer field and that the day when the two are equal is in the foreseeable future.

What will the future hold? In a field as fluid, not to say explosive, as electronics has proved to be, predicting what the future might bring is a risky business indeed. Who could have predicted in October 1906 that one month later Lee De Forest would give the world the vacuum triode, or who in 1947 could have predicted that the transistor would be born a year later? To be sure, if the doors of the research laboratories of the giants of the industry were all thrown open to a reporter, he could make quite some predictions (although they would, of course, not be predictions anymore). For many good and valid reasons, companies do not disclose the projects they are working on until the job is finished. To engage in flights of fancy has unfortunately never been a strong side of this reporter, and predictions will therefore be confined to things on which work is known to be in progress.

[1] The author is indebted to W. W. MacDonald, Editor of *Electronics*, who was kind enough to provide these figures.

Since tubes have been with us about 40 years longer than the recently developed semiconductor devices, it seems reasonable to expect more spectacular advances with the latter, (although the heaterless microminiature triode of General Electric described earlier makes one wonder what might happen in this field). Transistors of today are still low-power devices, just as the first vacuum tubes were. The material used in the present devices is either germanium or silicon. The reason for this is at least partly that these are simple four-valence elements, which made the development of the theory of the conduction of current through semiconductors less difficult. But attention is now being directed to more complicated materials, which may expand the range of the characteristics of these devices tremendously. In 1961 gallium arsenide has received the attention of semiconductor manufacturers. Theoretical considerations show that this material will be superior to silicon and germanium both for high frequencies and high-temperature performance, but at this time the preparation of the material in the required purity is more difficult. Experimental transistors have been built confirming the theoretical predictions. But the physicists are not stopping here; considerable work is being done on three element semiconductor materials, and their list is practically endless. One compound under investigation in the Bell Laboratories is silver antimony telluride ($AgSbTe_2$), which shows promise for use in thermoelectric refrigeration. Scientists have not stopped at these intermetallic compounds; work is even going on with organic semiconductors; at this time the subject of these studies is anthracene, one of the simplest organic compounds.

As far as industrial electronics is concerned, these future tools will provide the means to solve many more problems, and to conquer new fields. The semiconductor materials seem to hold the key to the direct conversion of heat into electricity, and the reverse, refrigeration by electricity. Ratings of SCR's will certainly go up, and turning the current off by gate action is a distinct possibility. As far as the consumer is concerned, the day when his car will be guided along the highway is perhaps quite near; RCA and General Motors are cooperating in the development of an electronic highway. Olson and Belar of RCA have presented a paper on a phonetic typewriter. From the General Electric laboratories comes a radically new system of recording called thermoplastic recording which possesses the advantage of almost instantaneous recording, and will produce pictures in either color or black and white which do not require chemical processing. For industry it may provide a new information storage system; it already can concentrate 100 times as much information in a given space as can magnetic recording. Truly, the future of industrial electronics looks more exciting than it ever did before.

Section 17

INFORMATION THEORY

Organized with the assistance of the IRE Professional Group on Information Theory

On Communication Before the Days of Radio*

COLIN CHERRY†

Summary—It is stressed that the really important factor in communication systems is not their speed of operation, but their reliability. This has been true throughout history. The earliest known postal system was that of the Persian empire which was later taken over by the Greeks and then the Romans, being used by them in their occupation of Europe 2000 years ago. The system which they adopted will be described briefly; it remained substantially unchanged in Europe until the later part of the Middle Ages. Every effort was made to make such systems reliable; it was of no value to Caesar to send messages if he could not be sure they would reach their destination.

These early systems were used by an aristocracy, and the most significant thing about modern telecommunication systems is their universal use by every man, woman and child. But their use is confined to areas of high literacy in the world, not so much because of wealth, but because of mental attitudes. The growth of modern postal services during the past 300 years is described and the effect of the coming of telephones and telegraphs commented upon.

WHEN modern communication systems and their influence upon us are discussed it is most usual that their speed is mentioned, whilst their effect upon "the tempo of life" (whatever that may mean) is stressed. In the author's opinion the really important social aspects of modern telecommunication derive, first, from their great reliability and, second, from their universality, by which is meant their use by every man, woman and child. These two factors have enabled the sheer size of societies to grow, from the village to the city-state, to the antique empires, until today our society, in the sense of areas of mutual dependence and cultural exchange, is being forced into global size. In early historic times there was indeed long-distance communication, covering empires and the known world, but it was confined to diplomatic, military and governmental use by emperors, generals, and other privileged aristocracy over the whole historic period from the Persian Empire to the coming of the railways. The structures of the dominant societies over the past two or three thousand years have taken a great variety of forms, but it is suggested that the types of civilized societies in which we live today are very much colored by the two factors mentioned—the utter reliability of our communication systems and their universal use by the "common man."

In today's world the use of our postal services, our telephones, telegraphs and all other forms of communication is confined to areas of high literacy. This is not only a matter of wealth, but one of mental attitude.[1] The same thing is true through history; literacy has always been an essential prerequisite for the adoption of all forms of communication service. An illiterate society is paralyzed; and we must not forget that even today nearly half the world's people are illiterate. Indeed we may argue that *societies can develop and "advance" only*

* Received by the IRE, September 18, 1961.

† Dept. of Elec. Engrg., Imperial College of Science and Technology, London, England.

[1] D. Lerner, "The Passing of Traditional Society," The Free Press, Glencoe, Ill.; 1958. (Reports of social surveys, in the Middle-Eastern counties concerning the effects of urbanization, increased literacy, mass-entertainment and communication upon social structure).

so fast as they can develop means of acquiring, recording and disseminating information.

Man's progress from the earliest communities, through the great empires of antiquity up to the gigantic and highly organized industrial societies of today, has been one long story of improved means of communication. As these have been developed, so the sizes of our communities have grown, giving all the advantages of the division of labor and production of wealth, of broadened cultural horizons, of control and protection of economies and other benefits of social *sharing*. The very word *communication* comes from the Latin *communico*, "I share."

But it is the *explosive* growth of all forms of communication which seems to be significant today. Print is only 400 years old, newspapers less than 200; telegraphs and telephones less than a century and radio half that time. Whereas the illiterate peasant of earlier centuries knew only of the doings of his own village (and indeed the same thing was true of a great bulk of our own people up to the days of our grandfathers), the thoughts of every literate man and woman nowadays are world-thoughts. Every time we open our morning newspapers we are confronted by the whole world's agonies. It is the suddenness with which this load has come upon our consciences which threatens to bring us to our knees.

Man is a communicating animal; he possesses the unique powers of speech and writing. Human experience is not a moment-by-moment affair, but has continuity; we have contact with our ancestors and our descendants, and we have a sense of history and tradition. All our communicable thoughts are bound by the structures of our languages and systems of symbolism, and these continually shift and change. The early Mediterranean scripts were in pictographs, ideographs and hieroglyphs, giving direct pictorial representation of objects and, by associations, of names, actions and ideas of all kinds. But it was the evolution of phonetic writing, during the Coptic period, which was the great step; speech and writing became closely linked. The civilizattions which did not adopt such techniques have been handicapped throughout their history, more so today than in earlier times.

We make relatively few *significantly* different sounds, when we speak, few phonetic symbols are needed. Phonetic writings settled down into sets of two or three dozen alphabetic letters, divided into consonants and vowels. This great simplification of script did not stop there. For example, ancient Hebrew script omitted the vowels. Deliberate condensation has been tried in numerous ancient writings; Church Slavonic abbreviated common words of religious texts, rather as today we write *cms* and U.S.A. for similar reasons of economy. Shorthand writing is another very early invention, said to be due to the Greek slave Tyro in recording the lengthy speeches of Cicero! This looks not unlike modern shorthand, but was based upon spelling; it continued to be used in Europe until the Middle Ages.

Together with these evolutionary steps and deliberate inventions for recording messages, other remarkably early developments took place for their transmission. Polybius, for example, is well known for his description of telecommunication techniques using torches and other visual means, and also for his coding of the alphabet.[2] The classical historians, Herodotus, Xenophon and Polybius, stress the need for communication services, based on central government, for control of scattered empires and for waging wars.

How were the great empires of antiquity held together? How did Caesar know what his generals were doing, or what revolts were taking place? By a system which was the forerunner of our present-day postal service. The earliest known postal service was that of the Persians, by which written letters were carried on horseback, in relays between established postal stations; this system was later introduced into Imperial Rome by Augustus. The Roman Empire was held together by an elaborate network of couriers, on horseback or in carriages, with relay posting stations at regular intervals along all main roads. Each postal station kept 40 horses and grooms,[3] and the speed was probably 50 miles a day[3-4] with accurate fixed times for collection and delivery. Similar systems were used by the Great Khan described by Marco Polo.

As I stressed at the opening of this article, it is *reliability* rather than speed which is the most important factor in communication. The various accounts of these ancient postal services all emphasize these points—precision of timing and certainty of arrival. Caesar had to know that his messages were received and when he could expect a reply. Truly we have increased the speed by one or two orders of magnitude in modern times, but of first importance is the fact that we have vastly increased the reliability and the scale. Caesar, in the form of authority, now appears in every home.

The Roman postal system was carried into Europe by the Roman occupation forces, and stayed there in various centers long after the collapse of their Empire. Indeed the system remained essentially the same until the late medieval period; at that early time several postal establishments in various states of Europe were run by the Universities! The University of Paris, for instance, ran a postal service[5] from the early 13th century until the 18th.

From the first days, postal services seem to have been organized under government control and used for carrying the kings' messages. It was not until the 17th century that the full social need for regular mail service to be used by government officials, merchants and all but

[2] Polybius, "The Histories," Book X. (2nd Century B.C.).

[3] H.-G. Pflaum, "Essai sur la *Cursus Publicus* sous le Haut-Empire Romaine," Imprimerie Nationale, Paris, France; 1950.

[4] A. M. Ramsay, "The Speed of the Roman Imperial Post," *J. Roman Studies*, vol. 15, pp. 60–74; 1925.

[5] The word *mail* comes from Old French *male* meaning a bag for travellers.

the "common people" was realized. In 1633 a weekly mail which took four to five days was set up between London, Antwerp and Brussels. By Cromwell's time (1599–1658) regular services were running on all main roads of England, under license from that remarkable new institution—the English Parliament.

Surprisingly enough there were several attempts by private persons to set up postal services. John Hill, for example, organized a *penny post* throughout England, but it was regularly attacked by Cromwell's soldiery. As early as 1680 a private London and suburban penny postal service was established[6]—with hourly collections, wall-boxes and sorting-offices which were later incorporated into the state system. Similar developments took place in America, at much the same date.

Throughout this early history of the postal services we read of constant efforts to improve the reliability of the mails, and of a steadily growing public confidence in this service. By Queen Anne's time, regular mails were reaching England's scattered colonial Empire and secrecy was (officially) ensured. Whereas under Cromwell's regime foreign mails were read by a board of examiners, now only Secretaries of State had authority to do this!

Attack and robbery on the highway was steadily defeated, by higher speeds and by arming the coach guards, so that public confidence increased. But it was the coming of the railways which finally established the postal services as part of the life of every man as something utterly reliable and always available—that which we accept today as our due, much as we take our air and water.

To quote figures from my own country, Britain, by the time of Sir Rowland Hill's Post Office Reform Bill (1837) letters in the mails totaled 88 million per annum (today there are 10,000 million packages a year of all kinds).

[6] By William Dockwra.

The railways were dominant too in the early history of telegraphy—theirs was the first "electronic" system of communication and the first alternative to the post. Although the suggestion that messages might be sent over a wire (by discharging a Leyden jar) was made[7] as early as 1753, the first really practical trials were made by the railway companies (starting in 1837) and for a long time telegraphs were confined to use by the railways. Even today we have the familiar sight of telegraph poles and wires through our railway carriage windows.

But soon companies were formed. In 1851 the stock exchanges in Paris and London were connected so that prices could be rapidly compared for the first time. Public confidence grew and people were prepared to pay high charges for telegrams; but for some years during the middle of the last century the service was confined to those towns and cities served by the railways.

Only brief reference is made here to the coming of telephones and radio, for their development falls within the lifetime of many of us or of our fathers. From the social point of view, their arrival had two major consequences. The telephone network required some form of *subscriber organization* for the first time (letters and telegrams are the affairs of individuals, but a telephone in your home requires a terminal connection, in your name, in the local exchange). The advent of radio brought mass-communication. Radio and television sets are now owned by people of all classes and conditions. The voice and vision of authority now appears in every home, to carry out education or propaganda, to inform or to deceive, to inflame or to hypnotise, to unify our language or to initiate clichés and slang, to model our heroes and set our norms. There is no doubt that radio and television will largely decide the road along which our society will pass into the future. They can be terrible instruments. We must watch ourselves.

[7] By an anonymous writer to the Scots Magazine (vol. 15, p. 73; 1753).

Noise and Random Processes*

J. R. RAGAZZINI†, FELLOW, IRE AND S. S. L. CHANG, FELLOW‡, IRE

Summary—Early investigators in the field of communications first realized that the presence of unwanted random noise was an important factor following the discovery that the maximum gain of an amplifier was limited by the discrete nature of currents in electron tubes. Called *shot effect*, this was first explained by W. Schottky and later by many other investigators. Much research on this problem during the second and third decades of the twentieth century finally led to the rigorous formulation of the phenomenon by B. J. Thompson and others in 1940. Concurrently, the problem of spontaneous thermal noise effects in conductors was studied and formulated. By 1940, the situation was developed to an extent that the application of mathematical statistics to explain and solve broader noise problems in systems was inevitable. About this time, the basic contributions of N. Wiener led to an understanding of the optimum linear filtration of signals imbedded in random noise. His work influenced the entire course of development of theory on the optimization of filters designed to abstract a signal out of its noisy environment.

I. Introduction

THE THEORY of noise and random processes in electronic devices was developed in two stages. The first stage spanning the two decades following 1918 brought about the understanding of the nature and effects of noise in vacuum tubes and circuits. The second stage, initiated by Wiener in the early 1940's, established the theoretical basis for the analysis and synthesis of systems subjected to random signals and noise. His rationale for optimization and design of optimum linear filters subjected to Gaussian signals and noise set the direction for much of the subsequent research on this subject. While a detailed discussion of optimum filtration and detection of signals embedded in random noise is treated in a separate paper on information theory, the subject is discussed briefly to demonstrate its link to the early work on the subject. The literature on noise and random processes is so extensive as to make a complete coverage impractical. Only highlights and significant stepping stones in the development of the theory are described.

II. Early Work on Sources of Noise

The first realization that unwanted random noise was a factor to contend with in the field of communications came during World War I when attempts were being made to design high-gain vacuum-tube amplifiers. It was soon found that there was a limit to the number of stages which could be cascaded in the quest for high gain due to an unacceptably high background noise which masked the weak signals being amplified. In his classic paper[1] Schottky first explained one of these effects and formulated the random component in the plate current of a vacuum tube.

Schottky ascribed the random fluctuations in the plate current to the fact that this current is composed not of a continuum but rather of a sequence of discrete increments of charge carried by each electron arriving at the plate at random times. The average rate of charge arrival constitutes the dc component of the plate current on which is superimposed a fluctuation component as each discrete charge arrives. He referred to this phenomenon as *schroteffekt* or *shot-effect* as we call it.

In arriving at his result, Schottky made a number of important simplifying assumptions. First, he assumed that the transit time of an electron from cathode to plate was infinitesimal so that the current pulse produced by each electron could be represented by an impulse or delta function. This immediately postulated that the random component of the current had an infinite frequency spectrum or was "white." This approximation is not a serious limitation except at extremely high frequencies when the transit time is a sensible fraction of a period.

The second and more limiting assumption is that the only force acting on the electron in transit is the electrostatic field between cathode and plate. The forces produced by other electrons in the space charge are neglected. This assumption is valid only when the plate current is temperature limited. The expression for the power spectrum of the random component of the plate current obtained by Schottky was

$$W(f) = 2eI \ \text{a}^2/\text{cps}, \tag{1}$$

where e is the charge of the electron in coulombs and I is the dc plate current in amperes.

For the first time a quantitive expression for the random component of plate current became available and estimates of the ratio of signal-to-noise power could be made. But the formula suffers from a very serious defect, namely, that it is applicable only when the plate current is temperature limited. In practice, vacuum tubes were then and are now being operated with plate currents far below temperature saturation and the measured values of random noise in the plate current were much lower than those predicted by the Schottky formula. For many years attempts were made to explain

* Received by the IRE, October 2, 1961; revised manuscript received, December 14, 1961.
† College of Engineering, New York University, N. Y.
‡ Dept. of Elec. Engrg., College of Engineering, New York University, N. Y.

[1] W. Schottky, "Theory of shot effect," *Ann. phys.*, vol. 57, pp. 541–568; December, 1918.

this reduction of shot noise,[2-5] but most were empiric and did not get to the basic explanation of the mechanism by which space charge could affect the result.

A major contribution to the understanding of the source of shot noise in space-charge-limited vacuum tubes was made in a set of papers by Thompson, North, and Harris.[6] Briefly, these investigators showed that a negative compensatory current pulse is generated each time an electron crosses the potential minimum between cathode and plate. Referring to Fig. 1, the region between cathode and plate is divided into an α and β region to the left and right of the potential minimum, respectively. As an electron e passes from the α to the β region, it increases the population of the β region, thereby depressing the potential minimum slightly. This inhibits the transit of a certain amount of charge out of the α region and returns it to the cathode. In effect, this is equivalent to a reverse coherent current pulse, as shown in Fig. 2, where i_n is the forward pulse of current produced by an electron entering the β region and i_c is the negative current pulse. The relation between the two is given by

$$i_c = \lambda i_n, \tag{2}$$

where λ is a constant dependent on the degree of space-charge saturation of the plate current. This modification results in a power spectrum for the random component of the plate current given by

$$W(f) = 2\Gamma^2 eI \ \text{a}^2/\text{cps}, \tag{3}$$

where Γ^2 is $(1-\lambda^2)$. For zero space charge, Γ becomes unity and the power spectrum reduces to Schottky's result.

Thompson, North, and Harris computed Γ based on theoretical considerations and proved the result by experiment. A typical plot of Γ vs cathode temperature is shown in Fig. 3. It is seen that the power spectrum of the shot noise in a vacuum tube can be as little as 4 to 5 per cent of the values given by the Schottky formula. The authors extended the theory to cover multi-electrode vacuum tubes. This work essentially completed the basic theory of random noise in vacuum tubes insofar as random noise was concerned.

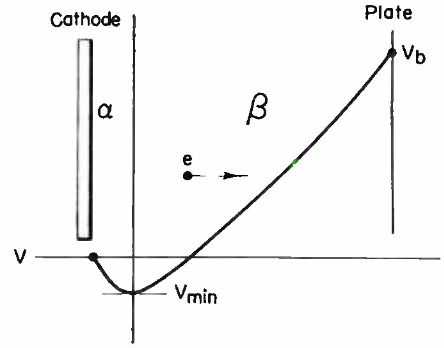


Fig. 1—Potential variation in space-charge-limited diode showing α and β regions.

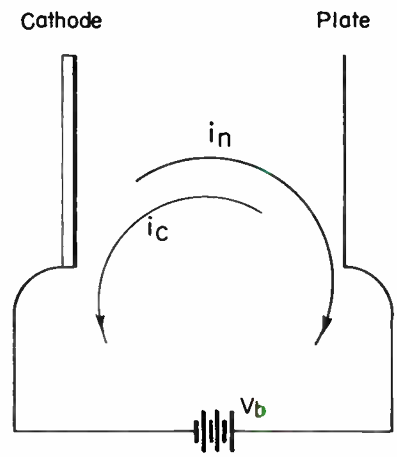


Fig. 2—Noise and compensating pulse produced in space-charge-limited diode.

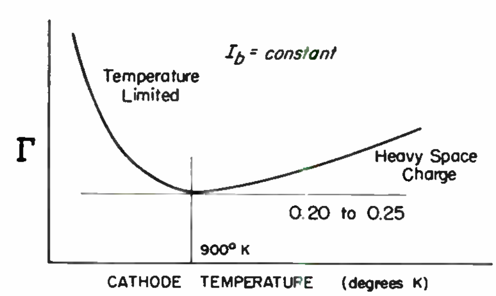


Fig. 3—Typical variation of Γ in a space-charge-limited diode with oxide-coated cathode.

In addition to fluctuation effects produced by vacuum tubes, it was found that random noise signals were generated in metallic resistors made of homogeneous materials. These effects were found to be temperature dependent and are known as thermal noise. A number of basic contributions to the understanding of thermal noise were made in the 1920's and 1930's[4,7,8] among which

[2] T. C. Fry, "Theory of shot effect," *J. Franklin Inst.*; 1925. Also, "Probability and Its Engineering Uses," D. Van Nostrand, Co., Inc., Princeton, N. J.; 1928.

[3] F. B. Llewellyn, "Study of noise in vacuum tubes and attached circuits," Proc. IRE, vol. 18, pp. 243–265; February, 1930.

[4] E. B. Moullin and H. D. M. Ellis, "The spontaneous background noise in amplifiers due to thermal agitation and shot effects," *J. IEE*, vol. 74, pp. 323–356; April, 1934.

[5] A. J. Rack, "Effect of space charge and transit time on shot noise in diodes," *Bell Sys. Tech. J.*, vol. 17, pp. 592–619; October, 1938.

[6] B. J. Thompson, D. O. North, and W. A. Harris, "Fluctuations in space charge-limited currents at moderately high frequencies," *RCA Rev.*, part 1 vol. 4, pp. 269–285; January, 1940; pt. 2, vol. 5, pp. 106–125; July, 1940; pt. 3, pp. 244–260, October, 1940; pt. 4, vol. 5, pp. 374–388; January, 1941; pt. 5, pp. 505–524; April, 1941.

[7] J. B. Johnson, "Thermal agitation of electricity in conductors," *Phys. Rev.*, vol. 32, 2nd ser., pp. 97–109; July, 1928.

[8] H. Nyquist, "Thermal agitation of electricity in conductors," *Phys. Rev.*, vol. 32, 2nd ser., pp. 110–113; July, 1928.

was the outstanding paper by Johnson in 1925. The source of thermal conductor noise was traced to the random excitation of the electron gas in the conductor in consequence of its existence in an environment of thermally-agitated molecules. The effect is similar to the Brownian movement of particles suspended in a liquid in which the thermally-agitated molecules of the liquid collide with the suspended particle and impart to it a certain amount of energy. Since the particle is cohesive, collision with any one of its molecules sets the entire particle in motion thereby resulting in random movements observable under a microscope.

The basic principle used to explain Brownian movement is that of equipartition of energy. The cohesive molecular group constituting the particle has imparted to it by collision an equal share of energy per degree of freedom just as the individual free molecules in the liquid have themselves equal shares of thermal energy. In the case of Brownian movement, the particle has three degrees of freedom thereby giving three independent energy storage modes, each of which stores a mean energy given by $\frac{1}{2}KT$ where K is the Boltzman constant and T the absolute temperature.

Analagously, in the case of the electron gas in a conductor, the same cohesive property that exists in the particle can be postulated. If one electron is set in motion by a collision with a thermally-agitated molecule, the entire electron gas in the circuit is set in motion by the suddenly changing magnetic field generated by the accelerating electron. As a result, the principle of equipartition of energy can be applied to this phenomenon and the electron gas derives a share of energy equal to $\frac{1}{2}KT$ for each energy-storage mode such as a condenser or inductor in the circuit. Based on considerations like these, the power spectrum of equivalent generator associated with each resistor is given by

$$W(f) = 4KTR \text{ v}^2/\text{cps} \tag{4}$$

where K, the Boltzman constant, is 1.372×10^{-32} joules/°K, T is absolute temperature in °K, and R is the resistance in ohms. The correctness of the hypothesis and the resultant formula for the power spectrum has been proved experimentally. Thus, all resistances in a circuit can be represented by a resistance and a white noise generator in series as shown in Fig. 4. Resistance components which are not homogeneous, such as composition resistors, generally produce unpredictable noise effects several times greater than homogeneous resistors, due to "crazy contact" effects. As in the case of tube noise, it is fair to assume that thermal noise spectra are white since the electron pulses produced by molecular collision are extremely short in duration.

Thus, by the late 1930's there was good understanding of the mechanisms producing random noise in electronic circuits and it was possible to relate these sources to the performance of a system. The link which made this possible is the fact that the power spectrum is related to the Fourier spectrum of the random signal by the following:[9]

$$W(f) = \lim_{T\to\infty} \frac{2}{T} \left| G(j\omega) \right|^2, \tag{5}$$

where $G(j\omega)$ is the Fourier transform of the random signal $g(t)$, and $\omega = 2\pi f$. Thus, if such a random signal is applied to a linear filter as shown in Fig. 5, it follows readily from the fact that the Fourier transform of the output is the product of the Fourier transform of the input and the transfer function of the filter that the relation between the power spectrum at the input and the output is given by

$$W_y(f) = \left| H(j\omega) \right|^2 W_x(f), \tag{6}$$

where $H(j\omega)$ is the transfer function of the filter or the Fourier transform of its impulsive response. Coupled with the fact that the power spectrum $W_x(f)$ is related to the mean-square value of the signal by the integral

$$\overline{x^2} = \int_0^{+\infty} W_x(f)df, \tag{7}$$

the net effect of noise signals generated spontaneously in electronic circuits can be calculated.

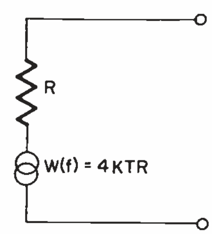


Fig. 4—Equivalent circuit of resistor including thermal-agitation noise-voltage generator.

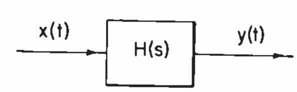


Fig. 5—Block diagram of linear filter with random input and output.

For convenience in analysis, the quantity spectral density $\Phi_{xx}(j\omega)$ is often used instead of power spectrum. Spectral density is an even function which extends from negative to positive infinite frequency and has values at each frequency half those of the power spectrum. Thus, the mean-square value of a signal is given by

$$\overline{x^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{xx}(j\omega)d\omega. \tag{8}$$

It should be noted here that noise sources so far considered are quite correctly assumed to generate white

[9] S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 23, pp. 282–332, 1944; and vol. 24, pp. 46–156, 1945.

noise spectra based on an understanding of the originating mechanisms. Until the 1940's a certain amount of optimization by adjustment of circuit parameters to obtain optimum signal-to-noise ratios was possible by straightforward analysis. It was later that fundamental work was done in optimum filtering of signals in the presence of noise which is not necessarily white as described in subsequent sections.

III. More General Statistical Tools

The techniques which were used in expressing and analyzing the effects of random noise produced by vacuum tubes and resistors were, in fact, applicable to general problems. For instance, electronic equipments became part of more comprehensive systems such as radar-tracking systems, fire-control systems, navigation systems, etc., during the late 1930's and early 1940's. Analysis and optimization of such systems were urgently required, especially under the stresses of World War II when extension of system capabilities became a matter of national survival.

In most cases involving more comprehensive systems, the mechanisms generating random signals cannot be clearly specified or postulated as they can for shot effect and thermal noise. As a result, the statistical properties of random noise must be determined by experiment from time records of typical situations. A direct determination of power spectrum from these data is not as convenient as through the indirect approach using the autocorrelation function. As far back as 1930[10] Wiener showed a simple interrelationship between the autocorrelation function and the power spectrum, a relation which was to prove most useful in experimental as well as analytical operations. This expression is known as the Wiener-Khintchine relation in recognition of later independent work done by Khintchine.[11]

In order for the Wiener-Khintchine relation to apply, a number of assumptions must be made regarding the random signal. First, the random process of which a particular time record is a member must be stationary; that is, its statistical properties are independent of the origin of time. Secondly, the random process must satisfy the ergodic hypothesis. This implies that the stationary random time functions, an ensemble of which constitutes the random process, possess the same properties whether averaged over time on any one function or over an ensemble of functions at any given time. This means that such characteristics as the autocorrelation function can be obtained by integrating over time as well as over the ensemble. Within these limitations, the autocorrelation function of a random time function $x(t)$ is given[12] by

$$\phi_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)x(t-\tau)dt. \quad (9)$$

The function is a measure of the dependence of the values the time function at any one time to another time displaced τ seconds. For instance, white noise is completely uncorrelated, the correlation function being zero for all finite values of displacement τ. Its mathematical expression is the impulse-response function $\delta(\tau)$ multiplied by a constant.

The Wiener-Khintchine relation states that the spectral density $\Phi_{xx}(j\omega)$ is the Fourier transform of the autocorrelation function[13]

$$\Phi_{xx}(j\omega) = \int_{-\infty}^{\infty} \phi_{xx}(\tau)e^{-j\omega\tau}d\tau, \quad (10)$$

and conversely, the autocorrelation function is the inverse Fourier transform of the power spectrum

$$\phi_{xx}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_{xx}(j\omega)e^{j\omega\tau}d\omega. \quad (11)$$

Similar expressions exist for the relation between two different random functions, $x(t)$ and $y(t)$:

$$\phi_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)y(t+\tau)dt \quad (12)$$

and

$$\Phi_{xy}(j\omega) = \int_{-\infty}^{+\infty} \phi_{xy}(\tau)e^{-j\omega\tau}d\tau \quad (13)$$

$$\phi_{xy}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \Phi_{xy}(j\omega)e^{j\omega\tau}d\omega, \quad (14)$$

where $\phi_{xy}(\tau)$ is known as the crosscorrelation function and $\Phi_{xy}(j\omega)$ the cross spectral density.

Eq. (6) can also be put in a more general form. Let y_1 and y_2 be the outputs from filters $H_1(j\omega)$ and $H_2(j\omega)$ with x_1 and x_2 as the respective inputs. Then

$$\Phi_{y_1y_2}(j\omega) = H_1(-j\omega)H_2(j\omega)\Phi_{x_1x_2}(j\omega). \quad (15)$$

In addition to its application in the analysis of systems in which random signals and noise exist, the Wiener-Khintchine relation is very useful for the experimental determination of pertinent properties of the signal. For instance, a computation and plot of the autocorrelation function can be made directly from the experimental records of a random signal. Usually, the procedure is to fit a convenient analytical expression to the experimental plot of the autocorrelation function and from this the power spectrum of the signal can be obtained. In cases where complex systems such as radar tracking systems must be optimized, information of this type is usually obtained in this manner.

[10] N. Wiener, "Generalized harmonic analysis," *Acta Math.*, vol. 55, pp. 117–258; 1930

[11] A. Khintchine, "Korrelationtheorie der stationaischen prozesse," *Ann. Math.*, vol. 109, pp. 604–615; 1934.

[12] H. M. James, N. B. Nichols, and R. S. Phillips, "Theory of Servomechanisms," McGraw-Hill Book Co., Inc., New York, N. Y.; pp. 271–273; 1947.

[13] *Ibid.*, p. 283.

In most systems which are of interest to the engineer and the designer, random processes are assumed to be Gaussian or at least assumed to be satisfactorily approximated by such a distribution. As discussed previously, a random signal or noise is usually generated by a large number of independent events (the shot-effect noise is generated by thermoinic emission of individual electrons, etc.). It follows from the central limit theorem in probability theory that, not only is the amplitude of such a signal normally distributed, but the joint distribution function of various amplitudes at various times is also a multivariate normal distribution function. A random signal with this property is called Gaussian. A significant implication of the above is that all statistical properties can be determined from a knowledge of the autocorrelation function or, equivalently, the power spectrum. Furthermore, it can be shown that Gaussian signals remain Gaussian after passing through linear filters. In practical problems, therefore, the measurement of the autocorrelation and its conversion into the power spectrum by Fourier transformation gives all the necessary statistical parameters needed for optimization of systems using procedures which were developed in the World War II years and later.

An outstanding piece of work in this regard is due to Rice.[9] Rice gave an estimate of the closeness to Gaussian distribution of a signal of the shot-effect type; and assuming the distribution to be Gaussian, he derived equations for a number of significant relations, for instance: the distribution of the total power of a random signal, the number of zero crossings and maxima and minima per second, and the power spectrum of the output signal from a nonlinear device.

IV. The Concept of Optimization

There are two types of problems in communication and control in which random signals play an essential role, namely: a) processing and separation of signal from noise, and b) detection of weak signals embedded in noise. Fig. 6 illustrates a problem of type a). The desired output $v(t)$ is the result of a specified linear operation on a signal $r(t)$. The available input signal $r(t)$ is contaminated with random noise $n(t)$. The problem is to design a filter for $i(t)$ so that the output $c(t)$ is closest to $v(t)$ in some well-defined sense.

A major contribution to this problem was made by Wiener[14,15] who, in 1942, worked out an optimization theory of such significance that its influence was to dominate at least the next two decades. He introduced a criterion for closeness of the signal filtered signal $c(t)$ to the ideal signal $v(t)$ on the basis of a least squares difference between the two. Perhaps both mathematical expediency and universality of Gaussian random noise (and signal) conspired to influence him and other pioneer investigators to select the least square error criterion as they did.

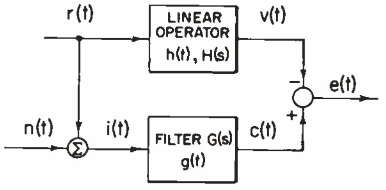


Fig. 6—Block diagram showing criterion for optimum filter.

The criterion for optimization in the case of random signal and noise is expressed as a cost $k(t)$ defined as

$$k(t) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} [e(t)]^2 dt. \tag{16}$$

The design of the so-called optimum filter $G(s)$ is dictated by a minimization of this cost function subject to the conditions that the filter be linear and physically realizable. The latter requires that impulsive response $g(t)$ be zero for negative time. By taking advantage of the fact that the noise and signal are Gaussian, and of the Wiener-Khintchine relation, Wiener solved the problem for the optimum linear filter.

Many modifications of his procedure[16,17] have appeared but, basically, the Wiener approach including the least-square-error criterion has not been altered fundamentally. Booton[18] has extended Wiener's theory to the case where both the signal $r(t)$ and the noise $n(t)$ are random but nonstationary. The difficulty with his result is that it resulted in no general analytic solution, although later investigators[19] obtained a solution which can be implemented only by the use of computers.

In the signal detection problem, many of the methods and criteria used by Wiener influenced the analysis of the problem of estimating the presence or nonpresence of a signal such as a pulse in the presence of random noise. A least-squares-error criterion was used by North[20] to maximize the predicted signal-to-noise ratio. Referring to Fig. 7, the impulse-response function of the predetection filter is denoted as $g(t)$. The problem is to decide whether a filtered output is due to noise alone $a_0(t)$, or due to signal plus noise $a_1(t)$. North solved this

[14] J. H. Laning and R. H. Battin, "Random Processes in Automatic Control," McGraw-Hill Book Co., Inc., New York, N. Y.; 1956.

[15] N. Wiener "The Extrapolation, Interpolation, and Smoothing of Stationary Time Series," Mass. Inst. Tech., Cambridge, Rept. of the Services 19, Res. Project DIC-6037; February, 1942. (Published by John Wiley and Sons, Inc., New York, N. Y.; 1949.)

[16] H. W. Bode and C. E. Shanon, "A simplified derivation of linear least square smoothing and prediction theory," Proc. IRE, vol. 38, pp. 417–425; April, 1950.

[17] L. A. Zadeh and J. R. Ragazzini, "An extension of Wiener's theory of prediction," *J. Appl. Phys.*, vol. 21, pp. 645–655; July, 1950.

[18] R. C. Booton, Jr., "An optimization theory for time-varying linear systems with nonstationary statistical inputs," Proc. IRE, vol. 40, pp. 977–981; August, 1952.

[19] K. S. Miller and L. A. Zadeh, "Solution of an integral equation occurring in the theories of prediction and detection," IRE Trans. on Information Theory, vol. IT-2, pp. 72–75; June, 1956.

[20] D. O. North, "Analysis of the Factors Which Determine Signal to Noise Discrimination in Radar," RCA Labs., Princeton, N. J., Rept. No. PIR-6C; June, 1943.

problem for white noise and later investigators[21–26] further enlarged on variations of this problem. The concept of autocorrelation and crosscorrelation function, its relation to the spectral density and the selection of a least-mean-square-error criterion have resulted in significant advances in the optimization of linear filters to which are applied Gaussian noise and/or signal.

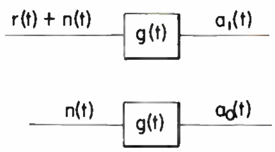


Fig. 7—Block diagram illustrating the decision to be made in an optimum detection problem.

[21] L. A. Zadeh, and J. R. Ragazzini, "Optimum filters for the detection of signals in noise," PROC. IRE, vol. 40, pp. 1223–1231; October, 1952.
[22] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypothesis," *Phil. Trans. Royal Soc. London*, vol. A231, pp. 289–338; 1933.
[23] J. L. Lawson and G. E. Uhlenbeck, "Threshold Signals," McGraw-Hill Book Co., Inc., New York, N. Y., ch. 7; 1950.
[24] M. Schwartz, "Statistical Approach to the Automatic Search Problem," Ph.D. dissertation, Harvard University, Cambridge, Mass., 1951.
[25] P. M. Woodward and I. L. Davies, "A theory of radar information," *Phil. Mag.*, vol. 41, series 7, pp. 1001–1031; October, 1950. See also PROC. IRE, vol. 39, pp. 1521–1524, December, 1951; and *J. IEE*, vol. 99, pp. 37–49; March, 1952.
[26] D. Middleton and D. Van Meter, "Detection and extraction of signals in noise from the point of view of statistical decision theory," *J. Soc. Industrial and Appl. Math.*, vol. 3, pp. 192–253; December, 1955; and vol. 4, pp. 86–119; June, 1956.

V. FUTURE TRENDS

As a result of intensive research in the past few decades, the field of noise and random processes is rapidly becoming mature. It is very unlikely that major contributions such as those made by Wiener are to be expected in the future. However, there are a few discernible trends which are noteworthy.

The understanding of the sources of noise in conductors, semiconductors and electron tubes is fairly well understood so that no important developments should be expected in these areas. On the other hand, in the field of optimization, developments on the concept, analysis and synthesis of adaptive systems which are responsive either to varying inputs or to environment are to be expected. The general subject of nonlinear optimum systems is still in a somewhat undeveloped state. The nonlinearity may already be present in an existing part of the system or may be deliberately introduced in the filter because the inputs are not Gaussian. Many special cases have been worked out, but results of a general nature are very few.

With the availability of fast computers of high capacity, it is expected that their employment as active in-line and off-line devices to implement optimum systems will increase. However, the underlying theory which postulates the employment of such computers in optimum systems is incomplete, and it is expected that research along these lines will be intensified.

Information Theory*

B. McMILLAN†, FELLOW, IRE AND D. SLEPIAN‡, FELLOW, IRE

***Summary*—During the last twenty years three theories dealing with the interaction of signals and noise in communication systems have come into being: detection theory; the statistical theory of filtering and prediction; and Shannon's information theory. They have developed rapidly and now play a key role in the communication engineer's understanding of his field. This paper presents a brief description of the central concepts of each of these theories, discusses their differences and common parts, and attempts to point out both their successes and shortcomings.**

* Received by the IRE, October 2, 1961; revised manuscript received, December 8, 1961.
† Assistant Secretary of the Air Force, Department of Defense, Washington, D. C.
‡ Bell Telephone Laboratories, Murray Hill, N. J.

I. INTRODUCTION

THE TERM "Information Theory" is used in the current technical literature with many different senses. Historically it seems first to have been generally applied to describe the specific mathematical model of communication systems developed in 1948 by Shannon [6].[1] In this pioneering paper, Shannon introduced a numerical measure, called by him and others entropy, of the randomness or uncertainty associated

[1] Indeed, Shannon himself in Ref. [6] uses the term (p. 20).

with a class of messages and showed that this quantity measures in a real sense the amount of communication facility needed to transmit with accuracy messages from the given class. He also showed (quite incidentally to his main argument) that this measure of uncertainty agreed *in certain aspects* with the common, vague intuitive notion of the "information content of a message." He accordingly used the words "information content" as a synonym for the precisely defined notion of entropy. As a result, his work and its immediate extensions became known as information theory. The other key concepts in Shannon's model, which will be discussed in detail later, are channel capacity and coding. Many workers today adhere to the following *Strict-Sense Definition: Information Theory is the study of the three concepts, entropy, channel capacity and coding, and of their interrelation either as abstract mathematics or in connection with their application to physical systems.*

As Shannon's work became widely known, a second more inclusive usage of the words came into being among electrical engineers. The exact boundaries of this wide-sense definition of Information Theory are as yet very fluid. For some engineers the words are used to describe all highly mathematical and theoretical investigations into the fundamental nature of the communication process. In this extended sense, information theory includes the study of filtering and prediction, signal detection, noise theory, signal representation, aspects of modulation theory and information processing systems, as well as the Shannon theory. The papers published in the IRE TRANSACTIONS ON INFORMATION THEORY serve as a guide to the subject matter covered by this definition. Other engineers, however, would restrict the usage of the words to those studies of the fundamentals of communications that involve probability theory and statistical notions so that, for example, signal representation and much of modulation theory would be stricken from the above list. These two versions of a wide-sense Information Theory sometimes pass under the names of Communication Theory and Statistical Communication Theory, respectively.

In recent years, particularly in Europe, a third even more extended meaning of the words Information Theory has become common. The term is used to describe any field or study in which "information" in any of its many shades of common meaning plays a role. In this widest sense, Information Theory covers a wide spectrum of disciplines. It includes parts of linguistics, cybernetics, physics, psychology, machine translation, statistics, pattern recognition, semantics, cryptography, data storage and retrieval, etc., as well as the subjects covered by the more restrictive definitions.

In this paper, we shall outline some of the more significant developments of the area referred to above as statistical communication theory. The subject can be divided into three parts: the theory of detection and signal extraction; the theory of filtering and prediction; and the Shannon model of communication systems. We hope to give the reader a general picture of each of these fields and to point out their differences and common parts. Limitations of space unfortunately restrict us to the most cursory of treatments: many important topics are slighted; credits for contributions are largely omitted. The few works cited in the Bibliography are only intended to give the reader some points of entry into the vast literature.

II. General Considerations in Statistical Communication Theory

There are two quite distinct fundamental reasons why probability theory plays such an important role in the description and understanding of communication systems. Both are associated with the engineer's lack of knowledge. The first arises from the fact that the designer of a communication system cannot know in full, microscopic detail the exact behavior, either of the parts of the system he is designing, or of the environment in which he will ultimately place it. At best, he can predict with reasonable accuracy only the average behavior of the system. Its departure from this average, he can only describe in probability terms. These fluctuations he calls "noise." Their importance as a limiting factor in the communication process was recognized at an early date.

The second reason for the appearance of probability theory stems from the fact that the designer cannot know in full detail the *exact* message for which he is designing the system. Indeed, if he knew the message beforehand, there would be little point in constructing the system. He can know, at most, only the class of messages from which the input to his system will be selected and the relative probabilities of selection of messages from that class. *It is this lack of detailed knowledge about the message that is the very reason for communication; it is lack of detailed knowledge of the communication system that limits the attainment of communication.*

As trivial as the point of the preceding paragraph may seem in hindsight, it was a long time in coming. Although hinted at many times earlier in the history of communications, it was not until the early forties, notably in the works of Wiener [11], that the importance and utility of regarding messages as stochastic in nature was clearly recognized. It was not until the work of Shannon that it was properly appreciated that *a priori uncertainty is the very commodity of communication.* The remarkable theoretical developments of the past twenty years are direct consequences of the emergence of this single idea.

Granted the foregoing description, modern communication theory can properly be considered to be an applied branch of probability theory. Its tools are those of the probabilist and statistician. Many of the tools were ready in the early forties, for probability theory, notably in the area of stochastic processes, was then undergoing a vigorous renaissance. (Perhaps this was the *reason* for the new look in communications.) Others developed later: the feedback of communication theory to probability theory is clearly discernible.

The probabilistic description and study of messages (=signals) alone or of noise alone falls in the province of noise theory which is treated in the article "Noise and Random Processes" by Ragazzini and Chang appearing in this issue on page 1146. The study of the *interaction* of signals and noise is the domain of statistical communication theory. Three quite distinct bodies of doctrine which answer this description are treated in the next sections of this paper. The first two, detection theory and filtering theory, treat specific tasks performed in certain parts of some communication systems. There is considerable overlap between these theories. Their classification as separate subjects, rather than as special cases of a larger theory, is now firmly entrenched, but appears to be for historical rather than logical reasons. The third subject, strict-sense information theory, treats the complete one-way communication system as a whole.

III. Detection Theory and Signal Extraction

An observer is presented with a number x known to be drawn from either of two populations—one P_0 is governed by a probability density $p_0(y)$; the other P_1 is governed by $p_1(y)$. Both functions, $p_0(y)$ and $p_1(y)$, are known to him. He must attempt to guess from which population his sample x was drawn. This situation, in its many generalizations and ramifications, forms the content of the subject that satisticians know as hypothesis testing. In communication work, it forms the core of detection theory. Here x (suitably generalized) repsents a received waveform, $p_0(y)$ describes a population of noise alone and $p_1(y)$ describes a population of signal and noise. The observer must hazard a guess to settle the quandary, "Is this wave-form noise alone, or is signal also present?"

Any rule that the observer arrives at for guessing can be described as a separation of all possible x values into two sets: the set for which he asserts P_0 and the set for which he asserts P_1. Following his rule, he can possibly make two kinds of errors: he may sometimes assert that the sample came from P_0 when it came from P_1; he may sometimes assert that the sample came from P_1 when indeed it came from P_0. How should he use his knowledge of $p_0(y)$ and $p_1(y)$ to make a good rule? The answer clearly depends on the relative cost to him of the two types of error just mentioned. It also depends on whether or not he has additional information, possibly in the form of *a priori* probabilities that the sample be drawn from P_0 or P_1. Finally, if the observer is allowed to take further samples from the unknown population, his rule for guessing may depend on the cost to him of making an additional observation.

The situation just described has been studied in considerable detail by statisticians under a wide variety of assumptions. However, the application of their results to the problems of the communication engineer is not in general a straightforward matter of substituting into existing formulae. The statistician has for the most part considered the sample x to be a single number, or at most a finite set of numbers. In many engineering problems, the observed sample, a waveform, is a continuum of numbers. Nontrivial limiting processes must frequently be investigated to extend the statistician's results to the engineer's needs. Then too the classical formulation appears deceptively easy: often relatively simple sounding engineering problems result in exceedingly complicated analysis, and one must be satisfied with approximate answers to the problems posed.

Technical details of the many engineering problems that have been treated by this extended theory of hypothesis testing are too involved for inclusion here. We mention only a few key concepts and results that have emerged. 1) Under a surprisingly large class of different assumptions regarding both *a priori* knowledge and the method of measuring cost of faulty decision, the form of the observer's best rule is the same. For his given sample x, he evaluates the "likelihood function" $l(x) \equiv p_0(x)/p_1(x)$. If this exceeds a critical number, k, called the threshold, he decides that x came from P_0; otherwise that it came from P_1. Differences in *a priori* assumptions and cost criteria only serve to alter the proper choice of k. 2) Many problems for which general solutions are not feasible in tractable form admit simple useful instrumentable results in limiting cases of very small or very large signal-to-noise ratios. 3) In the case of additive signals and noise, the assumption of very wide-band white noise yields great simplifications. Detectors, derivable by other means, emerge as optimum. Correlation detectors and matched filters play a key role here.

Closely associated with hypothesis testing is the theory of parameter estimation, or in communication language, the theory of signal extraction. In the statistical setting, an observer is given a sample x drawn from a population governed by probability density $p(x, \alpha)$ depending on a parameter α. The density $p(x, \alpha)$ is known to the observer as a function of x and α. He does not, however, know the value of α that maintained when his observation was drawn. He is to estimate this value from his observation x. To do this he forms an estimator function $\hat{\alpha}(x)$: evaluated at his observed sample x, $\hat{\alpha}$ is the observer's estimate of α. In communication applications, x is replaced by a received waveform consisting of a signal contaminated by noise. The observer knows the statistical structure of the noise and the analytic form of the signal. He does not know the precise value of certain parameters which enter into the description of the signal (*e.g.*, a shift of time origin or change in magnitude for a returning radar pulse). From this knowledge he can deduce a probability measure, the analog of $p(x, \alpha)$ just discussed, which describes the statistical structure of the possible received waveforms. This probability measure will depend on the literal parameters entering into the description of the signal. From examination of a particular received waveform, the observer wishes to deduce the probability measure describing the population from which it was drawn: that is, he seeks to determine the values of the signal parameters. As in hypothesis testing, application of the statistician's model to the communication engineer's needs calls for generalizing x from a number, or finite collection of numbers, to a function, or continuum of numbers.

If α can assume only one of two values, the situation just described is identical with that of hypothesis testing already discussed. If, as is usually the case, α can take a continuum of values, the situation is sufficiently different to warrant separate names and treatment for hypothesis testing and parameter estimation. In particular, the two types of error which play a key role in certain parts of hypothesis testing have no simple counterparts in parameter estimation.

What constitutes a good estimator $\hat{\alpha}(x)$ depends on how one measures the cost of guessing a faulty value for the parameter. One approach is to require that on the average $\hat{\alpha}(x)$ should yield the true value α of the parameter. Among such estimators, one of minimum variance is sought; that is, one seeks to minimize the average squared error of estimate. Such estimators do not always exist, however. Frequently the sample x is a sequence of numbers representing repeated independent samples from the population. In this case another weaker condition may be appropriate: one can ask merely that as the number N of samples increases the estimator $\hat{\alpha}_N(x)$ asymptotically have the unbiased and minimum variance properties just described. These and many other measures of effectiveness of estimators have been studied and applied to communication problems.

Detection and extraction theory have had many successes to date. Their earliest applications were to radar and various military systems. They are currently playing an important part in radio and radar astronomy and beyond-the-horizon radio communication systems. They can be expected to contribute significantly to problems of satellite communication and tracking, seismic analysis, and other systems operating with small signal-to-noise ratios.

The theories also have their limitations. We have already mentioned the difficulty of carrying out to completion solutions indicated theoretically. More fundamental are troubles of a conceptual sort found in specifying basic data required by the theory. The theories make use of *a priori* knowledge of signal and noise statistics to provide decisions which minimize some measure of the cost of faulty decisions. All too often in practical problems the *a priori* probabilities are poorly defined and the intelligent assignment of comparative costs for errors is difficult to make. Perhaps these difficulties should not be charged to the theory itself. Nevertheless, they limit its range of useful applicability.

IV. Filtering and Prediction Theory

The situation studied in what is commonly known as filtering and prediction theory is as follows. A received waveform $r(t)$ is available to an observer. It is known to be the *sum* of two parts, signal and noise: $r(t)=s(t)+n(t)$. The observer wishes to measure some attribute of $s(t)$ by performing a linear operation on $r(t)$. How should he choose the linear operation to best achieve this aim?

The problem can be regarded as a special case of estimation theory. The specialization lies in the assumed additivity of signal and noise and the restriction of the observer to linear operations. Historically this problem was treated in great detail [11] before the statistical theory of parameter estimation was carried over as a whole to communication problems. As a consequence it is generally regarded as a separate subject today.

The nature of the solution depends, of course, in the first place on what *a priori* knowledge the observer has about $s(t)$ and $n(t)$. If he has none, the problem is not very meaningful; if he knows either $s(t)$ or $n(t)$ exactly, the problem is trivial. A wide variety of interesting intermediate cases suggest themselves, and most have been studied. Typical are: 1) s and n are both continuous stochastic processes with known statistics; 2) n is stochastic and s is a polynomial of known degree but with unknown coefficients. The interval during which $r(t)$ may be observed can be finite or semi-infinite.

The nature of the solution also depends on what attribute of s is to be estimated as well as on the criterion by which errors of measurement are judged. As for the former, typical quantities to be estimated are: 1) s at some time instant past, present, or future, 2) the time derivative of s, or 3), in the case of a nonstochastic signal, some parameter entering the description of the signal. The error criterion most commonly considered is the mean squared deviation of the estimate from the true value of the quantity sought. This criterion is by far the simplest to deal with mathematically and in many situations is not unreasonable physically. For its computation, only second-order statistics of s and n are required and hence the linear operation (or filter) that yields the minimum mean squared error can be determined from the spectra and cross spectra of the signal and noise; these latter quantities are often available from engineering considerations. Some partial results are known for other error criteria.

Finally, the nature of the solution depends on restrictions that may be imposed on the class of linear operations within which a solution is to be found. In the case commonly treated, in which signal and noise are stationary, the filter is required to be physically realizable in the sense of circuit theory. It is the nature of the restrictions put upon the class of allowed solutions that to a large degree determines the mathematical difficulty encountered in finding a solution. In almost all problems of the class under discussion, the assumption of power spectra, for noise or signal plus noise, which are rational functions of frequency leads to significant simplification in the analytic difficulties encountered.

Experience has shown that in the neighborhood of the optimum filter, the mean square error is quite insensitive to (intelligent) changes in the filter design. It is true, then, that in practice very few filters are built following precisely the prescription of the theory: small changes can frequently effect considerable economy in equipment. Nevertheless, the theory has been an effective and valuable guide in the design of many practical systems and will undoubtedly continue to be of great value in the future.

V. Information Theory—Strict Sense

The mathematical models of the communication process sketched in the preceding paragraphs have tacitly fixed attention on what might be called the *problem of reception:* given a statistical, or parametric description of the class of signals, and of the noise, how to infer something desired about the original signal from a sample of noisy signal. The solution to a given problem of this kind describes, at least in mathematical terms, an optimum receiver for the particular classes of signals and noise involved. No attention is given in this model to the fact that often the engineer also has control, *e.g.*, by a choice of modulation scheme, over the class of signals for which his receiver is to be optimized. True, the model permits repeated optimization of the receiver for various classes of transmitted signals, so that one could in principle search for an optimum-optimum combination, but the theory offers no general guidance for this search nor does it give any description of the result of a successful search.

Shannon's theory, information theory in the strict sense, attacks the general problem of simultaneous design of transmitter and receiver for optimum communication of a given class of messages over a given medium. Furthermore the theory solves this problem and gives a quantitative description of the performance of the optimum combination. These are sweeping claims for a theory; our essay must now show both the scope and limitations of the model within which the theory operates.

Intrinsic to the strict-sense theory is a restriction to discrete or telegraph-like signals. Generalizations to signals taking a continuum of values have never successfully concealed this restriction; they have all made use of "quantizing," or reduction to a discrete situation. Shannon's model of an *information source* then begins with a discrete alphabet M of symbols, say $M_1, M_2, \cdots, M_K$. A simple case is that of telegraphy in which $K=2$, $M_1=0=$ space, $M_2=1=$ mark; this is usually called a binary source. At another extreme, $\log_2 K \doteq 40{,}000$ and M consists of all of the conceivable English paragraphs of fewer than 10,000 words.

An information source is just a stationary stochastic process whose sample sequences are strings of symbols out of the alphabet M. More mechanistically, an information source produces strings of text $\cdots m_{-1}, m_0, m_1, \cdots$ whose statistical properties are known. Here each m_i is drawn from the alphabet M, and one supposes that he can calculate all such probabilities as

$$\text{Prob}\left\{m_{-9} = M_{31} \text{ and } m_{-8} = M_9 \text{ and } \cdots m_{1057} = M_{93}\right\}.$$

Furthermore (stationarity), one requires that, *e.g.*, the probability above be equal to any one of the following:

$$\text{Prob}\left\{m_{-9+k} = M_{31} \text{ and } m_{-8+k} = M_9 \text{ and } \cdots \text{ and } m_{1057+k} = M_{93}\right\},$$

for $k=0, \pm 1, \pm 2, \cdots$. A time-shift does not change the statistics of the process.

The simplest kind of information source is one in which each M_i is drawn with probability p_i, $1 \leq i \leq K$, and successive choices are independent. The probabilities above would in this case all be the product $p_{31} \cdot p_9 \cdots p_{93}$. Such a source is called an independent source. The option to let M itself consist of very long sequences of text makes it possible to base the theory almost wholly on independent sources. In fact the known theory scarcely goes beyond this point, its greater generality being more in appearance than in fact.

Shannon shows that, when N is a large integer, long strings $m_1 m_2 \cdots m_N$ of text from the source have great statistical regularity. *E.g.*, for a binary source in which $p_1=0.99$ and $p_2=0.01$, it is highly probable that the proportion of spaces in a long string is very nearly 0.99. This numerically minute fraction of the possible long strings therefore exhausts most of the likely possibilities. In general, there is a number H, determined by the source, with the property that, given a large N, one can find approximately 2^{NH} strings $m_1 \cdots m_N$ which collectively exhaust almost all of the probability. More exactly, given two arbitrary small numbers $\epsilon > 0$, $\delta > 0$, there is a typically large $N_0(\epsilon, \delta)$ such that if $N \geq N_0$ then there is a set S of strings $m_1 \cdots m_N$ containing between 2^{NH} and $2^{N(H+\delta)}$ strings, such that the total probability of S exceeds $1-\epsilon$. We shall call S a set of "likely" strings of length N.

H is the *entropy* of the source. For an independent source, it can be shown that

$$H = \sum_{i=1}^{K} p_i \log_2 \frac{1}{p_i} \cdot$$

Always $H \leq \log_2 K$, and equality holds only when the source is independent and each $p_i = K^{-1}$.

There are exactly $2^{N(H+\delta)}$ distinct strings of binary symbols of length $N(H+\delta)$. One sees from the definition above therefore that there are enough such strings that one can establish a one-one correspondence, or code, between the likely source strings of length N and strings of binary symbols of length $N(H+\delta)$. Furthermore, one cannot succeed in this with fewer than 2^{NH} binary strings. Hence between H and $H+\delta$ binary symbols per symbol of text must be used for a one-one encoding of the likely source strings into strings of binary symbols. An alternative definition of H can be obtained from this fact. In any case, we see H as the minimum number of symbols of "standard" binary text needed, on the average, to represent one symbol of actual text. The relevance of such a measure of information rate to the design of commercial communication circuits should be clear.

Implicit in the concept of entropy or information rate is the notion of modulation—encoding the original text into symbols more suitable for transmission. H appears as the minimum amount of binary telegraphic capacity needed per unit of original text. Departures from this minimum may be needed to combat noise, as the theory will show.

The notion of a *noisy channel* requires two more discrete alphabets, say $A=(A_1, A_2, \cdots, A_F)$, an alphabet exhausting the possible or at least allowable input waveforms of unit duration, and $X=(X_1, \cdots, X_G)$, an alphabet of output waveforms. For realism, $G \geq F$, and typically $G \gg F$.

We think of each A_i and X_j as occupying a unit of time so that the channel takes as inputs the strings $\cdots \alpha_{-1}, \alpha_0, \alpha_1, \cdots$, and produces as outputs the strings $\cdots x_{-1}, x_0, x_1, \cdots$, where each α_i is from A and each x_i is from X. Noise on the channel is represented by the fact that, even when the whole input sequence $\cdots \alpha_{-1}, \alpha_0, \alpha_1, \cdots$ is given, the output sequence $\cdots x_{-1}, x_0, x_1, \cdots$ is not unique, and can only be described as a stochastic process—*i.e.*, in terms of probabilities like Prob $\{x_{-100}=X_{29}$ and $\cdots$ etc.$\}$. These probabilities depend, of course, on the particular input sequence assumed.

For realism, one must require of a channel that the statistics, say, of sequences $x_1, \cdots, x_L$ do not depend on the $\alpha_{L+1}, \alpha_{L+2}, \cdots$; *i.e.*, prediction is disallowed. For the theorems of the theory to be valid one must also require that the statistics of $x_1, \cdots, x_L$ do not depend in any material way upon the remote past—*e.g.*, upon $\alpha_{-R}, \alpha_{-R-1}, \cdots$ for large R; similarly, significant dependence on the remote future must be ruled out, for validity even if not for realism. Practically, these restrictions reduce the theory almost to considering only the *channel without memory:* this channel is described by a family $p(x, \alpha)$ of distributions already familiar to us: $p(x, \alpha)$ is the probability that x will be received when α is transmitted, x being drawn from the alphabet X, α from A. For this channel, the probability $p(x_1 \cdots x_N, \alpha_1 \cdots \alpha_N)$ that $x_1 \cdots x_N$ will be received when $\alpha_1 \cdots \alpha_N$ are transmitted is *by definition*

$$p(x_1 \cdots x_N, \alpha_1 \cdots \alpha_N) = p(x_1, \alpha_1)p(x_2, \alpha_2) \cdots p(x_N, \alpha_N),$$

independently of α_k for $k \leq 0$ and $k \geq N+1$.

The key to the strict-sense theory is the fact that the problem of estimation or reception becomes almost trivial when the integer parameter N which appeared above is large enough, *provided the engineer makes effective use of his option to select suitable transmitted sequences and in particular selects few enough of them.*

For channels without memory, and for certain simple generalizations of them, there exists a number C, called the *capacity* of the channel, whose properties are crucial. To describe these properties easily, we will introduce an auxiliary term: two input sequences $\alpha_1' \cdots \alpha_N'$ and $\alpha_1'' \cdots \alpha_N''$ will be called distinguishable at the level ϵ, where $\epsilon > 0$, if there is a set S_1 of sequences $x_1 \cdots x_N$ whose total probability under the distribution $p(x_1 \cdots x_N, \alpha_1' \cdots \alpha_N')$ exceeds $1-\epsilon$, and a set S_2 whose total probability under $p(x_1 \cdots x_N, \alpha_1'' \cdots \alpha_N'')$ exceeds $1-\epsilon$, such that S_1 and S_2 contain no sequences in common. This makes precise the idea that $\alpha_1' \cdots \alpha_N'$ and $\alpha_1'' \cdots \alpha_N''$ could with high probability be distinguished from each other at the output *provided no other inputs were allowed.*

A set of more than two input sequences will be called distinguishable at the level ϵ if each pair of them is so distinguishable. By restricting the channel inputs to sequences $\alpha_1 \cdots \alpha_N$ drawn from a set distinguishable at the level ϵ, the engineer simplifies almost to triviality the problem of "estimating" $\alpha_1 \cdots \alpha_N$ from $x_1 \cdots x_N$, and achieves an estimation, *i.e.*, communication, which is *exact* (without *any* error) with probability at least $1-\epsilon$.

The capacity C of a channel tells the engineer how many distinguishable sequences he can expect to find. Loosely, there are approximately 2^{NC} such sequences of length N. More exactly, given $\epsilon > 0$ and $\delta > 0$, there is an $N_0(\epsilon, \delta)$ such that, if $N \geq N_0$, there are at least $2^{N(C-\delta)}$ and no more than 2^{NC} sequences $\alpha_1 \cdots \alpha_N$ distinguishable at the level ϵ. The capacity C can be computed from the probabilities $p(x, \alpha)$.

Feinstein [2] gives a constructive process for finding approximately 2^{NC} distinguishable sequences. His method is of considerable theoretical importance, but unfortunately is too lengthy to be of use in the design of practical communication systems. The search for simple systematic methods to generate maximal sets of distinguishable sequences is called the coding problem. This problem has been studied at length for certain simple channels, and for moderate values of N some solutions [4], even practicable ones, have been found. It is the authors' impression that, even though the known solutions apply to highly idealized channels, their form has begun to guide the design of telegraph-like communication systems.

Given a source of entropy H per symbol, and a channel of capacity C per unit time, one can, with appropriate approximations as suggested by the discussion above, find a large L and a large N such that there are 2^{LH} likely sequences of text of length L from the source and 2^{NC} distinguishable sequences of length N available on the channel, and such that $2^{LH} \leq 2^{NC}$. Then one can imagine a code which establishes a uniquely determined and characteristic distinguishable channel input sequence $\alpha_1 \cdots \alpha_N$ for each likely sequence $m_1 \cdots m_L$ of text. This code describes the "modulator;" with the inverse "demodulator" or decoder at the receiver, it completes the "construction" of a communication system which averages H/C units of channel time to transmit one symbol of text from the source. Furthermore, it is only a rare sequence of text, or a rare event on the channel, which prevents *exact* recovery at the receiver of the text offered for transmission. The search for systematic rather than merely *ad hoc* "modulators" is sometimes included as part of the coding problem.

This completes our rather detailed and, hopefully, explanatory account of the elements of information theory in the strict sense. Our intent has been to show both its scope and its restrictions. In sum, it is larger in scope than the other theories here discussed because it includes explicitly the design of the "modulator" as well as of the "receiver." Nevertheless, it is restricted in its point of view to an idealization of the problem seen by the designer of a commercial communication service: performance is evaluated in terms of *channel time per unit of offered text*, and essentially exact transmission is the goal—there are no degrees or shades of "wrongness," all errors are equally serious, and it is the probability of error that one seeks to control.

We did not mention as restrictive the fact that the strict-sense theory as here described requires quantized signals and a quantized time parameter. Both of these elements of quantization are apparent in any extant version of the theory, and, practically, their presence is not really restrictive.

The authors tend to the belief that discreteness is intrinsic to the theory and cannot be eliminated. Attempts to conceal it are more for mathematical elegance than for increased scope.

One application of the theory to an apparently nondiscrete situation has become famous. Shannon calculates the capacity C of the band-limited channel with additive white Gaussian noise to be

$$C = W \log_2 \left(1 + \frac{S}{N}\right),$$

where W is the bandwidth in cycles per second, S is the average power of the transmitter in watts and N the average noise power in the band in watts. Specialized and idealized as this formula may be, it alone gives great insight into the quantitative factors that govern communication in the presence of additive noise.

The generality of the strict-sense theory is regarded by some as a weakness—it does not give specific hardware answers. This is, of course, because the model is not restricted to specific situations. In fact, specific situations can be treated within the framework of the model. The application then usually calls for computations of enormous complexity; in this the strict-sense theory is not unique.

The theory will probably always remain, like thermodynamics, conceptual and guiding rather than definitive. This is not because its calculations are difficult, but for a more basic reason: the theory presupposes something unattainable, namely, an exact knowledge in detail of the statistics of the source and of the channel. In this fact, again, the strict-sense theory is like most of the others here discussed.

Just because the theory has been largely guiding, it is difficult to point to specific accomplishments. The theory makes precise the idea that to combat noise one must use for transmission only a small fraction of the possible signals that the medium will support; nevertheless wide-band FM was invented, and even understood, long before the theory appeared. Similarly, the theory points strongly to quantization; nevertheless regenerative telegraph repeaters, and at least one independent invention of PCM, predate Shannon's papers. The theory also easily shows that (quantized) pulse-position modulation is as good as one can use for communicating at low ratios of signal to additive noise. Nevertheless, radar was invented almost two decades before this result. The least that one can say for the theory is that it has unified these, and many other, diverse appearing facts into a common conceptual framework. Perhaps the most that one can hope is that this framework will someday be enlarged to include without strain those problems, like that of radar, in which channel time per unit of text is not the quantity of primary or natural interest.

Bibliography

Strict-Sense Information Theory

[1] R. M. Fano, "Transmission of Information," John Wiley and Sons, Inc., New York, N. Y.; 1961.
[2] A. Feinstein, "Foundations of Information Theory," McGraw-Hill Book Co., Inc., New York, N. Y.; 1957.
[3] A. I. Khinchin, "Mathematical Foundations of Information Theory," Dover Publications, Inc., New York, N. Y.; 1957.
[4] W. W. Peterson, "Error-Correcting Codes," John Wiley and Sons, Inc., New York, N. Y.; 1961.
[5] F. M. Reza, "An Introduction to Information Theory," McGraw-Hill Book Co., Inc., New York, N. Y.; 1961.
[6] C. E. Shannon and W. Weaver, "The Mathematical Theory of Communication," University of Illinois Press, Urbana; 1949.

Detection, Signal Extraction, Filtering and Prediction

[7] W. B. Davenport, Jr. and W. L. Root, "Random Signals and Noise," McGraw-Hill Book Co., Inc., New York, N. Y.; 1958.
[8] C. W. Helstrom, "Statistical Theory of Signal Detection," Pergamon Press, New York, N. Y.; 1960.
[9] Y. W. Lee, "Statistical Theory of Communication," John Wiley and Sons, Inc., New York, N. Y.; 1960.
[10] D. Middleton, "Statistical Communication Theory," McGraw-Hill Book Co., Inc., New York, N. Y.; 1960.
[11] N. Wiener, "Extrapolation, Interpolation, and Smoothing of Stationary Time Series," John Wiley and Sons, Inc., New York, N. Y.; 1949.

Wide-Sense Information Theory

[12] J. R. Pierce, "Symbols, Signals and Noise," Harper and Brothers, New York, N. Y.; 1961.

Section 18
INSTRUMENTATION

Organized with the assistance of the IRE Professional Group on Instrumentation

Frequency and Time Standards*

L. ESSEN†

Summary—**The gradual development of quartz and atomic clocks has resulted in an improvement in the accuracy of frequency measurement by a factor of at least 10^4 in the past 30 years. The best quartz clocks drift upwards in frequency at an average rate of about 1 part in 10^9 per month and operate with a day to day stability of a few parts in 10^{11}.**

A caesium standard based on the magnetic deflexion of an atomic beam has been in operation since 1955 and recent models of this type of standard enable frequency and time to be defined in terms of atomic constants with an accuracy of a few parts in 10^{11}. It has revealed irregular and seasonal variations in the length of the day of about 1 msec.

The techniques of optical pumping and detection, of the use of buffer gases, and of wall coatings which enable atoms to be reflected without undergoing transitions have made it possible to make working atomic standards which are already of great stability and may well be capable of still further improvement.

I. Introduction

No branch of physical measurement has made more rapid progress in recent years than that of the measurement of time and frequency, and this advance has depended very largely on developments in the field of radio engineering. This review overlaps in time that written in 1955 by Lewis [1] but deals more particularly with the subsequent developments.

Thirty years ago at the National Physical Laboratory the standard of frequency was a 1000 cps tuning fork, which was maintained in vibration in a valve circuit. The fork circuit drove a phonic motor at 10 revolutions per second, and a reduction gear wheel produced electrical impulses at intervals of 1 second. These seconds intervals were compared on a chronograph with seconds marks obtained from a standard pendulum clock and in this way the frequency of the fork was measured with an accuracy of about 1 part in 10^5. The rate of the pendulum clock was measured in terms of the mean solar second by daily comparisons with observatory time signals. Today the standard is a quartz clock which operates with a day to day stability within 1 part in 10^{10} and which is calibrated whenever necessary by means of a caesium atomic beam standard with a precision of a few parts in 10^{11}. The quartz clocks are used to integrate atomic time and one of the applications of this atomic time scale has been to measure variations in the rate of rotation of the earth.

This development illustrates the fact that there is no essential difference between a standard of time and a standard of frequency. The passage of time is measured by counting the repetitions of some regular event, and whether the rotations of the earth, the vibrations of a piece of quartz or the waves emitted by an atom are used is simply a matter of convenience and of the precision required. When the regular events occur many times in a second it is more convenient to speak of the number per second or their frequency rather than the time between each event. The suggestion which is sometimes made that a unit of frequency should be defined independently of the unit of time is thus based on a misunderstanding. A unit of frequency is also a unit of time and if the sug-

* Received by the IRE, July 10, 1961; revised manuscript received, October 19, 1961.
† National Physical Laboratory, Teddington, England.

gestion were adopted there would be two units of the same quantity.

There is, however, one distinction which should be made. The axial rotation of the earth not only provides a standard of time interval but it also determines the time of day and the word "time" is in consequence used in both senses. Now that there is an alternative standard it is becoming important to distinguish between time interval and time because whereas the atomic clock can provide the former, the time of day can only be determined by astronomical measurements of the position of a point on the earth's surface relative to the sun.

II. The Unit of Time Interval and Frequency

It may be useful here to give the present definition of the unit of time interval and to explain how it is linked with the atomic unit, which is essential for the most precise measurements. The definition was adopted in 1956 because the variations in the length of the day, *i.e.*, in mean solar time, were causing increasing difficulties in a number of fields of work including the tabulation of the positions of the bodies of the solar system. The new unit, the second of Ephemeris Time (ET), is based on the revolution of the earth around the sun and is defined as the fraction 1/31 556, 925.9747 of the tropical year for 1900 January 0 at 12h Ephemeris Time [2], [3]. The value is chosen so as to equal the average value of the mean solar second over the period of about 200 years covered by Newcomb's Tables, and it is referred to a particular epoch because it is subject to a slow secular change.

The laws of motion of the planets are determined from observations extending over the same period of 200 years, and their positions are tabulated with this unit of time as the independent variable.

If the positions are measured in terms of mean solar time [also called Universal Time (UT)], a comparison of the measured and tabulated values gives the difference UT−ET. The correction UT−ET can be determined only in retrospect as it must be averaged over a considerable number of years to obtain the required accuracy.

The value of the caesium frequency in terms of ET determined [4] from measurements extending over three years is 9, 192, 631, 770 $\pm$ 20 cps and the atomic clock can therefore be used to make the second of ephemeris time immediately available, with an accuracy of $\pm$20 cycles or $\pm$22 parts in 10^{10}.

The atomic clock can be used with a precision of 1 part in 10^{10} and the value 9, 192, 631, 770 cycles has been used provisionally as a unit of atomic time, although this may not be the value finally chosen. From the point of view of the physicist the actual value chosen is immaterial (so long as it is stated) as the unit is a fundamental atomic constant. The value must be chosen so as to link atomic and astronomical time in a manner most convenient to the many users, whose requirements are not necessarily the same.

III. Quartz Clocks

The tuning fork was developed to a point at which it gave a stability of 1 part in 10^7 per week [5] and could have been improved still further. By this time however, the first quartz clock had been made by Horton and Marrison [6], and it seemed clear that quartz possessed many advantages. One fundamental advantage was the higher frequency of quartz vibrations. Frequencies of many millions of cycles per second were already being used for radio transmissions, and it was not very convenient to measure them in terms of a standard having such a low value as 1 kc. The problems of manufacturing quartz oscillators on a commercial scale for the control of transmitting stations and other radio applications have been solved step by step and high quality oscillators are now readily available for any frequency below 10 Mc. At present, effort is being devoted to making them sufficiently rugged to withstand rocket flight while still preserving a stability to 1 part in 10^9 per day [7]. The work on commercial oscillators though remarkably successful did not have as much effect on quartz standards of the highest precision as might have been expected. The 100-kc GT plate was, however, fairly widely used for this purpose and selected samples of the 1-Mc and 2.5-Mc plates [8] (Fig. 1) are probably as stable as any in use today. The drift rate of the 2.5-Mc oscillators ranges from 1 to 10 parts in 10^{10} per month. The oscillators in use at the Bell Telephone Laboratories have shown drift rates of about 6 parts in 10^9 in 2 years. Short time variations are of the order of 1 part in 10^{10} to 1 part in 10^{11} depending upon the method used for measurement. However, two standards designed quite early in the history of the development of quartz clocks and made in small numbers "by hand," have continued to give comparable performance. These are the 60- and 66.$\dot{6}$-kc bars developed at the Physikalisch Technische Bundesanstalt in Germany [9] and the 100-kc rings developed at the National Physical Laboratory in England [10]. One of the main practical problems of the lower frequency crystals is to hold the quartz firmly without damping its oscillations, and an overtone vibration having nodal points is therefore chosen. It is doubtful, however, whether there are any true nodes in a quartz vibration and it is always necessary to allow a certain amount of resilience in the support. The rings as first made rested on three points located in radial grooves cut in the lower face of the ring as shown in Fig. 2, but later versions now made at the British Post Office are supported, as explained by Lewis [1], by nylon threads. Threads are also used to support the bars and the latest form adopted is shown in Fig. 3. Fig. 4 shows the bar as mounted.

Both of these standards are operated at the temperature at which the temperature coefficient is zero and the other conditions of operation are all carefully controlled. The standards increase in frequency at rates which vary between a few parts in 10^{10} to a few parts in 10^9 per month. The results for a number of the rings operated at

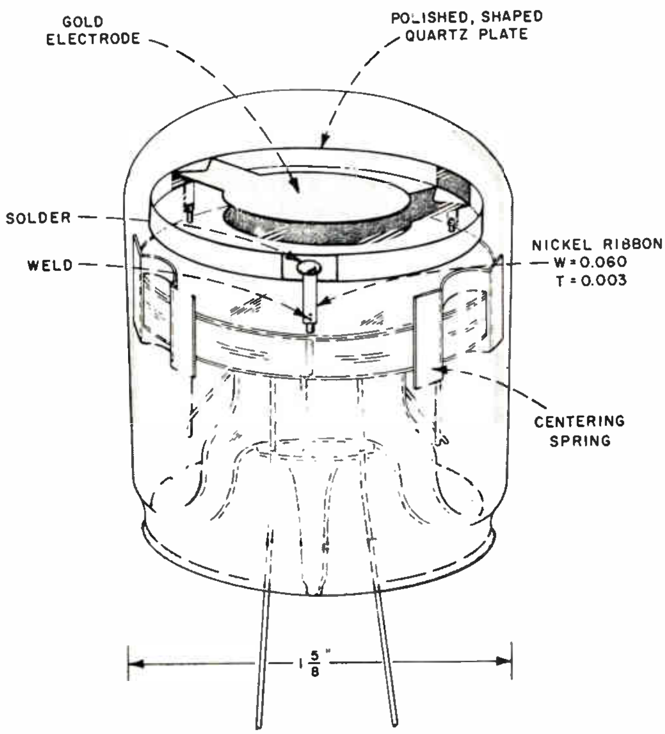


Fig. 1—Ribbon mounted 2.5-Mc quartz crystal unit developed by Bell Telephone Labs.

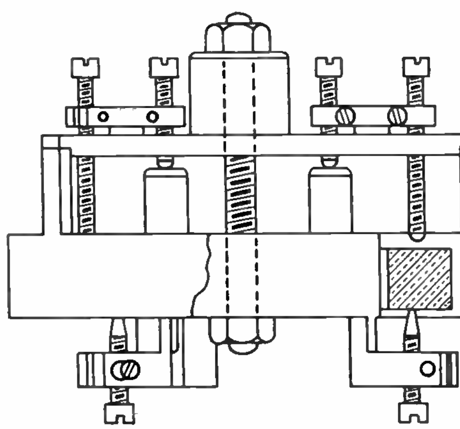


Fig. 2—Early three point mounting used on quartz rings made at the National Physical Lab. in England.

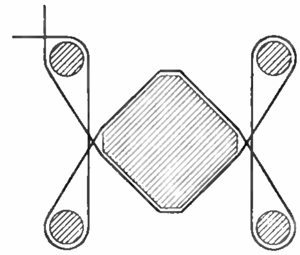


Fig. 3—Nylon thread mounting used on quartz bars of the Physikalisch Technische Bundesanstalt in Germany.

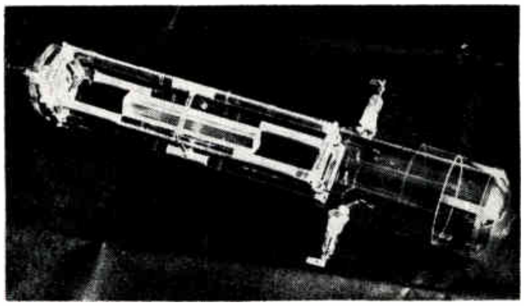


Fig. 4—Mounted quartz bar of the Physikalisch Technische Bundesanstalt.

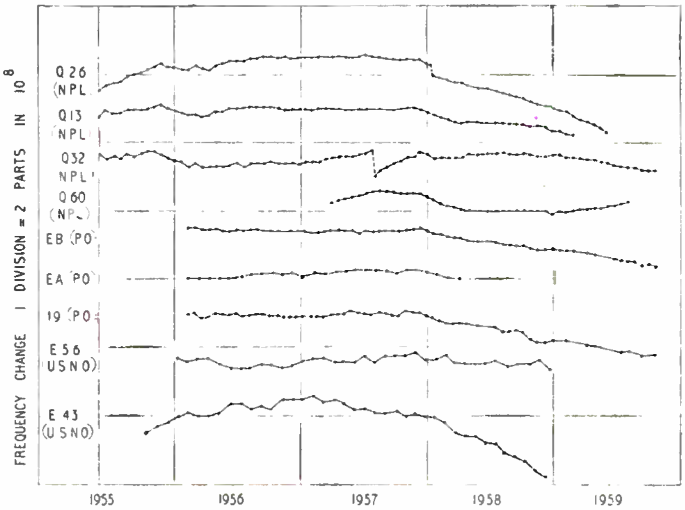


Fig. 5—Frequency stability of various quartz ring oscillators (linear frequency drifts have been removed).

the NPL, the British Post Office and the U. S. Naval Observatory are shown in Fig. 5 after linear drifts have been removed [11]. The somewhat erratic behaviour of the NPL standards at the ends of 1955 and 1957 may be due to failures in the mains electrical supply occurring at these times. Very similar results have been published for the PTB bars [12], which do not appear to have been used elsewhere. Several of the bars have drift rates of less than 1×10^{-9} per month, and further improvements to the circuits and mounting have been made [13]. The new bars operate with a day to day stability of ± 1 part in 10^{11}. The performance of quartz standards may be improved when they are maintained at a low temperature. A quartz ring designed by the British Post Office to have a zero coefficient at the appropriate temperature of about 10°C was housed at the bottom of a 50 ft bore hole and tested as a resonator. In a period of 3 years it drifted 18 parts in 10^9. Similar studies have been made at the NBS, Boulder, with GT plates and 5-Mc plates [14]. Tests have been made at Boulder and at the Bell Laboratories at liquid nitrogen and at liquid helium temperatures where a Q of 5×10^7 was achieved [15] with a short time stability of 1 part in 10^{11}.

The main requirement of a quartz clock when used in conjunction with astronomical measurements to provide a scale of time is uniformity of operation over periods of a year or more. The astronomical measurements have to be averaged over long periods and it is clear that if the rate of the quartz clock is not constant or varying in a uniform manner it is impossible to determine its value and then to use it for extrapolating time beyond the astronomical observations. The clocks themselves can be compared with a precision of 1 part in 10^{10} but although such comparisons help the user to decide which are keeping more constant they did not provide a more accurate time scale. This was limited to an accuracy of about 5 parts in 10^9 by the errors in the astronomical measurements. The position is completely changed now that atomic standards are available. These enable the quartz clocks to be calibrated in a few minutes with an accuracy of a few parts in 10^{11}, and what is required is a very high stability for the few days or weeks during which they may be used between calibrations.

IV. Atomic Clocks

It is obvious from the Bohr relationship $f=(E_2-E_1)/h$ between two energy states of an atom E_1, E_2 and the frequency f of the wave emitted or absorbed in a transition between them that any spectral line is a potential standard of frequency and hence of time interval. The author recalls that one of the pioneer investigators of quartz oscillators, D. W. Dye, considered the possibility of making a quartz plate so thin, and therefore of such a high resonant frequency, that it could be maintained in oscillation by a suitable light source. The gap from, say, 10^8 cps to 10^{15} cps was however far too great for the practical realization of this idea, but the recollection illustrates that radio engineers were alive to such a possibility. It was not many years before the gap was bridged and in 1934 Cleeton and Williams [16] at Michigan University excited a spectral line of ammonia at a frequency of 23,870 Mcs by a source of radio waves generated in the laboratory. The source used by them for exciting the transitions was a magnetron which generated a fairly wide band of frequencies and the ammonia was at atmospheric pressure, at which only a very broad resonance effect is observed. The width was mainly due to the effect of collisions and this can be reduced by reducing the pressure. Other broadening effects such as the Doppler effect then become important and under optimum conditions the line width remains at about 500 kc.

As the standard of astronomical time is known to about 5 parts in 10^9 it is clear that an atomic standard does not become useful until it at least equals this accuracy. Radical improvements were needed to the source of oscillations which must be stable and "monochromatic" to a few parts in 10^{10} and also to the atomic resonance the fractional width of which must be reduced to about 10^{-8} corresponding to 240 cps for ammonia.

A. Sources of Oscillation

The klystron valve can be made to oscillate in the region of frequencies of the order of 10^{10} cps in which some of the RF spectral lines are found and it has a much narrower frequency band than the magnetron. It can be further improved by a stabilizing circuit designed by Pound [17] which maintains it at the frequency of an external cavity resonator. Using a large high Q cavity, Essen [18] obtained a bandwidth of 1 cps with such an oscillator, and, with J. V. L. Parry, used it for the first precise measurement of the caesium hyperfine line now used in atomic clocks [19]. The frequency of the source could be varied over a wide range enabling the hyperfine line to be found and also enabling the neighboring Zeeman pattern lines to be investigated, but careful adjustment of the circuit was needed to obtain the narrow bandwidth required. Quartz oscillators can be varied by a few parts in 10^5 and are more convenient to use when the frequency is known within these limits. They have been used to control klystron oscillators and also by frequency multiplication to provide the exciting source directly. In the latter method the power that can readily be obtained is a few milliwatts at 10^{10} cps. If a wide frequency range is required, frequency synthesisers are now available which will produce from a quartz crystal any frequency up to say 10 Mc. This variable frequency can be included somewhere in the multiplying chain from a quartz oscillator of, for example, 5 Mc to give the required output frequency. One disadvantage of this method is the difficulty of keeping the output frequency free from sidebands which can produce errors of setting to the spectral line. The most reliable method is probably to multiply the frequency of a suitable quartz oscillator and to adjust it to give the spectral line frequency exactly. The oscillator is then measured by any convenient means in terms of the quartz standard which is thereby calibrated in terms of the atomic resonance. In this part of the measurement a frequency synthesiser can be used because sidebands do not effect the accuracy. At the National Physical Laboratory a number of schemes have been used [20] and the one which has proved the most convenient [21] is shown in schematic form in Fig. 6. A high-quality quartz oscillator is varied smoothly through a small frequency range centered on 5.0069 Mc and this frequency is multiplied by 2 and then by 306 by valve circuits with a final triode giving a power of 0.5 watt at 3064.2 Mc. The third harmonic generated by a germanium diode excites the atomic resonance and when the frequency of the quartz oscillator is adjusted to the peak of the resonance it equals 1/1836 of the atomic frequency. Instead of using a synthesiser to generate the frequency by which the second stage in the multiplying chain differs from 10 Mc this difference frequency is obtained from an auxiliary quartz oscillator after division by 17. The frequency of nominal value 10 Mc left after subtraction is compared in a phase comparator with the 100-kc standard. Any departure in phase generates an error voltage which is used to adjust the auxiliary quartz oscillator in order to restore the phase equality. The 10-Mc signal is thus maintained at a value equal to exactly 100 times that of the standard. The measurement consists in setting the variable oscillator

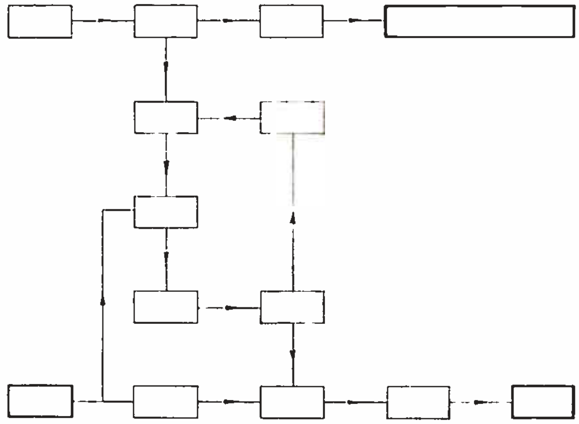


Fig. 6—National Physical Lab. scheme for exciting a caesium atomic beam.

to give the peak of the atomic resonance and then in measuring the frequency of the auxiliary quartz oscillator. This is done by comparison with the 100-kc standard, a small residual frequency being read off on a counter. If the frequency of the caesium resonance is taken as 9, 192, 631, 770 cps and the 100-kc standard has its nominal value, the counter reading will be 18, 826.7 cps. The difference from this, multiplied by 1.8357, gives the error of the 100-kc standard in terms of caesium in parts in 10^{10}.

This type of equipment is suitable when there are existing quartz clocks of high precision. The calibration in terms of the caesium only takes a few minutes and can be repeated as often as necessary—possibly two or three times a week. Atomic standards can, however, be made self contained by the incorporation of a quartz oscillator at a frequency of say 5 Mc and adjusting it by means of an electronic servo loop to keep a fixed relationship with the caesium resonance frequency. Such a scheme is used in the "Atomichron," a caesium atomic beam equipment [22] manufactured by the National Company, Inc. In this case (Fig. 7) the 5-Mc quartz oscillator is multiplied to 9180 Mc and also used in a synthesiser to generate a frequency of 12.63184 Mc which is added to 9180 Mc to give the caesium frequency. The division ratio, at the 1-Mc output is 9,192.631840 and as the frequency now being used for the caesium resonance is 9, 192, 631, 770 a small correction must be applied to the output frequencies of the Atomichron to make them correspond to this value of the unit of time.

B. Improvements to Intensity and Bandwidth

During the past decade there have been intensive and brilliantly successful programmes of work designed to increase the intensity and reduce the bandwidth of spectral lines. In the natural condition of a gas the two states that give a RF spectral line are almost equally populated. This can be changed and the intensity enormously increased by the process known as optical pumping [23]. A further improvement of intensity is

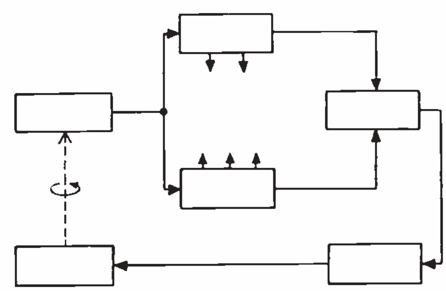


Fig. 7—Self-contained caesium beam standard using a servo controlled quartz oscillator, as in the "Atomichron."

afforded by using a measure of the light transmitted through the gas to detect the resonance. The Doppler width can be reduced and indeed almost eliminated by introducing a buffer gas [24], [25] or by coating the walls of the containing vessel. By using such methods atomic oscillators using the sodium resonance [26] at 1772 Mc, the rubidium resonance [27] at 6835 Mc, and the caesium resonance [28] at 9192 Mc have been constructed and a rubidium oscillator is in commercial production. Another approach led to the development of the ammonia maser [29] and quite recently to the hydrogen maser [30]. These developments are of great fundamental physical interest but have one drawback for frequency standardization in that the frequency is slightly affected by the devices used to secure the improvement in intensity and line width.

Up to the present, therefore, they have been used as constant oscillators but not for defining the natural Bohr frequency so closely that it can be used as a standard of frequency or time. This has however been accomplished in equipments based on the atomic beam method described in the next section.

C. Caesium Beam Standards

The difficulties of bandwidth and low intensity are most easily overcome by using the atomic beam magnetic resonance method developed at Columbia University by Rabi [31] and his co-workers. In this method, which can be used with atoms possessing a magnetic dipole moment, a beam of atoms (Fig. 8) passes to a detector through a system of magnets and a region of field alternating at the Bohr frequency. The magnets have a nonuniform field and deflect the atoms in one direction or the other according to which of the two energy levels they are in. When the frequency of the RF field is exactly equal to the Bohr frequency and is of the right amplitude transitions are induced and the deflections in the second magnet B are the opposite from those in the first magnet A and the atoms are thus focused on the detector. The width is governed only by the length of time the atoms spend in the exciting region in accordance with the Uncertainty Principle $\Delta f \Delta t = 1$, and the accuracy can be increased by increasing the length of path. This would have presented technical difficulties at the frequency of the caesium resonance (9192 Mc)

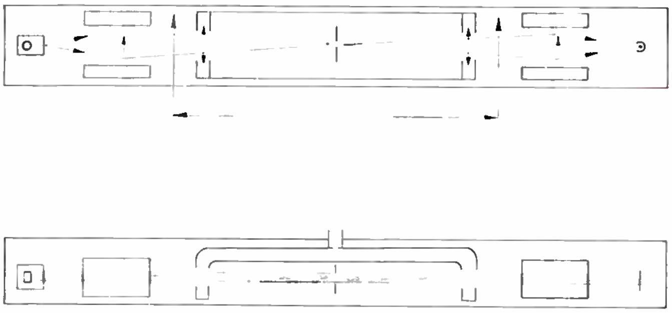


Fig. 8—National Physical Lab. caesium atomic beam chamber.

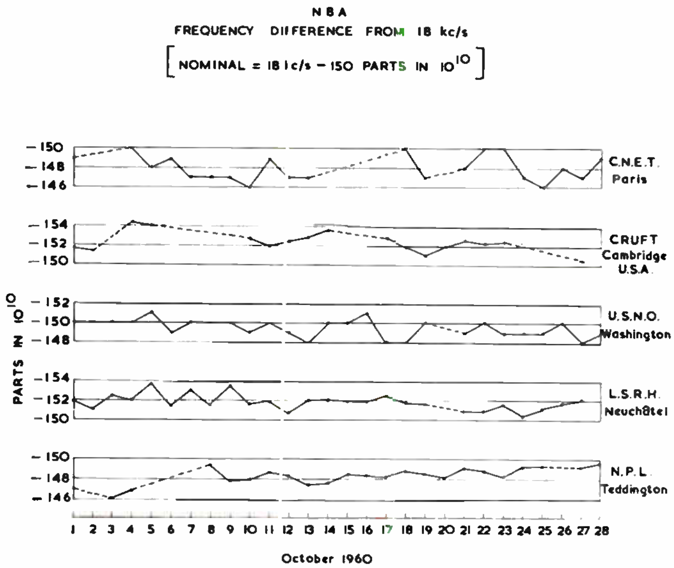


Fig. 9—The values of the frequency of the carrier wave of the station NBA as measured at a number of laboratories.

which is from many points of view the most convenient to use, but Ramsey made the very important suggestion that two short regions separated by a greater distance would be still more effective [32]. This configuration is ideal from the practical point of view as the short sections can consist of the standard waveguide used for such frequencies. Magnets of about 8000 Gauss with a similar gradient give deflections of the order of 1 mm so that wide slits and intense beams can be used. A strong signal is thus obtained at the detector and the radio frequency source can be set to the Bohr frequency with great precision and with very little experimental elaboration.

The development of an atomic beam standard was started by Lyons [33] at the National Bureau of Standards in Washington.

A thorough investigation of a beam standard constructed at the National Physical Laboratory by Essen and Parry [19] established that it could be used to define the Bohr frequency with a precision of 1 part in 10^{10}. A model built at M.I.T. by Zacharias [34] led to the construction of a commercial form known as the Atomichron, which was found to be in agreement [35] with the NPL standard to 2 parts in 10^{10}. The work on caesium standards was continued at the new laboratories of the National Bureau of Standards in Boulder by Mockler [36] and two equipments agreed in frequency to 2 parts in 10^{11}. Independently constructed standards have been completed at the National Research Council [37], the Laboratoire Suisse de Recherche Horlogeres [38], and the National Physical Laboratory [39]. The caesium beam standards of recent construction have a bandwidth of 50 cps or less and can be used with a precision of about 2 parts in 10^{11}. They are used to calibrate the working standards and measure the frequencies of a number of radio transmissions, which are controlled by high-quality quartz oscillators and are used for the widespread distribution of standard frequencies. The comparison of measurements made at a number of laboratories illustrates the state now reached in time and frequency measurements. Fig. 9 shows the different values obtained for station NBA operating at Summit, Canal Zone, during October 1960. They show that frequencies can be compared daily on a world-wide basis with an accuracy of ± 2 parts in 10^{10}.

Acknowledgment

This paper has been written as part of the research programme of the National Physical Laboratory and is published by permission of the Director of the Laboratory.

References

[1] F. D. Lewis, "Frequency and time standards," Proc. IRE, vol. 43, pp. 1046–1068; September, 1955.
[2] G. M. Clemence, "Standards of time and frequency," *Science*, vol. 123, pp. 567–573; April 6, 1956.
[3] W. Markowitz, "The photograph zenith tube and the dual rate moon-position camera" in "Telescopes, Stars and Stellar Systems," Kuiper and Middlenurst, Eds., The University of Chicago Press, Chicago, Ill., vol. 1; 1960.
[4] W. Markowitz, *et al.*, "Frequency of caesium in terms of Ephemeris time," *Phys. Rev. Lett.*, Vol. 1, pp. 105–106, 1958.
[5] D. W. Dye and L. Essen, "The valve-maintained tuning fork as a primary standard of frequency," *Proc. Roy. Soc. (London) A*, vol. 143, pp. 285–306; January, 1934.
[6] J. W. Horton and W. A. Marrison "Precision determination of frequency," Proc. IRE, vol. 16, p. 137; February, 1928.
[7] A. W. Warner and W. L. Smith, "Quartz crystal units and precision oscillators for operation in severe mechanical environment," *Proc. of 14th Annual Symp. on Frequency Control*, U. S. Army Res. and Dev. Lab., Ft. Monmouth, N. J.; 1960.
[8] A. W. Warner, "Design and performance of ultra-precise 2.5 mc quartz crystal units," *Bell. Sys. Tech. J.*, vol. 39, pp. 1193–1217; September, 1960.
[9] A. Scheibe and U. Adelsberger, "Quartz Clocks," *Hochfrequenz tech. u. Elektroakust*, vol. 43, p. 37; February, 1934.
A. Scheibe, *et al.*, "Konstruktion und Leistung neuer Quarzuhren der PTB," *Z. angew, Phys.*, Band 8, Heft 4, pp. 175–183; 1956.
[10] L. Essen, "A new form of frequency and time standard," *Proc. Phys. Soc. (London)*, vol. 50, p. 413; May, 1938.
[11] L. Essen, J. V. L. Parry and J. McA. Steele, "Frequency variations of quartz oscillators and the earth's rotation in terms of the NPL caesium standard," *Proc. IEE*, vol. 107, pp. 229–233; May, 1960.
[12] A. Scheibe, *et al.*, "Über die Qnarzuhrgruppe der P.T.B. und Vergleichsmessungen gegenüber Atomnormalen," *Z. angew. Phys.*, Band 11, Heft 9, pp. 352–357; September, 1959.
[13] G. Becker, "Konstruktion und Eigenschaften von Schwingquarzen hoher Gute," *Arch. Elektrotech. Übertragen*, Band 12, Heft 1, pp. 15–25; January, 1958.
[14] P. A. Simpson and A. H. Morgan, "Quartz crystals at low temperatures," *Proc. of 13th Annual Symp. on Frequency Control*, U. S. Army Res. and Dev. Lab., Fort Monmouth, N. J., p. 207; 1959.

[15] A. W. Warner, "Ultra-precise quartz crystal standards," IRE Trans. on Instrumentation, vol. I-7, pp. 185–188; December, 1958.
[16] C. E. Cleeton and N. A. Williams, "Electromagnetic waves of 1.1 cm. wavelength and the absorption spectrum of ammonia," *Phys. Rev.*, vol. 45, pp. 234–237; February, 1934.
[17] R. V. Pound, "Electronic stabilization of microwave oscillators," *Rev. Sci. Instr.*, vol. 17, p. 490; November, 1946.
[18] L. Essen, "A highly stable microwave oscillator," *Proc. IEE*, vol. 100, pt. 3, p. 19; January, 1953.
[19] L. Essen and J. V. L. Parry, "The caesium resonator as a standard of frequency and time," *Phil. Trans. Roy. Soc. (London) A*, vol. 250, pp. 45–69; August, 1957.
[20] L. Essen, E. G. Hope and J. V. L. Parry, "Circuits employed in the NPL caesium standard," *Proc. IEE*, pt. B, vol. 106, p. 240; March, 1959.
[21] E. G. Hope, "The comparison of highly stable frequency standards," to be published.
[22] "Atomichron," National Co., Inc., Malden 48, Mass., Instruction Manual; 1958.
[23] A. Kastler, "Les Methodes Optiques d'Orientation Atomique et leurs Applications," *Proc. Phys. Soc. (London) A*, vol. 67, p. 853; October, 1954.
J.O.S.A., vol. 47, p. 460; 1957.
[24] R. H. Dicke, "The effect of collisions upon the Doppler width of spectral lines," *Phys. Rev.*, vol. 89, pp. 472–473; January 15, 1953.
[25] H. G. Dehmelt, "Slow spin relaxation of optically polarized sodium atoms," *Phys. Rev.*, vol. 105, pp. 1487–1549; March 1, 1957.
[26] W. E. Bell, A. Bloom, and R. Williams, "A microwave frequency standard employing optically pumped sodium vapour," IRE Trans. on Microwave Theory and Techniques, vol. MTT-7, pp. 95–98; January, 1959.
[27] R. G. Carpenter, *et al.*, "A prototype rubidium vapor frequency standard," IRE Trans. on Instrumentation, vol. I-9, pp. 132–135; September, 1960.
[28] M. Arditi and T. R. Carver, "A gas cell atomic clock using optical pumping and optical detection," 1958 IRE National Convention Record, pt. 1, pp. 3–9.
[29] J. P. Gordon, H. J. Zeiger and C. H. Townes, "The maser, new type of microwave amplifier, frequency standard, and spectrometer," *Phys. Rev.*, vol. 99, pp. 1264–1274; August, 1955.
[30] H. M. Goldenberg, D. Kleppner and N. F. Ramsey, "Atomic hydrogen maser," *Phys. Rev. Lett.*, vol. 5, p. 361; October 15, 1960.
[31] I. I. Rabi, *et al.*, "The molecular beam resonance method for measuring nuclear magnetic moments," *Phys. Rev.*, vol. 55, pp. 526–535; March 15, 1939.
[32] N. F. Ramsey, "Molecular Beams," Clarendon Press, Oxford, Eng.; 1956.
[33] H. Lyons, "Spectral lines as frequency standards," *Ann. N. Y. Acad. Sci.*, vol. 55, pp. 831–871; November, 1952.
[34] J. R. Zacharias and J. G. Yates, "Atomic Beam Research," Res. Lab. of Electronics, M.I.T., Cambridge, Mass., Quart. Progress Rept., pp. 30–34; October 15, 1954.
[35] J. Holloway, *et al.*, "Comparison and evaluation of caesium atomic beam frequency standards," Proc. IRE, vol. 47, pp. 1730–1736; October, 1959.
[36] R. C. Mockler, R. E. Beehler and C. S. Snider, "Atomic beam frequency standards," IRE Trans. on Instrumentation, vol. I-9, pp. 120–132; September, 1960.
[37] S. N. Kalra, R. Bailey, and H. Daams, "Canadian caesium beam standard of frequency," *Nature (London)*, vol. 183, pp. 575–576; February 28, 1959.
[38] P. Kartaschoff, J. Bonanomi, and J. de Prins, "Etalons de frequence au cesium, description et resultats," *Helvetica Physica Acta*, vol. 33, pp. 969–973; December, 1960.
[39] L. Essen and J. V. L. Parry, "An improved caesium frequency and time standard," *Nature (London)*, vol. 184, pp. 1791; December 5, 1959.

The Measuring Devices of Electronics*

D. B. SINCLAIR†, FELLOW, IRE

***Summary*—As every branch of science develops, instrumentation especially adapted thereto develops with it. Electronics has been unusually fertile in spawning new ideas of all kinds, especially in measuring devices. This article discusses, in general, representative examples, as they relate to the needs that they satisfy or the novel concepts that they incorporate.**

Introduction

Significant measurements, in the early days of wireless, seemed not only hard to make but even hard to conceive. Maxwell's brilliant mathematical analysis had been completely verified by Hertz's experiments, and the fundamental interrelationships between electricity and magnetism thereby firmly established on a quantitative basis; but there yet remained an element of magic in electricity, safely contained within a circuit close at hand, suddenly appearing in another circuit remote from the original point of observation.

* Received by the IRE, July 25, 1961.
† General Radio Company, West Concord, Mass.

The concepts of inductance, resistance, and capacitance had been deduced from the behavior of circuits and components exposed to the influences of static charges and constant currents, and these concepts were extendable, without too much difficulty, to systems in which charges and currents varied relatively slowly with time. In wireless experiments, however, these quasi-steady-state conditions could not be assumed. Coils, which might be expected to behave like inductances, turned out to behave, instead, like capacitances; antenna wires acted sometimes like capacitances, sometimes like inductances, and sometimes like resistances; capacitances exposed to potential differences apparently well within their ratings inexplicably flashed over. These baffling phenomena were, of course, tangible manifestations of the three R's of this new art—resonance, radiation and reflections.

As these modes of behavior became understood, earlier concepts were interpreted in a wider framework to explain them. Thus, in particular, the concept of distributed parameters extended the ideas of inductance, resistance and capacitance to the analysis of antennas and transmission lines. Other concepts peculiar to the new devices, were found helpful—radiation resistance, propagation constant, characteristic impedance. These concepts and their definitions and measurement, illustrate the similarities and differences between what has, over the years, come to be known as electronics, and electricity or power. The basic quantities of electricity, their units of measurement, and their interrelationships with mass, length, time and the electromagnetic properties of free space, are well known. This article deals rather with those methods of measurement and those measuring devices that have arisen particularly from the problems and characteristics of electronic systems.

The Nature of Electronic Systems

The first question which one must ask is what these peculiar problems and characteristics are.

First and foremost, the power levels to be measured are frequently low. Although signals can be detected at great distances, these signals are weak, and the further one is away from the source the weaker they become. This led the Germans, who did much of the early work in radio-frequency systems, to characterize the entire field as *schwachstromtechnik*—weak current technology.

Second, the circuits used are frequently resonant or otherwise adjusted to a refined condition that is critically dependent upon the values of circuit parameters and sensitive to small disturbances of these values.

Third, the systems must respond over a band of frequencies so that information, as contrasted to power, can be transmitted from one point to another.

The implications of these characteristics are far reaching. The first demands high sensitivity in measurements of voltage, current and power; the second demands measuring devices that have negligible or highly predictable effects on the circuits in which measurements are to be made; and the third demands devices and techniques of measurement that can define behavior over substantial parts of the frequency or time domain.

Voltage and Power Measurements

Voltmeters

The device that first made electronics widely useful was, of course, the vacuum tube. Not only did it make possible the generation and reception of meaningful signals, but the very properties that made it useful for these functions made it equally so as a measuring tool.

Thus, its amplifying properties make it sensitive; its rectifying properties make it a splendid converter from hard-to-measure radio-frequency voltage to easy-to-measure direct current; its high input impedance makes its effect on circuits to which it is connected minimal; and it can easily be made to cover extensive frequency ranges without error.

It is interesting, in this connection, to note that the crystal rectifier, which had an honored position prior to the advent of the vacuum tube, has made a strong comeback in microwave applications. Neither the crystal diode nor the transistor, however, has yet become the basis of a classic instrument. The vacuum-tube voltmeter, on the other hand, has become one of the basic measuring instruments of electronics. It generally takes one of two forms, the peak-reading type[1] and the amplifier type.[2]

In the first type, a capacitor is charged up through a rectifier until the dc voltage across it equals the peak value of the applied ac voltage. This dc voltage is then amplified and used to actuate a dc meter. The rectifier circuit, which is widely used in "clamp" circuits, yields highly accurate, reproducible results at high voltages since departures from ideal performance cause variations only in the small *difference* between the applied peak voltage and the developed dc voltage. As the voltage to be measured decreases, however, this difference becomes larger and larger relative to the developed dc voltage, and at low voltages the response changes from "peak" to rms, or square-law and becomes critically dependent upon tube characteristics. This type of voltmeter is therefore best suited for voltage measurements above about one volt. It can be made to give accurate results at frequencies up to the hundreds of megacycles.

The second type of voltmeter uses a wide-band amplifier to make measurements at low voltages possible and, as a rule, uses negative feedback both to stabilize the amplifier and to "straighten out" the rectification characteristic. This provides a reliable "average" response that, in contrast to "peak," is often found useful when the waveform of the signal to be measured is not purely sinusoidal. Measurements can be made reliably, with voltmeters of this kind, at voltages of the order of 1 mv and above. However, the high sensitivity obtained with this instrument, when contrasted to that of the peak-reading type, is paid for by a decrease in the maximum frequency at which it is useful from the order of several hundred megacycles to the order of a few megacycles.

Standard-Signal Generators

For measurements at lower voltages and higher frequencies direct-reading voltmeters are no longer useful, and a completely different approach is to use a standard-signal generator[3] in a substitution procedure. This instrument, whose concept stems uniquely from the requirements of electronics, supplies a signal of known characteristics to compare with the unknown signal. From this comparison, a quantitative evaluation of the unknown can then be made.

A standard-signal generator comprises, basically an oscillator or oscillator-amplifier combination that can

[1] C. H. Sharp and E. D. Doyle, "Crest voltmeters," *Trans. AIEE*, vol. 35, pt. 1, pp. 99–107; February, 1916.
[2] S. Ballantine, "Electronic voltmeter using feedback," *Electronics*, vol. 11, pp. 33–35; September, 1938.
[3] L. M. Hull, "Rating of radio receivers," *Gen. Rad. Exper.*, vol. 3, pp. 1–2, 4; November, 1928.

be tuned to cover the desired frequency range, a voltmeter that measures the oscillator output voltage, and an attenuator that reduces this output voltage by a known amount. It is particularly useful with a sensitive receiver as a comparison device. By setting the output of the standard-signal generator so that both the standard-signal generator and the unknown signal produce the same receiver output, one can make the standard-signal generator read directly. An extension of this technique enables one to determine the limitations of the receiver itself that are imposed by noise and, thereby, to determine the corrections necessary to measure signals so small that they are comparable to system noise levels.

Thermal Devices

The measurement of a standard-signal-generator reference voltage, from which the output voltage is derived, can be made by methods that are free from some of the limitations imposed upon voltmeters for general use. Thus, in particular, the voltage-measuring system need not have a negligible effect upon the circuit to which it is connected. This freedom makes possible the use of power-measuring devices, which have been found satisfactory for measuring rms voltages at frequencies up to thousands of megacycles.

These devices actually measure the power dissipated within themselves, the corresponding voltages being deduced therefrom through knowledge of their impedances. Since these impedances are generally low and resistive, the devices can be made to cover wide frequency bands in lumped-parameter circuits and, in distributed-parameter configurations, can be adjusted to match transmission lines and waveguides and minimize reflections. They use different means to indicate the power dissipated, but all rely upon temperature rise as the measuring quantity. In bolometers and barretters the change in resistance with temperature is measured at dc or audio frequencies; in thermocouple systems and calorimeters the temperature rise is measured thermometrically; in hot-wire meters the temperature coefficient of expansion actuates a mechanical system; in incandescent lamps the change in brilliance of the filament is a very sensitive measure of small changes in power.

Power-Output Meters

The importance of impedance matching in electronic systems emphasizes a significant difference between them and power systems. In power systems efficiency is extremely important. An ideal system would be one in which there were no losses in the generating and transmission system and all the power was dissipated in the load. Even in a practical system the losses are small compared with the useful power. Any attempt to draw maximum power from the system by matching impedance would be catastrophic. Component parts are not designed to cope with such operation, and power into a matching load is not a proper system specification. In contrast, in communications systems generally, there is a need at every point to extract the absolute maximum power from any signal to avoid degradation of signal-to-noise ratio. Although the efficiency is only 50 per cent, maximum power transfer in linear systems occurs when the source and load impedances are conjugate in angle and equal in magnitude, and optimum results are obtained when the impedances are purely resistive.

Not all generating systems in electronics are linear, independent of load, and maximum undistorted output is not always obtained under matched conditions. Nevertheless the maximum power output is an important specification, generally, whether limited by clipping, in overloaded nonlinear systems, or by impedance matching in linear systems. Another specialized measuring device, the power-output meter,[4] has therefore been developed to measure this characteristic, consisting simply of a variable-impedance load and associated meter that reads the power dissipated therein.

Directional Couplers

Primarily useful at audio frequencies, the power-output meter is supplemented, at high frequencies, by the directional coupler.[5]

Still another instrument that has been developed to meet the special requirements of communications, the directional coupler measures power flow in a coaxial line or waveguide. In a simple form it consists of an auxiliary transmission line, terminated in a matching impedance at each end, which is coupled to the transmission line in which the power flow is to be measured. By proper choice of coupling to the electric and magnetic fields in the main line one can cause waves to be induced in the measuring line which travel to one end or the other, according to whether they are initiated by the incident or reflected wave in the main line. The power dissipated in each measuring-line termination is then proportional to the corresponding power flow in the main line and can be used as a measure of it.

Field-Strength Meters

Measurements of voltages in closed spaces, where they can conveniently be confined and defined do not, however, meet all the needs of electronics. At the highest frequencies, when the physical dimensions of the system under study become comparable to the wavelength of the signal, and at any frequency, when the signal is transmitted through space, it is necessary to analyze performance in terms of electric and magnetic fields. Measuring devices directly responsive to these quantities, however, have not been developed. Instead, per-

[4] John D. Crawford, "A power meter with a wide impedance range," *Gen. Rad. Exper.*, vol. 6, pp. 4–6; May, 1932.

[5] A. A. Pistolkors and M. S. Neuman, "Device for direct measurement of the coefficient of a travelling wave in feeders," *Elektrosvyaz'*, vol. 9, pp. 9–15; April, 1941.

formance of probes, loops and antennas has been calculated and used to relate the power picked up by them to the inducing fields. This power, when measured by the various methods previously noted, can then be used to indicate field strength. In illustration, the combination of a sensitive, stable receiver, a reference generator, and a loop antenna makes an excellent field-strength meter for radio-frequency measurements.[6]

Modulation and Waveform Measurements

Cathode-Ray Oscilloscopes

The essence of communication is change. If a given type of signal has existed from the remote past, is currently present, and continues on into the indefinite future, no information of any kind is conveyed thereby. Only if some property of the signal is modified as a function of time can intelligence be transmitted. In contrast to power-system requirements for measurements of sustained single-frequency voltages and currents, therefore, communication-system needs for methods of measuring changes with time have always been pressing. The cathode-ray oscilloscope has provided an excellent solution.

Coupled with the development of circuits to provide deflecting voltages that are accurately known as functions of time the low-inertia properties of the electron beam have made possible visual display of recurring waveforms at frequencies ranging upwards to the thousands of megacycles and photographic recording of single transients of very brief duration.

The requirement that a signal change as a function of time implies that the system in which it occurs must possess a finite bandwidth in the frequency domain, or be able to respond at a finite speed in the time domain. The cathode-ray oscilloscope can be used for analysis of either kind. In the frequency domain it can display sine-wave-modulation waveforms point-by-point, or swept; and in the time domain it can display pulse- or square-wave-modulation waveforms as recurrent phenomena, or the single-transient response to a step function.

Standard-Signal Generators

In addition to being a standard of voltage level, the standard-signal generator is a standard of modulation characteristics. Whatever the nature of the modulating voltage, the carrier signal must be varied in frequency or amplitude in accordance with it. The modulating system of the standard-signal generator supplies the known relationship between the modulating and modulated voltages; the response to the modulated signal of the system under test can be displayed as a function of time on a cathode-ray oscilloscope or measured in terms of its frequency components.

[6] C. R. Englund, "Note on the measurement of radio signals," Proc. IRE, vol. 11, pp. 26–33; February, 1923.

Wave Analyzer

For measurements of the Fourier-series components in a repetitive nonsinusoidal signal the wave analyzer offers high accuracy and convenience. It is, in common form, a highly selective heterodyne-type voltmeter which covers the audio-frequency range.[7] A variable-frequency local oscillator beats with the signal to yield a difference frequency that can be adjusted to pass through a narrow-band fixed-frequency filter. The filter bandwidth is chosen to be small compared with the spacing between frequency components that are to be separated, and individual measurements of their amplitudes can therefore be made. It is used principally to study distortion in sine-wave testing and frequency spectra in random-noise testing.

Spectrum Analyzer

Conventional wave analyzers have untuned input circuits and, because of this, their use is restricted to low frequencies. When they are used for measuring modulation characteristics, therefore, the signal must generally be rectified so that the modulation envelope can be examined rather than the modulated wave itself. When the actual modulated wave is to be analyzed it is necessary to use a highly selective voltmeter that can be tuned over a frequency range in the vicinity of the signal frequency.

For this kind of analysis the signal is considered as the combination of a carrier frequency and sideband frequencies that are created by the modulation process. A manually tuned, high-selectivity receiver can be used to separate these components, but it is frequently found more convenient to use a spectrum analyzer,[8] which, in effect, is a receiver of this kind that is swept electronically back and forth across the frequency band of interest. The output of a spectrum analyzer is presented, as a time-varying function, on a cathode-ray-oscilloscope screen.

Impedance Measurements

Direct-Deflection Methods

Early experimenters found the vacuum tube a good device to measure its own performance; they likewise found circuit behavior a good indicator of the circuit condition that determines it. Thus, not only is resonance a typical characteristic of high-frequency lumped-parameter circuits, but an understanding of resonance makes possible deduction of "unknown" parameters from the observed changes in behavior of resonant circuits when a "known" standard parameter is varied or connected in circuit.

For instance, it is easy to set maximum current in series-resonant circuits, or maximum voltage across

[7] L. B. Arguimbau, "Wave analysis," *General Radio Experimenter*, vol. 8, pp. 12–14; June–July, 1933.

[8] "Panoramic reception shows promise in radio navigation," *Electronics*, vol. 11, pp. 36–38; July, 1938.

parallel-resonant circuits, either by tuning a variable capacitor or by varying the frequency. It has been found that the value of the maximum current is inversely proportional to the effective resistance of the series-tuned circuit, and that the value of the maximum voltage is inversely proportional to the effective conductance of the parallel-tuned circuit. By comparing the change in maximum current or maximum voltage when an unknown impedance or admittance is connected to the circuit with the change that occurred when a known resistance or conductance is connected one can then deduce the resistance or conductance of the unknown. From the frequency and the capacitance change necessary to restore resonance when the unknown is connected one can deduce its reactance or susceptance. These methods have become known, respectively, as the resistance-variation[9] and conductance-variation[10] methods.

From the frequency and the shape of the resonance curve, when the tuning capacitance is varied, it has also been found possible to deduce the circuit losses. Both the resistance and reactance of an unknown impedance, connected in a series-resonant circuit, or the conductance and susceptance of an unknown admittance connected across a parallel-resonant circuit, can therefore be deduced from a single variable-capacitance standard. These methods have become known as the reactance-variation[11] and susceptance-variation[12] circuits.

It has also been found that the breadth of the resonance curve, when the frequency is varied, is inversely proportional to the storage factor, $Q=X/R=B/G$, of the tuned circuit, and that the ratio of the voltage occurring across one of the tuning elements in a series-tuned circuit to a voltage injected in series is also equal to Q. This principle of "resonant rise" is used in direct-reading q-meters for measuring coils.

Turning to distributed-parameter systems, one finds that not only is reflection of waves a typical characteristic but measurement, on a "smooth" line of known and constant characteristic impedance, of the distribution of the standing-wave pattern produced by a reflection yields all the information necessary to deduce the impedance mismatch that created it. A measuring tool of great use at the higher frequencies is therefore the slotted line, or slotted section, with which one can explore the distribution of standing waves in a coaxial-line or waveguide system.

Null Methods

In all the methods of measurement that fall in the general direct-reading category, the unknown parameters are related to standard parameters through observations of voltage or current ratios. Similar relationships can be established, with considerably higher precision, by balancing voltages to a null in the general manner of the Wheatstone bridge.

Bridges have traditionally been used for measurements of the highest accuracy, and attempts to extend low-frequency techniques to radio frequencies were made as early as 1897, when Nernst described bridge measurements at high frequencies using a spark-gap as an indicator of balance.[13] It was found very early in the game that excellent precision and resolution could be easily obtained, but that meaningful readings were frequently another story. The reason for the difficulties lies primarily in the relatively greater importance, as the frequency is raised, of the so-called residual parameters to be found, in combination with the desired parameters, in impedance standards and circuit components generally and in circuit connections and wiring which would ideally have no associated impedances or admittances at all. As our understanding of the nature, significance and causes of these has increased, better standards and improved configurations have made possible high-accuracy measurements with lumped-parameter bridges at frequencies up to the order of a few hundred megacycles. The use of distributed-parameter devices and connections has further extended null-method techniques to still higher frequencies.

Impedance Standards

A common denominator in all impedance measurements is the problem of securing accurate standards. The residual parameters, which may justifiably be listed as the fourth R of the electronic art, become progressively more and more serious as the frequency increases, usually as the square. There is therefore a great premium in selecting a standard that is as free from residuals as possible and in which the location and behavior of the residuals is best known. The problem is frequently compounded by the need for an adjustable parameter. A great deal of analysis and measurement has shown that the best lumped-parameter element for use at radio frequencies is the variable capacitor,[14] and it is used almost exclusively as the reference standard in impedance-measuring circuits of all kinds.

At microwave frequencies, interestingly enough, the problem becomes somewhat easier to solve. Inductance and capacitance are, dimensionally, specified by a length. At frequencies where physical dimensions are

[9] R. Lindemann, "Über Dämpfungsmessungen mittels ungedämpfter elektrischer Schwingungen," *Verh. der Deutsch. Phys. Ges.*, vol. 11, p. 28; January 15, 1909.

[10] D. B. Sinclair, "Parallel-resonance methods for precise measurements of high impedances at radio frequencies and a comparison with the ordinary series-resonance methods," Proc. IRE, vol. 26, p. 1469; December, 1938.

[11] V. Bjerknes, "Ueber electrische Resonanz," *Wied. Ann.*, vol. 55, p. 121; May, 1895.

See also R. von Traubenberg and B. Monasch, "Über die Verwendung kontinuierlicher elektromagnetischer Schwingungen bei Dämpfungsmessungen," *Phys. Zeit.*, vol. 8, p. 925; December 15, 1907.

[12] Sinclair, *op. cit.*, see p. 1467.

[13] W. Nernst, "Ueber die Verwendung schneller elektrischer Schwingungen in der Brückencombination," *Wied. Ann.*, vol. 60, p. 600; March 15, 1897.

[14] R. F. Field and D. B. Sinclair, "A method for determining the residual inductance and resistance of a variable air condenser at radio frequencies," Proc. IRE, vol. 24, pp. 255–274; February, 1936.

comparable to the wavelength, therefore, simple geometric configurations lead to simple boundary conditions and simple solutions of the field equations. As under steady-current and static-charge conditions, therefore, one can compute with accuracy the electrical behavior of a device. In particular, the characteristic impedance of transmission lines can be calculated, and the input impedance can then be set to yield accurately known resistance or reactance values by proper adjustment of the termination and line length. In further illustration one can compute, with very high accuracy, the resonant frequencies of cavities.

Frequency and Time Measurements

Wavemeters

The phenomenon of resonance, which is so characteristic of electronic systems, not only serves as a convenient indicator of circuit conditions for impedance measurement but also of the frequency of the signal. Again, the functional device itself provides the measurement solution. Different signals are transmitted at different frequencies; resonant circuits are tuned to these different frequencies to separate the signals; therefore, calibrated tuned circuits measure the frequencies in accord with the manner in which they are used.

At low frequencies, wavemeters, comprised of lumped elements, have gradually fallen into disuse as requirements for accuracy of measurement have demanded stability of calibration and precision of adjustment further and further beyond their capabilities. At microwave frequencies, on the other hand, cavities of simple geometry and construction provide one or two orders better performance and still serve a useful function.

Frequency Standards

The requirements for more and more accurate frequency measurements, which have stemmed from the highly sophisticated electronic systems of today, have brought the problems of the physicist and the electronic engineer closer and closer together. Thus, high-frequency stability is needed for color television and nuclear resonance measurements; very precise timing is needed for astro-navigation systems and computers.

The quartz crystal, like the internal-combustion engine, has been developed to a pitch of performance that would, no doubt, have been unbelievable to the original experimenters. Short-term stabilities of the order of 10^{-11} have been achieved, representing a degree of perfection of a high order. Long-term drift, however, remains a problem. The fundamental properties of matter, as exemplified by the emission or absorption spectra, have therefore been harnessed to provide long-term stability of frequency that is, by definition, perfect.

Here, again, is an interesting interdisciplinary extension into astronomy. Because frequency is really a measure of time interval, the fundamental standard has, in the past, been time, as determined from astronomical observations. The need for accuracy in space, however, has now outstripped the accuracy of our optical observations in space; and time, it is now proposed, should henceforth be determined from the frequency of the absorption line of cesium. As a practical matter, all international standards of frequency are now determined in terms of the cesium beam, and transmissions at very low frequencies of the order of 10 to 100 kc make them available for intercomparison throughout the world with a minimum amount of error from propagation anomalies.

Counters

As the stability and accuracy of frequency standards has improved over the years there has arisen a corresponding need for more and more precise comparison of an unknown frequency with the standard. Digital counters, which stem from the scalers of atomic energy research, have furnished an increasingly popular solution.[15] They incorporate a frequency standard, which establishes an accurate time interval, and count the frequency as a number of quantized events occurring during that time interval. Like all digital devices, they have a precision of ± 1 count in the least significant figure, but this inherent imprecision can be made arbitrarily small in determining over-all accuracy by adding as many decades as are required.

Pulse Generators

Early developments in electronics were nearly all explained and measured in terms of frequency. In more recent years, however, time has become an important basic quantity in radar, computer, and navigation systems. Not only have instruments like the counter therefore appeared on the scene, but also analytical techniques and devices to study circuit behavior in the time domain.

The prime illustration of this class of instruments is the pulse generator, which provides "rectangular" waves, as opposed to sine waves, that range from brief pulses to square waves. The transition times can be as short as a few nanoseconds. When used with cathode-ray oscilloscopes, pulse generators provide information on behavior entirely in the time domain; when used with wave analyzers they form a system that links behavior in the time and frequency domains.

Noise Generators

A common characteristic of electronic-system analysis and design is determination of, and design for, ultimate performance. Fundamental limitations, whether in parametric amplifiers, communication channels, radar systems, or computers, arise from random noise. Another basic instrument is therefore the noise generator,

[15] I. E. Grosdoff, "Electronic Counters," *RCA Rev.*, vol. 7, pp. 438–447; September, 1946.

designed to give the best approximation to white, Gaussian noise. Taken with the cathode-ray oscilloscope, this instrument, like the pulse generator, gives information in the time domain; with the wave analyzer it gives similar information in the frequency domain.

Conclusion

The measuring tools of electronics have developed at the forefront of a wide range of new technologies. Some of the specific devices have been touched on briefly here, but there have been many that have not even been mentioned. Electronic methods, for instance, have been widely used in measuring light, sound, temperature, pressure, velocity, acceleration—in fact, well nigh any physical quantity that can be segregated and assigned a number. Microphones are electronic devices; so are photocells, vibration pickups, accelerometers, Kerr cells, masers and Hall-effect transducers. What started out as a dimly understood, difficult art has become, over the years, one of the most rigorously developed, fundamental and well instrumented. Some years ago I propounded the idea that electronics, far from being a narrow specialty within the profession of electrical engineering, in fact covered a broad spectrum within which "electrical engineering" represented but a narrow band. Promptly dubbed "Sinclair's heresy" by my colleagues, this doctrine appears closer and closer to being definitive.

Digital Display of Measurements in Instrumentation*

B. M. OLIVER†, FELLOW, IRE

Summary—**Being a short chronicle of the advent of digital display in divers new electronic instruments, sundry remarks telling of the wonders they can do, together with a few words to the wise concerning pitfalls which may accompany their indiscriminate use.**

During the last twenty years a whole new branch of electronics has grown up, a branch concerned with the electrical generation, processing, recording, reproduction and display of data in digital form. As is usually the case, this art did not spring into existence suddenly out of one particular invention, but instead developed gradually as the cumulative result of many contributions. It was nourished by, and in turn helped support, adjacent fields such as digital computers and pulse code systems. But while there is no one birthday to celebrate, certain key developments deserve recognition.

In digital circuitry probably the most important single device is the bistable circuit or "flip-flop." Used today in innumerable modifications, the original form of this circuit was developed in 1919 by Eccles and Jordan. During the late 1930's flip-flops in tandem connection began to be used by nuclear physicists to scale down the counting rate from Geiger-Muller counters by factors of 2^n. These devices were all essentially binary counters in which the last "carry" pulse was used as the output to a mechanical register. In 1944, Potter[1] described a scale-of-ten counter in which feedback caused a binary counter to skip six of its possible 16 states. A neon lamp indicating the state of each flip-flop provided a binary-coded readout for each decade. In 1946, Grosdoff[2] described another binary counter with feedback to provide a scale-of-ten circuit. Grosdoff's circuit featured a direct readout obtained by a matrix of resistors arranged to light one of ten neon lamps for each stable state of the counter. By connecting such decades in tandem direct reading decimal counting registers of arbitrarily large capacity were achieved and the way was cleared for the widespread application of high speed electronic counting.

During this same period and in later years many devices were developed which inherently have ten stable states. Examples are the ten anode glow transfer tube,[3,4] the latching beam cathode ray tube,[5] and the

* Received by the IRE, July 28, 1961.
† Hewlett-Packard Company, Palo Alto, Calif.

[1] John T. Potter, "A four-tube counter decade," *Electronics*, vol. 17, pp. 110–113; June, 1944.
[2] I. E. Grosdoff, "Electronic counters," *RCA Rev.*, vol. 7, pp. 438–447; September, 1946.
[3] James J. Lamb and Joseph A. Brustman, "A poly cathode glow discharge tube for counters and calculators," *Electronics*, vol. 22, pp. 92–96; November, 1949.
[4] R. C. Bacon, "The dekatron," *Electronic Engrg.*, vol. 22, pp. 173–177; May, 1950.
[5] J. L. H. Jonker, A. J. W. M. van Overbeek, and P. H. de Beurs, "A Decade Counting Valve for High Counting Rates," Philips Res. Rept., Eindhoven, Netherlands, vol. 7, No. 2, pp. 81–111; 1952.

magnetron type beam switching tube.[6] Because of performance or price limitations these devices have been less widely used than the decade constructed of binary flip-flops, though in certain applications they are probably superior.

In 1950 the first commercial instruments appeared[7,8] combining an electronic counter with a precision time base and gating circuits. With these instruments, precise unambiguous measurements of frequency, pulse rates, period or time interval could be quickly and conveniently made. They revolutionized the whole art of frequency and time measurement and found very widespread use. Later, devices were developed to print out the digital readings and thus provide a permanent record of a series of measurements.

In 1953 Hagelbarger at Bell Laboratories built a readout device consisting of a neon tube with a single anode and ten stacked wire cathodes in the shape of numerals.[9] When current is taken from one of these cathodes, a glowing numeral an inch or more in height appears. Smaller versions of this tube were later developed commercially,[10] and provide a large in-line readout which is gradually replacing the ten neon glow lamp type. Other devices employing electroluminescent segmented numerals, stacked edge lighted plastic sheets with engraved numerals, and optically projected numerals have also appeared. In the most recently developed electronic counters, neon glow lamps driven directly by each flip-flop are used to illuminate a photoconductor matrix, which in turn not only drives a numerical indicator tube but also stores the previous reading while the decade is counting new information. For very fast digital data presentation special cathode ray tubes having beams shaped by apertures into symbols have been produced, but their expense has limited their use to large complex systems.

By 1954 the first commercial digital voltmeter had appeared.[11] This unit was electromechanical but was soon followed by all electronic designs based upon circuit techniques originally employed in pulse code modulation encoders. With the advent of these voltmeters it became possible to present in digital form any quantity which, by means of an appropriate transducer, could first be converted to a proportional voltage. Since this list includes most physical variables, it was suddenly possible to record and process all telemetered or locally sensed data from a large scale test directly in digital form. As a result, digital voltmeters found a large market in large instrumentation systems.

Digital techniques have also been extended to signal sources of various sorts. Here the digital information constitutes the input rather than the output. Examples are digital pulse delay generators, ac and dc voltage sources and frequency synthesizers. These devices will be recognized as the counterparts of the digital time interval counter, the digital voltmeter, and the digital frequency counter.

Virtues of Digital Display

Probably the greatest single virtue of digital representation is its ability to present data with practically unlimited precision; for example, eight digit data in which the smallest digit is significant is easily presented on a commercial counter. This is a precision of 1 part in 10^8. Assuming one can read a pointer position to within 0.01 in., comparable precision in a full range analog device would require a scale length of 10^6 in. or 16 miles. Thus digital readout is unusually necessary for very high precision. An exception is the reading or monitoring of a quantity which varies by only a small fraction of its total value. Here suppressed zero analog readout may suffice.

Another advantage of digital representation is the ease with which data in this form may be stored or processed without loss of accuracy. Data in digital form may be printed numerically, punched on tape or cards, or recorded magnetically; it may be fed through computers and normalized, linearized, or operated upon in any desired fashion. It may be stored and re-used or transmitted over long distances. In fact much of the popularity of digital instrumentation stems from the existence of digital computers to handle the data and of digital means for storing it.

Digital readout devices are less apt to be misread, particularly by unskilled personnel than are meter faces, for example. The digital device has no parallax error and for a given panel space is legible at a much greater distance. As a result digital devices have found application in production testing and other areas where neither high precision nor data processing is involved.

Vices of Digital Display

It is so easy to wax enthusiastic about the virtues of digital display that one is apt to overlook certain very real disadvantages. The first is the fact that digital devices are in general more complex and costly than their analog counterparts. Unless needed for any of the reasons stated above, digital devices may often be rejected on this basis alone.

For certain types of measurement it is hard to imagine a worse form of presentation than a digital one. As

[6] J. Bjorkman and L. Lindberg, "Development of trochtrons," *Trans. Roy. Inst. Tech.*, Stockholm, Sweden, no. 80; 1900.

[7] Saul Kuchinsky, "Multi-output beam switching tubes for computer and general purpose use," 1953 IRE National Convention Record pt. 6; pp. 43–54.

[8] A device known as the EPUT (Events Per Unit Time) meter was introduced in 1950 by Berkeley Instruments. The maximum counting rate was 100 kc. In early 1951 the Hewlett-Packard Co. introduced a combined frequency and period measuring instrument with a counting rate of 10 Mc (*Hewlett-Packard J.*, vol. 2, No. 5; January, 1951).

[9] Prior art anticipating Hagelbarger's work was found in H. P. Boswau, "A Signalling System and Glow Lamp Therefor," U. S. Patent 2142106; filed May 9, 1934; issued January 3, 1939.

[10] The "Nixie" tube developed by Haydu (now Burroughs) and the "Inditron" developed by National Union are examples.

[11] Introduced by Non-Linear Systems, this meter grew out of a model developed for the Naval Electronics Laboratories.

an example, consider the use of a digital voltmeter to read the detector output when lining up an IF amplifier. The commotion that appears on the voltmeter with each adjustment bears no direct relation to the sense of the tuning. True, one can stop and interpret the result, but this is an extra mental operation. The situation is even worse when seeking a null, for the digital voltmeter is apt to change ranges several times and the resulting bedlam is very confusing. By contrast a simple d'Arsonval movement clearly indicates a maximum, or a null.

A digital record is a rather poor way of presenting certain data, such as temperature or frequency drift, as compared, say, with a strip chart recording. For this reason some digital printout devices[12] are provided with analog output by means of which an analog recording of the data may be made simultaneously with the printed record. By synthesizing a voltage from only that group of digits having significant variation, a recording with appropriate zero suppression and hence easy readability is obtained.

When a production adjustment must be made to an accuracy no greater than about ± 2 per cent, much time will be wasted if a normal digital readout is used. Consider a power supply which must be set at 300 v$\pm$2 per cent. It is almost psychologically impossible for the production tester to leave the reading at 304 v, or worse, 296 v. There is a strong temptation to try for the round number, for the bull's eye. By contrast, few people are bothered at leaving a meter reading anywhere between two indicated limits on a scale.

As a final caution, it is probably not amiss to remind the reader that behind the digital display there is usually an analog device and that *precision of readability does not ensure accuracy of measurement.* The accuracy of a frequency counter is no better than the standard oscillator used, the digital voltmeter is no more accurate than its internal reference, its range dividers, its comparison amplifiers, and its thermal EMF's allow, regardless of how many digits of display are provided.

Conclusion

Instruments with digital display as developed over the last two decades provide a high precision readout. Often this precision is justified by the accuracy of measurement of which the instrument is capable, and which the application demands. Or the digital form of data out put may be desirable because of legibility, or required by automatic devices into which the data is fed for storage or processing. When these requirements do not exist and when simpler analog instruments are available they should be used. This is particularly true where the reading is to be watched by eye for extreme values, or set within easily discernible limits.

Lest the impression be left that digital display in electrical measurements is something entirely new to the last two decades, let us recall two venerable devices: the decade bridge and the integrating watt-hour meter. Like many of the newer digital readout devices the latter is driven by an analog measuring device, and it too can beguile the user into never questioning its apparent accuracy.

[12] Alan S. Bagley and Edgar A. Hilton: "A fast digital recorder with analog output for automatic data plotting," *Hewlett-Packard J.*, vol. 8, No. 8; March, 1957.

Section 19

MEDICAL ELECTRONICS

Organized with the assistance of the IRE Professional Group on Bio-Medical Electronics

History of Bio-Medical Electronics Art*

J. F. HERRICK†, MEMBER, IRE

Summary—**The almost phenomenal growth of the IRE Professional Group on Bio-Medical Electronics demonstrates the mutual interest that has been aroused both in engineering and in biology through new knowledge of living systems gained by instrumentation and technics recently available in this interzonal field.**

The engineer should develop a high-level respect for living systems at all stages of development, from that of unit cells to that of the amazing coordination of the living systems in the highly developed human being. An understanding of these systems will be advantageous to the physical sciences.

Engineers, physicists and biologists have learned that by serving each other they can serve their respective goals best.

The author believes that the science of electronics is the "liaison officer" responsible for converging the life sciences and the physical sciences into a single society. The organization of a society that would converge the man-made disciplines that cover the widely divergent fields reflected in the terms "bio-medical electronics," "biophysics" and "bio-medical engineering" was not achieved until the science of electronics was put to work in the life sciences.

ALMOST EVERY research scientist in the broad field of the life sciences today is turning his attention to the new methods and technics of investigations that employ the electronics art. The science of electronics is penetrating into every aspect of experimental biology and medicine, causing revolutionary changes owing to the stimulating and challenging technics for procuring new information under conditions that may not interfere with the living mechanisms.

The electronics scientist, on the other hand, is learning that living organisms have many engineering systems which, if understood, could be used to advantage in the further development of man-made systems of electronics. Thus, there are mutual rewards for both the research biologist and the research engineer in properly directed group studies in this very interesting interzonal field.

It must be made explicitly clear, first, that the field of bio-medical electronics is not limited to instrumentation; it is a separate discipline in science. Instrumentation is a fundamentally important requirement for obtaining desired information, preferably quantitative. However, the engineer's work should not be limited to that of a technician who may design, construct and maintain the instrument. Present types of instrumentation designed and developed by electronics engineers surpass the most fantastically imagined types of only a few years ago, but instrumentation is only a means to an end. The solution of a problem by electronic technics usually requires the engineer to understand the problem to be investigated and to take an active role in the design, application and performance of the electronic equipment as well as the

* Received by the IRE, August 7, 1961; revised manuscript received, October 30, 1961.

† Vista Laboratory, California Institute of Technology, Pasadena, Calif.

retrieval of data. The man who is responsible for the instrumentation should be the man who uses it in the research problem, at least in the early applications. Such participation will ensure a better solution of the research problem. The engineer will function also as a valuable assistant in data analysis.

Second, it should be made clear to the doctor of medicine that the nonmedical members of the research team respect highly the art of medicine—that aspect of medicine which is so necessary (and often confidential) for treating the patient as well as the disease. The engineer or physicist will not presume to enter this important realm of medicine. Obviously, bio-medical electronics is concerned with the purely objective scientific aspects of the research problem.

New knowledge or devices in the nonmedical sciences often raise the general question: Can these new agents be added to the armamentarium of the practicing physician? Unfortunately, historically the experimental use of some of these by those who had very limited fundamental understanding of them and who were bold enough to exploit them before fundamental studies were made placed many such procedures in the category of quacks and quackery. Needless to say, these circumstances did not command the respect of those whose interest and active participation were needed in order to place this wedding of medicine and physics at the proper level to attract the attention of our best scientists.

Some pioneer biophysicists defined this interzonal field as the application of the theory, laws and technics of physics to the broad field of biology. Knowing little about the hidden and intricate mechanisms of living things they expected to find obedience to the laws of nonliving things. They knew well the laws of nonliving things and found great satisfaction whenever they discovered conformity to these laws. Perhaps, in some instances, such obedience occurred when the conditions of experimental investigation altered the natural function. Unfortunately, the natural function was not always known and hence the interference with it by their experimental procedures was not recognized fully. Some investigators assumed that they had not altered the naturally performing functions significantly.

To search for the application of laws originating from inanimate things when beginning a study of living organisms seems to be investigating with preconceived ideas, which may impede the discovery of new laws—nature's laws for living things. The true biophysicist wants to obtain new information about the physics of living things, rather than to relate the conformity of living organisms to the already well-known laws of the nonliving. The pure scientist is more interested in new knowledge than in instances of conformities to old knowledge. Biophysics (bio-engineering or medical electronics) will emerge from adolescence when physical societies voluntarily invite sessions on the physics of living matter into their programming, thus recognizing the contributions to the field of physics.

Unfortunately, biologists have not demonstrated enthusiasm in trying to find new laws that might resolve some of their most basic problems; instead, they have been willing to accept the present-day laws of physics and chemistry. With introduction of the new electronic technics and with more rigorous thinking, perhaps they will seek more energetically for formulation of new laws.

Investigation of the biologic effects of various physical agents, an ever-present division of research in biophysics, has been very important therapeutically as well as academically. It has led to outstanding contributions to the life sciences, some of which have earned the Nobel prize for the investigator. The results of research on the biologic effects of the various types of electromagnetic radiations, namely microwaves; infrared, light and ultraviolet rays; X-rays; gamma rays; and secondary cosmic rays, have all been more interesting to the biologist than to the physicist or engineer. This type of research has added little new knowledge to the fields of physics and engineering. However, it has demonstrated clearly to the biologist the way of attacking a biologic problem by physicists or engineers who have been well trained in the fundamental methods for obtaining a quantitative understanding of complex problems in the physical sciences.

Pitfalls in the Development of a New Interzonal Field

All the man-separated disciplines in the broad field of the sciences are gradually coalescing. The life sciences seem to be entering the center of the stage, and because of the more penetrating studies in the dynamic aspects of the life sciences, the physical sciences are called in to play basic roles. In the achievement of such a coalescence, certain pitfalls must be avoided. One pitfall is that of the investigator who is competent in one discipline of science and assumes that he can just "read up" in another discipline and proceed to enter upon a research problem; this has occurred in biophysics particularly. Before schools for studying biophysics had been established, the physicist who had decided to become a biophysicist assumed that he could begin a research project by bringing his tools (instruments, measurement technics, and theoretical as well as experimental methods of attack) to the biology laboratory. Since some biologists were not competent in the fields of mathematics, physics and engineering, it seemed that fairly rapid progress could be made in this interzonal field when the self-styled biophysicist entered. However, the true researcher soon realized that the problem was more difficult than he had anticipated, that the many variables in the biologic fields could not be controlled so readily, and that he could not learn biology by just "reading up" on it. He gradually realized that the biologist was doing exceedingly well with what appeared to be less modern technics.

This early method of attack found more than one opportunist working in a "no-man's-land," where the biologist could not understand the biophysicist because of his limited training in physics and where the physicist could not understand the biophysicist because of his limited training in biology. Such an opportunist would conduct an investigation either theoretically or experimentally and submit for publication a manuscript that was difficult to review critically because reviewers equally competent in both biology and physics were not readily available. Such a state of affairs placed the science of biophysics on an insecure basis which it has taken years to correct.

Very few scientists today are competent to do research single-handedly in interzonal fields without reaching their limit in any of the various disciplines required for the complete and correct solution of the given problem. The solution of an interzonal problem, therefore, demands group research. Group research when well organized and properly directed can achieve remarkable results. Although original ideas do come from individuals, the dissemination of these ideas through the generous sharing of them enriches the world of science. Selfish individuals do not belong in the great field of group research—selflessness is needed for objective team work. Nor can the situation be dominated by a "master mind"—an individual who presumes to have superhuman insight and who regards other members of the team as his assistants, his technicians. Such conditions attenuate or suppress completely the potentialities of the talents and responsibilities of the "assistants," who will not cooperate under these circumstances in the search for pay dirt in the digging processes. A most necessary requirement for group research, however, is a competent coordinator with true wisdom and with ability to finger-point highly significant contributory facets to the new knowledge, thereby gaining a quantitative understanding of biologic systems. The correct coordination of these facets may lead to the recognition of the particular biologic system.

Instrumentation

The list of present-day instruments for research in bio-medical electronics is truly fantastic. The complete history of almost any component part of these instruments would prove to be intensely interesting to any reader when written by a person gifted for describing details in an entertaining manner. The history of such component parts as the vacuum tube, the transistor, waveguides, magnetic tape, transducers, thermistors, capacitors, transformers, relays, delay lines—to mention only a few—would proclaim the ingenuity, know-how skill, and the many years of drudgery in research, academic and applied, of many who played a role in the evolution of today's product. Electronic instrumentation demands high-level competency in circuitry, in lay-out design and in meticulous insight and care in construction. The final instrument is truly a work of art.

Computers that simulate memory and intelligence, electroencephalographs that measure the electrical characteristics of the central nervous system, electrocardiographs together with the most recently developed ballistocardiographs that measure the performance of nature's amazing circulatory pump, electromyographs that portray fundamental information relative to living servomechanisms—these are only a few of the instruments that are being applied in the search for a quantitative understanding of living systems.

Some research biologists may not be favorably disposed toward these apparently complex instruments. They prefer to use simpler and more direct methods. Obviously, a simple and direct method is the ultimate goal in experimental procedures, and major contributions to science have been made by such methods. However, the phrase "simple and direct" is relative and depends upon the attitude of the researcher and the problem to be investigated. The living organism itself is the most complex system in any biologic research problem. This living organism must dominate the experiment at all times and not play a secondary role when intricate manipulations are required for operating the electronic instruments.

Instruments, even the most complex, are only a means to an end, and represent a wasted effort if not used. The instrumentologist's responsibility is not terminated when he delivers his instrument to the biology laboratory. He should be eager to know that it is performing its duty efficiently in the retrieval of knowledge. The proper use of any instrument should be the major concern of the manufacturer. An electronic instrument is the culmination of the efforts of many great scientific minds coordinated and determined hopefully to add new knowledge for a higher level of civilization. It is not static but is ever improving according to new component parts and materials. Electronic instruments become obsolete with the advancement of new materials, new discoveries that may replace the older component parts, new ideas for circuitry, and the like.

Medical Electronics Advantageous to the Electronic Engineer and (as Well as) the Biologist

The advances in the physical sciences through research into nonliving phenomena have been so outstanding that whatever contributions have come from the biologic sciences to the physical sciences seem to have received little emphasis.[1] After all, it was the lowly

[1] K. S. Cole, "Four Lectures on Biophysics," Instituto de Biofisica, Universidade do Brasil, Rio de Janeiro, pp. 7–17; 1947.

frog that demonstrated a physiologic characteristic that triggered the idea for the galvanic or voltaic cell (the early source of continuous electromotive force). Two other examples of contributions to the engineering sciences (not directly to bio-medical electronics) are countercurrent exchange and the law of flow in tubes. The former phenomenon is nature's feat for cooling arterial blood. This mechanism permits the penguin to stand foot deep in ice water for prolonged periods without losing body heat. Before the arterial blood reaches the feet it loses heat to the proximal veins which are returning blood to the trunk. Engineers use an analogous system for cooling liquids flowing in pipes. The law of Poiseuille was born from a study of the circulation of blood in the capillaries of the mesentery of small animals. Poiseuille was a doctor of medicine who was a physicist at heart. His investigations prompted him to go to small glass capillary tubes for further studies. The absolute unit of viscosity, the poise, honors his name. He gave a fundamental law to physical sciences.

The animal world contains amazing systems which are attracting the attention of our competent engineers and physicists. If these systems could be incorporated into man-made systems, certain innovations would be achieved. A few very interesting living systems that I would mention are:

1) The sonar systems of bats and moths. These systems surpass the best sonar system that technologists have produced so far. Bats can detect their specific signals in the midst of intense noise.
2) The radar system of the electric fish, which deserves study for innovations in man-made radar systems.
3) The computer systems of several animals, particularly the amazing computer system of man, which are receiving intense study at the present time.
4) The communication system of many living infrahumans, which is attracting the attention of numerous investigators trained in the engineering and physical sciences.
5) Nature's semiconductors, which are demonstrating new aspects advantageous to the physical sciences.
6) Nature's amazing transducers such as the eye, the ear and the skin, which challenge the physical scientist who is attempting to make better transducers.

One of our leading American biophysicists has described one over-all contribution of the life sciences to the science of electronics very adequately in the following sentence: "The requirements of meaningful biological measurements have continually stimulated the development of electronic components and systems to new levels of sensitivity, stability, accuracy and general sophistication."[2]

Superficial review of the contributions of electronics to the life sciences indicates that the major contribution has been in the form of instruments. As stated previously, instrumentation is only a means to an end. Placing an instrument in a biology laboratory should not terminate the responsibility of the true instrumentologist who is seriously interested in the application of his instruments for obtaining reliable data—data that cannot be retrieved by other devices. He entertains hopefully a desire to make a contribution for beneficent purposes.

Living organisms possess complex living systems. Engineers are trained in systems engineering, so biologists have turned to the engineer to search for a quantitative understanding of these complex living systems. The engineer as a co-scientist with the biologist will bring a new way of thinking to biology. He must become a full-time member of the group engaged in the solution of problems that require his skill and know-how for the correct and complete resolution of the given problem in the life sciences.

Organizational Difficulties

The science of biophysics is as old as, or older than, the man-named discipline, physics. Many pioneer physicists made notable contributions to the life sciences. However, throughout these many years a society of biophysicists was never achieved until the recent organization of the Biophysical Society. Perhaps this prolonged delay in organization was caused by the lack of a unifying factor, that is, lack of a common denominator for the widely divergent activities of scientists who called themselves biophysicists. Another fundamental factor which may have contributed to the delay in organizing a society was the aloofness of the medical profession toward technical assistants (their "nonprofessional" associates).

An organization that preceded development of the Biophysical Society by about 5 years is the IRE Professional Group on Medical Electronics (presently named the Professional Group on Bio-Medical Electronics). The rapid growth of the IRE Professional Group on Medical Electronics, which was conceived in 1951 and born in 1952, indicates the intense interest and general activity in this interzonal field of science. The membership in this professional group today exceeds that of the American Physiological Society, which held its semicentennial celebration in 1938. The innovation of forming professional groups within the Institute of Radio Engineers was a profound and brilliant idea for maintaining the togetherness of a society which today

[2] Personal communication from K. S. Cole, Natl. Institutes of Health (NINDB), Bethesda, Md.

is rapidly approaching a membership of more than 85,000. The historical background of this idea is described well by Van Atta.[3] The origin of the Professional Group on Medical Electronics has been described completely by Montgomery,[4] who is to be commended highly for organizing this group. Perhaps organization of the IRE Professional Group on Medical Electronics may have functioned as a catalyst in bringing about the subsequent organization of other societies (in this interzonal field) which were so difficult to organize previously. For our readers who may wish to have more detailed information concerning the convergence of the two widely divergent fields of knowledge, namely physics and biology, the author highly recommends some intensely interesting papers written by Schmitt.[5–7]

A broad view of the history of bio-medical electronics points to electronics as a common denominator that has been a dynamic agent for converging the divergent sciences of physics, chemistry, engineering, biology and medicine. The science of bio-medical electronics has a brilliant future.

[3] L. C. Van Atta, "The role of professional groups in the IRE," Proc. IRE, vol. 38, pp. 1124–1126; October, 1950.

[4] L. H. Montgomery, "The origin of the professional group on medical electronics," Proc. IRE, vol. 47, pp. 1993–1999; November, 1959.

[5] O. H. Schmitt, "The Emerging Science of Biophysics," unpublished data.

[6] —— "Where is medical electronics going? A symposium in prediction," 1956 IRE National Convention Record, pt. 9, pp. 107–112.

[7] —— "Biological servomechanisms and control circuitry," 1954 IRE National Convention Record, pt. 9, pp. 34–41.

Electronics in Clinical Research*

R. STUART MACKAY†, SENIOR MEMBER, IRE

Summary—**Engineering training can profitably be applied to medical problems, both for better understanding of biological mechanisms and for development of more effective clinical methods. Also from clinical investigations, one can derive concepts applicable to nonbiological systems. Examples of developments from clinically-oriented research are cited to show some of the possibilities and problems in this field.**

Introduction

ELECTRONICS as a tool (the instrumentation aspect) enters most modern technology. In recent years, it has come to be considered as proper for electrical engineers to apply their talents and training to the problems of biology and medicine as to, say, those of computation or communication. In addition to serving as a useful tool, understanding of basic biological processes can be furthered by those with training in medical electronics or biophysics. Thus, generation of nerve impulses or the processing of sensory information seem best understood by those with some such training. A more specific example is how feedback theory gives us an understanding of the clinical observation that strychnine produces convulsions by depressing post-synaptic inhibition, and thus removes stabilizing negative feedback. Conversely, useful concepts in the technology of computation and complex system engineering have originated after study of biological systems. We will have some distance to go before we can approach the perfection of the frequency-modulated sonar of some bats, or the subtlety of performance of the human nervous and other regulatory systems. It seems likely that it is in some of these hybrid fields that the largest steps and most rapid progress are to be expected.

By electronics, we may mean circuits containing components in which the mean path of the current carriers is comparable with the physical dimensions of the elements. These can amplify (*e.g.*, thyratrons, transistors, tubes) and give rapid responses. Thus, they are useful components of systems for monitoring electrical signals accompanying various bodily processes. Also, many physiological variables are best measured using electronic techniques in conjunction with transducers of pressure, blood velocity, temperature, gas concentration, etc. Clinical procedures can often be considered either diagnostic or therapeutic, and there are also therapeutic methods which depend upon the effect of electricity, radiation, or ultrasonics on the human body.

In this short space, it is possible only to cite a few examples and problems from this diverse field, and to present some of the exciting future possibilities. Results of investigations most directly applicable to the treatment of patients will be stressed. Of course, more basic

* Received by the IRE, July 5, 1961. The work reported in this paper was supported in part by Public Health Service Grant B-2333.
† University of California, Berkeley, Calif.

investigations often require the most sophisticated instrumentation and observation interpretation, with results becoming incorporated into routine procedures. For example, a method for measuring the electrical negative resistance of nerve fibers gives one a tool for investigating the effects of poisons or anesthetics, and possibly eventually a clinical measure of nerve function in special cases.

Tonometry and Pressure Measurement

Of all people over 40, several per cent are presently going blind from glaucoma, a major cause of irreversible adult blindness. Here, an increase in pressure within the eye, if not detected and treated, will be followed by permanent blindness. Devices for detecting this increase in pressure are called tonometers, and the introduction of electronic methods seems likely to revolutionize their use and spread their application to a larger section of the population. The difficulties have been that the procedure is cumbersome and somewhat inaccurate, and furthermore, anesthetics had to be applied to the eye in order that the patient tolerate the hardware placed upon his eyeball.

There are two new principles. The first might be called acoustic tonometry. This is suggested by the observation that increasing tension in a drumhead causes a rise in frequency. Similarly, increasing pressure in the human eye might give it a higher characteristic frequency if set vibrating. One might measure the acoustic impedance of the eye by placing into contact with it a core driven by a coil fed from an ac source. From acoustic impedance, one obtains two numbers corresponding to real and imaginary parts. One is related to the springiness of the eye and the other to viscous and other losses within the eye and associated vessels. The latter might be correlated with clinical abnormalities, but the former is of interest here.

The magnitude of this "springiness factor" can be estimated as follows.[1] If a flat plate is pressed against the surface of a sphere (the eye), then an increase in area of contact proportional to deflection results; for a small deflection, there will be little increase in pressure within the sphere. Simple geometry indicates a force increasing directly proportional to deflection, and thus release of the plate should yield simple harmonic motion, with the eye acting as a perfect spring. The frequency of oscillation should be approximately independent of vibration amplitude or initial pressure:

$$F = \sqrt{\frac{RP}{2\pi M}}$$

where M is the mass of the flat-ended plunger momentarily pressed against the eye, R is the radius of the eyeball, and P is the pressure within the eye.

In Fig. 1, this is seen to give the correct order of magnitude in the normal range of pressures, with some spread in the readings. Each point shows the frequency of maximum vibration amplitude, for a given pressure, as the exciting frequency to a solenoid driven plunger was varied. Response amplitude as a function of frequency at any one pressure indicates a Q factor of approximately four. There are deviations from ideal performance due to damping and nonlinear bending forces, and so this method seems most applicable to basic studies of the visco-elastic properties of the cornea.

Classic tonometers, and their electronic relatives using a transducer rather than a lever to measure motion, have either measured the force required to flatten a predetermined area of the eye against a transparent plate, or else have measured the indentation of the cornea by a plunger of known weight. The latter procedure doubles the pressure within the eye. Sources of error in tonometers are variable corneal stiffness opposing probe approach, tissue tension tending to expel a plunger, surface tension of tears, variable eye size, the variable curvature of astigmatism, corneal scarring, etc. None of these affect the reading of the instrument

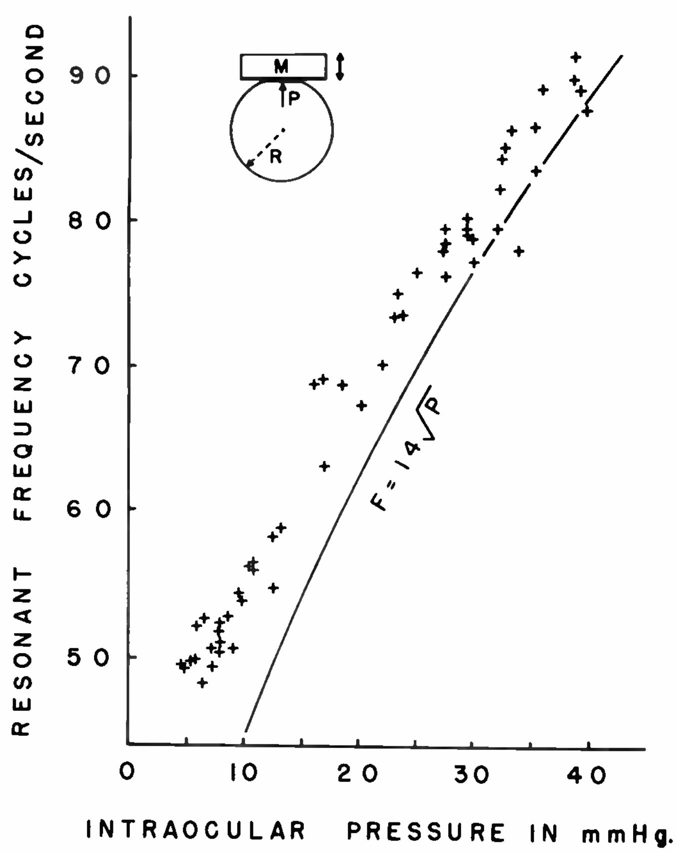


Fig. 1—Experiment of plate vibrating against eye. Plot of characteristic frequency vs intraocular pressure observed with five rabbit eyes of approximately uniform size. An idealized theoretical curve neglecting the damping term and elastic forces is included. Adjustment of initial force and correction for finite Q improves fit.

[1] W. K. McEwen, D. M. Maurice, R. S. Mackay, M. Margolis, R. Vreeland, R. St. Helen and M. J. Hogan, "Investigations of the Mechanical Properties of the Eye, Part I. Acoustical Tonometry of Enucleated Rabbit Eyes," unpublished study, 1954–1958.

of Fig. 2, according to first-order theory.[2] Here, the eye is momentarily flattened to beyond the pressure-sensitive center of the probe (1.5-mm diameter), and the only force exerted on the plunger is the intraocular pressure, if the plunger end does not deviate from the plane of the surrounding ring. Any minute motion of the plunger is sensed by the variable inductance displacement detector, and this signal is amplified and fed back to an electromagnet to supply an increasing force to return the plunger. Magnet current is a direct measure of applied force and, if the gain of the feedback system approaches infinity, then the resulting displacement will approach zero. The measurement is absolute and, since there is no plunger motion, field nonuniformities do not introduce nonlinearities in the scale, and friction problems are minimized. Recent models have used a sensitive differential transformer as the displacement detector (deflection of a half micro-inch per millimeter mercury pressure) and a relatively stiff spring rather than feedback to oppose motion. Silicone rubber here pots the transformer core within the transformer to supply support, bearing, and spring. Acceleration insensitive units also have been constructed from a small quartz crystal mounted flush in a plate. The entire observation can be made in any orientation in a second without anesthetic (a thin layer of liquid remaining between probe and eye probably reduces sensation). The last advantages are important for drug-testing animals and for use with children.

Too much probe force raises the eye pressure, but there must be enough force to flatten the cornea. As the probe advances against the eye, the output signal is observed to rise and then drop somewhat before rising again. If the cornea suddenly buckled (like a bistable oil-can bottom), then the intraocular pressure would be measured by the first crest height. Experiment proved

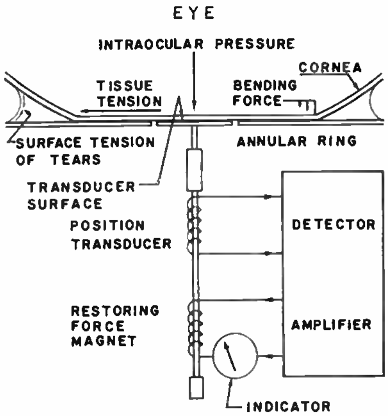


Fig. 2—Flattening the eye to beyond the pressure sensitive plunger assures intraocular pressure alone being measured. Tendency toward motion is opposed by the feedback current, which thus measures the pressure.

[2] R. S. Mackay, E. Marg and R. Oechsli, "Automatic tonometer with exact theory: Various biological applications," *Science*, vol. 131, pp. 1668–1669; June 3, 1960.

the alternative explanation, involving the interaction of the bending forces before the flattened area extends beyond the pressure sensitive area. Pressure then is measured by trough height, and the excess height of the first crest measures corneal stiffness. Information about plastic effects and fluid outflow rate from the eye's anterior chamber is provided by comparing corresponding points on the curve produced while the probe is advancing and then receding.

Ease of fluid outflow from the eye can be inferred by comparing the indication above with a measure of area of flattening (the force on the handle or on the "guard ring" alone). With the newest methods, the noise level before contact is limited by brownian motion in the transducers.

Application off of center allows measurement while looking into the eye. Pressing variably until one sees the pulse start and stop (ophthalmodynamometry) allows estimation of the arterial blood pressure in the retina, which is close to the pressure in the internal carotid artery, of interest for cardiovascular diagnosis. A difference between eyes indicates occlusion in the vessels.

This principle is applicable to the measurement of other pressures such as within the head of a reclining infant, inside a bladder, uterus,[3] or within an intact blood vessel. The thickness of the wall to be flattened determines the probe diameter. In cases of enlarged spleen, this technique may approximate the pressure in the portal system, which observation presently necessitates a serious surgical procedure. It may prove useful to evaluate muscle tone by pressing the tonometer against the bundle. While with A. A. Ismail in Egypt, some observations with a larger unit were made on the abdomen and its ascetic fluid. People there suffer from bilharzia and may display a distended abdomen due to the presence of a dozen liters of fluid, releasable by puncture. Such a unit applied to the top and bottom of the abdomen of a seated subject indicates an appreciable difference in reading due to the column of fluid. On the other hand, a subject suffering from heart failure may show a distended abdomen due to lack of rapid elimination of swallowed air (nitrogen). Then the two tonometer readings are about the same, and the subject should not be punctured, but rather placed in an oxygen tent. In this application, nonuniformities in the abdomen wall (fat lumps) may give an elevated reading. A plunger with a swivel action with respect to the plane of the surrounding ring may overcome this difficulty. All these applications use the consequences of the simple fact that the pressures on two sides of a flat surface are the same.

The feedback tonometer, like any regulator, can be adjusted to oscillate. It then acts as a negative resistance coupled to the eye through a bilateral force transducer. Oscillation frequency is observed to increase with intraocular pressure, as expected.

[3] C. N. Smyth, "The guard-ring tocodynamometer," *J. Obstet. and Gynaecol. Brit. Empire*, vol. 64, pp. 59–66; February, 1957.

Oculography and Hangovers

The human eye acts like a battery in that there is a relatively steady dc voltage existing between the cornea and the retina, making the front positive with respect to the back. Thus, if a person rolls his eyes to the left, the left side of his head will become a few hundred microvolts positive relative to the right side. One has the action of a potentiometer, the magnitude of these voltages being proportional to the angle of the eyes from straight forward. By recording these voltages from surface electrodes placed around the eye sockets, one can determine where a person is looking. The method has been used to study the fashion in which people observe a scene, and the regions of special importance, in subjects ranging from advertising posters, circuit diagrams, and airplane cockpits, to pictures of pretty girls.

Some of the more remarkable applications of the technique have been in clinical and basic studies of the central nervous system and the effects of drugs thereon.[4] Nystagmus is a vibratory motion of the eyes, which may rotate to one side and then snap back cyclically (*e.g.*, when an inexperienced person spins on the dance floor). It is observed as a result of vestibular lesions. Another form follows consumption of alcohol. This last usually occurs with closed eyelids, but by this electrical method one can tell where a person "is looking" though the eyes are shut. A typical experiment involves drinking a measured quantity of whiskey and then laying down on one's back with head turned to one side. Soon it is recorded that the eyes beat to the right when the head is to the right, and vice versa. This continues for about three hours, followed by one to two hours of no abnormal tracings. Then, for 10 to 20 hours (long after all alcohol has left the body), a second phase of vibration is found. Jerking of the eyes now is in the opposite direction relative to head position. It is at this time that the person reports the subjective feeling of a hangover. Drugs that interfered with this second stage of eye activity were tested[4] and found to interfere with the sensation of hangover; *i.e.*, unpleasant symptoms became less marked. The effect of a new alcohol dose ("pick-me-up") was studied similarly.

Circuits are being designed to analyze automatically a tape-recorded nystagmus pattern for degree of activity. They will be used to study the site and mechanism of action of certain drugs on the central nervous system. One of the major problems for workers in medical electronics is the reduction and understanding of large masses of data.

The effects of drugs on motor activity and the impairment of coordination due to hypoxia, etc., can be studied using classic tests such as standing steadiness, finger-finger touching,[5] etc. These can all be made more sensitive by incorporating electronic measurements and, if the signals are recorded in electrical form (magnetic tape), then results are subject to automatic analysis.

In connection with the human eye, other work might be mentioned. Various television systems have been designed to measure pupil diameter. It is a significant clinical entity, since the pupil pathway goes through many important areas of the brain. It has been known that a spot of light imaged on the edge of the iris of the eye produces oscillation of pupil size. Recently,[6] this has been analyzed in servomechanism terms. Oscillation frequency depends on certain physiological variables, and such an effect might be used in testing drugs, quite aside from its basic interest. Similarly, one might investigate oscillations produced by feeding back eye position to an oscilloscope spot being observed, after administration of drugs. Biological observations and theories often yield practical medical applications.

Radiology

Of the clinical specialties, possibly radiology is most susceptible to physical analysis and reduction of empiricism by the application of information theory and biophysical investigation. This applies to the production of a radiogram to give maximum information with minimum dose for diagnostic purposes. The analysis of a radiogram is an extremely complicated procedure, and even the radiologists themselves cannot say exactly what they do when they, for example, read a chest film. It is much more than, say, comparison of the two lungs. Even following the pattern of eye motion during reading does not give too much help, and conception of an automatic chest film reader is an extremely difficult problem in pattern recognition. (Some success with the cancer cell screening problem has been achieved.[7]) However, the production of the radiograph itself could probably be made automatic with the present state of our knowledge.

Conditions for minimum dose are 1) a constant intensity image,[8] and 2) either a monochromatic X-ray source or a color display or equivalent[9] (all require electronic methods). Intensifiers have absolute limits that prevent indefinite improvement.[10] Electronic, and especially television techniques, have given the radiologist a step forward in recent years. Television techniques allow the production of bright multiple images

[4] G. Aschen, M. Bergstedt and J. Stahle, "Nystagmography," *Acta Oto-Laryngol.*, suppl. 129; 1956.

[5] L. Goldberg, "The Effects of Drugs on Motor Activity, from Quantitative Methods in Human Pharmacology and Therapeutics," Pergamon Press, New York, N. Y., pp. 88–97; 1959.

[6] L. Stark, "Stability, oscillations, and noise in the human pupil servomechanism," Proc. IRE, vol. 47, pp. 1925–1939; November, 1959.

[7] R. C. Bostrom, H. S. Sawyer and W. E. Tolles, "The Cytoanalyzer—An automatic prescreening instrument for cancer detection," in "Medical Electronics," C. N. Smyth, Ed., Iliffe and Sons, London, England, pp. 479–487; 1960.

[8] R. S. Mackay, "Constant intensity image for minimum dose radiography," *Brit. J. Radiol.*, vol. 31, p. 642; November, 1958.

[9] R. S. Mackay, "X-ray visualization and analysis of multicomponent subjects," *Science*, vol. 128, pp. 1622–1623; December 26, 1958.

[10] B. Jacobson and R. S. Mackay, "Radiological contrast enhancing methods," in "Advances in Biological and Medical Physics," J. H. Lawrence and C. A. Tobias, Eds., Academic Press, New York, N. Y., vol. 6, pp. 201–261; 1958.

remote from scattered radiation. By viewing the fluorescent screen on which a fluoroscopic image appears with a television system, it is possible to observe finer detail with more sensitivity without dark adaptation, and thus with less dose to patient and physician. This is not only an aid in routine clinical procedures, but is of help to clinical instruction where a whole class can follow the course of an examination on remote multiple viewing screens.

Furthermore, the image can permanently be recorded in the form of an X-ray motion picture. Since television output brightness can be increased as desired, the radiation exposure of the patient can be determined by the fineness of detail to be observed, rather than by the sensitivity of available photographic emulsions.[11] A movie is extremely useful for instruction, research, and in normal practice. It is impossible to look everywhere at once, and the rerunning of the moving picture of an examination tends to bring out some new aspects on each successive viewing. The recording of the video image on a magnetic tape can also result in less patient dose than if an ordinary movie camera were used to photograph a television monitor.[11] The possibility here of immediate viewing and reviewing of the sequence is important. Storage of the image in electrical form is also useful in that appearance can be modified during playback to the optimum by altering contrast, etc. The single-frame viewing possibility of the newer helical-scan recorders is very useful in radiology. An application example is the use of mouth and throat pictures to aid a subject in speech therapy, to determine advisability of surgery for malformation, and to make a normal bodily function graphic in anatomy classes.

A film or recording can contain subtle pieces of information unnoticeable to the human eye, which requires brightness differences of roughly one per cent for detection. Simple contrast enhancing circuits can unveil these. Mixing with a signal some of its own derivatives emphasizes edges or contours. Restricting the low-frequency response of the television system increases the contrast in fine structure without slow variations in intensity driving part of the over-all system out of the useful range (equivalent to "logetronics"). These methods[10] have proven clinically useful. Actually, a kind of edge enhancement due to interaction between adjacent receptors may normally sharpen all senses to their stimuli. One attempts the opposite of edge enhancement in "removing" spots from clothing. A slight motion of an object can make structure more noticeable because of the effect on the retinal receptors.

The graph of the variable density pattern in a single subject plane (single television line display)[11] may help clinical accuracy and be useful in making radiology more quantitative. Also, one can see smaller density variations as a deviation in the density plot than one can observe directly on a television monitor or radiograph. To do this, one must restrict the high-frequency response with a condenser to reduce noise; this also cuts out the finest detail, which provides another example of the general rule[10] that one can trade resolution in brightness for geometrical resolving power (by integration over a different area size). Contrast enhancement is effective only if there is no detail of interest finer than a specified amount, for similar reasons.

Some blurring is inevitable in forming any image. If one knows the way in which any point on the object becomes blurred in the image, and if one has the final image, then by a suitable mathematical transformation, one can say what the object actually looked like. This deblurring process does not disobey any entropy or information theory considerations, and it can be approximated in some cases by mixing with a signal a suitable selection of its derivatives. Thus, one might build a high intensity X-ray tube using a broad focal spot that would yield somewhat blurred images. These images would be blurred in a known way and thus, in principle, could be rendered sharp by after-treatment. However, one would pay a high penalty in signal-to-noise ratio because of the emphasis by this process of high-frequency noise.

Besides edge emphasis and deblurring, such procedures might also be used to emphasize structure in the image in a particular range of sizes. One can do this using television techniques, but it is interesting that these procedures can also be carried out by optical methods that draw on electrical engineering filtering concepts. An image (subject) placed in the optical system of Fig. 3 is considered as made up of various Fourier components in the vertical direction. One is shown acting as a diffraction grating. The high-frequency components corresponding to fine detail deviate the output monochromatic light beam more than the low-frequency components. A diaphragm or stop placed before the final camera can be used to reject all structure of a particular spacing or size, *i.e.*, to act as a filter for a particular object periodicity or frequency. A suitable image there could constitute a matched filter for a given subject. By using long focal length lenses, these diffraction effects can be observed with normal-sized photographic images. One can approximate deblurring action by the edge emphasis mentioned above, and this has been tested in preliminary experiments by placing before the camera a graded stop that is relatively dark at the center, and which gradually drops off in density toward the edge. The radial rate of dropoff of opacity is the same in some cases as the frequency response of the corresponding differentiating circuit in a television system. Optical filters can be more arbitrarily determined than electrical ones, though negative signals can cause problems.

Viewing a television image of a rotating object with an extra time delay to one eye gives a true three-

[11] R. S. Mackay, "Television cineroentgenography: Dose and contrast factors," *Am. J. Roentgenol., Radium Therapy and Nuclear Med.*, vol. 85, pp. 342–351; February, 1961.

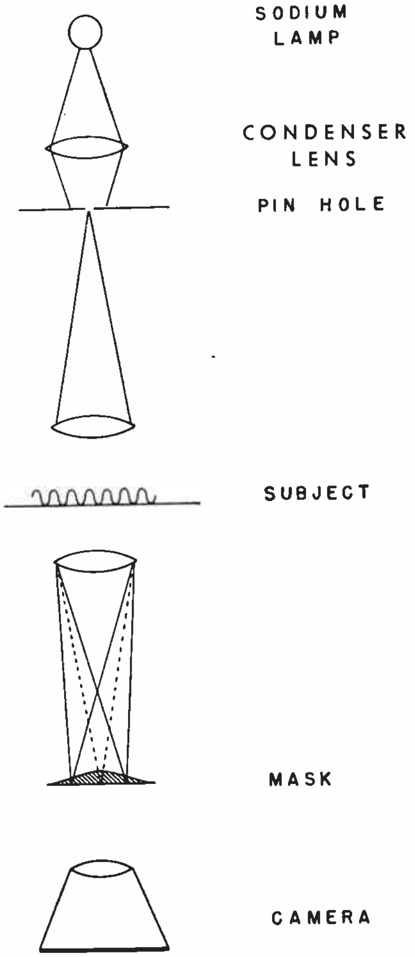


Fig. 3—Optical system capable of filtering out selected frequency components in a subject to emphasize edges or elements in a particular range of sizes. The camera photographs the object, one frequency component of which is shown producing diffraction.

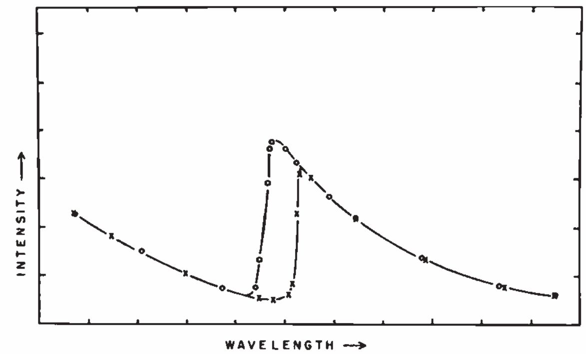


Fig. 4—Intensity of X rays transmitted by one or the other of two elements. For iodine and barium, the K edges fall at 0.373 and 0.331Å (33.2 and 37.4 kv), respectively. Falloff outside this region is largely determined by the spectrum emitted by the tube.

dimensional image, and this can be done by simply placing a neutral filter before one eye.[12] The accompanying random noise pattern then takes on a depthy swirling-cloud appearance. This may provide a way to study either the statistics of noise, or binocular depth perception using patterns without cues other than binocular parallax.

The presence and distribution of one element may be important, *e.g.*, of iodine introduced into a body cavity to outline it, or pulsed into the heart for visualization of blood flow pattern. Special X-ray equipment using the spectral information in an image, in addition to the usual lumped absorption pattern, permits working with smaller quantities. The graph of absorption of an element as the incident X-ray frequency (kilovoltage) is increased displays a sudden drop in transmission as the quantum energy becomes sufficient to dislodge K electrons (Fig. 4). Two elements can be adjusted so that their absorptions coincide everywhere except between their K edges.[13] An image of an object formed from an X-ray beam filtered by one element will differ from the image formed by filtration with the next higher element, mostly in those regions where there was present some of the first element, since all other elements in the subject will have similar absorptions in the two transmitted spectra.[14] For example, the distribution of iodine within a subject might be obtained by subtracting images produced with iodine and with barium filtration. Spectral information is not wasted as in ordinary radiography.

Two photographic images are not sufficiently comparable, and a detector-memory combination of high precision seems needed for subtraction. A television tube can be used as the detector (Fig. 5) with a spinning filter pair, and then any given detection point need only maintain constant sensitivity for the time of a half-revolution (say 1/30th second). A property of the visual system allows dispensing with an electronic storage system and differencing mechanism for the two images. Flicker is seen in two light distributions alternating at low speed in those regions where the intensity is different. Alternatively, the two images can be compared by conversion to color, where local differences in intensity appear as a shift in hue. In preliminary experiments, differentiation by color was demonstrated, though no evaluation of ultimate sensitivity has yet been made. The filter pair consisted of solutions of barium chloride and sodium iodide to avoid possible diffraction effects, the second ion having remote characteristic absorptions and apparently tolerable unbalancing action. Three filters might similarly effectively isolate two wavelengths for a three-color display.[14] For better resolution and lower dose, the X-ray tube can be placed to one side to produce characteristic secondary radiation that then traverses the subject from a fluorescent source incorpo-

[12] R. S. Mackay, "Three-dimensional movies without special equipment," *Science*, vol. 119, pp. 905–906; June 25, 1954.

[13] P. Kirkpatrick, "Theory and use of Ross filters," *Rev. Sci. Instr.*, vol. 10, pp. 186–191, June, 1939; vol. 15, pp. 223–227; September, 1944.

[14] R. S. Mackay and C. C. Collins, "Color X-ray images and enhanced contrast," *J. Biol. Phot. Assoc.*, vol. 25, pp. 114–120; August, 1957.

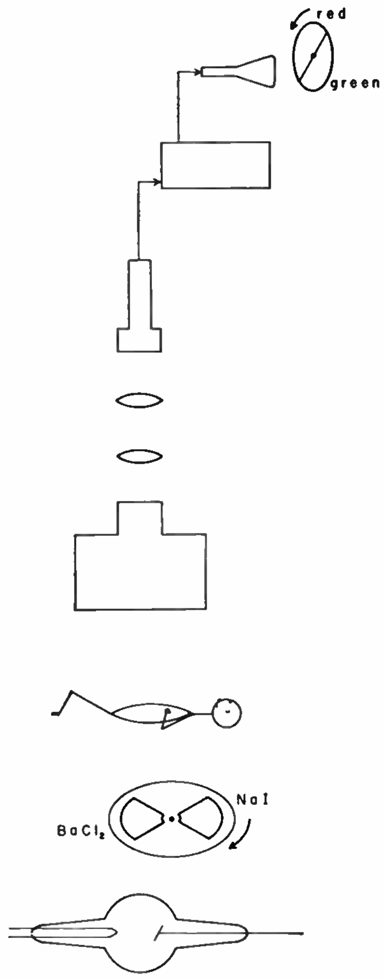


Fig. 5—Television system for visualizing the distribution of iodine (not necessarily radioactive) in the body. The source preferably has characteristic lines on both sides of the iodine absorption edge.

rating cesium, barium, or lanthanum; a special tube with a barium sulphide target would have a similar effect.

The efficient television detectors that are now almost exclusively used show storage, since at each point the effect of each interacting quantum is added to that of all the rest that arrive there in the interval since the previous frame scan passed that point. Then, the total signal is suddenly released as scanned. In Fig. 5, the change of filters, flyback, and then pulsing the X-ray source must be almost simultaneous to avoid mixing "illuminants"; source pulsing might be replaced by having the line of scanning out information from the orthicon just follow the image of the line of demarcation between X-ray wavelengths as it advances over the subject. Intensifier persistence and noise both limit performance.

The advantage of this method is that special equipment over that already found in hospitals is minimal. A lower radiation dose method using a special X-ray tube with rotating secondary emitters to produce two alternating monochromatic beams, has been demonstrated and it gave 100 times the usual radiographic sensitivity to the presence of iodine,[10] while also providing ***in vivo*** quantitative chemical analysis.

In the future, remote vision and hearing might actually be augmented by remote feeling, in the fashion of remote manipulators, to allow remote diagnosis and therapy. Certainly, a central pool of specialists for referral is possible, the most immediately practical one being of radiologists to read films presented over a television link.

The Heart and Learning Filters

Electrical signals which can be picked up from the surface of the body accompany the action of muscles. For example, the electrocardiogram is a display of the voltage picked up by electrodes as the heart muscles work.[15] Ramifications employ various electrodes, three-dimensional displays of the potential pattern or vector displays of the equivalent dipole, displays of derivatives or integrals of the potential, autocorrelation functions, etc. Abnormalities warn of arythmias or malfunction. The foetal heart deserves special attention, not only as an index of well-being, but because the signal may warn of a multiple pregnancy.

It is now possible to study the onset of conditions such as heart failure, or in general, to make a continuous recording of the start of an event that may only happen once at an unknown time. A patient with expected difficulties can be fitted with electrodes and a frequency-modulated transmitter to telemeter his electrocardiographic potentials to a laboratory for recording as he moves about freely. The first of two tape recorders runs continuously and has a loop of tape corresponding to a time of a few minutes per turn; this passes through an erase head and then a nearby record head. At all times on this loop of tape are recorded the previous few minutes of heart action. At any time, if irregularity (for example, a significant variation in frequency) is noted, this first recorder is stopped and the unmodified second recorder is simultaneously started to carry out a prolonged recording of the subsequent action in the usual way. From these two recordings, one can study the onset and subsequent history of any irregularity.

Heart action often is studied by listening to chest sounds with a stethoscope. Fidelity is improved by employing an electrical transducer and amplifier. Thus far, acceptance of this has been poor, apparently not because of greater complexity or cost, but because too much unfamiliar signal is produced. Another method of circulatory system study relates to the fact that everyone at some time has lain in bed and heard the springs squeak in time to his heart beat. (Some then decide they are having heart trouble, and the squeak becomes faster and louder.) Quantitating this effect with motion

[15] O. H. Schmitt, "The biophysical basis of electrocardiography," *Proc. 1st Natl. Bio-Phys. Conf.*, Columbus, Ohio, March, 1957, H. Quastler and H. J. Morowitz, Eds., Yale University Press, New Haven, Conn., pp. 510–562; 1959.

transducers has opened the field of ballistocardiography.

In both the electrical and the acoustical situation, the filtering out of extraneous noise is helpful. One might consider the optimum linear filter for these purposes, or better, the optimum nonlinear filter. Gabor has discussed filters that match themselves to a type of waveform through a learning process.[16] In essence, the matching process samples the sample functions at the Nyquist interval, *e.g.*, with a tapped delay line or a multiple-head tape recorder. A function involving internal machine coefficients and the various samples is formed. Internal coefficients are varied in a systematic way until the mean-square-error difference between the function and the desired output (the task) is a minimum, this being taken over the entire ensemble of inputs. The process produces a match.

The nonlinear filter might adjust itself to some abnormality. In principle, one might take a long record of heart signals (sound or electrical, on tape) and produce the final message (task desired) by filtering out everything but the arythmias. The filter will learn, to the best of its ability, to give a response when the arythmia comes along. In the above double tape recorder method, the observation of the waveform and the decision that an irregularity is present can be made by a human operator. This is a normal function of the cardiologist who studies his patient's electrocardiographic records and then proposes suitable therapy. However, it seems possible that these filters could be designed to recognize the various waveforms and distinguish abnormal from normal ones. These would be useful both in clinical diagnosis, and to initiate the above-mentioned switching operation. In building any filters, one is not only limited by the technical possibilities of design, but one will inevitably build in his own bias as to what constitutes a normal and an abnormal pattern. In some cases, experienced clinicians cannot agree, and so the problem is not always a simple one.

If a variety of abnormal waveforms were repeatedly played into a learning filter, it would gradually (subject to certain orthogonality conditions), so to speak, pick out what was statistically similar in all of them and arrange itself to respond to this similarity. Besides adjusting itself to a match, it might actually extract some common factor or statistical regularity between the various records (input teaching samples) that would be overlooked by a human making judgments.

Recently, two groups described experimental work on adaptive waveform recognition and self-optimizing nonlinear filters.[17,18] Such a filter might also in part substitute for the eternal search for the transfer function of a person, or his parts, by acting as a human simulator. To cause such a filter to simulate a person, one would repeatedly play in suitable functions and simultaneously the person's actual response as the target function (task). After the filter had adjusted itself to simulate the response of the person, then other inputs could be tested. It might be noted that the transfer function of a person is not constant. The adjustment of a human transfer function to a task is demonstrated by engraving or dissecting under a microscope, wherein stability results though the gain may be 100 times higher than in normal freehand dissection or engraving.

These considerations apply to the recording of any of the other electrical potentials generated in the course of body functioning, whether they result from uterine contractions, brain activity, the response of the retina to a flash of light, or etc. Some such potentials have magnitudes less than the noise level (biological or amplifier) of the experiment, but if they are evoked by a controlled stimulus, then addition of a number of responses (*e.g.*, on a storage tube) gives a clear signal with the noise averaged out.[19]

The human heart is a large muscular structure of four chambers; it displays certain difficulties that can be treated electrically. An experiment to visualize one problem employs a jellyfish with its central part cut out to leave a ring of muscle and nerve. Stimulation of one point produces a muscular impulse traveling out in both directions around the circle to meet and cancel as each runs into the refractory state left behind the other. Squeezing the ring prevents passage of one of the two impulses, and this temporary block can be removed before the other impulse arrives there. Such a ring left sitting in a pan of seawater can conduct an impulse around and around for a week. A temperature rise increases velocity and diminishes the number of similar impulses that can simultaneously travel around the ring in one direction. Such a system is one of the newest suggestions for a computer storage and logic element,[20] using a thread of nerve analog.[21] In the heart, nonuniformities can result in small waves propagating themselves around the various chambers in a so-called circus movement that prevents a coordinated contraction from pumping blood. Either the auricles or ventricles can show this fibrillation, the chambers of the human heart being large enough to sustain such continuously varying circular paths. Ventricular fibrillation is fatal and can be induced by some anesthetics or by chest

[16] D. Gabor, "Communication theory and cybernetics," IRE Trans. on Circuit Theory, vol. CT-1, pp. 19–31; December, 1954.

[17] C. V. Jakowatz, R. L. Shuey and G. M. White, "Adaptive Wave Form Recognition," Res. Lab., General Electric Co., Schenectady, N. Y., Rept. No. 60-RL-2435E; May, 1960. Also presented at 4th London Symp. on Information Theory; 1960.

[18] D. Gabor, W. P. L. Wilby and R. Woodcock, "A Self-Optimizing Nonlinear Filter, Predictor, and Simulator," presented at 4th London Symp. on Information Theory; 1960.

[19] Communications Biophysics Group and W. M. Siebert, "Processing Neuroelectric Data," Technology Press, Mass. Inst. Tech., Cambridge, Tech. Rept. No. 351; 1959.

[20] H. D. Crane, "The neuristor," IRE Trans. on Electronic Computers (Correspondence), vol. EC-9, pp. 370–371; September, 1960.

[21] R. S. Mackay, "What is a nerve?" IRE Trans. on Medical Electronics, vol. ME-7, pp. 94–97; April, 1960; see also R. S. Mackay, "Negative resistance," *Am. J. Phys.*, vol. 26, pp. 60–69; February, 1958.

surgery irritation, and it often is the cause of death following electric shock. Though not spontaneously reversible, it can be interrupted with a massive electrical shock to the heart, following which, normal beating is usually resumed. Proper values of current and voltage are known and, with an exposed heart, "plugging into" the 110-volt, 60-cycle line for a fraction of a second is appropriate. Successful impromptu emergency defibrillation in a hospital has been produced by wrapping the exposed ends of a light fixture cord around surgical instruments applied to opposite sides of the patient's heart, and inserting the plug into the wall for an instant.[22] Closed chest defibrillation requires correspondingly more power.[22]

With heart block, the contraction of the auricles does not spread down to the ventricles, and the latter, perhaps contracting at their own intrinsic slow rate, may inadequately pump blood. The heart can be stimulated into periodic contractions by cyclically applying small shocks from a so-called pacemaker. This can be implanted, but if powered by a mercury cell, would require changing every five years.[23] Success in recharging endoradiosondes without interrupting transmission suggests powering a pacemaker by a rechargeable nickel-cadmium cell. Then it would only be necessary for the subject to sleep in the vicinity of a suitable radio-frequency oscillator once a week for battery recharge. A signal from the auricles might trip a relaxation oscillator that would pulse the ventricles, thus bringing them into synchronization, to match body needs. Investigation of a self-powered unit, in which a small transducer attached to the auricles would directly generate the voltage for the ventricles, has thus far yielded no results, nor have more complicated arrangements in which the relatively powerful contraction of the ventricles would store energy to be used in generating the pulse to initiate the next contraction of the ventricles.

Instantaneous Frequency Determination

In observing heart rate, muscle or nerve impulses, bat sonar, etc., one may need instantaneous frequency indications rather than averages over several cycles. The start of a period can commence a condenser discharging and the end terminate it. Discharge through an element having current proportional to voltage squared yields a final condenser voltage linearly proportional to frequency (inversely proportional to time). Resistively shunted thyrite or a biased diode network can display suitable nonlinearity. This has proven to be one of the more useful methods.[24] The principle, which has general frequency modulation detector applications, has been used to produce an oscilloscope display of reciprocal time.[25]

Seemingly, the initial condenser voltage must be infinite, which is inconvenient. Instead (Fig. 6), the initial condenser voltage V_1 corresponds to the minimum expected period T_1 or the maximum frequency. Discharge is delayed by a monostable circuit for this time, during which the previous reading is recorded and the condenser recharged to V_1. This circuit is sensitive while its memory is undergoing readout and resetting. With aperiodic cycles (*e.g.*, velocity inferred from time for labeled flow to traverse a fixed distance), the rate is stored for interrogation until the last possible moment during the start of the succeeding operation.

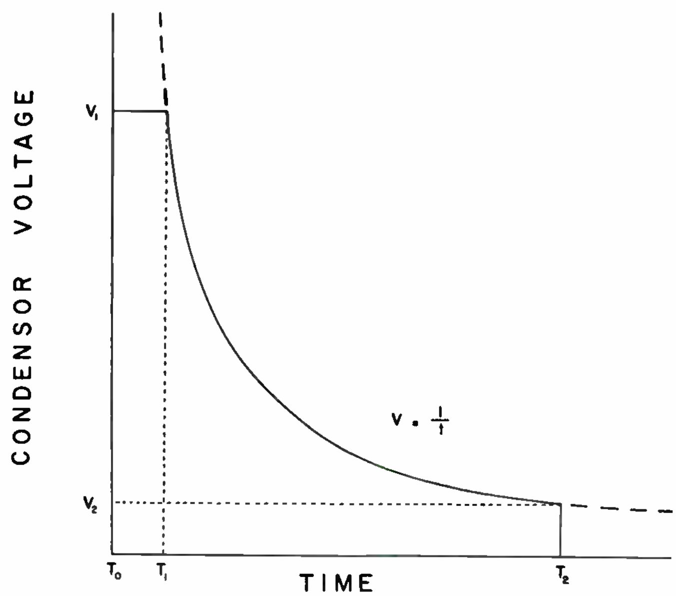


Fig. 6—If the time for one cycle of an event is T_2 and condenser discharge starts at T_1 from V_1, through a square-law device, then frequency indication is linearly proportional to the final condenser voltage V_2. The period T_0 to T_1 can be used for readout and resetting from the previous cycle.

Radio Telemetering to and From Within the Body

One can transmit signals into the body. Passive receivers (a tuned circuit plus diode) have been implanted to stimulate the brains of experimental animals in several laboratories. A heart pacemaker can similarly be powered. Such a receiver in a tooth could pick up the signal from a pocket-carried hearing aid and invisibly deliver the signal, via a tooth, by bone conduction. Also, the release of drugs from a capsule could be controlled externally.

Recording of various internal physiological variables with minimum surgical or other interference with the normal state is desirable. An approach places within

[22] R. S. Mackay, "Ventricular defibrillators," *J. Am. Med. Assoc.*, vol. 154, pp. 1421–1422; April 24, 1954.

[23] W. Greatbatch, W. M. Chardack and A. A. Gage, "A transistorized implantable cardiac pacemaker," 1960 IRE International Convention Record, pt. 9, pp. 107–115.

[24] R. S. Mackay, "Linear scale velocity meter," *Electronics*, vol. 25, pp. 158–161; October, 1952.

[25] E. F. MacNichol and J. A. A. Jacobs, "Electronic device for measuring reciprocal time intervals," *Rev. Sci. Instr.*, vol. 26, pp. 1176–1180; December, 1955.

the body a small transducer-transmitter combination. Swallowable radio transmitters called endoradiosondes thus far have performed satisfactorily while traveling the entire length of the gastrointestinal tract, and also after insertion into the bladder. A smaller transmitter has been designed for insertion into the anterior chamber of the eye to study pressure fluctuations and glaucoma. Related pressure transmitters have application to studies of teeth clenching and standing steadiness. The devices presently further animal studies, drug research and testing, routine human observation, etc. Some variables transmitted are pressure, temperature, oxygen concentration, acidity, and radiation intensity.

The frequency characteristic of a tuned circuit can be measured by bringing near it a grid-dip meter. Thus, the first experiments of this general nature attempted to measure the motion of an iron core near a tuned circuit in response to pressure changes. Signals thus generated seemed weak and unreliable, and so when junction transistors were perfected, passive transmitters were temporarily abandoned, and active ones containing a battery were tested.[26] Independent work from various parts of the world was described.[27,28] Passive transmission has been reported[29] and suggested for a two-valued variable.[30] There are differences in requirements for active and passive transmitters, and these are mentioned in a general summary of this field.[31]

One transmitter is shown in Fig. 7, and its signal pattern in Fig. 8. The transistorized Hartley oscillator is frequency-modulated by core motion, most of the radiation coming from the coil. Restoring force to the core is supplied by air pocket compressibility, and it is possible to balance the effect of temperature on this against that on the ferrite, tuning capacitor, and transistor. Temperature sensitivity can be enhanced by adjustment in the other sense. Variable resistance sensors of temperature and of pressure (liquids with a void between electrodes, or conducting epoxy) also worked. Pressure-responsive modulation by a light aluminum armature in place of the ferrite core of Fig. 7 seems advisable in variable g situations (rocket or satellite experiments), and in restraint-of-pill experiments by external magnets; transmitters of other variables are relatively insensitive to gravitational and magnetic fields. In a variable acceleration situation, a projecting ring prevents direct pressure by the body on the diaphragm. Transmission frequency is usually in the range of $\frac{1}{2}$ to 10 Mc.

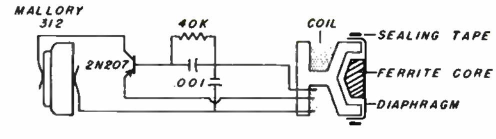


Fig. 7—Frequency-modulated radio transmitter to be swallowed for telemetering pressure or temperature. This unit is 9 mm in diameter, being determined by battery size.

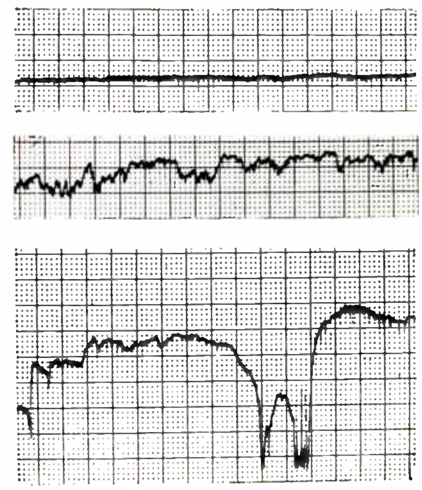


Fig. 8—Spontaneous restoration of peristalsis in the large intestine (cecum) after cessation following the trauma of abdominal surgery (removal of gall bladder). The small regular fluctuations are pressure changes due to breathing. Major vertical divisions are 20-second intervals. Top: after one hour, no activity. Center: 10 minutes later, activity starts. Bottom: three hours after center. At top and center, the deflection times 1.6 gives pressure in millimeters mercury; at bottom, multiply by 3.2. From a study by Dr. John Carbone and the author on the effect of the drug Cozyme or Ilopan in hastening intestinal activity in the post-operative period. (The above records are without drug.)

Methods of reception and demodulation have been given.[31] Two methods for overcoming loss of signal with pill reorientation are essentially either to switch cyclically between three perpendicularly (or otherwise) oriented receiving loops, or else to take the signal from the three receiving loops and feed each to a frequency doubler before addition and presentation to a receiver. The resulting signal is frequency-modulated and does not vanish for any capsule orientation.[32] Booster transmitters, using a frequency shift between input and output to prevent oscillation, can be used to relay signals. Multiple transmitters record gradients. Subject freedom is important in animal experiments where there is no discussion or cooperation.

Reversible mechanical expansion has been used to sense acidity[33] and oxygen concentration[32] changes, but electrochemical methods now seem better. With nonlinear condensers as modulators, one obtains the 100-megohm input resistance required by a glass electrode

[26] R. S. Mackay and B. Jacobson, "Endoradiosonde," *Nature*, vol. 179, pp. 1239–1240; June 15, 1957.

[27] J. T. Farrar, V. K. Zworykin and J. Baum, "Pressure-sensitive telemetering capsule for study of gastrointestinal motility," *Science*, vol. 126, pp. 975–976; November 8, 1957.

[28] M. von Ardenne and H. B. Sprung, "Über Versuche mit einem Verschluckbaren intestinalsender," *Naturwiss.*, vol. 45, pp. 154–155; April 1, 1958.

[29] J. T. Farrar, C. Berkeley and V. K. Zworykin, "Telemetering of intraenteric pressure in man by an externally energised wireless capsule," *Science*, vol. 131, p. 1814; June 17, 1960.

[30] M. Marchal and M. Marchal, "Procédé et dispositifs d'exploration radio-physiologique," French Patent No. 1200043; 1959.

[31] R. S. Mackay, "Radio telemetering from within the body," *Science*, vol. 134, pp. 1196–1202; October 20, 1961.

[32] R. S. Mackay, "Endoradiosondes: Further notes," IRE Trans. on Medical Electronics, vol. ME-7, pp. 67–73; April, 1960.

[33] B. Jacobson and R. S. Mackay, "A pH endoradiosonde," *Lancet*, p. 1224; June 15, 1957.

to make pH measurements.[31,34] A representative problem for such a unit would be to see if vitamin D administration to a child with rickets changes the pH in the intestine due to the effect on calcium. A problem for further investigation is the detection of the site of blood leakage into the gastrointestinal tract (or even to distinguish stomach from duodenum). Blood can show a dielectric constant of several thousand, but a capacitor probe does not make a satisfactory *in vivo* detector.[31] An earlier method[35] was to label red corpuscles with radioactive iron and to form the radio capsule into a low-voltage radiation detector by placing a fluorescent screen around the exposed transistor. Greater sensitivity is sought.

Sound and Ultrasound

The detection of auditory malingering and psychogenic deafness is a subtle problem. In the former, a person may feign deafness and learn to overcome normal responses to startling sounds. Electronics can help. A time delay for acoustic signals is a tape recorder playing a loop in quick succession through an erase, a record, and a readout head. It is a remarkably startling effect to read aloud from a book and have a throat microphone signal played back into the ears after a quarter-second time delay. The speed of reading *will* diminish or halt[36] if the person can hear.

Possibly related to this is the observation that playing a loud, steady noise into a stammerer's ear may relieve his affliction.[37] Recent reports of painless dentistry employing loud noises may have a limited nonphysiological basis.[38] Adequate validation is a major problem here, as with any innovation.

Extremely intense sounds affect the ear, as judged by the reports of deafness and loss of equilibrium of early jet fliers. But then Ménière's disease, which has been treated by surgical destruction of the labyrinth, might be alleviated by intense sound application; this apparently is being tested in several laboratories.

Audio methods can indicate when a pair of grasping forceps comes into contact with a hard foreign object.[39] A sound transducer mounted beyond the handle of forceps will generate an audible signal whenever the tips of the forceps contact and grasp a hard object. This has been used to grasp gall stones in a common duct operation without tearing loose normal tissue, and in detecting and grasping kidney stones, some of which cannot be seen in an X-ray image and hence are called "pale stones." Radio transmission of these signals is quite convenient.

K. Mooslin has noted that a suitably placed stethoscope reveals a harsh sound during micturization if a urethral stricture is present. An electronic extension of the method to blood flow may warn of impending vascular difficulty (bruits of aneurisms are already observable).

Bringing an extremely simple innovation into a foreign study can often provide a major contribution. Thus, neurology may be helped by a vibrator constructed from a modified earphone. Peripheral neuritis can result from vitamin deficiency, chronic alcoholism, pernicious anemia, etc., and its first symptom is early loss of vibration sense. This is possibly dependent on bone motion and transmitted as a deep sensation over separate channels (posterior column of spinal cord) in the nervous system. The condition is accompanied by loss of equilibrium and a lack of knowledge as to where the feet are, thus making it difficult to walk at night. The vibration sense is tested by applying the handle of a tuning fork over various bony regions (standard sites); 64 cycles is a standard frequency, but 50 to 60 are as good. The test is sometimes considered unreliable and unreproducible because of the lack of constancy of the vibration of the tuning forks, both in application at a single site and when making comparisons at corresponding sites on opposite halves of the body. A vibrator driven from the ac line can maintain steady amplitude and give the subject time to decide on responses. If vibration is perpendicular to the subject probe, then application variations will not affect vibration amplitude.

To digress momentarily, sonic techniques are not always best for locating foreign bodies. Metal fragments can be localized by passing near them a probe containing suitably energized coils. Thus, a balanced differential transformer will give an unbalance audio signal if any metal is brought near its end; a complete localizer, including power source, can be formed into a probe the size of a large pen. This is useful for eye splinters, bullet fragments, broken hypodermic needles, etc. It will detect either iron or nonferrous fragments, and one can construct electromagnets that will attract and possibly withdraw copper, lead, etc., as well as iron.[40]

An interesting problem recently referred to the author by a foreign police is of a clinical type, though having to do with foreign bodies, camels, and the law. The problem was how to detect from a herd one camel into whose stomach had been loaded 30 pounds of polyethylene-encased marijuana for smuggling purposes. An economical solution seems to be the use of a conductiv-

[34] H. S. Wolff and R. S. Russ, "Constructional aspects of radio pills suitable for mass production," *Proc. 3rd Internatl. Conf. on Med. Electronics*, London, England, IEE; 1961.

[35] R. S. Mackay, "Radio telemetering from within the human body," IRE Trans. on Medical Electronics, vol. ME-6, pp. 100–105; June, 1959.

[36] W. R. Tiffany and C. N. Hanley, "Delayed speech feedback as a test for auditory malingering," *Science*, vol. 115, p. 59; January 18, 1952.

[37] E. C. Cherry, B. M. Sayers and P. M. Marland, "Experiments on the complete suppression of stammering," *Nature*, vol. 125, p. 874; November 5, 1955.

[38] W. J. Gardner and J. C. Licklider, "Auditory analgesia in dental operations," *J. Am. Dental Assoc.*, vol. 59, pp. 1144–1149; December, 1959.

[39] R. S. Mackay, "Foreign body and kidney stone localizer," IRE Trans. on Medical Electronics, vol. ME-7, pp. 74–76; April, 1960.

[40] W. V. Lovell, "Electromagnet removes nonferrous metals," *Electronics*, vol. 28, pp. 164–166; September, 1955.

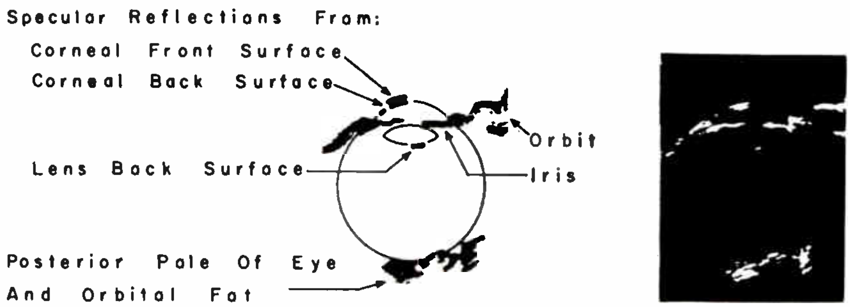


Fig. 9—Sound image of cross section of living human eye using 15-Mc sound pulses. The simple scan mode here does not visualize the lens.

ity test of the abdominal region. It may prove feasible to do this without contacting the animal by using a nonmetallic mine detecting circuit. Sonic exploration would also work.

In recent years, sonar techniques have been used to explore the body by scanning with a piezoelectric crystal which is periodically pulsed and receives returning echoes from interfaces separating various organs and structures. In this way, a detailed map in the form of a cross section can be built up. Ultrasonic echoes have recently been used to detect shifts in the midline of the brain due to tumors or injuries.[41] Abnormalities from cancers to foreign bodies have been visualized by Wild, Howry and other workers.[42] Fig. 9, made with E. Marg and R. Oechsli, shows the resolution in the eye obtainable with 15-Mc sound pulses. Nonmechanical scanning is being investigated. Acoustic attenuation increases with frequency; lower frequencies are needed to penetrate large structures, while higher frequencies increase angular resolving power. Sound at low intensities does not seem to have undesirable cumulative effects that may be characteristic of X-rays.

Another sonic technique involves irradiating the subject by a continuous sound source and placing beyond the subject a biconcave polystyrene lens to focus an image, in sound waves, of the subject onto a television-type tube.[43,44] This tube is fronted by a quartz sheet, and the piezo voltages are scanned off to convert the sound image into a video signal for television viewing. Diffuse sound illumination and frequency modulation (or source pulsing) should be used here to remove specular and diffraction effects respectively. Momentarily pulsing on the source and then the detector (point-by-point, or all-over as an electron mirror) should allow imaging a single layer (depth), and without glare or dazzle from strong surface reflections. Some of these tubes, if slowly discharged, have the potentially useful property of displaying only moving objects.[43] Recently Jacobs, Cugell and Phan have built an ultrasonic vidicon, showing storage, using phonon-induced conductivity in cadmium sulphide. (Note: more recently see Hutson, *et al.*, *Phys. Rev. Lett.*, vol. 7, pp. 237–239; September 15, 1961.)

More intense ultrasonic beams have been used to perform brain surgery that has helped cases of Parkinsonism and of phantom limb pain.[45] Zinser has used a long, thin, flexible magnetostrictive "drill" that is passed up the urethra past the bladder to fragment kidney stones stuck in a ureter; yet ultrasound can clean teeth safely.

Gas Analysis and Control

Respiratory function has many aspects. Gas flow velocity is one, and it can be measured in the lung.[46] Gas composition is important. Deep sea divers must control oxygen concentration to avoid both anoxia and oxygen poisoning; computers are being designed to study these problems and then to control gas mixtures. Even at atmospheric pressure, premature infants placed in an incubator with improper oxygen control can be blinded by retrolentalfibroplasia. Oxygen gas is usually analyzed by its magnetic properties, or by polarography after diffusion through a teflon membrane into a test region.

Another important normal component of the breath is carbon dioxide. One method of analysis employs a nondispersive infrared gas analyzer, which performs the interesting function of quantitatively detecting the presence of an organic gas after a sample of it is placed in the instrument. One form functions as follows. A nichrome wire heated by a current emits infrared energy through a thin mica window into a chamber with a flexible back that acts as a condenser microphone. To

[41] S. Jeppsson, "Echoencephalography," *Acta Chir. Scand.*, suppl. 272; 1961.

[42] J. M. Reid, "Diagnostic applications of ultrasound," Proc. IRE, vol. 47, pp. 1963–1967; November, 1959.

[43] C. N. Smyth and J. F. Sayers, "Ultrasonic image camera," *The Engineer*, pp. 348–350; February 27, 1959.

[44] W. Freitag, H. J. Martin and G. Schellbach, "Descriptions and results of investigations of an electronic ultrasonic image converter," in "Medical Electronics," C. N. Smyth, Ed., Iliffe and Sons, London, England, pp. 373–379; 1960.

[45] W. Fry, "Intense ultrasound in investigations of the central nervous system," in "Advances in Biological and Medical Physics," J. H. Lawrence and C. A. Tobias, Eds., Academic Press, New York, N. Y., vol. 6, pp. 281–348; 1958.

[46] J. B. West, "Bronchial flow meter for measuring air flow inside the lung," *Proc. 3rd Conf. on Med. Electronics*, London, England, July, 1960, IEE; 1961.

detect carbon dioxide, place carbon dioxide gas in this chamber to expand under the heating action of the infrared source. Only those wavelengths characteristic of carbon dioxide will be absorbed from the continuum emitted. Gases breathed through the space between source and chamber will generally absorb at other wavelengths and not diminish the heating action; any carbon dioxide present will reduce the signal from the microphone. The heat source is periodically interrupted at, say, 10 cps, and the ac signal used as the output to avoid drifts due to ambient temperature changes.

A clinical test using a gas analyzer involves holding a breath containing $\frac{1}{2}$ per cent carbon *monoxide*. One measures how much CO is removed in the exhalate, relative to the original. This allows a measurement of diffusion rate into the lung capillaries, which can change with respiratory ailments. Dilution effects in this method can be calculated by noting the simultaneous effect on some other gas, such as helium, which can better be analyzed by thermal conductivity or acoustic velocity methods. In connection with such investigations, it might be noted that digital computers are useful for discrete variables and data storage (*e.g.*, diagnosis and symptoms), but analog computers are often more convenient for visualizing function. One of the first analog computers was probably a collection of tubes used to study breathing function.

The administration of anesthetic gases brings in another variable. Depth of anesthesia has been controlled automatically by a regulator circuit that senses certain components in the electroencephalographic pattern and increases or decreases anesthetic gas administration accordingly.[47]

[47] J. W. Bellville and G. M. Attura, "How electronics controls depths of anesthesia," *Electronics*, vol. 32, pp. 43–44; January 30, 1959.

Conclusion

Changes are to be expected in all these fields, especially when there is "cross fertilization" between the physical and biological sciences. As an example of the great changes within the realm of possibility, the often discussed proposal for electrical anesthesia was recently tested in the successful performance of major surgery with its use.[48] Similarly, blood clotting can either be hastened or retarded by a direct electric current of suitable polarity.[49] Furthermore, it apparently has become possible to produce selected sensations such as pleasure by direct electrical stimulation of selected parts of the brain.

It is hoped that the entrance of more physical scientists into biological fields will result in the selection of more suitable variables to measure and the design of more decisive experiments, and that attempts at computer diagnosis will force a greater uniformity and precision upon medical nomenclature.

In closing, it should be said that in the allotted space, it is impossible to mention most of the possibilities included in the title. The selection was obtained in the simplest possible way, being merely items with which the author has special interest or familiarity. For a more complete listing of workers and works, the reader is referred to the "Bibliography on Bio-Medical Electronics," presently occupying three volumes, prepared by The Medical Electronics Center of the Rockefeller Institute, and published by the Professional Group on Bio-Medical Electronics of the Institute of Radio Engineers.

[48] J. D. Hardy, L. W. Fabian and M. D. Turner, "Electrical anesthesia for major surgery," *J. Am. Med. Assoc.*, vol. 175, pp. 599–600; February 18, 1961.

[49] P. N. Sawyer and S. A. Wesolowski, "Studies on direct-current coagulation," *Surgery*, vol. 49, pp. 486–491; April, 1961.

Data Handling, Computers and Diagnosis*

LEE B. LUSTED†, FELLOW, IRE

Summary—**As quantitative bio-medical science develops, the amount of data to be handled will increase. Already the investigator frequently finds that the analysis of the data he has collected is beyond his computational ability and he is forced to turn to automatic data processing. The bio-medical investigator must have access to a variety of both analog and digital computing equipment if progress is to be made in biomedical research.**

Automatic data-processing techniques and equipment need to be introduced to the areas of public health and patient care. Projects dealing with medical records, medical literature retrieval, automatic chemical analysis, the recording of signs, symptoms, nurses' notes, etc., must all be tied together in regional medical computing systems as soon as possible. Particular attention must be given to this systematic collection and analysis of medical data before regional computing systems can help physicians with diagnosis and treatment.

I. Introduction

ONE of the most important changes ever to take place in bio-medical science is now in progress. Previously, work was primarily of a qualitative and descriptive character, whereas present and future efforts will be to develop a quantitative bio-medical science. Part of this change is represented by the expanding fields of biophysics and bio-medical engineering and by the current interest in bio-medical computing.

The development of instrumentation and the application of mathematics and statistics to bio-medical problems have greatly increased the amount of data to be handled. The invesigator frequently finds that the analysis of the data he has collected is beyond his computational ability as well as his patience, and he is forced to turn to automatic computation methods. It is the author's belief that, if we are to make progress in bio-medical research, the investigator must have ready access to a variety of both analog and digital computing equipment.

There are, for the investigator, at least two major categories of activity which may prove to be even more important than the increased speed and accuracy associated with automatic computing. These are 1) the revelation of hidden correlations which cannot be extracted by other than computer techniques, *e.g.*, on-line recording and analysis of psychophysiologic data, and 2) the design of models for testing their potentiality for prediction. These contributions should help increase the amount of investigation and should open new areas of research.

In the recent past only a few bio-medical research projects have used automatic data processing. The National Institutes of Health Advisory Committee on Computers in Research has been making a study of this field and has estimated that, at present, no more than 1 project in 20 is seriously utilizing computers. The funds allocated to such work are probably less than 0.3 per cent of the bio-medical research funds. It seems likely that in five to ten years, the use of automatic computing and mathematical analysis will have influenced bio-medical research, so that 75 per cent of the projects will be using mathematics and data processing, and at least 50 per cent of all research projects will require automatic computation.

The study of the biomedical computing field made by the NIH Advisory Committee on Computers has revealed that an effort must be made to develop each of the categories listed in Table I if bio-medical computing is to progress.

TABLE I

A. General Facility Developments
- 1) Large computing-center development.
  - a) Single-purpose geographically oriented bio-medical computing center.
  - b) Partial use of institution-wide large computing center.
- 2) Small bio-medical computing centers (serves medical school or department complex).

B. Laboratory Facility Developments
- 1) Small general-purpose computer facilities for individual laboratories.
- 2) Special-purpose computer facilities for specific laboratory functions.
  - a) On-line use of general-purpose computer.
  - b) Adapted, altered or specifically constructed computers.
- 3) Peripheral computer facilities.
  - a) Analog-digital conversion facilities.
  - b) Data-handling facilities.

C. Computer Use Projects
- 1) Investigations of possible computer uses.
  - a) Utilizing bio-medical computing facilities.
  - b) Purchasing time on already available institution or industry computers.
- 2) Investigations of mathematical and other techniques which may lead to computer use.

D. Computer Technology Projects
- 1) Development of computers (digital and analog) to meet the needs of the bio-medical field.
- 2) Development of accessory equipment such as special input-output devices for the bio-medical investigator.
- 3) Investigations of programs and programming.

E. Training in Bio-medical Computing

If bio-medical computing is a very important part of the current bio-medical revolution, then it is obviously in the best interests of the country to develop bio-medical computing and facilities as rapidly as possible. Fortunately, the National Institutes of Health, the National Science Foundation, the National Academy of Sciences—National Research Council and a panel of the

* Received by the IRE, June 29, 1961.
† Dept. of Radiology, University of Rochester School of Medicine and Dentistry, Rochester, N. Y.

President's Science Advisory Committee have taken an active interest in the problems of the bio-medical computing field. The interests and activities of these groups must be coordinated if bio-medical computing is to be developed to an optimum degree in the United States.

From the general categories of Table I three subjects have been selected for further discussion, namely, 1) training, 2) computer technology projects, and 3) the geographically oriented bio-medical computer center.

II. Training in Bio-medical Computing

In the past biologists and medical investigators have not had extensive training in mathematics and quantitative methods [20]. Therefore, the success of the change to a quantitative bio-medical science depends upon retraining existing scientists and the training of future scientists. The number of people who have had experience with automatic computing is small, and there is great need for training programs which will acquaint the bio-medical investigator with automatic data processing.

One of the questions now under discussion is whether the biologist and the medical scientist need more extensive mathematic training in order to make the best use of automatic data processing. The importance of this question is emphasized by a study made recently by Hopkins and Berry [1] on the mathematics background of medical students. They state:

> From these results, even allowing for imperfections in test procedures, we conclude that on the average these students generally, though sufficiently apt and scholarly in quantitative relations, are not sufficiently equipped with basic mathematics to master an introductory course in statistical methods. With this deficiency standing in the way of mastery of the fundamental philosophy and techniques of science, we should think again when we talk about producing physicians who are scientists as well as physicians.

If we believe that computers will extend the ability of the investigator to solve problems and also to seek new ways to extend his problem solving ability [2], then it may be very important to ask that the biologist and medical investigator be well acquainted with mathematics.

At the present time it seems highly inefficient and practically impossible for a single individual to execute all the different aspects of bio-medical computing, such as planning, logical analysis, coding and debugging. Therefore, it probably will be necessary to have teams of workers among whom communication will be fairly easy. This implies that the bio-medical computing worker must be facile not only with computer programming and operation but also with bio-medical concepts. He will require a special kind of training, and a high priority must be placed on the development of bio-medical computing training programs. Scientists who attended the course have been enthusiastic about the M.I.T. Research Training Conference on Computer Techniques for Biological Scientists directed by Dr. T. T. Sand and held at M.I.T. during the summer of 1961. More courses of a similar nature are needed.

III. Computer Technology Projects

A great deal of thought needs to be given to the specifications for bio-medical computing equipment. Some of the early computers which were developed for business applications or the physical sciences are not well suited for solving problems in biology and medicine. The fact seems to have escaped the attention of many computer designers that the computer should be considered in the context of the environment in which it operates. This applies also to input-output equipment design. The concept of man-computer symbiosis [3] is one of particular importance to the future of bio-medical computing. The best attitude at present is one of flexibility, which means that computers of different designs and different sizes will be tried by investigators in an attempt to determine what type of machine gives the most help in solving certain classes of problems. This does not mean, however, that we should allow the effort to proceed in an unorganized manner. Each investigator should be free to choose the type of automatic data-processing equipment he prefers, but he should be asked to report back to a central clearing house at the end of each year on how well the equipment worked in solving his specific problems. Thus we could readily accumulate a great deal of information which would be helpful in designing future machinery for bio-medical computing.

This effort should be organized on a national basis because of the great importance of automatic data processing to the bio-medical sciences and because of the amount of money involved in the development of computing equipment, displays and controls.

Let me cite a problem which has convinced me that an organized development of the bio-medical computing field is necessary. This problem concerns the standardization of recording techniques and equipment for analysis of neurophysiologic and electroencephalographic data.

The amount of automatic data processing using special-purpose analog devices and large general-purpose computers has been steadily increasing in such fields as brain research, psychophysiologic research and cardiovascular research. An investigator in one of these fields may wish to compare his data with that of a colleague who is doing a similar set of experiments, but suddenly he finds to his dismay that the colleague is using a different type of data-accumulation system and it is difficult, if not impossible, for them to compare data. Of course the investigator is disturbed, and he may request help from the electronic manufacturer who supplied his data-accumulation system. The manufacturer may wish to help the investigator, and may even try to make a survey, hoping he can determine the specifications for an "optimal" computer which will meet the needs of a large number of investigators. But the manufacturer usually becomes quite bewildered by the conflicting requests for different sampling rates, for the number of channels for magnetic-tape recording, and for the bandwidth requirements for the data to be handled, etc.

Thus far several preliminary attempts to get investigators to resolve these differences of opinion have failed.

To help solve this kind of problem and that related to input and output equipment development, etc., the NAS-NRC Committee on the Use of Computers in the Life Sciences, and the NIH Advisory Committee on Computers in Research can do much by inviting groups of investigators and engineers to meet regularly and to work out specifications.

IV. The Geographically Oriented Bio-medical Computing Center

When Zworykin [4] suggested in 1956 that automatic data processing could prove extremely useful in reducing the case load of the physician, he stirred up a storm of protest from the medical profession.[1] However, as Zworykin [5] observed three years later in opening the First Conference on Diagnostic Data Processing at the Rockefeller Institute, "We understand now that this criticism was due mostly to the misconception that we were trying to replace the doctor by a cold, hard calculating machine. Instead, of course, we are trying to introduce a new scientific tool to aid the diagnostician." In 1959 Ledley and Lusted [6] proposed the idea of a health computing system as an aid to physicians, and the Jones Committee of Consultants on Medical Research [7] has recommended that regional computing centers be established for bio-medical data processing. As a result of the interest in the "center idea," the 87th Congress appropriated a sum of 5 million dollars to be used for bio-medical instrumentation and bio-medical computing centers.

Although preliminary plans have been completed for several regional centers, none are in operation. There are a number of reasons for starting the regional centers as soon as possible.

During the last four years, a time when the number of college graduates has been increasing, the number of college students applying to medical schools has dropped. This trend has occurred at the same time that the population in this country has been increasing and constitutes a serious threat to an increase in the number of physicians. To keep the present physician/population ratio of 141/100,000, the United States must increase the annual number of graduates from the 1959 level of 7400 physicians to a level of 11,000 by 1975 [8]. Further estimates indicate that to reach this level we would need 16 new 2-year medical colleges and 18 new 4-year medical schools. Some grave doubts have been expressed that this goal can be reached, or indeed that it need be attained, since it may be unnecessary to maintain the present physician/population ratio. If we can supply the physician with instrumentation and computer aids for diagnosis, he should be able to practice more efficiently and see more patients than he did formerly. A regional bio-medical computing center for processing the medical data on all patients in the area would be an important diagnostic aid for physicians. Because the regional center will serve many physicians and hospitals it is obvious that the cooperation of medical schools, the medical profession, computer engineers, hospital administrators and health organizations will be required to establish and maintain it.

The regional bio-medical computing center of the future probably will carry on research activity as well as service activity, and both of these, at least in the early stages of development, will be concerned with problems of medical diagnosis. Before we can decide what kind of computing equipment and input-output devices are needed, we need to know more about the kind of data we are working with.

Let us consider the physician taking care of the patient in the hospital since this represents a more complex situation than the physician's office.

The data on the patient comes from the history, physical examination, laboratory tests, special examinations and the nurses' notes. There are problems in collecting data from each of these sources, and we have just begun to try to solve them. For instance, an easy method has not been devised for recording history and physical examination data. Some experiments have been tried using "mark-sense" cards and "mark-sense" sheets are being developed. Progress has been made in the development of automatic laboratory test equipment [9], and it should not be difficult to transmit the output to the bio-medical computing center. Blumberg [10] has recently pointed out that 13 per cent of nursing time is spent on charting nurses' notes, and about 1/6 of the nursing budget is spent for nurses' time on clerical tasks concerned with recording and performing medication orders. This is certainly an attractive area in which to introduce automatic data processing. Blumberg has made some tests with a system he calls the Hospital Indicator for Physician Orders (HIPO), which could transmit data directly to the central data processor.

I have purposely avoided a discussion of the computer equipment in the bio-medical computing center. At present we can only guess at what types of machines will be needed. Probably we shall need some machines capable of doing scientific computing and others capable of storing a sizeable amount of data from medical records. They should be expansible and able to handle multiple inputs of analog and digital data simultaneously. The output data should be presented in a form most suitable to the person needing the data. Obviously, we must wait on the advancement of computer technology and the

[1] A recent statement by Dr. Peter Forsham, Prof. of Medicine, University of California, San Francisco shows the extent to which some physicians have become convinced about the future role of computer aids to medical diagnosis. Dr. Forsham says, "By the year 2000, both diagnosis and treatment will be largely turned over to technicians equipped with electronic devices. The physician's job will then be essentially what it was in the nineteenth century—to serve as friend and confessor to the patient, counsel him on emotional problems, and help him adjust to diseases the machines are still unable to lick." (*Med. Economics*, vol. 38, p. 71; May 8, 1961.)

development of better displays and controls before we can meet all of these requirements.

The physician would like to ask questions of the regional computing center such as: From the patient data which I just sent you what is the differential diagnosis (please list the most probable diseases)? What further information do I need in order to make a definite diagnosis? What is the latest information on the treatment of this disease? On the basis of the patient's response to this particular treatment during the past week, what is the prognosis (what can be expected in the future)?

The medical information needed to answer these questions has not been collected in computer usable form. We need to undertake the collection of the necessary data by teams of physicians as soon as possible. Otherwise, we will have developed excellent data-processing equipment, but we will not have reliable data to process, and we will be operating, as one observer recently noted, in the "disposall mode," *i.e.*, garbage-in processed garbage-out, with output dependent on a "packing factor."

When groups of physicians do start to collect and to process quantitative medical data, they are going to be confronted with a number of problems. First, in many areas of medicine the signs and symptoms have never been rigorously defined [11]. Second, physicians have not studied carefully for some fields of medicine the accuracy of medical diagnostic procedures and the effect of observer error on the final diagnosis [12]. And third, the internal inconsistencies in the body of medical knowledge have not been searched for carefully. It would seem that the real revolution in clinical medicine will begin when physicians start to work on these problems. The study will be an enormous task. There is already an indication of the magnitude of the task from preliminary work done in hematology [13]. For this single medical specialty it took over 5 man-years to examine carefully the literature, the signs and symptoms, and the diagnoses. Many internal inconsistencies in the body of hematology literature were found, and the investigators have concluded that it will be necessary to collect a new body of hematology data using carefully defined signs, symptoms and laboratory tests.

On the basis of the hematology study, a rough estimate of the amount of work required to examine the entire body of medical literature and to collect new data on the basis of carefully defined standards would be 500–600 man-years. This does not seem unreasonable since it is in the same order of magnitude as the amount of work required to develop the SAGE radar system. Until this work is done, it seems quite certain that the physician will not be helped by "computer-aided" diagnosis.

For the past five years, Dr. Ledley and the author have studied the role of computers in medical diagnosis. Our attention has been concentrated on mathematical models of various aspects of the medical diagnostic processes, which can form the basis of computer programs. In addition, we have considered problems that arise in the compiling of statistics that relate symptom-disease combinations and in the accumulating and recalling of the desired aspects of a patient's total medical record, including individual bio-chemical and physiologic norms [6], [14]. We believe that these mathematical techniques will be helpful to the physicians who undertake the systematic collection and analysis of medical data in the future.

The mathematical models which we have presented represent a more precise formulation of the reasoning foundations of medical diagnosis where sequential decision techniques use 1) logical analysis to determine possible alternative diagnoses, 2) probabilistic analysis to determine the probabilities of these alternatives, and 3) value theory to assist in the choice of treatment plan. In practice, the physician first utilizes an iterative procedure of logical analysis and diagnostic testing, employing in each cycle tests of an increasingly specific nature for narrowing the alternative possible diagnoses. When alternatives still remain, Bayes' formula presents important quantitative insight into the relative roles played in the diagnostic decision by both the disease-symptom relationships of medical knowledge and the circumstantial factors involved in the location, season, epidemics, etc. At present the primary purpose for making a diagnosis is to guide the processes of treatment-plan formulation, but perhaps in the future, with computer help, the diagnosis step can be omitted and the treatment considered directly. Frequently, there is a choice among alternative treatments, which choice can involve decisions under certainty, under risk, and under uncertainty. In those cases for which the treatment plan must be formulated in the absence of a precisely defined diagnosis, a diagnosis-analysis treatment-formulation iterative cycle, in the dynamic programming sense, can be utilized [15].

If such procedures are to be applied in practical situations, the use of high-speed electronic computers is certainly indicated. Recent work by Warner [16] on congenital heart disease cases and by van Woerksom and Brodman [17] on the use of the Cornell Medical Index has demonstrated the help that conditional probability and Bayes' formula calculations can give the physician. Some questions have been raised concerning the size of the computer memory which may be needed. The first projects will probably be concerned with the amount of data in a medical specialty or subspeciality. This might range from 200 symptoms and 300 diseases to 800 symptoms and 100 diseases [15], [18]. Suppose we have a textbook of 400 pages for a medical specialty; the book would then contain approximately 1.2×10^6 characters. By allowing symbolic logic a 10-to-1 reduction in characters with six bits per character, we should be able to record the book in 720,000 bits, which is easily attainable on a modern drum memory. Ledley has made the following estimate [18]. Suppose there are

20,000 possible symptom-disease complexes associated with 300 diseases and 200 symptoms. Then, allowing 500 bits for each symptom-disease complex, $500 \times 20,000 = 10 \times 10^6$ bits would be required. There are many manufacturers today who make 10,000,000-bit drums. However, if we assume that these 300 diseases enter into only 1000 different disease complexes, and that 20 bits are sufficient for each disease complex, then, on a 10,000,000-bit drum, there will be 9,980,000 bits remaining for symptom complexes. Thus a 10,000,000-bit drum would easily record about 50 symptom complexes for each disease complex, or 50,000 symptom-disease complexes at 200 symptoms per complex.

A few of the problems which must be solved for the regional bio-medical computing centers have been outlined in general terms. No plea for the development of these centers could be more eloquent than a quotation from a recent editorial in the *Journal of the American Medical Association.* [19]:

> Scientific data are accumulating at an ever increasing pace, which threatens to outdistance the ability of the physician to keep abreast of new knowledge. Medicine faces a difficult problem, since there is an urgent need for some device to make pertinent information available promptly to the physician. Without some device for making the information available to the physician of today and tomorrow he will be less competent in the care of the patient than formerly.

References

[1] C. E. Hopkins and V. B. Berry, "On mathematical literacy of medical students," *J. Med. Educ.*, vol. 36, pp. 338–341; April, 1961.

[2] L. Fein, "The computer-related sciences (synnoetics) at a university in the year 1975," *Am. Scientist*, vol. 49, pp. 149–168; June, 1961.

[3] J. C. R. Licklider, "Man-computer symbiosis," IRE TRANS. ON HUMAN FACTORS IN ELECTRONICS, vol. HFE-1, pp. 4–11; March, 1960.

[4] V. K. Zworykin, "Medical electronics will provide technical facilities with which life scientists will implement their work," 1956 IRE NATIONAL CONVENTION RECORD, pt. 9, pp. 99–102.

[5] V. K. Zworykin, "Welcome—introduction to the conference (on diagnostic data processing)," IRE TRANS. ON BIO-MEDICAL ELECTRONICS, vol. BME-7, p. 239; October, 1960.

[6] R. S. Ledley and L. B. Lusted, "The use of electronic computers to aid in medical diagnosis," PROC. IRE (Bio-Medical Electronics Issue), vol. 47, pp. 1970–1977; November, 1959.

[7] "Federal Support of Medical Research," Rept. of the Committee of Consultants on Med. Res. to the Subcommittee on Depts. of Labor and Health Education and Welfare of the Committee on Appropriations, U. S. Senate, Washington, D. C.; May, 1960.

[8] U. S. Dept. of Health, Education and Welfare, Public Health Service, "Physicians for a Growing America," Rept. of the Surgeon General's Consultant Group on Med. Education (The Bane Report); 1959.

[9] R. H. Muller, Ed., "Automatic Chemical Analysis," *Ann. N. Y. Acad. Sci.*, vol. 87, pp. 609–951; July 22, 1960.

[10] M. S. Blumberg, "Hospital Automation—Needs and Prospects," address delivered at the Annual Mtg. of the California Hospital Assoc., Santa Barbara; October, 1960. (To be published in *The Modern Hospital.*)

[11] S. Cobb, "On the development of diagnostic criteria," *Arthritis and Rheumatism*, vol. 3, pp. 91–95; February, 1960.

[12] L. B. Lusted, "Logical analysis in roentgen diagnosis," *Radiology*, vol. 74, pp. 178–193; February, 1960.

[13] C. Berkley and M. A. Atamer, Rockefeller Institute, New York, N. Y. Personal communication.

[14] R. S. Ledley and L. B. Lusted, "The reasoning foundations of medical diagnosis," *Science*, vol. 130, pp. 9–21; July 3, 1959.

[15] R. S. Ledley and L. B. Lusted, "Medical Diagnosis and Modern Decision Making," presented at Annual Mtg. of the Am. Math. Soc., New York, N. Y.; April 6, 1961.

[16] H. R. Warner, *et al.*, "A mathematical approach to medical diagnosis-applications to congenital heart disease," *J. AMA*, vol. 177, pp. 177–183; July 22, 1961.

[17] A. J. van Woerkom and K. Brodman, "Statistics for a diagnostic model," *Biometrics*, vol. 17, pp. 299–318; June, 1961.

[18] R. S. Ledley, "Letter to the Editor," IRE TRANS. ON BIO-MEDICAL ELECTRONICS, vol. BME-8, p. 263; October, 1961.

[19] Editorial, "Advanced Medical Ignorance," *J. AMA*, vol. 176, pp. 804–805; June, 3, 1961.

[20] Lee B. Lusted, "Quantification in the life sciences," IRE TRANS. ON BIO-MEDICAL ELECTRONICS, vol. BME-9, pp. 1–3; January, 1962.

Additional Source Literature on Bio-medical Computing

[A] IRE TRANS. ON MEDICAL ELECTRONICS, vol. ME-6; September, 1959. Papers from the 11th Annual Conf. on Electrical Techniques in Medicine and Biology, Minneapolis, Minn.

[B] IRE TRANS. ON MEDICAL ELECTRONICS, vol. ME-7; October, 1960. Proceedings of the Conf. on Diagnostic Data Processing, Rockefeller Institute, New York, N. Y.

[C] PROC. IRE, vol. 47, November, 1959. (Bio-Medical Electronics Issue.)

[D] R. S. Ledley, "Report on the Use of Computers in Biology and Medicine," Natl. Acad. Sci.—Natl. Res. Council; May, 1960.

[E] R. S. Ledley, "The Use of Computers in Biology and Medicine," McGraw-Hill Book Co., Inc., New York, N. Y. (To be published, 1962.)

Education for Engineers in Bio-Medical Research*

SAMUEL A. TALBOT†

Summary—**Current bio-medical research is not only more demanding instrumentally, but inquires into biological phenomena which are far more complex physically and chemically than twenty years ago. It results that the instrumental requirements for significant observations are often misinterpreted by biologists because of inadequate concepts of the systems involved. Consequently, the engineer himself must understand the bio-systems and the design of the experiment to assess the adequacy of instrumentation and interpretation of readings. So complex are most bio-systems that this, in general, requires the most advanced engineering concepts.**

Education in engineering with some physiology to the master's level may yield adequate technical results, but does not produce or earn a satisfactory scientific relationship. For this a much higher scientific competence is needed in engineering, in biological understanding and in the interdisciplinary area. The latter is the special application of engineering science to bio-systems: a body of knowledge still barely begun, but rapidly growing. Discussed herein is the nature of such engineering, biological and interdisciplinary training, which should lead to a satisfying career as an engineering scientist in bio-medical research. For good reasons, this is seen as a collaborative relation.

TWENTY YEARS AGO, the engineer associated with medical research was mainly occupied with designing equipment. Bio-amplifiers became specialized; versatile stimulators, optical and mechanical components—all presented problems suited to the engineering training of the day. The bio-medical problems were simpler too: description, mapping, recording simple functions in one or two channels, taking perhaps a derivative or integral, and passive electric analogs of bio-systems satisfied the bio-scientist.

Today the demands of measuring devices are increasingly difficult. Inertialess implants, versatile catheter tips, broadcasting, sensors and drivers in new ranges of sensitivity and precision, in multiple channels and scanned arrays, must still be developed. Besides an observable, its variances and several derivatives as to time and space are needed. Stimuli are subtle, involving time and space patterns. With miniaturization and multiplexing of sensors and equipment, one must detect ever finer detail: millimicrons, milli g's, millidynes, millidegrees, microwatts, microlumens, at ever faster speeds—micro- and nanoseconds. Correlations, transforms and digitization are common.

Such discriminative, multiplex and sensitive instrumentation indicates that the biologist's problems are much more complex today. His hypotheses and models have become sophisticated. The cell membrane is a complex mechanism for which simple concepts from chemistry and physics do not suffice. Neural summation is not just addition. Arterial pulsation fits neither an RC or resonator model. Scores of bio-control systems emerge, involving multiple sensors and actuators, parametric modulation, nonlinear ranges and adaptive transforms. In short, modern instrumentation reveals performance for which the physiologist has inadequate physical concepts, but which the engineer without detailed physiology cannot even recognize, much less analyze. Today the biomedical engineer is needed as much for experimental design, reduction and interpretation of observations, as for instrumentation.

It is fortunate that with this revolution in biological research has come a revolution in engineering science. The modern engineer, with advanced training, encounters coupled electrical, mechanical, hydromechanical and chemical systems, almost complex enough to use as building blocks for bio-systems. The organism has routinely used pattern-transforms in time and space domains, only now being discovered in signal theory. Switching networks, statistical storage, combined digital and analog processes are ancient in the brain. In mechanics, shear and viscoelastic propagation are just what is needed for heart-sounds. In chemical engineering, the nonequilibrium chain reaction is reached just before the student's Ph.D., but goes on in each of his cells. In a word, engineering science is just beginning to produce concepts adequate for biological models and experiments, as similarly, twenty years before, the engineering art was barely able to handle bio-measurements.

Apparently several *kinds* of bio-medical engineers are needed. To assemble existing components into special systems, to modify measurements, to help with complex data processing in real time, needs practical instrument engineers. With fair success, men with experience in industrial research can grow to fill this requirement.

But to design the specifications of today's bio-instruments may mean knowing the physical aspects of a bio-problem better than the biologist. For example, the biologist calls a "sonic catheter" a device whose threshold is far above vascular sound. To evolve a nerve-impulse analyzer requires not only a critical knowledge of analytic methods but clear understanding of activity patterns in a nucleus or bundle. Today's designer of bio-instruments must know a great deal about a special biological field. For this requirement, a master's training in engineering, a study of biological instrumentation

* Received by the IRE, July 10, 1961.
† School of Medicine, The Johns Hopkins University, Baltimore, Md.

in a special field (such as neuro or cardiovascular physiology), and some understanding of current research in this area can produce an instrument designer of great value in biomedical research.

Our present problem, however, is education for engineers in bio-medical *research*. By this we mean one who seeks to understand and formulate biological behavior and processes in terms of the engineering theory of systems organization. He differs from the biomedical scientist in that he brings to his problems analytical methods and concepts derived from the advanced engineering sciences. He differs from the bio-physicist in that he is familiar with the analysis and synthesis of complex dynamic systems, whether physical, chemical or psychological.

For instance, he asks whether a membrane, a hormone system or a flow-pattern operates as it does, through a balance of forces, a stationary entropy state (open or cyclic), or by dynamic error-minimizing. If the latter, what are the dynamic characteristics of the sensing mechanism, in what order and range do the parametric actuators operate, what is the range of stability and what pathologies disturb it and how? In dealing with the electrocardiogram, the bio-engineer seeks not only its minimal eigen vectors, but totally new representations that carry more clinical information. In respiratory or muscle physiology he seeks models of increasing competence, but with observables chosen so as to keep them tractable. In this way, working closely with the bio-specialist, he contributes significantly to the pattern of scientific understanding. But also, he builds a scientific foundation for the applications engineer to design his bionics and medical research instrumentation.

The need for such engineering scientists arises 1) in aerospace research, where biological problems cannot be isolated, but appear always in systems context, and often combine observations at high (cognitive, affective) and low (physiological) levels of organization; 2) in commercial bio-instrumentation to plan the validation and evaluation of research and clinical devices; 3) in bio-medical research to study alternative ways of data processing, and in designing experiments to test models of bio-systems. Bio-medical engineers are especially needed to formulate the dynamics and stability of biological control and goal-seeking systems. This is an area of research inaccessible to biologists alone, but of great bio-medical importance. 4) In universities, both to teach and to systematize knowledge in this growing field.

We must face frankly the problem of status. In the academic world, as in scientific societies, the doctorate becomes essential to acceptance or advancement. The experienced instrument designer's status in a group of bio-scientists is not high. The remedy is straightforward: today, the bio-medical engineer can make unique contributions to the *science* of bio-systems, if he will acquire the knowledge and experience to do it.

Training for engineers in bio-medical research raises questions of depth and breadth in engineering, in biology and in the interdisciplinary area. Engineering faculties have found that most of their sciences at the graduate level have biological application. Especially on the conceptual side, advanced thermodynamics, fluid and solid dynamics, electrical systems and signal theory, transforms, nonlinear and complex functions, matrix and numerical analysis have proved a foundation for the science of biological systems. Particular engineering applications such as solid-state devices, plastics, transport processes, may be sacrificed, if anything. Two solid years of such advanced engineering studies are not excessive. It is debatable whether "modern physics," except for its mathematical training, relates closely to the levels of organization composing bio-systems. But the theory of mechanics, electricity and physical chemistry are all heavily involved; a doctorate in any of these fields would lead to broad areas of biological research. A composite emphasis on "systems engineering" where possible, would sensitize the student to the cybernetic problems in the strange mixture of chemical, mechanical, electrical and informational processes found in biology.

The present impetus in training for biomedical engineering in electrical engineering departments rises partly from the wide use of electrical methods in transducing and recording, in data transforms, handling, storage and computation. Partly, it comes from interest of E.E.'s in control systems and nets, in communications, and the advanced analysis of performance and stability. Partly too, the electrical manifestations and mechanisms of living organisms have attracted study from the days of Galvani, not to mention electrodiagnosis and therapy. Electrical observations are easier than mechanical or chemical, more accessible and versatile. However, to current bio-medical science, electrical observations are less interesting than many chemical processes. The bio-medical engineer cannot afford to be ignorant of the hormone systems, immunologic reactions, junctional transmission, anabolism and metabolism, as well as acoustics, hydrodynamics, and transmembrane thermodynamics. If we identify bio-medical engineering with electrical techniques and problems, we lose a great deal of interest by and contact with the biological sciences.

Generality of engineering principles and training is increasing in modern engineering schools. Increasingly it is recognized that many special sciences are but particular aspects, and engineering departments now offer more joint and generalized courses. *Thermodynamics* now covers heat engines, open nonequilibrium chemical systems, mechanical, chemical and electrical crosscoefficients in a unitary way. Its new notation applies as well to every level or organization, not only the physicochemical. The bio-medical engineer needs these powerful general concepts, transcending the equations of motion, conservation of matter and energetics. *Signal* *theory* is no longer restricted to electrical filters and time series. Its generalized form encompasses any observable in *n* space, with limited aspects in the space defined by each observing or analyzing system. This is excellent for the bio-medical engineer, who cannot neglect the manifold manifestations needed to measure a complex observable such as anxiety in an astronaut. As the engineer distinguishes his physicochemical phenomena from his generalized mathematical vehicles, the unity of nature and the imperfection of his tools become clear. So, in his preparation, the bio-medical engineer with an eye to the future should seek those graduate schools where engineering is taught in its greatest generality, though without neglect of special applications. Fortunately some of these schools offer bio-medical engineering too.

What *biological* education should be undertaken? To understand the problem and communicate with biologists there can never be too much—some advocate equality. To answer this question, we may consider the need to know current research. His colleagues in biology can do better with the bio-medical literature, but the literature of engineering theory and practice pertinent to biosystems is growing by leaps. This suggests a collaborative responsibility: the engineering scientist must know biological science in general, his field of biological problems well, but the engineering sciences as an expert. Even Helmholtz today could not be equally a biological and physical scientist. Admitting some limitation, we ask, "How much bio-science is really necessary?"

Should we say: enough to have perspective, to see the operative factors, to be communicative and responsive with bio-scientists, and to be able to shift one's activity as opportunity offers and knowledge opens? A mere survey can give a false sense of understanding. Further, if one is not to be at a disadvantage among bio-medical scientists, and if one is to know what they take for granted, the course in medical physiology would seem to be the minimum for biomedical engineering research. First, one obtains for the first time an appreciation of the complexity of the working bio-system, the tentativeness of inferences, conclusions from conflicting evidence, the design of biological experiments, multiple variation, the expertise needed to avoid artefact, faulty preparation, and the time factor. The delightful childlike simplicity of physical experiments becomes appreciated. Second, the nature of biological materials is sensed, through sight, touch, sound and smell. Their fragility, intricacy, inaccessibility become clear through experience. The immanency of injury and death bear home. Responsibility is a part of every experiment—hours of preparation and after-care are involved. This is a new world, to some. All these attitudes and sensitivities are important for research among bio-scientists.

The full course in medical physiology will prove difficult for those without an excellent memory for unrelated facts. Nevertheless, it gives a foundation on which specialization in any kind of bio-medical problem is quickly built, with the guarantee that thereafter no system will be regarded as alone. The research engineer needs this breadth and adaptability if he works as one of his kind in a medical research center. Naturally for his research field he eventually must go much deeper in biology, whether in formal courses or by reading.

Physiology was mentioned as minimal. Some biophysicists will urge the indispensability of cytology, since the cell is the basic bio-system; and, for the specialist in bio-chemical systems, this is essential. Psychologists similarly will urge an introduction to psychophysics and operant behaviorism—important to understanding the sensory systems. Furthermore, essential prerequisites to physiology are anatomy and physiological chemistry, both loaded with memory work. Some compromise can be made here by dropping their laboratories and concentrating on theoretical aspects. Experience shows that a single anatomical technic can later be learned if needed. While there is no substitute for anatomy in the round, in bio-medical research one seldom works alone.

Medical biochemistry is essential, because of the enormous importance of its control systems, which tie in closely with the neural mechanisms, not to mention the circulatory, muscular and sensory. Physical in viewpoint, systematic and thermodynamic, it is less of a departure. But, like all bio-sciences it abounds with minute facts one cannot neglect. (One need not be bothered because general biology and organic chemistry are pre-entry requirements. They are offered in every summer school.) Discussions with numerous bio-scientists about bio-medical engineering have shown that the foundation here described should be adequate. It adds to about a solid half-year of biological science at the graduate level, which should be creditable as one minor toward the engineering doctorate.

Thus far we have spoken of *bio-sciences* and *engineering sciences*. There remains the interdisciplinary science of "bio-medical engineering" itself. Nuclear engineering is more than a set of standard engineering courses with nuclear physics added. Similarly, bio-medical engineering has its own body of combinational knowledge, now being assembled. The propagation of elastic waves through tissue cannot be described by the equations of any problem in inanimate mechanics: the terms and parameters needed in this case only occur here. Although multipole radiation theory does apply in the body, the field is neither "near" nor "far," but intermediate, a case rarely needed inanimately. In the nervous system the "noise" of an incoming time-series is converted from the frequency domain to a spatial domain. Bio-systems encompass enormous dynamic ranges by shifting mechanisms. The "velocity-constants" in biochemistry turn out to be hormone-operated valves in systems with neural components. Bio-transducers have modulatable gage-factors. And the adaptive, navigational, goal-seeking, anticipatory aspects of even simple processes are striking. In both analytic and synthetic aspects, a bio-system raises new physical problems not familiar in inanimate engineering. Their solutions constitute the interdisciplinary science.

Educationally speaking, then, there already exists *interdisciplinary* knowledge which might be surveyed in half a year. We have a fairly extensive quantitative and theoretical formulation of the basic physical properties and effects of bio-materials and structures, as to electromagnetic, sonic, mechanical (including fluid) coefficients: absorptive, transmissive, and conductive. The photic, radiation and nuclear aspects are well explored. Secondly, a number of bio-processes have been described in mechanistic terms between input and output; many simple modular fragments and chains of action are now in significant analytic form, transcending mere families of curves. These raise at once challenging problems of formulation. Thirdly, there is the integrative aspect: what we know at present of bio-systems (control, adaptive, goal-seeking, self-organizing) at various levels from intracellular to organismic behavior. Not a single problem is completely solved, in the sense of explaining each function in the flow-diagram, but the loops and nodes are known in a few cases. Many are partly explored and dynamically modeled in segments. Fourthly, the bio-medical engineering scientist must readily present his observations into computable form. Training in the variations of parameters by analogs, and with transforms and combinations by digitals, is now part of the art of experimentation, as well as interpretation, which he can contribute to a bio-research team. Biological data (especially clinical) is so massive and digital computers still so awkward in handling it, that one must keep up in this field. The representation of multidimensional biological data, and its optimization is also our task. Lastly the biological system itself as a computer also warrants intensive study. Its powers exceed and its methods excel our present computing approaches.

In sum, we have just sketched a solid year of bioscience and of its interaction with advanced engineering science. To do this and qualify for the engineering doctorate in some schools may require an extra year. A compromise may be possible in the *dissertation*. A biological problem from an engineering point of view need not involve the anatomy of living tissue; or, if so, this part may be collaborative. The conceptual, analytic and experimental design, the observations and analyses may suffice for the engineering degree. Can such a graduate turn at once to productive work? All our experience with doctoral education says that he must first gain experience and knowledge of practical ways. In industry this is systematic, and should also be so in bio-medical engineering. A wide range of biomedical methodology is not hard to acquire by "rotating internship" in a well-organized training program. During this time, one can also dig deep in a special field—nervous, cardiovascular, behavioral, endocrine—and gain that expertise by which one becomes self-supporting and self-generating of research results.

A man so educated is a generalist. He can move where opportunity offers. Not only can he supervise the design of *instrumentation* from the applications point of view, but also he can contribute valuably to design of the biological experiment in terms of observations which employ mathematically feasible models and hypotheses which connect the biological and physical science. Therefore he has a future not only as a consultant, but also an educator, which qualifies him for academic appointments in the bio-medical milieu to the highest level of his ability.

Many engineers in biomedical research are not yet convinced that there exists such a field of science. Or if there does, that the time has not yet come for an organized academic pattern. This is a matter of temperament and judgment. The fact remains that such programs are now widely being established, and, by cooperation, are organizing the interdisciplinary knowledge. By the *standards of competence and excellence* they set, the graduates in biomedical engineering can gain respect and recognition worthy of the reputation of their schools.

Section 20

MICROWAVES

Organized with the assistance of the IRE Professional Group on Microwave Theory and Techniques

Survey and History of the Progress of the Microwave Arts*

GEORGE C. SOUTHWORTH†, FELLOW, IRE

Summary—**This is a review of the history and technical progress of the microwave techniques beginning with fundamental research and continuing to present-day practical applications. Particular emphasis is placed on the evolution of techniques that are now of great practical importance. Included are the ordinary 2-wire transmission line and the useful tuning devices that it has provided and also the more recent waveguide techniques which have added not only a new medium of transmission but new antennas for radiating and receiving electromagnetic waves, and a new type of circuitry for dealing with microwaves. Reflecting discontinuities, sometimes conductive and sometimes reactive, when judiciously located inside a waveguide, may perform numerous useful functions such as matching transformers, frequency filters and special networks that very knowingly allow wave power to pass in one direction and not in the reverse. The present paper is a greatly condensed portion of certain chapters of a new book entitled "Forty Years of Radio Research."[1] The latter tells of the author's personal experiences, not only in connection with microwaves, but in connection with the development of much of the rest of the radio spectrum as well.**

Introduction

MICROWAVES represent the newest and possibly the most useful band of the entire radio spectrum. Beginning at roughly one thousand million cycles, known as one kilomegacycle (sometimes called one gigacycle) and abbreviated 1 kMc, it extends upward to a frontier now temporarily poised at about 30 kMc. These limits correspond to wavelengths of 30 cm and 1 cm, respectively. Current research extends far beyond, possibly to 100 kMc or even higher. Included in this new band are all of the known forms of electrical communications—telegraphy, telephony, television and radar.

Microwaves are characterized by various new techniques. In place of the coils and condensers of conventional radio, which upon approaching this frequency range were becoming vanishingly small, equivalent results in this new band are accomplished by properly spaced discontinuities. Also in this band, radio directivity, one of the time-honored objectives of radio engineering, is at its best. In this band too, another of the objectives of electrical communications is also at its best; that is, the transmission from one point to another, almost as a single package, of a maximum of information. For example there may be transmitted over the same facility, alternatively, either several television channels, or many hundreds of telephone channels, or a very great number of telegraph channels. Although this is the proud accomplishment of a radio system in the microwave range, there is in prospect in this same range of frequencies, a waveguide transmission line that has far wider bandwidth potentialities. Thus, in this upper register of frequencies both radio and guided

* Received by the IRE, January 3, 1961; revised manuscript received, November 20, 1961.

† Bell Telephone Laboratories, Inc., Murray Hill, N. J.

[1] G. C. Southworth, "Forty Years of Radio Research," Gordon and Breach Scientific Publishers, New York, N. Y.; 1962. (350 pp. 60 Figs.)

waves may become worthy competitors for similar jobs. These and other features of this new band of frequencies will be discussed below.

It is not known for certain when the term *microwaves* originated. It is known, however, that a somewhat related term *microrayons* was used about 1933 by Andre Clavier and his associates of *Les Laboratories Le Materiel Telephonique* in describing a newly developed directive radio system for communicating across the English Channel. The frequency was about 1.75 kMc ($\lambda = 17$ cm). The available power was roughly 1 watt. Directive antennas used at each end, having power gains of 2000 (33 db), gave the 1 watt of power the effect, by ordinary radio standards of the time, of 4 kw. These directive gains are comparable with present-day microwave practice. This was a truly pioneering venture as it pointed the way to the rich rewards that come from the use of the higher radio frequencies.

Beginnings of Microwaves

Like radio itself, and like many other branches of electrical technology, microwaves had their origin in the fundamental electromagnetic theory. One group of mathematical physicists, bent on explaining the behavior of conventional two-wire lines, adapted the electromagnetic equations to the cases of two parallel conductors and also to two coaxial conductors. From these deductions developed much of the basic theory underlying ordinary transmission lines and also the various standing-wave phenomena now so widely used in microwave circuitry. Somewhat later, another group of mathematicians fitted the same electromagnetic equations to hollow metal pipes and to dielectric wires and, by invoking similarities with ordinary two-wire lines, they extended the theory to waveguides generally. These two techniques, coaxial and waveguide, remain as important methods for dealing with microwaves. Often the roles are complementary. Sometimes they appear in the same piece of apparatus, often in a way that makes it difficult to distinguish between the two. In recent years, a very useful microwave technique has grown from a special adaptation of two-conductor theory. In this case two flat conductors are mounted on opposite sides of a thin insulator. This is referred to as strip-line technique.

Early History of Waveguides

Waveguides, as contrasted with coaxial lines, have an extremely interesting past. That it might be possible to transmit electromagnetic waves through hollow metal pipes must have occurred to physicists almost as soon as the nature of electromagnetic waves became fully appreciated. That this might actually be accomplished in practice was probably in considerable doubt, for certain conclusions of the mathematical theory of electricity seemed to indicate that, without a return conductor, it would not be possible to support inside a hollow conductor the lines of electric force of which waves were assumed to consist. Heaviside was one of the early doubters.[2]

Perhaps the first analysis suggesting the possibility of waves in hollow pipes appeared in 1893 in the book "Recent Researches in Electricity and Magnetism" by J. J. Thomson.[3] This book, which was written as a sequel to Maxwell's "Treatise on Electricity and Magnetism," examined mathematically the hypothetical question of what might result if an electric charge should be released on the interior wall of a closed metal cylinder. Even now, this problem is of considerable interest in connection with resonance in hollow metal chambers. A much more significant analysis, relating particularly to propagation through dielectrically filled pipes, both of circular and rectangular cross section, was published in 1897 by Lord Rayleigh.[4]

As regards experimental verification, it is of interest that Sir Oliver Lodge[5] as early as 1894 approached, but probably did not quite realize, actual waveguide transmission. Of much greater significance were some experiments reported a year later by Viktor von Lang[6] who repeated for electric waves the interference experiment that had been performed for acoustic waves by Quincke some years earlier. Other similar experiments were performed later by Drude[7] and Weber.[8] All these experiments were done with the damped waves from spark discharges.

In about 1913 Professor Zahn[9] of the University of Kiel became interested in this problem and assigned certain of its aspects to Schriever and Reuter, two young candidates for the doctorate. They had barely started when World War I broke out, and both left for the front. Reuter was killed at Champagne in the autumn of 1915, but Schriever survived and returned to complete his thesis in 1920,[10] using for his source the newly available Barkhausen oscillator. Schriever's work was aimed specifically at dielectric wires, and was the first to use continuous waves.

The contributions of Thomson, Rayleigh, and their followers were, of course, purely mathematical. Those of von Lang, Drude, Weber, and Schriever were experimental, but they were of rather limited scope. The con-

[2] O. Heaviside, "Electromagnetic Theory," reprinted by Dover Publications, Inc., New York, N. Y., vol. 1, p. 399; 1893.

[3] J. J. Thomson, "Recent Researches in Electricity and Magnetism," p. 344; 1893.

[4] Lord Rayleigh, "On the passage of electric waves through tubes or the vibrations of dielectric cylinders," *Phil. Mag.*, vol. 43, pp. 125–132; February, 1897.

[5] O. Lodge, *Proc. Roy. Inst.*, vol. 14, p. 321; 1894.

[6] V. von Lang, "Interenzversuch mit Electischen Wellen," *Sitzber, Ges. Wiss. Wien, Abt. II*, vol. 104, 1895; p. 989, 1896. Wiedemann, *Ann. Physik und Chemie*, vol. 57, p. 430; 1896.

[7] P. Drude, "Über die Messung Elektrischen Wellenläugen mittels der Quickéschen Interferenzrohre," *Ann. Physik und Chemie*, vol. 65, p. 481; 1898.

[8] R. H. Weber, "Electromagnetic waves in metal pipes," *Ann. Physik*, vol. 8, pp. 721–751; July, 1902.

[9] H. Zahn, "Detection of electromagnetic waves on dielectric wires," *Ann. Physik*, vol. 49, pp. 907–933; May, 1916.

[10] O. Schriever, "Electromagnetic waves in dielectric conductors," *Ann. Physik*, vol. 63, pp. 645–673; December, 1920.

cept of the hollow pipe as a useful transmission element, for example as a radiator or as a resonant circuit, apparently did not exist at these early dates. Nothing was yet known quantitatively about attenuation, and little or nothing of the present-day experimental technique had yet appeared. At this time, the position of this new art was perhaps comparable with that of radio prior to the time of Marconi. The reader who wishes to learn more about the early history of microwaves should consult standard textbooks on the subject.[11]

Later History of Waveguides

The history of waveguides changed abruptly in about 1931 when it was shown for the first time that they could be put to practical use. Several patent applications were filed, and numerous scientific papers[12–15] were published. More recently, a great many papers have appeared; too many, in fact, for detailed consideration at this time.

The author's interest in guided waves stems from some experiments done while working with Lecher wires in a trough of water about 1920. In one case there were found, superimposed on the guided waves that might normally travel along two parallel conductors, other waves having a velocity that somehow depended on the dimensions of the trough. These may now be identified as the so-called dominant or TE_{11} type. Eleven years later, this work was resumed, and since that time a continued effort has been made to develop a useful technique for dealing with microwaves from fundamental principles of waveguide transmission.

The earliest experiments consisted of transmitting electromagnetic waves through tall cylinders of water. Thus, it became possible to set up, in the relatively small space of one of these cylinders, many of the wave configurations predicted by theory. In addition it was possible, by producing standing waves, to measure their apparent wavelength and thereby calculate their phase velocity. Also, by investigating the surface of the water by means of a probe, the directions and also the relative intensities of lines of electric force in the wave front could be mapped. It is probable that certain of these modes were observed and identified for the first time. Figs. 1 and 2 show the nature of this early apparatus and the types of wave configuration that were mapped. A more complete account of this early work may be found in an early article by the author.[14]

[11] G. C. Southworth, "Principles and Applications of Waveguide Transmission," D. Van Nostrand Co., Inc., New York, N. Y.; 1950.

[12] J. R. Carson, S. P. Meade, and S. A. Schelkunoff, "Hyperfrequency waveguides—mathematical theory," *Bell Sys. Tech. J.*, vol. 15, pp. 310–333; April, 1936.

[13] G. C. Southworth, "Hyperfrequency waveguides—general considerations and experimental results," *Bell Sys. Tech. J.*, vol. 15, pp. 284–309; April, 1936.

[14] G. C. Southworth, "Some fundamental experiments with waveguides," Proc. IRE, vol. 25, pp. 807–822; July, 1937.

[15] G. C. Southworth and A. P. King, "Metal horns as directive receivers of ultra-short waves," Proc. IRE, vol. 27, pp. 95–102; February, 1939.

Shortly afterwards (1933), sources giving wavelengths in air of 15 cm became available, and the experimental work was transferred to air-filled copper pipes only 5 in in diameter. At this time, a 5-in hollow-pipe transmission line 875 ft in length was built, through which both telegraph and telephone signals were transmitted (Fig. 3). Measurements showed that the attenuation was relatively small.

It was recognized at an early date that a short waveguide line, with suitable modification, might function as a radiator and also as a reactive element. Most obvious were the electromagnetic horn and the resonant cavity. These properties were likewise investigated experimentally, and numerous useful applications were proposed. Descriptions of these early methods may be found in another early article by the author[13] as well as in several of the early patents. As will be noted, modern waveguide circuitry had its beginnings in the efforts to obtain a more efficient transfer of microwave power from a source to a waveguide transmission line, thereby providing the elements of a transmitter, and again the efficient recovery of the microwave power at the receiving end, thereby providing the elements of a receiver.

As might be expected, a great many people contributed in one way or another to the early success of this venture. Particular mention should be made of the very important parts played by the author's colleagues, A. E. Bowen and A. P. King, who, during its early and less promising years, contributed much toward transforming rather abstract ideas into practical working equipment, such as frequency meters, standing-wave detectors, and terminations. Much of this equipment was found to be of important military use immediately upon the advent of war. Also of great importance were the parts played by the author's colleagues, Dr. S. A. Schelkunoff, J. R. Carson, and Mrs. S. P. Meade, who, in the early days of this work, provided a substantial segment of mathematical theory that previously was missing. During the succeeding years, Dr. Schelkunoff, in particular, made invaluable contributions in the form of analyses which, in some cases, indicated the direction toward which experiment should proceed and, in others, merely confirmed experiment, while in still others, gave answers not readily obtainable by experiment alone. A particularly important result of this analysis was the discovery that, for one of the types of waves that may be propagated through a circular metal pipe of given diameter, the attenuation decreases with increasing frequency. It is expected that this principle will play a very important role in the future of microwaves.

Beginning sometime prior to 1936, Dr. W. L. Barrow, then of the Massachusetts Institute of Technology, also became interested in this subject and, together with numerous associates, made very substantial contributions particularly in the direction of determining the best proportions for electromagnetic horns. No less than eight scientific papers were published covering

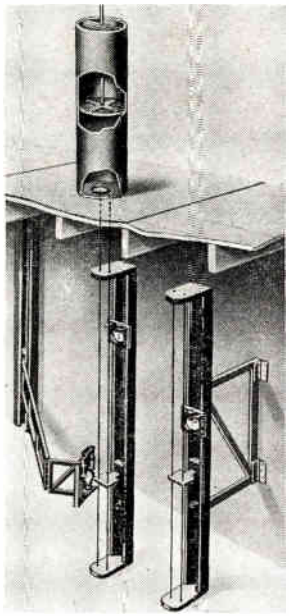


Fig. 1—Apparatus used to identify guided waves. Relative velocity and cutoff frequency were measured. A probing crystal detector and meter made it possible to identify various modes.

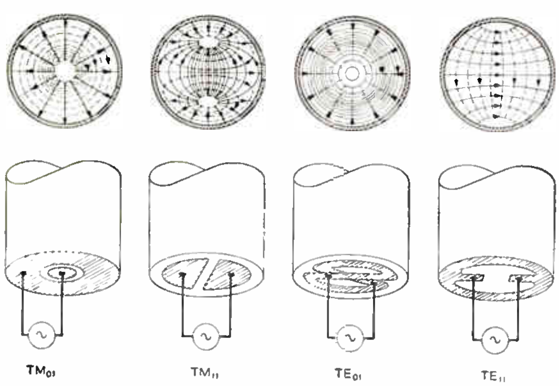


Fig. 2—Schematic of various launching devices used to set up various modes.

special features of hollow-pipe transmission lines and electromagnetic horns.[16–18] For several years the work at the Massachusetts Institute of Technology and at the Bell Telephone Laboratories probably represented the major portion of, if not indeed the only, work of this kind in progress, but with the advent of World War II, hundreds or perhaps thousands of others entered the field. For the most part, the latter were workers on various military projects. Beginning with the considerable accumulation of unpublished technique which was made freely available to them at the outset of the war, these workers, along with others in similar positions

[16] W. L. Barrow, "Transmission of electromagnetic waves in hollow tubes of metal," Proc. IRE, vol. 24, pp. 1298–1328; October, 1936.

[17] W. L. Barrow and F. M. Greene, "Rectangular horn-pipe radiators," Proc. IRE, vol. 26, pp. 1498–1519; December, 1938.

[18] W. L. Barrow and L. J. Chu, "Theory of the electromagnetic horn," Proc. IRE, vol. 27, pp. 51–64; January, 1939. Also, *Trans. AIEE*, vol. 58, pp. 333–338; July, 1939.

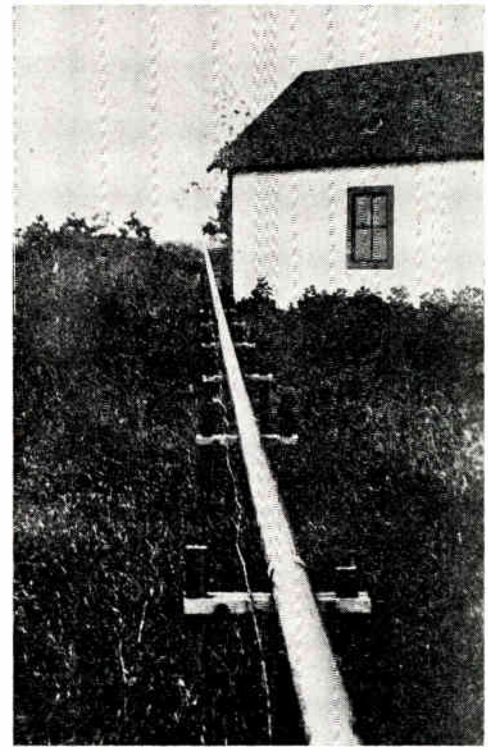


Fig. 3—A section of the first waveguide transmission line built August, 1933.

elsewhere in this country and in Europe, have helped to bring this technique to its present very satisfactory state of development.

More About the Prewar Development

The progress made before World War II is conveyed in part by Figs. 4–9. Fig. 4 shows numerous pieces of measuring equipment fundamental to the development of microwaves. It is representative of techniques as of about 1934. Figs. 5 and 6 show a platform as it appeared for the presentation of two demonstration lectures before the IRE on February 2, 1938, and again on February 1, 1939. In the first, four of the more important modes were produced in the apparatus shown, and their orientations were plotted for the audience on a blackboard placed over the several sources. As an added feature, 3-kMc ($\lambda = 10$ cm) waves were produced and were propagated through a 3-in metal pipe and received in a loosely-coupled resonant chamber.

The second demonstration, in 1939 (Fig. 6) had two objectives. In the first objective, a match termination was developed by a step-by-step process as the audience observed the gradual disappearance of the standing wave. By this time the standing-wave detector had become a standard piece of measuring apparatus much as we know it today. The second objective of this demonstration was to plot for the audience the directive patterns of several electromagnetic horns. Both lectures were repeated before various sections of the IRE.

A third demonstration lecture was planned which was aimed particularly at reactive elements in a waveguide,

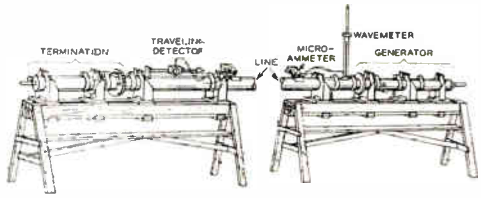


Fig. 4—Optical benches containing microwave equipment typical of 1934.

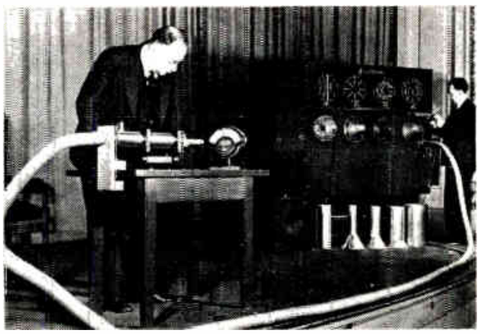


Fig. 5—First demonstration of waveguides before the IRE on February 2, 1938. This emphasized particularly, the different modes of transmission and their respective cutoff frequencies.

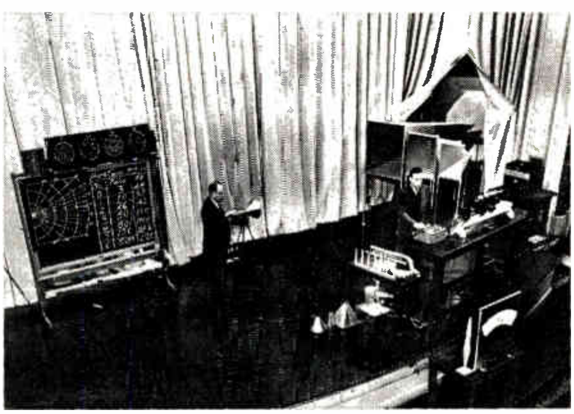


Fig. 6—Second demonstration of waveguides before the IRE on February 1, 1939. This emphasized measuring techniques as well as the electromagnetic horn.

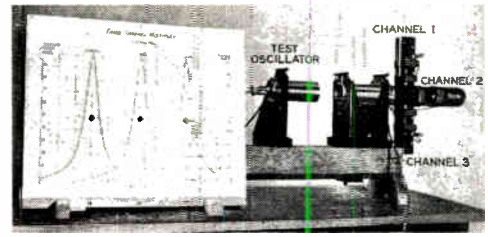


Fig. 7—An early form of three-channel multiplex and its measured frequency characteristic as of 1940.

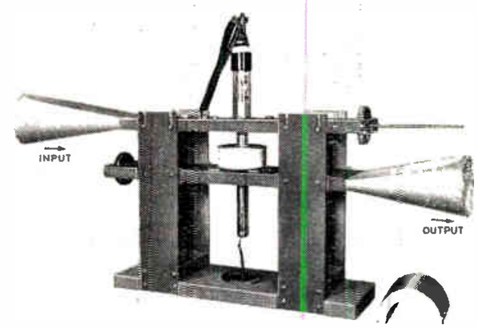


Fig. 8—An early adaptation of the klystron amplifier to waveguide use, 1939.

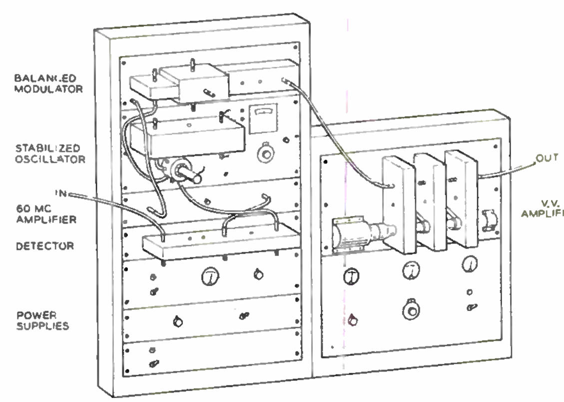


Fig. 9—Microwave apparatus as of the summer of 1941. Sketch from the notes of a project engineer for a proposed 3000-Mc ($\lambda = 10$ cm) microwave repeater for relay use. The apparatus was built, but its assembly and testing was interrupted by war effort.

including such composites as filters, transformers, and waveguide matching devices, but the gathering war clouds over Europe made its presentation appear unwise.

The principle of multiple reflection from discontinuities and the associated principle of cavity resonance played an important part in microwave development. In some cases they were used to match a source of power to a waveguide. In others they served to match the waveguide to a receiver, perhaps a crystal detector. In still others they served to pass freely a band of frequencies, perhaps a television channel, while discriminating sharply against adjacent frequencies. Together these principles formed the foundations of microwave circuitry.

The evolution of microwave circuitry was greatly aided by mathematical theory. The latter showed that the various reflecting discontinuities, such as irises and semiconducting bodies could produce certain impedance effects not unlike those of familiar radio practice. The relationships, however, were rather complicated and in extreme cases involved hyperbolic functions, a subject which many engineers find difficult. Engineering analysis was greatly facilitated by certain graphical analyses devised by Phillip H. Smith.[19] Not only were laborious calculations avoided, but, while the work was in progress, the engineer could visualize the step-by-step processes under way. Indeed, in some cases he could look

[19] P. H. Smith, "Transmission line calculator," *Electronics*, vol. 12, pp. 29–31; January, 1939. "An improved transmission line calculator," *Electronics*, vol. 17, pp. 130–133, 318, 320, 322, 324–325; January, 1944.

forward to the answer. Few gadgets of microwave circuitry have been more useful than the Smith diagram.

Special Materials

There developed at an early date a need for special materials. Not only was there a need for metals of good conductivity, for example, silver, but there was a need for special dielectrics and also a need for magnetically permeable materials. The latter were not readily achieved, but ultimately they too appeared. Not always was there a need for low-loss materials, but indeed sometimes the need was for materials with substantial loss. Early forms were resistance films and also absorbing plugs placed in the path of guided waves. As they were made more stable, some became standards of attenuation and others became terminators for microwave transmission systems. Such materials have played a very important part in the development of methods of microwave measurements.[20]

Sometimes the need was for rather unusual materials. One such need was for a material whose resistance varied nonlinearly with amplitude. Such a material already existed in the form of the crystal detector of early radio practice. Starting from these humble beginnings, the crystal detector was greatly improved by Russell Ohl of the Bell Telephone Laboratories and many others, and became a thoroughly stable rectifier. With the advent of double detection in microwaves followed by certain related modulation processes, the crystal detector became an extremely important device, in both the modulation and the demodulation of signals. An even more important use of this principle made it possible to generate harmonics and thereby extend materially the frequency frontier of microwave research. A great deal of research has gone into the improvement of good nonlinear materials. It is especially interesting, too, that this research triggered off the chain reaction that led ultimately to both the transistor and the solar battery, as well as to the important branch of physics known as physics of the solid state.

Another important microwave device of the prewar era was the thermistor. It depends for its success on the properties of certain materials, notably the oxides of manganese, cobalt, nickel, and copper, to change resistance with temperature. A tiny bead made from a mixture of these materials, when matched to a waveguide and heated by received wave power, provides a measure of microwave power. It first came into use in 1936. The wavemeter and the standing-wave detector shown in Fig. 4, together with the stabilized detector, the thermistor, and the standard decibel attenuator provided many of the measurements necessary for the early development of microwaves. Today all remain important elements in microwave measurements.

In the late summer of 1940 a visiting mission from England brought to America their best radar techniques, including the pulsed magnetron, a device giving not only greatly increased power, but an output signal that was well adapted to radar. At the Bell Telephone Laboratories, in particular, the visitors saw the latest in microwave techniques. These later proved of great value in pushing radar to the higher frequencies needed for accurate bombing and for the precise directing of naval gunfire.

[20] G. K. Teal, M. D. Rigterink and C. J. Frosch, "Attenuator materials, attenuators and terminations for microwaves," *Trans. AIEE*, vol. 67, pp. 754–757; August, 1948.

Wartime Development

The development of microwaves was greatly accelerated in the late summer of 1940 when the National Defense Research Committee was formed, with radar as one of its principal objectives. This organization promptly set up, both at the Massachusetts Institute of Technology and at Columbia University, special laboratories to develop and apply microwave techniques to radar problems. This grew rather rapidly to a huge organization of several thousand top-flight scientists marshalled mainly from universities of the United States and Canada. In addition, very important but less publicized parts were played by other organizations.

Microwave research continued at the Bell Telephone Laboratories, and shortly there followed radar designs for field use, for use in fire control on ships and for use on both submarines and airplanes. Comparable research and design came from the Radiation Laboratory of the Massachusetts Institute of Technology. Composites of the best designs were subsequently manufactured in quantity. Fig. 10, selected largely for its ready availability, is representative of the end results of this work, and is important because it represents early fruition of microwave techniques.

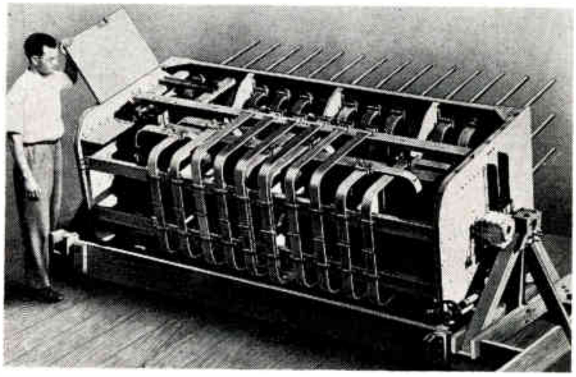


Fig. 10—An early form of microwave radar utilizing waveguide techniques. Not only were microwaves guided to the various antennas, but the phase was varied progressively, thereby causing the beam to scan a distant landscape.

The fundamental building blocks comprising microwave techniques were by no means complete before the War. A great many came during the War and a few followed the War. Because of wartime restrictions and the necessity of close team work, individual credits were often sacrificed. Mention should be made, in particular, of devices by which phase could be added progressively to a waveguide line. This was a product of A. G. Fox of the Bell Telephone Laboratories in 1941.[21]

Also, there were junctions of four waveguides so arranged that when power was introduced in one branch it divided equally between two alternate branches while being balanced at the fourth branch. This had a counterpart in the hybrid balance of ordinary telephone practice. It was proposed by W. A. Tyrrell of the Bell Telephone Laboratories in 1941.[22] Almost as important was a particular form of the balance in which there was an uneven balance between adjacent arms while a balance was maintained at the fourth arm. The latter became known as the directional coupler. Its origin is somewhat obscure because of the many people in war work at the time. However, one early description was given by Mumford in 1947.[23] All of these devices promptly found practical use.

During the War, it was discovered by Robertson and King[24] that falling rain could absorb and scatter large amounts of passing wave power at all frequencies above about 10 kMc. Also it was predicted by Van Vleck and verified by Becker and Autler that there was an absorption band for water vapor (not water droplets) at 22.3 kMc. Other absorption bands were predicted by Van Vleck and verified by Beringer for oxygen at 60 kMc and 120 kMc as well as many others at still higher frequencies. Rain absorption in particular seemed very important for it posed a considerable threat for the future use of the higher frequencies for purely radio uses.

The absorption bands due to water vapor and oxygen came as no surprise, for Professor Williams and Dr. Cleeton[25] of the University of Michigan, several years before, had found that ammonia gas absorbed a narrow band of frequencies centering around 45,000 Mc. This was the forerunner of some very fruitful molecular and atomic research in which the microwave techniques were to play a very important role.

It was in this period of microwave research, too, that it was discovered that measurable amounts of microwave radiation could be received from the sun.[26] This discovery, together with Jansky's discovery of noise from interstellar space, ushered in modern radio astronomy. Although initiated in America, this work was quickly endorsed in both Australia and England, and impressive research programs were set up promptly. Radio astronomy now represents one of the more important areas of fundamental investigation, with budgets amounting to millions. It is the proud accomplishment of microwaves development that it should be able to aid in the exploration of both the infinities of outer space and the infinitesimal interior of the atom.

In this wartime period microwaves virtually leapfrogged their way across the frequency spectrum from 3 kMc to possibly 30 kMc, all in a period of two or three years. In each case new sources of power were needed to make a new beachhead while the needed measuring apparatus was being developed, preliminary to the evolution of new electronic devices. To break an apparent stalemate and supply, when needed, adequate research tools, the expedient was used of generating, by means of crystals, the harmonics of such frequencies as were readily available. Steps from 3 kMc to 9 kMc to 27 kMc were typical. Good measuring equipment including wavemeters, standing-wave detectors, power-measuring devices, and standard attenuators were built for all of these ranges prior to 1943. Today harmonic production remains one of the favored methods of exploring the frequency frontier. Microwave measurements were naturally an essential part of microwave research.

Microwaves in the Post-War Period

Although the number of fundamentally new principles of the microwave techniques diminished after the War, the number of applications increased markedly. Released from urgent wartime responsibilities, engineers turned naturally to peacetime uses. Communication, as contrasted with radar, was perhaps the most obvious alternative. Research had already evolved various microwave sources and also before the War had come an amplifier for such signals. The postwar period was to see a marked extension of its frequency range. The klystron, the first practicable electronic device to break from conventional principles, was a most welcomed prewar entrant to the microwave field. Soon came the closely-spaced triode, an ingenious adaptation of the space-charge principle to the new microwave task. Shortly, there came the traveling-wave amplifier, a device that could amplify extremely wide bands of frequencies. All three devices could assume the form of oscillators as well as amplifiers, thereby providing sources of microwave power at ever increasing frequencies. Details of the development of electronic devices are covered in other articles of this issue. These have been important in the history of microwaves, for they have paved the way to wider use.

Microwaves were aided tremendously in the postwar period by the development of new materials. Of particular importance, were the ferrites, insulating materials that exhibited pronounced magnetic properties. Particular ferrites were able to produce, in the microwave range, the effect of Faraday rotation so familiar in optics. One result of this development was a new family of microwave circuit components, each capable of a

[21] A. G. Fox, "An adjustable waveguide phase changer," Proc. IRE, vol. 35, pp. 1489–1498; December, 1947.

[22] W. A. Tyrrell, "Hybrid circuits for microwaves," Proc. IRE, vol. 35, pp. 1294–1306; November, 1947.

[23] W. W. Mumford, "Directional couplers," Proc. IRE, vol. 35, pp. 160–165; February, 1947.

[24] S. D. Robertson and A. P. King, "Effect of rain upon the propagation of waves in the 1- and 3-centimeter regions," Proc. IRE, vol. 34, pp. 178–180; April, 1946.

[25] C. E. Cleeton and N. H. Williams, "Electromagnetic waves of 1.1 cm wave-length and the absorption spectrum of ammonia," *Phys. Rev.*, vol. 45, pp. 234–237; February, 1934.

[26] G. C. Southworth, "Microwave radiation from the sun," *J. Franklin Inst.*, vol. 239, pp. 285–297; April, 1945.

variety of useful functions. In one particular case, a ferrite device passed microwave power freely in one direction, while discriminating sharply against power in the reverse direction. All of this may be accomplished without the loss of one half of the transmitted power that is inherent in most wave-balancing systems. (This material is described more completely in other articles of this issue.)

Other devices playing an increasingly important part in microwaves resulted from the further study of the general subject of solid-state physics, which has already led to such important devices as tunnel diodes, parametric amplifiers, masers, and to low-noise receivers, generally. The latter are particularly useful in radio astronomy and also offer promise of being useful in the intercontinental communications by way of reflecting satellites. It is too early to properly evaluate the importance of such devices, but it is quite obvious that they will be very important.

Some Postwar Research and Application

While the development of microwave radar has continued unabated for both military and peacetime uses, many other applications have appeared. In particular, microwaves have become very useful in the radio relaying cross-country of broadbands of visual, audio, and telegraph signals. Continuing a program started in 1941 but interrupted by the War, the Bell System resumed the task in 1943 and in 1947 set up a thoroughly practicable radio relay system between Boston and New York on a frequency centering of about 4 kMc. Repeater stations located at intervals of perhaps thirty miles picked up attenuated signals, demodulated them to a relatively low frequency, and after remodulation and amplification transmitted them onward at a slightly different frequency. The system often referred to as TD-2 was soon extended from New York to Chicago and ultimately to the Coast. It is now a vast network estimated in 1960 at 40,000 route miles comprising nearly 40 per cent of the total Bell System intercity circuit mileage. In the meantime similar systems have been developed by others in America and elsewhere. In particular, considerable use of microwave radio relay is also being made by the U. S. Government, the Western Union Telegraph Company and the various pipe-line companies and Turnpike Commissions. The combined use of such systems is estimated (1960) at 32,000 route miles.

Improvements in microwave relay systems have made it possible, by using frequencies as high as 6 kMc, to transmit over the same system alternatively either 10,800 telephone channels or perhaps a half dozen television channels. This newer application is sometimes referred to in Bell System circles as the TH system.

In 1960 experiments were started looking toward microwave communications over transoceanic distances, through the intermediary of a reflecting balloon orbiting at perhaps 1000 miles above the surface of the earth. Satellite periods of two hours were typical. Though the useful period of a single orbiting reflector is, according to this plan, definitely limited to a few minutes for each transit, it is expected that several such reflectors will provide practically continuous service. Frequencies of the order of 7 kMc are contemplated. Although the research leading to this rather unusual relay system, has, for the most part, made use of conventional methods, there are notable exceptions. Like the receivers of radio astronomy which look toward the low temperatures of interstellar space, these receivers may also profitably use low-noise detectors. This has entailed an interesting bit of solid-state research on the fringe of microwave research.

Much of the early research with microwaves centered around the so-called dominant or TE_{11} mode, from which have flowed most of our present-day microwave applications, both in radio and radar. The dominant mode is but one of the double infinity of modes or configurations that are possible in a waveguide. There is another mode of great practical interest, particularly in cases where it is desirable to guide microwaves to great distances. In this mode, which is known as TE^0_{01}, the lines of electric force are everywhere coaxial circles, parallel to the walls of the circular conducting pipe and for this case the attenuation, since it involves only tangential currents, progressively decreases as the frequency is indefinitely increased. During the thirty years since this principle was first understood, there has been the constant hope that a microwave transmission line might sometime be evolved that would incorporate bandwidth with relatively low attenuation.

After years of painstaking research, the practical realization of this objective now seems assured. Attenuation of the order of 2.5 db per mile is typical with bands corresponding to perhaps 200,000 speech channels. This assumes pipes to be perhaps 2 in. in diameter. Thus future communications seem to be leading not only toward radio but toward a form of the time-honored wire line. Only time can tell how the work-load will be divided between the two media. It is likely that both will be used. Radio will probably be preferred both for the mobile services and for radar. But guided transmission will probably be preferred in cases where either privacy or freedom from interference is paramount.

Microwave Measurements*

HAROLD A. WHEELER†, FELLOW, IRE

Summary—**There are presented briefly the measurement techniques that are most closely associated with microwaves (the frequency range of 0.3 to 300 Gc or the wavelength range of 1 meter to 1 mm). After a logical classification of such measurements in confined space (such as waveguides) the subject is introduced by a review of the principles peculiar to microwaves. Typical of this frequency range is the standing-wave technique and the reflection chart, utilized for evaluating and presenting the "normalized" value of any impedance relative to the wave impedance in a waveguide. More conventional methods are employed for evaluating transmission properties (attenuation and phase shift) by comparison with primary or secondary standards. Various special methods are discussed, including some of those most in need of further development in the future.**

I. Introduction

The subject of microwave measurements will be presented briefly with reference to its general scope and some of its more interesting features. This article is one of the several sponsored by the Professional Group on Microwave Theory and Techniques (PGMTT) for the Fiftieth Anniversary Issue of the Proceedings of the IRE.

In the preparation of this brief article, it has been necessary to rely on a thin sampling of the great variety of techniques that are now associated with the subject. The emphasis will be placed on those techniques of measurement which are peculiar to the field of microwaves, or at least are most closely associated with this field. Also the attention will be directed mainly to laboratory tests essential for design and for proof of performance, rather than operational tests suitable for monitoring the behavior of equipment in the field.

Table I is a brief outline of some aspects of microwaves and their measurement. It indicates in some degree the limits that have been adopted for this article. The frequency range of 0.3 to 300 gigacycles per second corresponds to a wavelength range from 1 meter down to 1 millimeter, being roughly the limits over which a hollow waveguide is useful in single-mode propagation. Over this same range, space-charge electronics is utilized mainly in microscopic dimensions in the solid state (crystal rectifiers, etc.) while free electrons in a vacuum are utilized in terms of their transit time (magnetrons, klystrons, etc.). Measurements and standards are patterned after ac measurements at lower frequencies, but are typically based on electric and magnetic fields rather than voltage and current.

Frequency standards and measurements are omitted here because they are to be covered in a separate article sponsored by the Professional Group on Instrumentation (PGI). Free-space measurements are omitted because they fall in the scope of the Professional Group on Antennas and Propagation (PGAP) rather than the PGMTT.

TABLE I
Microwaves—Scope of this Article

- Frequency Range.
  - Frequency 0.3–300 Gc.
  - Wavelength 1 meter—1 mm.
- Location of Typical Phenomena.
  - Confined space.
    - Waveguides, hollow.
    - Rectifiers, crystal.
    - Electronics, transit-time.
  - Free space (outside of scope of article).
- Measurements and Standards.
  - Frequency, wavelength (EM field).
  - Modulation, spectrum.
  - Power (EM field products).
  - Impedance (EM field quotients).

This presentation will rely on outlines in tables for broad coverage, and these are intended to be self-explanatory. The text and illustrations will be directed to a few topics selected for their importance or particular interest in the field of microwaves. A carefully selected list of references is appended for further information.

II. Classification of Microwave Measurements

The next few tables give some outlines indicating the scope of certain major parts of the entire field.

Table II is an outline of measurements in confined space, analogous to wired circuits at lower frequencies. Here they are typically based on electric and magnetic (EM fields) rather than voltage and current. These fields are usually confined by conducting walls that serve as a shield. Otherwise, this outline is mainly taken over from lower frequencies and included here as a reminder.

Table III is an outline of some concepts that are common in the measurement of linear impedance with respect to waves in a guide (or in free space). It is customary to define the wave impedance of a square area of wavefront, or of the bounded wavefront in a waveguide, and then to use this value as the reference for "normalizing" any other impedance, such as an obstacle in a waveguide. The latter is then expressed simply by a ratio.

Implicit in these outlines are the phenomena and practices that are peculiar to microwaves and their measurement, as will here be discussed further.

* Received by the IRE, August 30, 1961; revised manuscript received, November 15, 1961.
† Wheeler Laboratories, Great Neck, N. Y. (A subsidiary of Hazeltine Corp.)

TABLE II

MICROWAVE MEASUREMENTS—CONFINED SPACE

Power (EM field products).
- Strength, level; average or peak power.
  - Strong (transmitter).
  - Weak (receiver).

Impedance (EM field quotients).
- Linearity, dependence on power level.
  - Linear, independent of power level.
  - Nonlinear, dependent on power level.
    - Rectifying.
    - Limiting.
    - Switching.
- Reciprocity, dependence on direction of transmission.
  - Reciprocal, independent of direction (uniport or multiport).
  - Nonreciprocal, dependent on direction (multiport).
- Activity, conversion of local power.
  - Passive, power gain impossible.
  - Active, power gain possible.
- Frequency dependence.
  - Incidental, nonselective, wide-band.
  - Essential, selective, narrow-band.

TABLE III

MICROWAVE MEASUREMENTS—LINEAR IMPEDANCE

Dimensions.
- EM quotient.
  - Impedance (Z).
  - Admittance (Y).

Ratios (in waveguide).
- Normalized Z or Y (relative to wave Z_0 or Y_0).
- Reflection coefficient (related to SWR).
  - Return loss.
  - Return angle, phase shift.
- Transmission coefficient.
  - Gain or loss, attenuation.
  - Angle, phase shift.

Parts, Real and Imaginary.
- Rectangular, in-phase and quadrature.
- Logarithmic, magnitude and angle.

III. PRINCIPLES PECULIAR TO MICROWAVES

In microwaves, the wavelengths (1 meter to 1 mm) are such that components of wavelength dimensions can be made and used in the laboratory and in operational equipment. In this respect, microwaves in the electromagnetic medium, free space, are comparable with sonic and ultrasonic waves in acoustic mediums (gases, liquids, solids); however, free space is unique in its monolithic set of electromagnetic properties, as distinguished from the idiosyncrasies of material mediums.

Microwave components, being of wavelength dimensions, have peculiarities in which they differ from components for the lower or higher frequencies.

At lower frequencies, a reactive component is specialized as to its properties of electric or magnetic energy storage, whereas a microwave component typically has a mixed field. At lower frequencies, connections are made by wires, and a pair of terminals is required at a "port," whereas microwave connections are typically made by pipes and flanges ("plumbing") and only an opening is required at a port [29], [30].

At higher frequencies (exemplified by light waves) the reverse is true. Components are many wavelengths in size, and intercoupling is accomplished through the mechanism of waves. The smallest waveguides (glass fibers) still support propagation in many modes. The highest degree of reflection is far inferior to that obtainable for microwaves.

The most distinctive technique of microwaves is the measurement of reflection coefficient by the use of a slotted section, and its presentation by plotting on the reflection chart. This is enabled by the convenience of wavelength dimensions in the laboratory. The following section is devoted to this subject.

Microwave signal sources rely on free electrons, as a rule, in applications utilizing their transit time. At lower frequencies, electron tubes utilize space charge, and there are various other kinds of sources. At higher frequencies, there are no sources of similar character, and coherent sources are just being discovered for certain discrete frequencies.

The use of transit time in microwave sources enables "voltage tuning" by control of electron velocity. Even in the complicated integral structure that is typical of microwave oscillators, there are designs in which voltage variation can sweep the frequency over a ratio of 1:2.

A microwave signal detector of simplest form is the crystal rectifier. It operates by virtue of electron space charge in a microscopic space in the solid state.

While any particular indicator for lower frequencies is usually responsive to voltage or current, the typical microwave indicator is responsive to power. It is customary to "tune" the load to a "matched" condition in which it takes all the "available power."

Where microwaves are propagated through a waveguide having lateral dimensions exceeding the wavelength, there are possible several modes of propagation. In such "multimode" behavior, each mode is analogous to a separate circuit. This is comparable with the series and parallel modes of a pair of wires over a ground plane.

Departure from reciprocity in two-port microwave devices is analogous to that at lower frequencies, but usually is manifested in a different kind of environment. The electronic amplifier relying on grid-voltage control of space current at lower frequencies is supplanted by one utilizing electric-field control of beam velocity. The familiar electromechanical gyrator (exemplified by the loudspeaker voice coil) has an analog in the magnetoelectronic electron-spin gyrator in ferrites at microwave frequencies. Such principles have been utilized in a variety of nonreciprocal devices that are convenient only in the field of microwaves.

The analysis of multiport networks in the terms of determinants, tensors and matrices has been extended to the form known as the "scattering matrix" [18], for which the writer prefers the more descriptive designation, "wave-transmission matrix." This form, which may have originated in the M.I.T. Radiation Laboratory, is a matrix of transmission coefficients among a set of ports. From one port to the same port, in the main diagonal of the matrix, this becomes the reflection coefficient. In the ideal form, the coefficients are all voltage ratios, or all current ratios, referred to a constant wave resistance which is the same at all ports. This emphasizes the applicability of such a matrix to power transmission among several ports in waveguide. It is this utility that is responsible for its common use in microwave devices, both reciprocal and nonreciprocal.

Two of the most common waveguide junctions offer the best examples of the wave-transmission matrix. One is the hybrid junction (exemplified by the "magic T") which has one pair of waveguides in series and parallel connections with another pair. The other is the directional coupler between a parallel pair of waveguides, each having ports at both ends. Their respective matrices have the following forms:

$$\begin{bmatrix} 0 & 0 & a & b \\ 0 & 0 & -b & a \\ a & -b & 0 & 0 \\ b & a & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & a & jb \\ 0 & 0 & jb & a \\ a & jb & 0 & 0 \\ jb & a & 0 & 0 \end{bmatrix}.$$

In either case, $a^2+b^2=1$, for conservation of power. In the simplest form, $a=b=1/\sqrt{2}$. The zeros in the main diagonal indicate nonreflecting or matched terminations. The zeros in the minor diagonals indicate decoupled or conjugate pairs. The latter are analogous to an ordinary bridge circuit in a balanced condition.

IV. Standing-Wave Technique and Reflection Chart

The technique of measuring reflection in a waveguide by observing a standing-wave pattern is well adapted to the wavelengths of microwaves. If the sole cause of the reflection is the load impedance at the output end of the waveguide, this impedance can be computed from the measured complex reflection coefficient in the manner proposed by the late P. S. Carter about 1937 and published in 1939 [23]. This is a practical method in which the waveguide is a primary standard of impedance, based on mechanical dimensions, and the standing wave provides a comparison ratio relative to the standard.

Fig. 1 shows the well-known slotted section of rectangular waveguide for observing the standing-wave pattern. The two usual observations are the distance (d) from the load back to the point of minimum voltage, and the standing-wave ratio (SWR) of maximum over minimum voltage (VSWR). The latter is conveniently expressed in decibels (db) as indicated. The distance is expressed as a number of wavelengths (λ_g) in the guide.

In practice, there may be several wavelengths of waveguide between the slot and the load. This separation becomes immaterial if, at the same frequency of measurement, a reference observation of this distance (d_0) is made with a short-circuit (SC) reflecting plane substituted for the load. Then the load impedance is determined by the shift of the minimum ($d-d_0$).

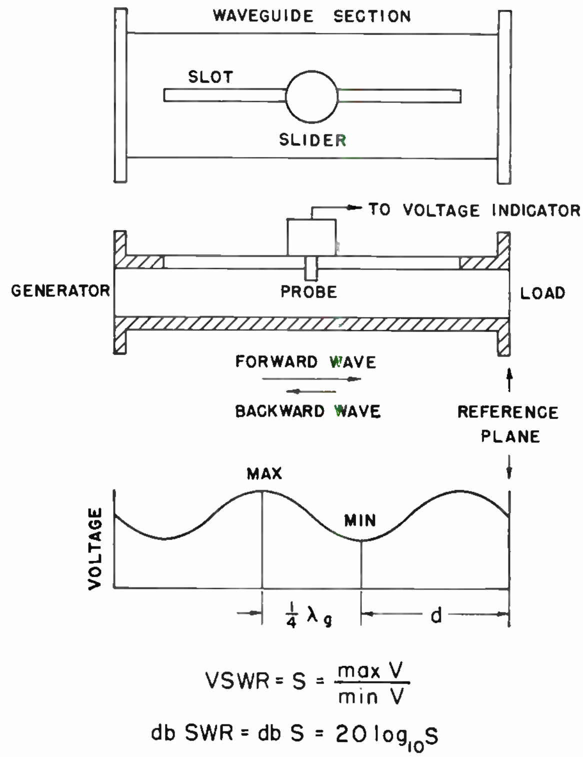


Fig. 1—The slotted section used for standing-wave measurements in a waveguide.

The complex voltage reflection coefficient ($VRC=\rho$) is determined from the observations as follows, in terms of its magnitude and angle:

$$|\rho| = \frac{S-1}{S+1}; \qquad \angle\rho = \theta = \frac{d-d_0}{\lambda_g}360° \pm 180°.$$

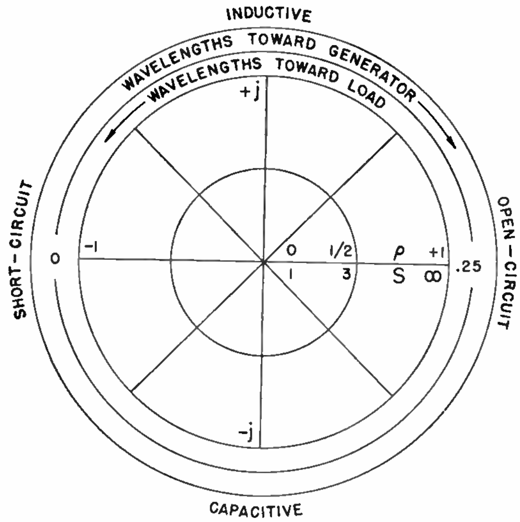
The load impedance (Z) is then computed relative to the guide wave impedance (Z_0) as follows:

$$\frac{Z}{Z_0} = \frac{1+\rho}{1-\rho}.$$

This ratio is the "normalized" load impedance. The computations by the above three formulas are usually avoided by the use of the reflection chart to be described.

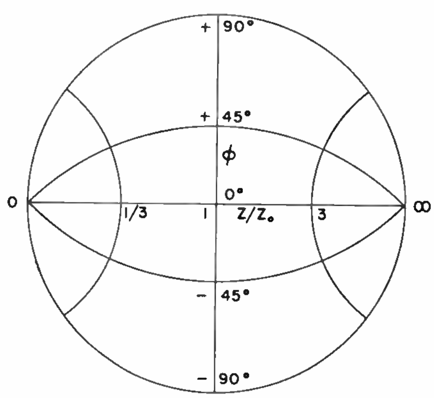
Fig. 2(a) shows the essentials of the reflection chart, in an arrangement which has been tentatively selected by a committee of Electronic Industries Association (EIA). The voltage reflection coefficient is plotted on the complex plane in the conventional manner, so all values fall within the unit circle [18], [23]–[25], [34]. The contours of magnitude and angle are respectively concentric circles and radials. The outermost notations refer to the character of the load impedance corresponding to various parts of the chart. The wavelength scales, having ½ wavelength around the circle, indicate the relation between angle and length of waveguide.

The contours of SWR also are concentric circles. Therefore it is customary to plot a point on the chart

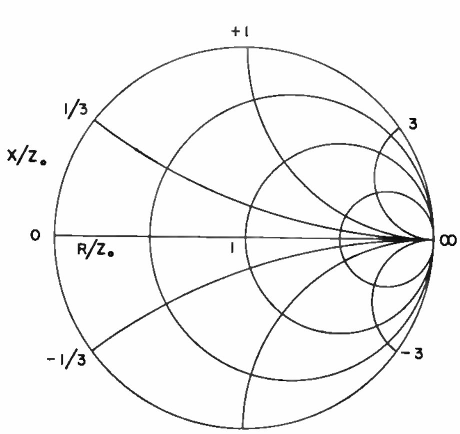


(a) Voltage reflection coefficient, $\rho = |\rho| \exp j\theta$; $|\rho|$ circles, θ radials.

$$\text{VSWR} = S = \frac{1 + |\theta|}{1 - |\theta|} \text{ circles.}$$



(b) Impedance, $Z = |Z| \exp j\phi$; $|Z|$ latitude, ϕ longitude. (Carter chart)



(c) Impedance, $Z = R + jX$; R circles, X arcs. (Smith chart)

Fig. 2—The reflection chart.

with the aid of a radius arm having a scale of SWR (most conveniently, db SWR).

Other sets of contours on the reflection chart may be used as graph paper or as transparent overlays, if the load impedance is to be reduced to numerical components. Fig. 2(b) and (c) show the forms commonly used for this purpose. Fig. 2(b) shows the Carter chart [23], having contours of impedance magnitude ($|Z/Z_0|$) and phase angle (ϕ). Fig. 2(c) shows the Smith chart [24], having contours of the real and imaginary parts of impedance; these are essential for adding impedances in series. A rotation of this form by 180° gives the parts needed for adding admittances in parallel.

There are variations of the reflection chart, both circular and otherwise, that are adapted for special purposes. A notable one of these is the "projective" chart of Deschamps [18], which retains the same circular rim but has a different radial scale.

The standing-wave technique may be simplified if it is required only to measure the magnitude of reflection coefficient or SWR. This can be measured by a two-way directional coupler and a ratio meter (forming a reflectometer). The ratio can be plotted instantaneously against a sweeping frequency on a scope. This method is most useful if the objective is merely to show that the reflection is within a tolerance.

V. Test Equipment

Table IV is intended to show some of the principal types of test equipment that are essential in microwave measurements. These are largely parallel to those for lower frequencies, but with emphasis on some that are adapted to waveguide techniques as distinguished from wired circuits.

VI. Signal Sources and Indicators

For any part of the frequency spectrum, signal sources and indicators are related in several respects. The source power must exceed the indicator power level of detection, preferably by several orders of magnitude. Furthermore, this relationship must be obtained for some practical kind of modulation, if any is required to increase the sensitivity of detection.

Table V lists some typical sources useful for microwave measurements [17]. The frequency ranges and power levels are estimates and intended merely to indicate the order of magnitude. The listed devices are those likely to be used for laboratory tests, as distinguished from high-power transmitters, but naturally the latter are required for some tests.

Table VI lists some typical indicators useful for microwave measurements. Here also the frequency ranges and power levels are estimates intended merely to indicate the order of magnitude. The listed devices are those likely to be used for laboratory tests, as distin-

TABLE IV

MICROWAVE MEASUREMENTS—EQUIPMENT

- Signal Sources (generators, oscillators, transmitters).
- Signal Indicators (meters, detectors, receivers).
- Spectrum Analyzers.
- Wavemeters, Resonators.
- Standards in Waveguide.
  - Wavelength.
  - Wave impedance.
  - Wave reflection.
  - Wave transmission.
    - Attenuation.
    - Phase shift.
- Comparators in Waveguide.
  - Balance type.
    - Hybrid junction.
  - Standing-wave type.
    - Slotted section.

TABLE V

LABORATORY SIGNAL SOURCES

- Electron tube (space charge).
  - 0.3–3 Gc.
  - 1 watt.
  - Mechanical tuning (wide range).
- Tunnel diode (crystal).
  - 0.3–3 Gc.
  - 10^{-3} watt.
  - Mechanical tuning.
- Reflex klystron (electron velocity).
  - 1–60 Gc.
  - 0.1 watt.
  - Mechanical tuning (10–20 per cent).
  - Voltage tuning (about 1 per cent).
- Backward-wave oscillator (BWO) (electron velocity).
  - 1–100 Gc.
  - 1–10^{-3} watt.
  - Voltage tuning (1:2 range).
- Harmonics
  - 30–300 Gc.
  - 10^{-2}–10^{-6} watt.
  - Oscillator harmonics or crystal multipliers.

TABLE VI

LABORATORY SIGNAL INDICATORS

- Classification:
  - Type of Input Circuit.
    - Minimum detectable power.
    - Device in input circuit (various kinds).
      - Frequency range.

- RF Amplifier.
  - 10^{-18} watt.
  - Electron tube (space charge).
    - 0.3–30 Gc.

- RF Converter (frequency changer).
  - 10^{-15} watt.
  - Electron tube (space charge).
    - 0.3–1 Gc.
  - Crystal diode (rectifier).
    - 0.3–30 Gc.
    - Up to 300 Gc, less sensitivity.
  - Parametric amplifier (crystal diode).
    - 0.3–10 Gc.

- RF Rectifier (detector).
  - 10^{-9} watt, dc output (no modulation).
  - 10^{-12} watt, ac output (modulation).
  - Electron tube (space charge).
    - 0.3–3 Gc.
  - Crystal diode (rectifier).
    - 0.3–30 Gc.
    - Up to 300 Gc, less sensitivity.

- RF Thermal Effect.
  - 10^{-6} watt, dc output (no modulation).
  - 10^{-9} watt, ac output (modulation).
  - Resistance variation.
    - 0.3–100 Gc.
    - Barretter (metal conductor, fine wire).
    - Thermistor (semiconductor, small bead).
  - Gas expansion (Golay cell).
    - 100–300 Gc.
    - Resistance sheet (thin film).

guished from sensitive receivers, but naturally the latter are required for some tests.

VII. SPECTRUM ANALYZER

The principle and utility of the spectrum analyzer are general in their application over the entire radio-frequency spectrum. However, it happens that the development and application of this device have been most closely associated with microwaves, because it was made available to meet some specific needs which arose in this field [10].

The spectrum analyzer is essentially a narrow-band receiver provided with frequency sweep and instantaneous display of amplitude frequency spectrum on a scope. Fig. 3 shows the resulting patterns for some typical cases. In the display for a pulsed signal, there is a comb of vertical lines, as shown in Fig. 3(a), (b) and (d); the spectrum is the envelope of these lines. In the display for a steady signal, or for random noise smoothed by a filter, the spectrum is simply plotted as shown in Fig. 3(c) and (e).

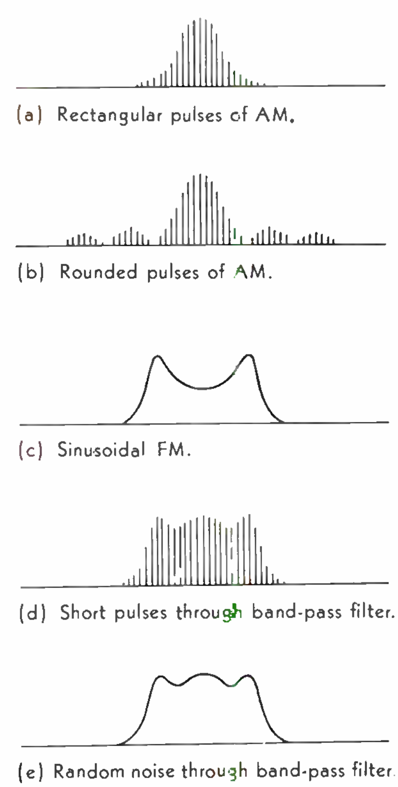


Fig. 3—The frequency-spectrum display for typical cases.

VIII. Frequency Standards and Wavemeters

As mentioned above, standard-frequency generators are outside the scope of this article. However, the use of a hollow-cavity resonator as a wavemeter is peculiar to microwaves and is here included in the form of one example.

Fig. 4 shows a resonator bounded by a right-circular cylinder of conducting walls (preferably silver-plated). One end of the active space is a piston supported without conductive contact to the cylinder, and provided with a micrometer drive and scale. The diameter is somewhat greater than the cut-off wavelength (λ_c) and the length is $\frac{1}{2}$ the guide wavelength (λ_g). The pattern of the magnetic field (H) is indicated. This cavity is coupled with one or two connecting waveguides through small holes, not shown, and resonance is indicated by a transmission or absorption method.

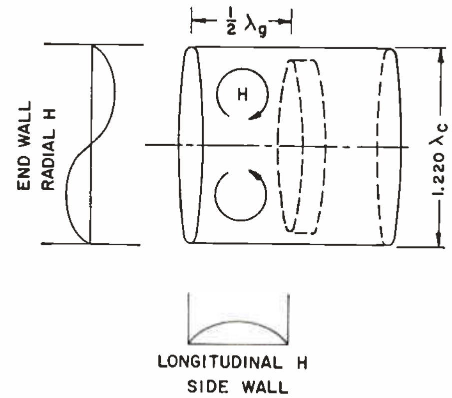


Fig. 4—A wavemeter utilizing the TE-01 mode in in circular waveguide.

IX. Power Measurement and Calibration

Microwave power measurement usually relies on the heating of a load impedance by all the power or by a small fraction of the power. This is distinguished from the low-frequency indication of the product of voltage and current in a wattmeter. Reference is made to some of the indicators listed in Table VI.

The thermal measurement of average power in some devices offers the advantage of direct comparison between RF power to be measured and dc power which can be directly metered. The comparison is nearly independent of the properties of the load resistor. This principle is developed to greatest refinement in the self-balancing "bolometer bridge," where the RF power is read directly on a dc meter. The required temperature-sensitive resistor is made of a fine wire of metal ("barretter") or a small piece of semiconductor material ("thermistor").

X. Impedance Standards and Comparators

The customary standard of impedance is the wave resistance presented in a rectangular waveguide supporting a single-mode propagation. The walls (of high conductivity) are nearly perfect reflectors, so the dimensions determine the impedance as closely as they can be measured. Therefore the waveguide provides a primary standard for common use, which is remarkable in the practice of measurements. (The air has a slight effect which can be taken into account if significant.) Usually this impedance is not expressed numerically, but any measured impedance is expressed relative to this standard (Z/Z_0).

In this situation, self-impedance is identified with reflection coefficient in the manner explained above with reference to the reflection chart. Also the standing-wave technique is a direct method for comparing an unknown impedance with the waveguide as a standard.

Fig. 5 shows some standards of reflection (or impedance) that are adapted for inclusion in a waveguide. The fixed standard Fig. 5(a) is stable and closely computable, so it approaches the status of a primary standard.

An adjustable standard of reflection (or impedance) is more difficult because it involves motion in two dimensions and a corresponding calibration in two parameters (such as R/Z_0 and X/Z_0) for each frequency of operation. Fig. 5(b) shows one form that is suitable for rectangular waveguide, and is useful over a wide range of values. The reference plane is stationary (as at the mid-position of the center of the sliding insert) so the motion enables all phase angles of reflection coefficient. The sliding insert may be made of metal like the waveguide walls, or may be made of dielectric material to avoid the problems of contact at the edges of the slot.

Primary and secondary standards of transmission properties are more highly developed and more diversified. A number of types are listed in Table VII. Attenuation or irreducible loss has an absolute significance, relative to zero in a lossless waveguide. Angle or phase shift has only a relative or differential significance, since it is present in all waveguide connections. In each case, there are some primary standards dependent on dimensions rather than materials, and some secondary standards dependent on material properties such as dissipation or dielectric constant.

Transmission measurements are made by comparing the unknown with a standard by one of the usual methods. If the differential or substitution method is used, the RF circuit to be tested may be followed by a superheterodyne receiver, and its transmission properties measured by an equal change in the IF circuit of the receiver. This procedure is well adapted for a piston attenuator and a phase rotator in the IF circuit.

In the use of either reflection or transmission standards, a balancing method with a null indicator is capable of the highest degree of refinement. In waveguides, the ideal comparator is the symmetrical hybrid-T ("magic-T") junction, shown in Fig. 6 [35]. The indicated external connections to the four arms are those used for comparison of reflections (or impedance), in which case the junction performs the function of the

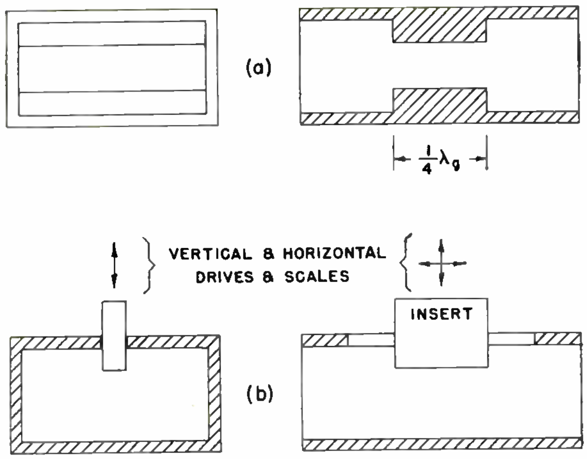


Fig. 5—Fixed and adjustable standards of reflection in waveguide.

TABLE VII

TRANSMISSION STANDARDS

Attenuation (irreducible loss).
Primary standards (computable from dimensions).
Piston attenuator (waveguide beyond cutoff).
Scale linear in decibels.
Rotation in circular waveguide.
Independent of frequency.
Hybrid junction (symmetrical).
Fixed ($\frac{1}{2}$ power, about 3 db).
Secondary standards.
Directional coupler.
Dissipation (insert in waveguide).
Angle, Phase Shift (differential).
Primary standards (computable from dimensions).
Rotation in circular waveguide.
Independent of frequency.
Waveguide section.
Trombone.
Secondary standards.
Dielectric (insert in waveguide).
Distortion of waveguide ("squeeze" section).

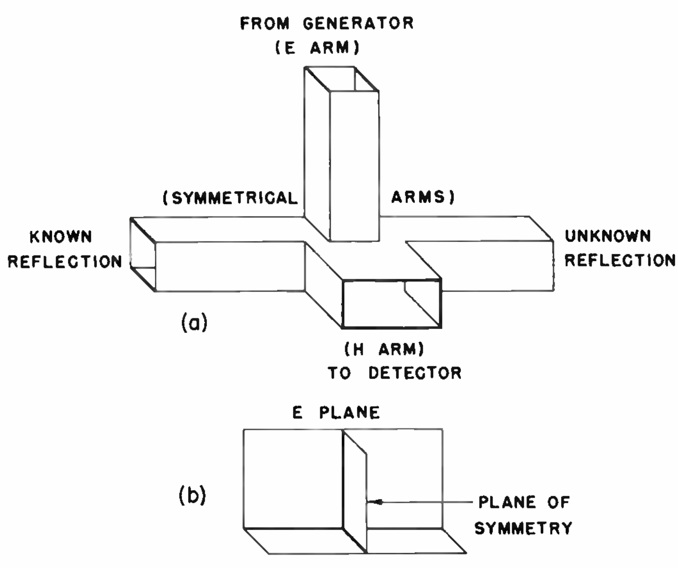


Fig. 6—The hybrid-T junction used as a symmetrical comparator.

pair of equal arms in a Wheatstone bridge. The equality of these arms is assured by mechanical symmetry of structure.

XI. MICROWAVE SPECTROSCOPY

Microwave spectroscopy [14], [17], [19], like its counterparts at wavelengths from infrared through X rays, consists basically of some observation of the frequencies of emission or absorption of energy by the molecules of a particular substance, usually in gaseous form. Each "line" which lies in the microwave portion of the spectrum represents an energy transition between two of the various rotational or vibrational states of a molecule; therefore the materials studied are usually those having a polyatomic molecular structure that is mechanically well defined.

A typical microwave absorption spectrometer comprises a signal source with the required frequency variation, a long waveguide filled with a gas under test, and a signal indicator that is sensitive to small changes in output power. Unusual refinements are required in frequency stability and measurement.

Important applications of microwave spectroscopy are found in the "ammonia clock," and the "maser," the optical maser or "laser," and the use of the hydrogen line in radio astronomy.

XII. DEVELOPMENTS NEEDED FOR FUTURE

In general, there are great opportunities for refinement of the techniques described and of some others closely related. A few of the needs for improvement may be mentioned here.

Comparison standards of reflection and impedance are needed in adjustable form, with convenient calibration for use in a balance method.

Microwave beam deflection in an oscilloscope has been demonstrated, but needs development into a convenient laboratory form.

Multimode techniques in waveguide are steadily progressing to a higher degree of understanding and utility.

For the shortest (millimeter) waves, there will always be an increasing need for signal sources and indicators; also further opportunity for refinement of single-mode waveguide techniques and for alternative methods related to optical practices.

The rating of a multiport device in terms of the wave-transmission matrix is due for increasing use, especially with further experience in the required methods of measurement.

XIII. CONCLUSION

An attempt has been made to present in a compact form some of the more interesting and significant developments in the field of microwave measurements. These reflect remarkable progress over 25 years of activity, stimulated especially by the needs of radar, guided missiles and communication relay networks.

ACKNOWLEDGMENT

It is a pleasure to acknowledge the writer's reliance on 20 years' experience with engineering groups active in microwave measurements, first in Hazeltine Corporation during World War II, and subsequently in Wheeler Laboratories. In particular, he is grateful to his associates, P. A. Loth and I. Koffman, for their helpful cooperation in the preparation of this paper.

REFERENCES

General

[1] IRE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES.
[2] PROCEEDINGS OF THE IRE.
[3] IRE Standards.
[4] *Microwave Journal.*
[5] *Bell System Technical Journal.*
[6] "Radiation Laboratory Series," McGraw-Hill.
[7] Electronic Industries Association, Standards.
[8] Hewlett-Packard Application Notes (esp. No. 21, 38, 39, 40); *H-P Journal* (Many articles on microwave measurements).
[9] National Bureau of Standards, publications.

Books

[10] C. G. Montgomery, "Technique of Microwave Measurements," Rad. Lab. Ser., vol. 11, McGraw-Hill Book Co., Inc., New York, N. Y.; 1947. (Comprehensive treatment of developments at MIT during World War II.) (See other volumes of this series for measurements of particular components.)
[11] G. C. Southworth, "Principles and Applications of Microwave Transmission," D. Van Nostrand Co., Inc., New York, N. Y.; 1950.
[12] H. M. Barlow and A. L. Cullen, "Microwave Measurements," Constable and Co., Ltd., London; 1950.
[13] D. D. King, "Measurements at Centimeter Wavelength," D. Van Nostrand Co., Inc., New York, N. Y.; 1952.
[14] W. Gordy, W. V. Smith, and R. F. Trambarulo, "Microwave Spectroscopy," John Wiley and Sons, Inc., New York, N. Y.; 1953.
[15] M. Wind and H. Rappaport, "Handbook of Microwave Measurements," 2 vols., Polytech. Inst. of Brooklyn, N. Y.; 1954.
[16] "Modern Advances in Microwave Techniques," MRI Symp. Proc., vol. 4, Polytech. Inst. of Brooklyn, N. Y.; 1954. (Emphasis on measurement techniques.)
[17] C. H. Townes and A. L. Schawlow, "Microwave Spectroscopy," McGraw-Hill Book Co., Inc., New York, N. Y.; 1955.
[18] "Reference Data for Radio Engineers," 4th ed., I.T.&T. Corp., New York, N. Y.; 1956. (Ch. 22, G. A. Deschamps, scattering matrices, reflection chart of "projective" type with "hyperbolic protractor.")
[19] D. J. E. Ingram, "Spectroscopy at Radio and Microwave Frequencies," Philosophical Library, New York, N. Y.; 1956.
[20] E. L. Ginzton, "Microwave Measurements," McGraw-Hill Book Co., New York, N. Y.; 1957.
[21] F. J. Tischer, "Microwave Measurements," Springer, Verlag, Berlin, Germany; 1958. (In German.)
[22] "Millimeter Waves," *MRI Symp. Proc.*, vol. 9, Polytech. Inst. of Brooklyn, N. Y.; 1959. (Emphasis on measurement techniques.)

Selected Articles

[23] P. S. Carter, "Charts for transmission-line measurements and computations," *RCA Rev.*, vol. 3, pp. 355–368; January, 1939. (The basic reference on the standing-wave technique for measuring impedance; reflection chart with $|Z|,\phi$ contours.)
[24] P. H. Smith, "Transmission line calculator," *Electronics*, vol. 12, pp. 29–31; January, 1939. (The basic reference on reflection chart with R, X and G, B contours.) Also vol. 17, pp. 130–133, 318–325; January, 1944.
[25] D. F. Bowman, "The notch wattmeter for low-level power measurement of microwave pulses," *Proc. Natl. Electronics Conf.*, vol. 2, pp. 361–371; October, 1946. (For comparison between peak power and constant power level.)
[26] A. L. Samuel, "An oscillographic method of presenting impedances on the reflection-coefficient plane," PROC. IRE, vol. 35, pp. 1279–1283; November, 1947. (Frequency sweeping for instantaneous plotting on reflection chart.)
[27] H. A. Wheeler, "Geometric Relations in Circle Diagrams of Transmission Line Impedance," Wheeler Monographs, No. 4, Wheeler Labs., Great Neck, N. Y., July, 1948. (Background and various forms of reflection chart.)
[28] ——, "The Piston Attenuator in Waveguide Below Cutoff," Wheeler Monographs, No. 8, Wheeler Labs., Great Neck, N. Y.; January, 1949. (Primary standard of attenuation, up to high values.)
[29] ——, "Measuring the Efficiency of a Superheterodyne Converter by the Input Impedance Diagram," Wheeler Monographs, No. 9, Wheeler Labs., Great Neck, N. Y.; March, 1959. (Loss circle on reflection chart, applied to frequency changer.)
[30] ——, "The Transmission Efficiency of Linear Networks and Frequency Changers," Wheeler Monographs, No. 10, Wheeler Labs., Great Neck, N. Y.; May, 1949. (Loss circle on reflection chart.)
[31] A. F. Harvey, "Instruments for use in the microwave band," *Proc. IEE*, pt. 2, vol. 98, pp. 781–792; December, 1951. (Brief, many references.)
[32] D. Dettinger and R. D. Wengenroth, "A standard waveguide spark gap," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-1, pp. 39–48; March, 1953.
[33] N. A. Spencer, "Cavity-resonator design charts," *Electronics*, vol. 27, pp. 186–188; May, 1954. (In terms of normalized dimensions.)
[34] H. L. Bachman, "A waveguide impedance meter for the automatic display of complex reflection coefficient," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-3, pp. 22–30; January, 1955. (With frequency sweeping, the locus on reflection chart is displayed instantly on oscilloscope.)
[35] P. A. Loth, "Recent advances in waveguide hybrid junctions," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-4, pp. 268–271; October, 1956.
[36] D. Dettinger, "The advantages of expressing standing-wave ratio in decibels," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-5, p. 218; July, 1957.
[37] D. S. Lerner and H. A. Wheeler, "Measurement of bandwidth of microwave resonator by phase shift of signal modulation," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-8, pp. 343–345; May, 1960.
[38] S. J. Miller, "The traveling wave resonator and high power microwave testing," *Microwave J.*, vol. 3, pp. 50–58; September, 1960. (Theory, experiment, references.)
[39] "URSI Radio Measurement Methods and Standards," *J. Res. NBS*, vol. 64D, pp. 591–606; November–December, 1960. (Review for U. S., 1957–60; includes microwaves, many references.)

The Future of Microwave Communications*

S. E. MILLER†, FELLOW, IRE

Summary—The future of microwave communications is foreseen by noting the status of advanced research and then deducing its logical impact on future systems. The trend toward higher frequencies is expected to result in millimeter waveguide systems, and perhaps in communication using visible radiation, with an order of magnitude more channel capacity than the largest existing commercial systems. Satellite radio communications are expected to provide versatile broad-band international communications in an era when tying together the world's major population centers is of first importance. Point-to-point radio-relay systems are expected to provide the largest volume of domestic long-distance circuits, with explosive growth in the near future. New solid-state devices have appeared and should lead to systems with lower power consumption, longer component life, and eventually lower operating cost.

WHAT IS THE *future* for *microwave* communications?[1] We shall try to foresee the future of microwave communication technique in the sense that communication research unfolds the future. By contrast we shall not attempt a prediction of the business trends which influence the magnitude of commercial communications systems, important as these trends certainly are. We must also avoid the folly of trying to predict the unknown. The present task only becomes feasible to the extent that we can proceed from existing new components and ideas to the probable influence on future communication systems. We might hope to find a situation analogous to that wherein this year's birth rate largely determines the size of the kindergarten classes five years from now, and the size of the college entrance group 18 years from now.

I am a firm believer that many, if not most, *major* communication system advances come as a consequence of new communications components, which in turn come directly or indirectly from improved knowledge in physics. During the last ten years new results from physics have created revolutionary opportunities for communication engineers. First consider, for example, the *nonreciprocal* components based on the ferrites or garnets, and second, the semiconductor devices giving gain with very little power consumption. Remarkable changes in communication systems resulted. Are such changes at an end? I believe not, and will try to lead the reader to that conclusion.

What is included in *microwave?* There appears to be no concise definition in IRE literature, but an accepted lower limit is around 1000 Mc. For the upper limit we are more completely on our own. Regarding microwave as synonymous with micro-meter, or micron, we arrive at the region near 3×10^{14} cps, or 300 teracycles. This is just below the visible spectrum where, indeed, communication techniques are appearing, and it seems logical to include this region in the present discussion.

Potential Communication Needs

One of the motivations for advancements in the communications art is an unsatisfied need for service. We can point to a number of such needs as a prelude to anticipating the future trend of the microwave communications art.

There is as yet no satisfactory method of sending live television across the ocean. Few doubt that an economically attractive method would experience widespread commercial usage. Systematic expansion of the present types of services—telephone, data transmission, etc.—will also increase the need for overseas facilities as business and social ties become more international.

Similar expansion of the existing communication needs occurs in the domestic area where the population growth and associated expansion of business indicates an expanding demand for service. The existing types of microwave systems, point-to-point radio relay and tropospheric scatter, will continue to provide valuable communication service but there is simply not enough available radio spectrum (unimpaired by atmospheric transmission difficulties) to take care of the foreseeable needs for communication service of the present commercial types. In addition, we may expect the creation of new types of services, which would necessitate additional expansion of communication facilities. For example, almost any form of person-to-person television would make existing commercial communication systems totally inadequate.

Going to Higher Frequencies

Viewing the past 40 years most generally, communication advances have formed a pattern of going to higher frequencies. Why? Principally to fulfill communication needs, for at the higher frequencies more bandwidth is available. A complementary point of view is that by going to higher frequencies more bandwidth is technically feasible, and this feasibility itself serves as a stimulus for additional communications usage. That trend is expected to continue.

In the microwave region, the earliest services were provided at the lower frequencies followed by exploitation of common carrier bands in the 4000, 6000 and 11,000 Mc regions. Above about 11,000 Mc, rain attenuation and other atmospheric propagation effects

* Received by the IRE, July 5, 1961; revised manuscript received, November 24, 1961.
† Guided Wave Research, Bell Telephone Laboratories, Red Bank, N. J.
[1] The emphasis is on *future* and *microwave*, as suggested by the Anniversary Issue Editors.

begin to become serious. However, in shielded media, research indicates the likelihood of commercial utilization of the 35,000 to 100,000 Mc band, with higher bands as a research possibility. Recently, an initial milepost has been reached in the 300 teracycle region—the creation of coherent sources of power. Thus both in the 100,000 Mc region and in the teracycle region, the frequency frontiers of communication technology are being pushed back.

In going to higher frequencies, we must mate a suitable transmission medium with suitable techniques of signal generation, amplification, modulation and the usual signal handling components such as hybrids, filters, multiplex equipment, etc. Invention, innovation and pioneering development will be required. Next we consider how a few of the new results seem likely to project importantly on communication in the microwave region.

Optical or Teracycle Region Systems

Two coherent sources of power in a teracycle region have been created; using pulsed operation due to power limitations, a ruby maser-type oscillator has been operated in the region near 433 teracycles. Also, a CW helium-neon gaseous maser with power output of a few milliwatts has been operated near 260 teracycles. In the initial experiments, the stability of output frequency vs time or stability of amplitude vs time leave much to be desired from a communications standpoint. Nonetheless, the accomplishments are exciting in their implications. A 10 per cent band, which is normally achievable with relative ease in microwave communication components, would make available a bandwidth of 30,000,000 Mc for communication!

What would the medium be for such frequencies? One is tempted to think in terms of conventional radio-type transmission because some of the numbers appear so attractive. For example, a five-inch antenna (telescope) operating at 300 teracycles would have a half-power beamwidth of 1.6 seconds of arc, and at a distance of 25 miles (present-day radio-relay-station spacing) the diameter of the beam would only be of the order of one foot. Two such five-inch antennas having a spacing of 25 miles are not in the far field zone and would have an insertion loss in the order of 10 to 15 db in vacuum. However, when both ends of the circuit are on the ground, atmospheric propagation effects (notably refraction and rain attenuation) are almost certain to be very serious. One is inclined, therefore, to think in terms of some kind of a pipe as a shelter for the medium with suitable mirrors or lenses to guide the beam between repeater points. For space communication, more conventional radio methods may be attractive.

What modulation and multiplexing systems will be feasible? What kinds of detectors and modulating devices will such a system use? How is signal-to-noise to be defined and calculated in this region where the quanta is large and $h\nu > kT$? It is apparent that we have here a fertile field for research, the outcome of which may be revolutionary.

TE_{01} Waveguide Systems

Conventional types of microwave services will continue to expand and this will inevitably lead to serious problems of mutual interference and frequency occupancy. The severity of these problems is greater in more densely populated areas, and already congestion of the common carrier bands is threatening cities like New York and Los Angeles. On going to higher frequencies to seek relief, one encounters serious atmospheric absorption and diffraction effects. One is forced to seek transmission media other than the atmosphere at frequencies above 15,000–30,000 Mc. Spurred on by this need, research is well advanced on the communication use of the TE_{01} wave in round guides, and future commercial application appears virtually certain.

A two-inch ID perfect copper tube has a theoretical loss of 3 db per mile at 35,000 Mc and $\frac{3}{4}$ db per mile at 90,000 Mc. Recent experimental results at the Bell Telephone Laboratories on 400-foot to 2000-foot lengths have shown practical loss values around 3.5 db per mile at 35,000 Mc and 1.7 db per mile at 90,000 Mc. Because hundreds of modes can propagate and multipath transmission is experienced, PCM is likely to be used as the signaling method. A single pipe, operating in the 35,000–90,000 Mc region, is expected to be able to handle 100,000 two-way voice channels, 100 standard black-and-white or color television channels in two directions, or the equivalent bandwidth of other grades of service.

One key to the achievement of attractive transmission losses has been the development of multimode waveguides with controlled characteristics for the undesired as well as desired modes. In the 2-inch ID guide referred to above, hundreds of modes can propagate. Fortunately, only a few are important when the structure is carefully controlled, and these few modes can be given approximately-predetermined loss and phase constants. Even in components such as hybrids and channel filters, it has been advantageous to use multimode structures to achieve broader bands or lower losses. These techniques have become familiar in research, and may be expected to find engineering application in fields outside millimeter wave communication as well as for that purpose.

The most fundamental limitation in the transmission medium appears to be straightness deviations having a wavelength equal to electrical beat wavelength, $2\pi/\beta_1-\beta_2$ where β_1 is the phase constant for the circular electric wave and β_2 is the phase constant for any of several low-order parasitic modes coupled to the circular electric wave by the straightness deviation. For the 2-inch ID guide at 35–90 kMc, this critical straightness wavelength is on the order of one to five feet, and it is necessary to maintain manufacturing deviations and installation deviations of that particular period at a small value. However, it is noteworthy that smooth changes of direction on the order of 1° per foot can easily be accommodated if not performed too frequently, and one-fifth that value on a continuous basis with tolerable penalty. Abrupt changes of direction can be made by tapering to a smaller diameter, or by using mirrors. Thus, the installation problem for long-distance TE_{01} waveguides at the present state of the art is uncritical with regard to following hills and dales and changes in direction horizontally. Two miles of helix waveguide, a preferred form of mode discriminating waveguide, have been fabricated in the research laboratory, and in the immediate future a study will be made of the correspondence between a theoretical model of the multimode medium and the experimentally observed transmission properties. Looking farther ahead, the future prospect for the present medium lies in moving the operating frequency upwards in the spectrum where even more bandwidth and less delay distortion are theoretically available; the prospect of devising simplified forms of mode discriminating waveguides is also promising.

Really long haul (4000-mile) application of the waveguide medium calls for acceptance of a relatively imperfect loss frequency characteristic due to higher-order mode interactions. The use of PCM with regenerative repeaters effectively accommodates these medium imperfections, simultaneously relieves repeater-circuit tolerances (another necessity at the carrier frequencies employed), and is economical in the use of signaling power. Extensive electronics research and development work over a 15-year period has succeeded in producing the necessary signal sources, amplifiers, regenerators, and communication circuitry for this frequency region. Experimental models have been demonstrated on a component basis in the center of the 35–90 kMc region, and a research model of a complete PCM repeater is in the final stages of preparation for over-all performance evaluation.

An important communication problem, not characteristic of waveguide and microwave alone but perhaps appearing here for the first time, is the technique of operating a nationwide network of digital signals, maintaining the various channels which are to be interleaved with sufficient accuracy of pulse-to-pulse timing so that the interleaving may be done on a time division basis. This represents a challenge whose solution lies somewhere in the future. With the marked increase in digital communication it is believed to be important.

Viewed broadly, waveguide systems can now be engineered with specific commercial objectives, and the initial steps for such development are being undertaken in several countries. The engineering challenge is to create complete designs and establish economically optimum manufacturing techniques. The tempo of such work will be tied in closely with over-all service demands. The cost per channel in a fully loaded waveguide is expected to be appreciably lower than that of existing communication systems, but the magnitude of initial investment is such that a minimum service requirement of the order of 10,000–20,000 equivalent voice circuits is estimated to be necessary to make the system prove in. Thus the waveguide system will be a superhighway of communications, which incidently will increase the demand for lower-capacity feeder routes via microwave radio relay and other techniques.

From the research point of view waveguide communication is a dynamic field with more sophisticated media, more sophisticated solid-state components and more sophisticated modulation systems lying ahead of us.

One immediate challenge in the research area is signal generation and amplification in the region above 100,000 Mc. Solid-state techniques are beginning to appear and seem almost certain to attract additional pioneering work. Tunnel-diode oscillators have been made at the microwatt level at 100,000 Mc and tunnel-diode amplifiers with 15–18 db noise figure have been made experimentally in the region near 55,000 Mc. A variable capacitance parametric amplifier with double sideband noise figure around 6 db was made at 17,000 Mc, and a similar unit operated without precision measurements at 30,000 Mc. The future should bring advances in these arts. Greater power capacity is needed in these solid-state devices which, if achieved, would make feasible all-solid-state repeaters.

Just as research work to date has established a good way to make a waveguide system, future research is expected to lead to more advantageous solutions to critical problems.

Satellite Radio Communications

A new "transmission medium" recently became available—earth satellites. Radio communication from one ground point to another has been carried out by passive reflection from a 100-foot diameter balloon and preliminary experiments are being conducted to employ a belt of metallic dipoles as a semi-continuous reflecting layer for communication signals. Alternatively, work is progressing on an active system using a solar power operated repeater on the satellite itself. By operating in the 1- to 10-kMc regions, background noise from the sky is very low and the very attractive noise figures of masers can be used to advantage. System input-noise temperatures below 30°K have been achieved.

These early experiments certainly predict a substantial expansion of overseas communication via satellites. New demands appear on many of the old components as a result of this system development. Antennas with the highest practical gain and lowest possible sidelobes are needed, with consequent challenges in mechanical design of the very large but necessarily precise structures which result. The very highest power CW microwave sources are needed in passive systems in order to provide adequate SNR's. There is a great premium on very low-power consumption in the repeaters to be used in the satellites and on obtaining long life due to the very high cost of replacing repeaters. These are old requirements in the submarine cable business but represent new standards in the evolution of microwave techniques. The first proposed satellite repeater systems call for vacuum tubes (broad-band low-noise traveling-wave tubes) in the repeater, but in the future there certainly will be a strong attempt to do the job with low-power solid-state devices.

The microwave and other portions of the satellite repeater will experience new environmental requirements due to radiation in the Van Allen belts. Once again this has brought physics more closely into the engineering activities.

The proposed satellite systems have created a use for an old invention—frequency modulation with feedback, proposed prior to 1940. FM with feedback did not find commercial use in point-to-point radio-relay systems but its ability to improve the noise breaking point for an FM circuit is at the present state of the art a very valuable contribution to system feasibility. Future work is needed to improve our understanding of FM with feedback, and to try to achieve even better power efficiency.

We anticipate a need somewhat later for satellite-to-ground communications of a broad-band nature in which orientation of the satellite antennas will be used. The small amount of atmosphere between the satellite and ground station may permit the use of the millimeter wave or teracycle regions with the associated reduction in power loss for limited satellite antenna size, provided it is acceptable to be out of service during occasional bad weather. Satellite-to-satellite communication on paths wherein there is no atmosphere at all may be needed, and there the principal problems will be to provide equipment with very low-power drain and a high order of reliability using antennas of limited size.

These notes on space communication barely hint at the fascination which the subject holds for those working on it. The new technical challenges are combined with the more subtle glamour of space activities generally which captures the imagination of technical and nontechnical men alike. Even the most pedestrian of communication techniques applied between two manned space vehicles will be exciting. But, in addition, satellite communication systems should provide versatility in international broad-band communications in an era when tying together the world's major population centers is of first importance. We can anticipate unparalleled consumer interest in this field as well as genuine technical challenges and resulting contributions.

Conventional Microwave Communication Systems

Despite the glamour of satellite communication realities and of optical communication potentialities, the work horse of domestic long-distance communications for the next ten years is going to be conventional point-to-point microwave radio relay. Some people have referred to this situation as the forthcoming "microwave explosion." But the systems used will not necessarily be present-day systems. Viewing the spectrum broadly, increased congestion will necessitate improved frequency stability, improved antenna directivity, and efficient multiplexing methods. The solid-state art is only now beginning to be felt in this region. Parametric amplifiers and tunnel-diode devices are in various stages of research or exploratory development but have not yet had a real impact on complete systems. There are great prospective advantages for these newer techniques. Perhaps most outstanding is the hope that lower power consumption and longer life for the solid-state elements will lead to eventual lower operating costs. In addition, very low-noise figures have been achieved, less than 1 db (double sideband[2]) in the 6000 Mc region with refrigeration of the diode and 3.5 db with 1000 Mc bandwidth at 11,000 Mc without refrigeration. Techniques for thermoelectric refrigeration of the input circuits are in exploratory phases and might be expected to appear in finished systems in the years ahead.

A Note on a Possible Future Design Philosophy

In order to meet current communication needs, a system designer must employ components which exist—components whose characteristic behavior is sufficiently well understood to permit system interactions between components to be engineered with assurance. This is important. The man who insists on using existing components is not against progress—he merely recognizes that he cannot provide communication service with *ideas* alone. But with the component art moving ahead at such a rapid rate, new and improved systems are becoming feasible at shorter time intervals. To me as an individual the answer would appear to be the design of systems with evolution in mind—evolution to meet communication needs with the latest advances in component art. This is not as simple to do as one might presume at first thought. The more completely one optimizes all the system design interactions, the less the flexibility which remains in the final layout. A system aimed at planned evolution will show long range performance advantages to offset shorter range off-optimum design choices. But it must be done in a hard-headed way to avoid paying seriously for "future gains" which never materialize.

Mobile Communications

It seems highly probable that the future will see an explosive growth of mobile communications. Personal paging, land vehicular, air-to-ground, satellite telemetering and control, air-traffic control—these and other mobile services are in growing demand. It is possible but not certain that microwaves will be used to satisfy them. One requirement is for low-power-consumption miniaturized receivers which can now be realized more easily at frequencies below 1000 Mc, but the pressures of increased demands for frequency space may force some of these services above 1000 Mc.

[2] With signal received on both sidebands, as in radioastronomy.

Microwave Local Transmission

In the millimeter-wave region, antenna beams of 1° or less can be obtained with an aperture of one foot. If solid-state microwave sources should become feasible *and inexpensive* at power levels of a few milliwatts, radio transmission in this frequency region might become useful as a short-distance local plant facility. Atmospheric losses, which are too great to permit long distance transmission, can be tolerated for distances of a mile or so. Broad radio bands not otherwise employed might find use, and the local communication plant might get a wide-band service capability not easily obtained by other methods. The essential low-cost property may follow feasibility without great delay due to the large volume requirements associated with local plant equipment. This is speculative, but possible.

Other Services

Special services such as military communications and tropospheric scatter microwave radio will profit from the same microwave advances noted above and might be expected to evolve along similar lines.

Conclusion

The future for microwave communications is indeed bright. We can see in today's new developments the seeds of much future progress in the technique of communication, and there is apparent a growing service demand which will require these advances.

Spanning the Microwave Infrared Gap*

PAUL D. COLEMAN†, MEMBER, IRE

Summary—**The basic problems encountered in both classical and quantum electronics when attempting to produce coherent radiation in the microwave-infrared gap region extending from roughly 100- to 1000-microns wavelength are reviewed. Brief historical accounts are given of past work leading up to the submillimeter wave frontier. Some selected examples of current work are described to indicate the direction research in this area is taking. While it is the conclusion that a real breakthrough in the microwave infrared problem is not imminent, it is felt that the new background knowledge and the increased applied physics experience now acquired by the ultramicrowave engineer will hasten the day when a solution to the problem will be obtained.**

Introduction

Basically the radio, electronics, and communication engineer is interested in problems of generation, transmission, detection, and modulation of coherent electromagnetic radiation. Also the physicist has long employed electromagnetic radiation as a diagnostic tool for studying matter.

Useful amounts of RF coherent power became available with the invention of the triode by de Forest[1] in 1907, although the first feed-back oscillator of Meissner's[2] did not appear until 1913. By 1936, a microwave electronic tube, the magnetron,[3] had been extended to a wavelength of 6 mm. However, in the last 25 years microwave electronics devices have been pushed only about a factor of three further, so that today one frontier of the microwave-infrared gap stands at the order of 2 mm wavelength.[4]

Recently, the achievement of an optical maser,[5,6] the laser, has produced coherent power in the visible range of the spectrum. Thus the microwave-infrared gap now stands bracketed from below by both classical and quantum electronics devices and from above by quantum electronics devices.

Care must be exercised, however, in jumping to the conclusion that since the gap region has been bracketed, it is about to be spanned.[7–10] While all ultramicrowave engineers firmly believe the problem will be solved, there still remains much hard work and origination of ideas before the bridge is completed.

Classical electronics devices have utilized lumped circuits, transmission lines, and waveguides as the fre-

* Received by the IRE, June 19, 1961; revised manuscript received, November 20, 1961.

† Ultramicrowave Group, University of Illinois, Urbana, Ill.

[1] L. de Forest, U. S. Patent No. 879,532; January, 1907.

[2] A. Meissner, German Patent No. 291,604; April, 1913.

[3] E. E. Cleeton, and N. H. Williams, "The shortest continuous radio waves," *Phys. Rev.*, vol. 50, p. 1091; 1936.

[4] A. Karp, "Backward-wave oscillator experiments at 100 to 200 kMc," Proc. of IRE, vol. 45, pp. 496–503; April, 1957.

[5] T. H. Mainman, "Stimulated optical radiation in ruby," *Nature*, vol. 187, pp. 493–494, August, 1960; *Brit. Commun. and Electronics*, vol. 7, pp. 674–675, September, 1960.

[6] A. Javan, W. R. Bennett, and D. R. Herriott, *Phys. Rev. Letts.*, vol. 6, pp. 106–110; February, 1961.

[7] P. D. Coleman and R. C. Becker, "Present state of the millimeter wave generation and technique art—1958," IRE Trans. on Microwave Theory and Techniques, vol. MTT-7, pp. 42–61; January, 1959.

[8] I. Kaufman, "The band between microwave and infrared regions," vol. 47, pp. 381–396; March, 1959.

[9] I. P. French, *et al.*, "The Radio Spectrum Above 10 kMc," RCA Victor Co., Ltd., Montreal, Canada; Rept. No. 7-400; July, 1959.

[10] B. Epsztein, "Millimeter Waves," PIB-MRI, Brooklyn, N. Y., Res. Rept. No. 840-60; July, 1960.

quency limit has moved upward to its present frontier of around 150 kMc. The difficulty of extending microwave electronics devices further strongly suggests that the new marriage of electronics should be with physical optics structures. This does not mean that microwave electronic tubes cannot be pushed still further, anymore than that triode tubes cannot be pushed to *S*-band, but it does mean that the difficulties may increase to the point that it is desirable to look for other techniques.

Quantum electronics devices initially were built at *K*-band.[11] Here microwave structures were employed to contain the maser material. However, the laser, in leaping over the microwave-infrared gap up to the visible part of the spectrum, utilized a physical optics structure, the Fabry-Perot resonator.[12] Given the appropriate maser material, there is no reason to believe that a maser system using a Fabry-Perot resonator cannot span the entire microwave-infrared gap. This does not mean that all other schemes should be abandoned anymore than microwave tubes should have been discarded when the microwave maser was invented.

The question of whether a classical electronics device can successfully invade the submillimeter range has yet to be answered. If the answer is negative, then the submillimeter area represents a demarcation line between classical and quantum physics.

In the last 50 years, the radio engineer has successfully met the challenge of producing higher and higher frequencies. The transition from lumped circuits to waveguides required learning field theory. Combining semiconductor, ferrite, and plasma devices with microwave circuits required learning solid-state and plasma physics. Microwave masers required a knowledge of quantum mechanics. Hence, today the ultramicrowave radio engineer is essentially an applied physicist. It is this training, this new experience, this broader outlook, that will enable him to successfully conquer this microwave-infrared-gap ogre and exploit it for the benefit of man.

Basic Interactions

Electromagnetic energy can only be produced in essentially two ways:

1) The classical method of retarding electrons by electromagnetic fields, and thereby converting their kinetic energy into electromagnetic energy.
2) The quantum method of permitting a system to interact with an electromagnetic field to undergo a transition from a higher energy state to a lower energy state, thereby converting its internal energy into electromagnetic energy.

The classical electronics problem involves the solution of Maxwell's and Newton's equations simultaneously in the case of a self-excited system or separately in the case of a driven system. A breakthrough in the submillimeter wave problem from a classical electronics point of view would mean finding the appropriate medium and boundary conditions that would yield practical design values and useful power outputs.

In the quantum electronics problem, a knowledge of the induced and spontaneous emission probabilities of the energy level scheme is required. In theory, these quantities could be calculated from Schrodinger's equation, but in practice they are usually experimentally measured. Thus, success in this area is closely associated with problems and knowledge of materials.

In attempting to invade the submillimeter region, both classical and quantum electronics devices encounter the same fundamental problems:

1) a coherence problem,
2) a field containment problem,
3) an energy conversion problem,
4) a characteristic frequency problem.

Coherence in classical schemes requires electron bunching or collective interaction as opposed to individual behavior. Coherence in quantum schemes requires locating the molecules in a coherent electromagnetic field to induce transitions that are phase-related to the inducing field.

In both the classical and quantum systems a coherent electromagnetic field is needed to exist somewhere in space. Classically, the electrons must work against an external field or in the case of accelerated motion against its own field. Both microwave masers and tubes, of course, use microwave waveguides or cavities for field containment. The frequency limit of such devices is, therefore, directly dependent on the frequency limits of microwave structures.

All oscillators, by definition, are energy converters, their aim being to transform readily available energy into the desired coherent energy of a given frequency. Availability, efficiency, coupling the energy in and out of the system, saturation, etc., are practical problems that must be considered.

Finally, an oscillator must have some type of characteristic frequency associated with it to determine its operating point. This frequency can be a resonant frequency, a frequency for synchronization, a cyclotron frequency, a frequency determined by an energy level, a harmonic frequency, a plasma frequency, a Doppler shifted frequency, etc.

At the moment, classical electronics devices are mainly limited by the field containment problem. However, several selected examples will be discussed later to suggest how this problem may be solved. Quantum electronics devices for the microwave-infrared gap are presently limited by the characteristic frequency problem in the form of a search for suitable maser materials.

[11] J. R. Gordon, H. J. Zeriger, and C. H. Townes, "Molecular microwave oscillator and new hyperfine structure in the microwave spectrum of NH_3," *Phys. Rev.*, vol. 95, pp. 282–284; July, 1954; vol. 99, pp. 1264–1274; August, 1955.

[12] W. Culshaw, "Resonators for millimeter and submillimeter wavelengths," IRE Trans. on Microwave Theory and Techniques, vol. MTT-9, pp. 135–144; March, 1961.

Classical Electronics

The last twenty years have seen the rapid development of microwave electronics sources such as the magnetron, klystron, and traveling-wave tube. In all these devices, the microwave circuit, the field containment structure, is an integral part of the tube. Progress toward higher frequencies with these tubes has principally involved refinement and development of better microwave circuits and the forming and focusing of dense, sharply defined electron streams. While a better microwave circuit relieves some of the requirements on the electron beam and better electron beams[13] relieve some of the requirements on the circuit, it now appears that the practical limits of devices of this type are around 1 to 2 mm wavelength.

The state of the art of conventional backward-wave oscillators is perhaps best represented by the *O*-type Carcinotrons of Compagnie Generale de Telegraphie Sans Fil (CSF). Proposals[14] aimed at 600 kMc with power outputs approaching a milliwatt are being considered. Remarkable as this achievement would be, it represents extending microwave-type techniques beyond their normal practical limits because of a lack of appropriate new ideas for the gap region.

If microwave field containment structures represent a major obstacle in pushing on to higher frequencies, how then might electron beams be used in physical optics type of devices which are appropriate to the submillimeter part of the spectrum?

There are two basic ways of achieving a structure having a phase velocity v_p less than that of light c. The first is to use a periodic metallic guide; the second is to use a material medium having an index of refraction greater than one. This second approach naturally leads to the study of the Cerenkov[15] effect, *i.e.*, the radiation of electromagnetic energy by a charged particle q moving through a material medium with a uniform velocity v greater than the phase velocity v_p of light in the medium.

Since the ionization losses of a charge passing through a medium would be excessive, the practical realization of the Cerenkov effect requires that the charge pass near or through a hole in the medium. Also, to achieve coherent Cerenkov radiation having a line spectrum, a bunched electron beam must be used as in any microwave tube. Fig. 1 illustrates a typical Cerenkov configuration wherein a bunched, cylindrical beam having a velocity v, is passed through a cylindrical hole in an infinite dielectric medium having a dielectric constant ϵ.

Coherence requires that the bunched beam move the distance AC in the time the Cerenkov radiation moves the distance AB and that the time phase difference between points A and B is a multiple of 2π.

Thus

$$t = \frac{AC}{v} = \frac{AB}{v_p} \quad \text{or} \quad \cos\vartheta = \frac{AB}{AC} = \frac{v_p}{v}, \tag{1}$$

and

$$\omega t = \omega\left(\frac{2\pi}{\omega_0}\right) = 2\pi n \quad \text{or} \quad \omega = n\omega_0, \tag{2}$$

where ω is the frequency of the Cerenkov radiation, ω_0 is the electron beam fundamental bunching frequency and n a positive integer.

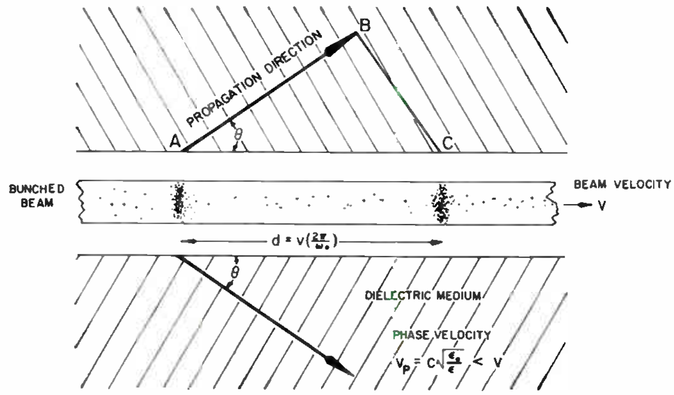


Fig. 1—Cerenkov radiation produced by a bunched electron beam in infinite, nondispersive, dielectric medium.

The Cerenkov radiation propagates away from the electron beam in a cone of semi apex angle θ and contains frequencies corresponding to each current harmonic frequency of the bunched beam.

Thus in this Cerenkov system, the coherence problem is solved by the beam bunching which at the same time determines characteristic frequencies $n\omega_0$. Since no resonance effects are required in the dielectric medium, the physical size of the dielectric can be arbitrarily large with no mode problems in the microwave sense. The field containment problem involved reduces to a proximity problem in that the beam must pass near the dielectric or through a hole in it.

The z component of the electric field for the case being considered varies as $I_0[\omega r/v(1-\beta^2)^{-\frac{1}{2}}]$ where I_0 is the modified Bessel function of the first kind. As the beam velocity $v = \beta c$ approaches c, the velocity of light, it is seen that the argument of I_0 becomes smaller because of the $(1-\beta^2)^{\frac{1}{2}}$ factor, resulting in the electric field decreasing less rapidly away from the dielectric surface. Thus the requirements on the proximity problem can be relaxed at relativistic velocities.

Finally, how is the energy coupled in and out of the structure? This can be easily done by shaping the dielectric from geometric optics ray tracing considera-

[13] Z. S. Chernov, "Methods of focusing electron beams in modern microwave devices," *Radio Engrg. and Electronics*, (*USSR*), vol. 3, pp. 1227–1242; October, 1958.

[14] "Millimeter waves, the new frontier," *Raytheon-CSF Bull.*, Waltham, Mass.; October, 1960.

[15] J. V. Jelley, "Cerenkov Radiation," Pergamon Press, New York, N. Y.; 1958.

tions as shown in Fig. 2.[16] The angles α and ϕ of the dielectric cone are chosen so that the coherent Cerenkov radiation emerging at the characteristic angle ϑ is internally reflected by surface 1, strikes surface 2 at the Brewster angle and then escapes parallel to the cone axis. The dielectric cone acts both as a beam-coupling structure and an aperture antenna.

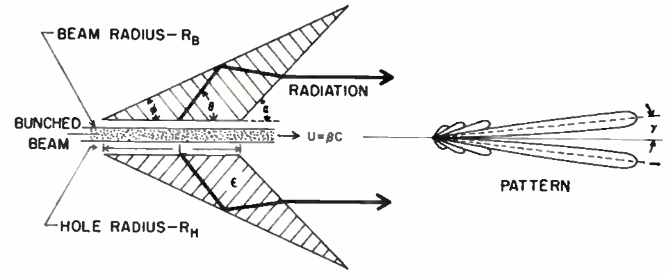
The Cerenkov cone represents a Maxwell problem in that the bunched beam (charge ρ and current $\bar{J}$) is given with the problem to compute the fields and the interaction. It is assumed that in the case of a megavolt bunched beam, the fields generated by the beam are not of sufficient intensity to affect its motion (the constant current generator approximation).

Producing a high-energy bunched beam[17] represents a Newton problem. Here the appropriate field of frequency f_0 must be determined to exert forces on the electrons to accelerate and bunch them in the desired fashion. If this bunching can be done with a field at a fundamental frequency of 35 kMc and if appreciable harmonic currents up through the 50th can be obtained, then coherent power up to and beyond 1750 kMc could be produced.

Work on this type of Cerenkov coupler has already produced 8-mm power at the watt level. Investigation of an optical type of feedback is underway to further increase the power output.[18] Also, anisotropic media, wherein there are strong, high Q absorptions, should be expected to couple strongly to a beam at frequencies in the vicinity of the resonance. In this area, maser materials and classical Cerenkov couplers may have overlapping interests.

Since the Fabry-Perot resonator has been used with great success in the laser, consider how this genuine physical optics structure can be used with an electron beam. A physical optics system, being in general very large compared to the operating wavelength, usually involves essentially plane waves wherein the electric and magnetic fields are transverse to the direction of propagation. For an electron beam to interact with this TEM wave, it must interact obliquely so as to obtain a component of the electric field in the direction of the beam as illustrated in Fig. 3.

The synchronism condition for the beam and the wave to maintain constant phase relationship is readily seen to be $v \cos = v_p$ which means $v_p < v < c$ for a practical case. Hence the plane wave has to exist in the appropriate dielectric medium. Assuming the electric field $\bar{E}$ and the current density $\bar{J}$ shown, the average power-per-unit length P/D delivered to the wave is $-\frac{1}{2}E_0I_0$ tan, where I_0 is the current amplitude.



CERENKOV CONE COUPLER

Fig. 2—A megavolt electronics—Cerenkov radiator employing geometric optics techniques.

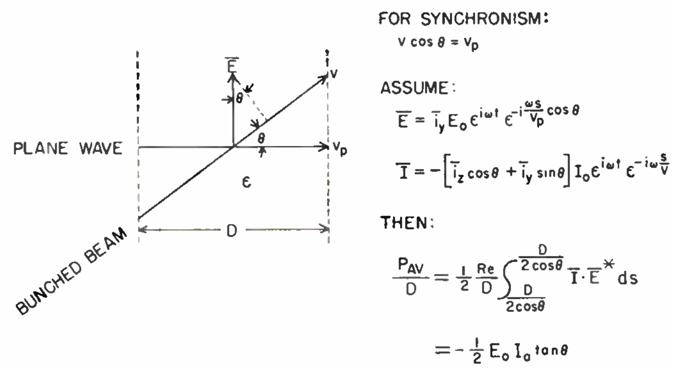


Fig. 3—Interaction of a bunched electron beam with a pulse wave at oblique incidence.

If these ideas[19] are applied to a Fabry-Perot resonator, an arrangement such as that indicated in Fig. 4 is obtained. Here the resonator consists of a large teflon cylinder of cross sectional area A and length D between two metal plates. One of the metal plates has a circular metal grating to permit the polarized radiation to be focused by a lens. Using a 14 ma, 0.88 Mev, S-band Rebatron beam, a power output of 150 mw[20] was obtained at 36.9 kMc for the condition $A=46$ cm^2 and $D=8.5$ cm in excellent agreement with theory.

A final example of the most elementary classical electronics scheme which neatly solves all the four basic problems but suffers somewhat from efficiency is the coherent radiation annihilation method[21] shown in Fig. 5. Here, a bunched electron beam of velocity $v=\beta c$ strikes a metal target where it is abruptly brought to rest and disappears. If ω_0 is the fundamental bunching frequency of the beam, I_n the current amplitude of the nth harmonic, the angle between the beam direction and a point of observation, and η the free-space impedance, then the power radiated into the hemisphere (2π steradians) is given by the expression[21]

[16] P. D. Coleman, and C. Enderby, "Megavolt electronics Cerenkov coupler for the production of millimeter and submillimeter waves," *J. Appl. Phys.*, vol. 31, pp. 1695–1696; September, 1960.
[17] P. D. Coleman, "Theory of the rebatron," *J. Appl. Phys.*, vol. 28, pp. 927–936; September, 1957.
[18] C. Enderby, "A Cerenkov Radiator for the Production of Millimeter and Submillimeter Waves," Ultramicrowave Group, University of Illinois, Urbana, Tech. Note No. 1, WADD Contract AF33(616)-7043; December, 1960.
[19] P. D. Coleman, "Electron devices for the millimeter-infrared Gap," 1961 IRE INTERNATL. CONVENTION RECORD, pt. 3, pp. 54–63.
[20] M. D. Sirkis, R. J. Strain, and W. E. Kunz, "Electron beam excitation of a Fabry-Perot interferometer, *J. Appl. Phys.*, vol. 32, p. 2055; October, 1961.
[21] B. Hakki, and H. Krumme, "Coherent generation of microwave power by annihilation radiation of a prebunched beam," PROC. IRE (Correspondence), vol. 49, p. 1334; August, 1961.

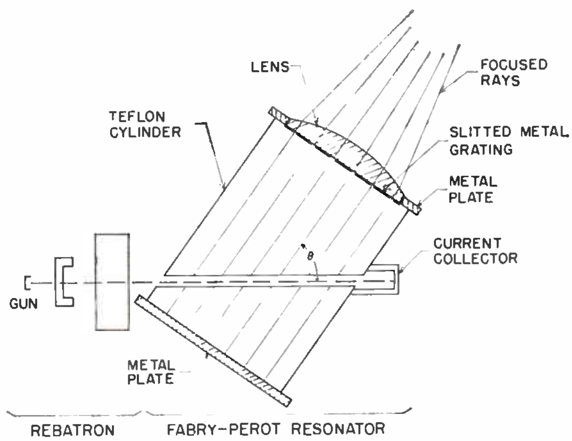


Fig. 4—Excitation of a Fabry-Perot resonator with an electron beam. An example of optical electronics.

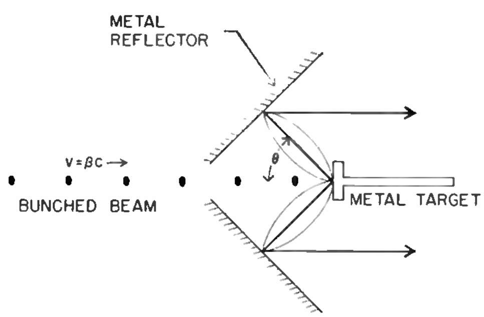


Fig. 5—Production of coherent annihilation radiation by a bunched megavolt beam.

$$P(n\omega_0) = \int P(n\omega_0, \bar{r}_0)d\Omega = \int \frac{\eta I_n{}^2\beta^2 \sin^2\vartheta}{8\pi^2(1-\beta^2\cos^2\vartheta)^2}\,d\Omega,$$

which gives

$$P(n\omega_0) = \tfrac{1}{2}I_n{}^2\left\{\frac{\eta}{4\pi}\left[\frac{1+\beta^2}{2\beta}\ln\left(\frac{1+\beta}{1-\beta}\right)-1\right]\right\},$$

where $d\Omega$ is the element of solid angle. It will be observed that the interaction resistance (term in curly brackets) is independent of the frequency. For a $\beta=0.97$ (2-Mev beam), this resistance has the value 96 ohms. If the metal target is small, power will also appear in the second hemisphere so that the resistance can be approximately doubled if so desired. In this case a current amplitude I_n of 0.1 amperes would yield 1 watt of power.

The point of the preceding selected examples is to give support to the idea that classical electronics has not reached a barrier around one mm wavelength if microwave field containment structures are abandoned.[22] What is needed is a change in viewpoint from extending techniques appropriate to one spectrum region into another spectrum region where they are not suited. It will, of course, be difficult for engineers, who live with deadlines and target dates, to venture into unknown areas filled with always expected hazards. However, it is hoped that a few hardy pioneers will make it.

[22] G. E. Weibel, "PIB MRI," *Symp. Series*, Polytechnic Press, Brooklyn, N. Y., vol. 8, April, 1958.

Quantum Electronics

Quantum electronics might be said to have begun in 1917, not long after the founding of the IRE, with a paper by Einstein[23] on stimulated and spontaneous emission of radiation by molecules. However it was not until 1950 that knowledge and background experimental techniques, particularly in the field of microwave spectroscopy, had reached the stage where the time was ripe for various people to begin to formulate ideas[24] and proposals in the direction of molecular amplification. It required another four years before the first ammonia maser was operated by Townes[11] and associates. In 1956, the three-level, CW, solid-state maser was proposed by Bloembergen[25] and built by Scovil[26] and associates in 1957. Finally in 1960, Maiman[5] successfully demonstrated an optically-pumped, three-level, ruby maser (the laser) in the visible part of the spectrum. Thus, the last five years have seen the rapid application of basic physics principles in quantum mechanics to the problems of generation and amplification of coherent radiation. These principles, experimental techniques, and knowledge of materials had been developing over a period of thirty years.

Since the maser has been successfully operated on both sides of the microwave-infrared gap, what are some of the problems involved in covering the entire region?

First consider a three-level maser illustrated in Fig. 6(a). Energy is pumped in at frequency ν_{13} and coherent radiation produced at either frequency ν_{23} or ν_{12}. The drive power must be great enough to overcome thermal relaxation effects so that roughly[27]

$$P > Nh\nu_{13}p_{31} = Nh\nu_{13}/T_1,$$

where N is the number of molecules, p_{31} is the transition probability and T_1 the relaxation time.

Typical numbers for a microwave case are $N\simeq 4\times 10^{19}$, $T_1\simeq 0.2$ seconds, $h\nu_{13}\simeq 6\times 10^{-24}$ joules, yielding a pump power in excess of 1.2 mw. In the ruby laser $N\simeq 10^{18}$, $T_1\simeq 10^{-5}$ seconds, $h\nu\simeq 3\times 10^{-19}$ joules, yielding pump power of the order of 30 kw. Optical power outputs of this magnitude can be obtained by pulsing a flash tube at the megawatt level.

The problem of the pump frequency always being greater than the signal frequency in the three-level

[23] A. Einstein, "Zur Quantentheorie der Strahlung," *Physik, Z.*, vol. 18, p. 121; 1917.

[24] Townes's ideas presented by A. H. Nethercot, Millimeter wave Symp., University of Illinois, Urbana, Ill.; May, 1951.

[25] N. Bloembergen, "Proposal for a new type solid state maser," *Phys. Rev.*, vol. 104, p. 324; 1956.

[26] H. E. Scovil, G. Feher, and H. Seidel, "Operation of a solid state maser," *Phys. Rev.*, vol. 105, pp. 762–763; January, 1957.

[27] J. P. Wittke, "Molecular amplification and generation of microwaves," Proc. IRE, vol. 45, pp. 291–316; March, 1957.

scheme, causes some difficulty even when the system can be pumped with incoherent optical power. Suppose the wavelength of interest were 0.56 mm (*i.e.*, 10^3 larger than the 5600 A pump line in ruby). Assuming one could devise a scheme that yielded one photon out for each pump photon in, the efficiency could not exceed 10^{-3} because of the difference in energy of the photons. In practice, many other competing decay processes would occur to further reduce the efficiency.

To avoid the high-frequency pump problem, schemes such as the two-level pulsed field maser and harmonic pumping have been considered.

Fig. 6(b) schematically representes a paramagnetic material where the energy level spacing is a function of the applied magnetic field H. The idea[28] is to invert the spin population at a low-frequency ν_{12}, then increase H from H_i to H_f in a time short compared to the relaxation time and obtain oscillation at the higher-frequency ν_{34}. Unfortunately, the wavelength obtained is given approximately by the relation λ (mm) H (gauss)$\simeq 10^5$. Hence, to obtain wavelengths below one millimeter, magnetic fields in excess of 100,000 gauss are required.

To understand the idea of harmonic pumping, consider the Hamiltonian of a molecule in the presence of a radiation field. The energy will consist of three terms

$$H = H_M + H_R + H_I$$

where H_M is the energy of the molecule, H_R the energy of the radiation field alone, and H_I the interaction energy. In the typical maser case H_I is small compared to the other two terms. However, suppose in the case of the transition in ammonia at 24 kMc the energy density $E^2/2$ in the electric field is made the order of the dipole interaction energy density $N\mu E$. Let $N=3\times10^{19}$ molecules/cm³, then for $\mu=10^{-18}$ esu, the electric field intensity E equals 60 esu or 18000 v/cm. For field strength of this order, H_I will be comparable with H_R and thus lead to higher order (harmonic terms) in the perturbation theory. In effect, the belief[29] is that the maser level can be pumped at a subharmonic of the level spacing frequency ν as indicated in Fig. 7 where the third subharmonic is shown. While this scheme has not as yet been proven to work, it suffers from the problem of high input power requirements and most likely the difficulty of all practicable realizable nonlinear systems in that the efficiency of conversion to higher harmonics falls off rapidly with harmonic number.

The submillimeter maser problem is quite different from the classical electronics problem of building a tube for 5 cm wavelength after having built tubes for 10 cm and 1 cm. Much hard work remains before the maser will invade the microwave-infrared gap. Also problems of modulation and tunability of masers are very young and seeking ideas for solution.

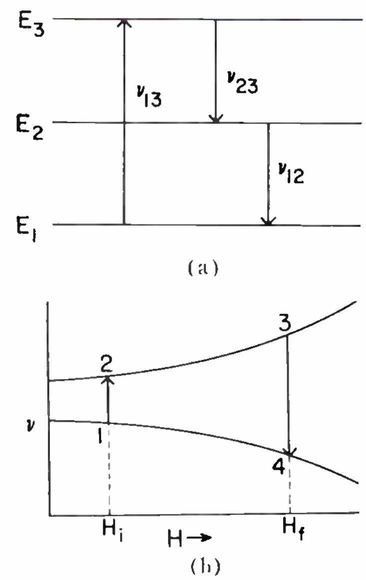


Fig. 6—(a) Three- and (b) two-level maser energy levels.

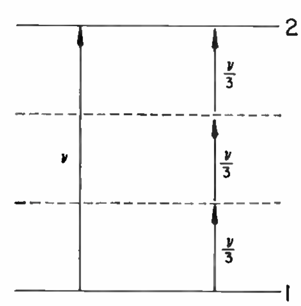


Fig. 7—Harmonic pumping of a two-level maser.

CONCLUSIONS

It was the intent of this paper to show that while rapid progress has and is being made, a real breakthrough on the microwave-infrared gap problem has not been achieved. Many difficulties still plague both the classical and quantum electronic engineer in trying to span this gap region with a coherent radiation source.

The detector problem, while not mentioned in this paper, appears to be of comparable difficulty to the source problem. In the near future, as work on sources progresses, more effort will undoubtedly have to be shifted into the detector area.

Real progress has been made on the transmission problem by Goubau[30] and associates with their beam waveguides.

Someday soon, it is hoped that a sufficient number of pieces will fall into place to permit a practical submillimeter wave system to be built. In the meantime, for those persons who enjoy a challenge, this gap region has much to offer.

[28] S. Foner, *et al.*, "Pulsed field millimeter wave maser," in "Quantum Electronics," C. Townes, Ed., Columbia University Press, New York, N. Y., p. 487; 1960.

[29] J. Fontana, R. Pantell, and R. Smith, "Harmonic Generation by Means of Multiple Quantum Transitions," paper delivered at 2nd Internatl. Conf. on Quantum Electronics, Berkeley, Calif.; March 23–25, 1961.

[30] G. Goubau and F. Schwering, "On the guided propagation of electromagnetic wave beams," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-9, pp. 248–256; May, 1961.

Microwave Interaction with Matter*

M. L. STITCH†, SENIOR MEMBER, IRE

Summary—**The interaction between microwaves and matter is illustrated by two classes of phenomenon: those in which as a result of the interaction the microwave field loses energy, and those in which it gains energy. Among the first class, there are two ways of marking the interaction: one is by detecting the radiation, the other is by detecting the matter. The first way is explained by the technique of microwave spectroscopy, electron spin resonance spectroscopy, and optical pumping. The second way is explained by the technique of molecular beams. Then the second class is introduced by the ammonia maser. Next it is shown how the maser principle led to the solid-state maser amplifier, and how it was extended into the infrared and optical region to bring about the laser. The current state-of-the-art is illustrated by the CW gas laser and the atomic hydrogen maser. Throughout these explanations the methods by which the resolution is increased is examined.**

Introduction

THIS TOPIC of Microwave Interaction with Matter is so broad and well documented that it would be impossible to give anything but the barest statements within the allotted space if we did not restrict the range of our subject matter in some fashion. Likewise, even with a limited selection of topics if any attempt were made to delineate all the key advances together with the contributors to these advances, we should require many additional pages to list the bibliography.

Therefore, we shall limit the discussion to microwave radiation interactions that occur with neutral matter. The subject of interaction with plasmas and electron streams would be of greater interest to our colleagues of the professional groups on Nuclear Science and Electron Devices. We shall further exclude the interaction of matter with electromagnetic fields produced by high-velocity electrons such that the velocity exceeds the velocity of light in the material media; *i.e.*, Cerenkov radiation. This will be discussed by Prof. Coleman elsewhere. Within the scope of neutral matter lies ferromagnetic materials. But the subject of ferrites has been covered exhaustively in the pages of the IRE PROCEEDINGS. Indeed it has been the subject of a special PROCEEDINGS IRE issue. We shall therefore not cover this topic. We shall mention only the principal investigators who presided at the founding of a field or technique.

The States of Matter

Neutral matter, by which we mean bound electron states of matter, can exist in either the ground state or excited states. Though it is neutral, it can possess both electric and magnetic dipoles or higher-order multipoles. The mechanism of energy transfer between matter and radiation is most commonly by coupling between the electric component of the radiation field and the electric dipole, or between magnetic field component and magnetic dipole. The energy transfer may occur over a broad band of frequencies or it may be a sharp "line" resonance transfer.

If, as a result of the interaction, the microwave radiation field loses energy to the material medium, absorption has taken place and we shall classify the material as passive. If the converse is true, emission has occurred and we shall call the material active.

The transfer of energy between a radiation field and atoms or molecules which can undergo transitions between an excited state of energy E_e and a ground state of energy E_g must obey the Bohr condition on frequency:

$$E_e - E_g = h\nu,$$

where ν is the frequency. If E_e and E_g are sharp states, ν is sharply defined for each molecule, though owing to the Doppler effect the frequency "seen" by an individual molecule in motion is not necessarily that which appears on the fixed instruments in the laboratory. If E_e and/or E_g occupy a band of states, then a corresponding band of radiation field frequencies can interact with the molecule or atom.

At steady state, and at temperatures T, the ratio of atoms or molecules N_e occupying an excited state of energy E_e to those, N_g, occupying the ground state of lower energy E_g is given by the Boltzmann factor

$$R = N_e/N_g = \exp - (E_e - E_g)/kT. \quad (1)$$

Under normal finite positive temperature conditions, it is evident that R is less than one. It is then not surprising that the earliest work in the field involved an absorptive interaction.

Microwave Spectroscopy

Work on microwave spectroscopy was done by Cleeton and Williams [1] at the University of Michigan around 1934. They filled a rubberized cloth container with NH_3 gas at a slight excess over atmospheric pressure, and then directed the output of an early magnetron type oscillator on to an optical-style echelette grating of 18 elements. The reflection of appropriate order was then directed on the ammonia sample and the transmitted intensity measured by means of a thermoelectric element. The absorption vs frequency was plotted. This exhibited a characteristic peaked resonance response; the peak lay near 24 kMc, with a bandwidth at half power of approximately 12 kMc. With a Q of approxi-

* Received by the IRE, August 16, 1961.
† Hughes Aircraft Company, Culver City, Calif.

mately 2, this could hardly be called a sharp resonance but it was a beginning.

At the close of World War II the general availability of reasonably monochromatic, tunable, CW microwave-signal generators in the form of klystrons, invented by the late Varian brothers,[1] opened up the era of modern high-resolution microwave spectroscopy.

Three groups, working independently, set the stage for a period of intensive exploration in this field: Bleaney and Penrose [2] at Oxford University; Good [3] at the University of Pittsburgh, and later at Westinghouse Laboratories, and Townes [4] at Bell Laboratories and later at Columbia University. Like the earlier Cleeton and Williams experiment, the sample used was ammonia, and the methods used to study the interaction phenomenon was to examine the effect of the interaction on the microwave radiation.

The Good experiment will serve as an example of the technique. See Fig. 1. The microwave radiation emitted by the klystron is divided between the waveguide cell containing ammonia and the attenuator. The two outputs are rectified and presented at the input of a differential amplifier which feeds the vertical plates of a CR tube. Ideally the outputs are balanced so that at a frequency off resonance, there is no output at the differential amplifier. Then if the cell arm and the attenuator arm were reasonably nonresonant structures, and the klystron was frequency modulated by a sawtooth voltage applied simultaneously to the horizontal plates of the CR tube, any resonant absorptions in the ammonia gas would be displayed as an absorption vs frequency line on the face of the CR tube.

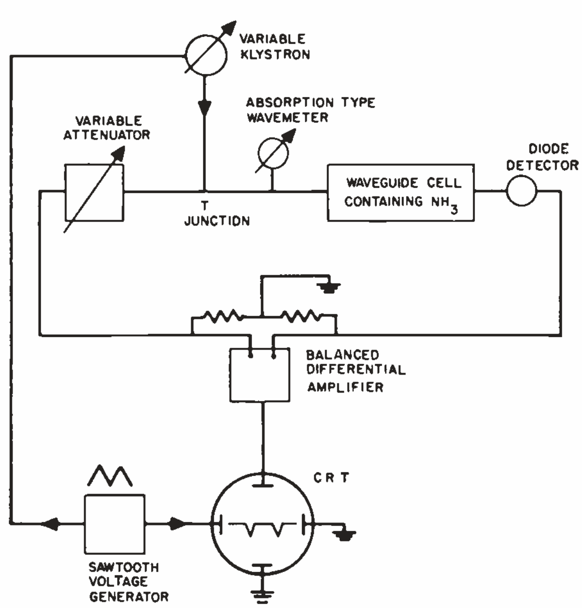


Fig. 1—Simplified block diagram of a microwave spectrometer of the type first used by W. Good.

[1] Sometime after this paper was submitted the author was saddened to hear of the accidental death of Sigurd Varian, who so follows his brother.

Attainment of Narrow Lines

This and the other two experiments were performed at pressures considerably lower than used by Cleeton and Williams. Both Good and Townes worked at under 0.1 mm Hg pressure, and Bleaney and Penrose at 0.2 mm Hg. For reasons that will be explained later, this gave very narrow lines. Indeed Townes, who was first among the three to discuss the role of saturation [4] as a line broadening mechanism, obtained in those early experiments a half-power line width of only 200 kc. This corresponded to a Q in excess of 10^5, a vast increase over the days of Cleeton and Williams.

The ammonia resonance, called an "inversion" resonance, is a complex quantum mechanical phenomena which involves the ammonia molecule turning itself inside out and will not be discussed here. However, it may be said that the resonances were sharp enough in these early experiments to detect the so-called hyperfine structure caused by magnetic-dipole interaction between the N^{14} atomic nucleus and internal molecular magnetic fields. To understand the reasons for the sizeable increase in resolution, which is to say the decrease in line width achieved by these and later experiments, we review briefly the mechanisms of absorption and line broadening.

The microwave radiation induces the ground-state molecules to undergo transition to the appropriate excited state, and vice-versa. Probabilities for such transitions are identical in either direction and are dependent on the microwave field strength. Let p be the probability per unit time of a transition in either direction; and N_g, N_e, respectively, are the molecular populations in the ground and excited states. Then by definition of probability, it is evident that

$$dN_e/dt = -dN_g/dt = (N_g - N_e)p. \qquad (2)$$

From (2) above we see that absorption of radiation and the consequent transition of molecules from ground to excited states is dependent on maintaining a net excess of molecules in the ground state. At first glance these seem to be a logically inconsistent set of conditions: a sizeable number of transitions from the ground state to the excited state should soon reduce the excess to zero, and thus eliminate absorption. The resolution of this apparent paradox is that there exist relaxation mechanisms by which the molecules return to the ground state without re-emitting energy back to the microwave field.

For gases this relaxation mechanism is provided by collisions. These serve to re-establish temperature equilibrium and reassert the population distribution between ground and excited states expressed by (1). A semiclassical way to look at this is to assume that the collisions interrupt the absorption process. Thus one has absorption given by a pulsed sinusoid with an envelope of width τ, where τ is the mean time between collisions. This leads in the well-known way to a Fourier transform with a spectral width which is reciprocal to τ. Since collision rate is proportional to pressure p, one has for line width, $\Delta\nu_p$, due to "pressure broadening,"

$$\Delta\nu_p \propto p$$

as mentioned above.

We might well inquire if one continuously reduces the pressure, where is the limit to line width? This limit is provided in two ways. First, as pressure decreases the time between relaxation collisions increases and one starts to get into the region of saturation. That is: at the center of the absorption line, where the absorption rate is the greatest, the process expressed by (2) is not sufficiently offset by relaxation with the result that the population departs markedly from the Boltzmann distribution (1), in that the excess in the ground state is greatly reduced (*i.e.*, the equivalent temperature T in (1) is increased). This reduces the absorption intensity. In the "wings" of the absorption line, where the absorption is less, this saturation effect is less and therefore the reduction in absorption intensity is less. As a result the absorption intensity is reduced mostly in the center of the line. Thus the half-power points occur at lower absorption corresponding to a greater frequency width. See Fig. 2.

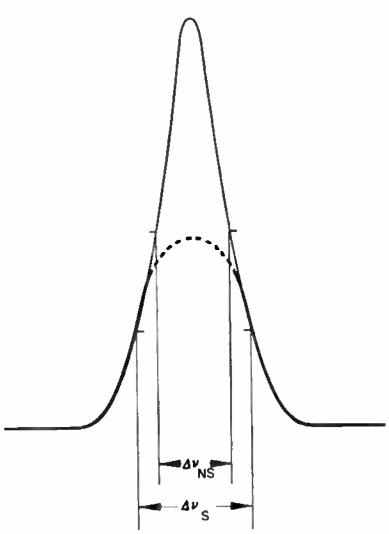


Fig. 2—Illustration of how any mechanism which decreases the intensity at the center of a resonance increases its apparent half-power width.

Saturation broadening can be alleviated by decreasing the microwave power as much as the sensitivity of the detector will allow. A second limit on line width vs pressure decrease is collisions with the walls. This can be decreased by increasing the dimensions of the sample cell or lowering the temperature. Another lower limit is set by so-called Doppler broadening. This is based on the above-mentioned fact that the rapidly moving molecules "see" different frequencies from that emitted by the microwave generator, and as the frequency of the incident wave is varied, different velocity classes of molecules "see" the appropriate frequency. From kinetic theory the velocity dependence and therefore the Doppler width is proportional to $(T/M)^{1/2}$, where M is the molecular weight. For NH_3 at room temperature $(\Delta\nu)_{Doppler}$ is approximately 70 kc. Methods which reduce Doppler width will be discussed at appropriate places.

The floodgate was opened by the early high-resolution microwave spectroscopy work. Many ingenious ways to decrease the line width and increase the resolution, to extend the list of materials which could be examined, to extend the frequencies up to millimeter waves, etc. could be chronicled. As a result of this work much was learned about molecular and nuclear parameters, collision processes, and valence theory, as well as other fields.

Electron Spin Resonance

A varient of the microwave spectrometer is the electron spin resonance spectrometer. In this device, the matter is frequently in liquid or solid form. Further this matter must be restricted to atoms or molecules in which an unpaired electron exists; *i.e.*, paramagnetic material or free radicals. The reason for this is evident from an explanation of its principle of operation. An electron in a magnetic field assumes one of two orientations. Either its magnetic moment axis is parallel or antiparallel to the applied field. The energy difference between these two states is proportional to the applied magnetic field. The resonance frequency for the transition is given by

$$\nu(\text{mc}) \approx 2.8 \times 10^6 H \text{ (gauss)}$$

for a free electron.

The presence of other contributions to the local magnetic field at a bound electron caused by chemical binding and nuclear perturbations in effect makes the paramagnetic electron a minute probe of the molecular and nuclear parameters as well as a tool for exploring relaxation processes. The technique is more sensitive at microwave frequencies corresponding to higher applied external magnetic fields than at radio frequencies. This is because $E_e - E_g$ is greater in (1) leading to greater ground population excess. Experimentally, electron spin resonance is much like microwave spectroscopy except that an external magnetic field is used. Since the microwave frequency can be fixed and the magnetic field varied, many of the problems caused by the necessity for broad-band operation of microwave spectrometers is avoided here. The higher concentrations of the condensed phase of matter increase sensitivity of this method. Early work was done by groups in the U.S.S.R. [5], the United States [6], and Great Britain [7].

Double Resonance and Optical Pumping

One problem that microwave spectroscopic techniques have is that the fractional excess of molecules is given by the Boltzman distribution of (1), corresponding to microwave frequencies and room temperatures. Thus, at K band the excess is ≈ 0.4 per cent. This excess we recall is what provides a net absorption when stimulated by the presence of resonant microwave radiation.

A technique for altering the distribution of population so as to increase the disparity between two states involved in a microwave transition and thus increase the magnitude of the microwave resonance is the method of double resonance. The philosophy here is to cause a simultaneous resonant interaction between one of the two states and a third state while observing the microwave resonance interaction between the original two. One such method involves a close lying state, a nuclear interaction state, and requires an RF signal. This method, invented by Feher [8] at Bell Laboratories in 1956, is called the ENDOR (Electron Nuclear DOuble Resonance) technique. It provides an enormous increase in resolution by permitting one to sweep the second resonance frequency. Another technique which we shall briefly describe is the method of "optical pumping" or optical orientation. This grew out of the work of Bitter [9] at M.I.T. and Kastler [10] and Brossel [11] in Paris.

For an illustration, we will make use of the simplified scheme in Fig. 3. Optical radiation at frequency ν_{23} is supplied to the system. This stimulates atoms in State 2 to absorb radiation and jump to State 3. Within $\approx 10^{-8}$ seconds the excited atoms spontaneously emit with equal probability at frequencies ν_{32} and ν_{31} which carries them in equal numbers back into states No. 1 and No. 2. The net effect of this process is to increase the population in State 1 at the expense of State 2, and so microwave absorption at frequency ν_{12} is enhanced.

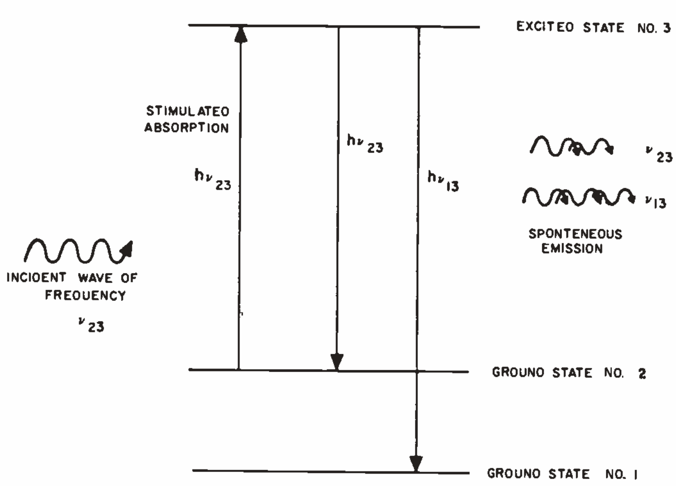


Fig. 3—Simplified optical pumping scheme (not drawn to scale, typically $\nu_{13}/\nu_{12} \approx 10^5$).

A second accomplishment of this technique is that no longer is collision required as a mechanism for relaxation since the pumping radiation maintains the population disparity and thus performs this function. If a way could be found to prevent the atoms from interacting with each other, a very sharp resonance could be achieved. Typically the atoms, alkali metal vapor, are placed in a transparent sealed container to be irradiated with light and microwaves. The pressure of the alkali metal vapor is determined by controlling the temperature of the container.

How can the atoms be isolated from each other? One must decrease the vapor density of alkali atoms until the average time between interatomic collisions is increased sufficiently. Unfortunately collisions with the container walls, which now predominate, still set an upper time limit to the length of the microwave train $\approx 10^{-4}$ to 10^{-3} sec. Not only does this give rise to undesired broadening of the microwave resonance line, but the relatively short time the atoms spend between collisions in the weak optical radiation available is insufficient to effect a sizeable population shift. In other words, the amount of population shift (sometimes called the "degree of optical alignment" or atomic "orientation" or "polarization" because the effect of optical absorption and re-emission is to transfer net angular momentum and hence magnetic moment to the atoms of alkali metal vapor) increases with the intensity of optical resonance radiation and the duration of uninterrupted (by collisions) optical absorption by the atoms.

Dicke [12] of Princeton University proposed an ingenious solution to these difficulties: In addition to the dilute alkali metal vapor at a pressure $\approx 10^{-7}$, a so-called buffer gas composed of magnetically inert atoms such as argon or helium at a relatively high pressure of ≈ 1 mm is put into the cell. The actively absorbing alkali metal atoms, too dilute to collide with each other, collide with the numerous neutral buffer gas atoms. These collisions can be shown to cause almost negligible perturbations of the state of the alkali metal atoms. As a result, the alkali metals perform a random walk and thus spend appreciable time (≈ 1 sec) in both optical and microwave radiation fields. Dicke showed further that there results a sort of "collision narrowing" in which the Doppler width is replaced by a line width due only to the finite time spent in the interaction field, *i.e.*, the Fourier transform of a pulse of duration τ.

A block diagram of an optical pumping apparatus is shown in Fig. 4. The resonance is monitored by the optical rather than the microwave radiation. An examination of Fig. 3 will make it apparent that at steady state a single photon absorption at microwave frequency at resonance is accompanied by a single photon absorption at optical frequency. Since typical optical detectors will detect of the order of an optical photon whereas typical microwave detectors require several hundred microwave photons to overcome noise, there is an enormous gain in sensitivity in optical detection. An obvious application of the foregoing technique is in the commercially available Rb gas cell frequency standard which combines better than 1 part in 10^9 stability with small size and high portability.

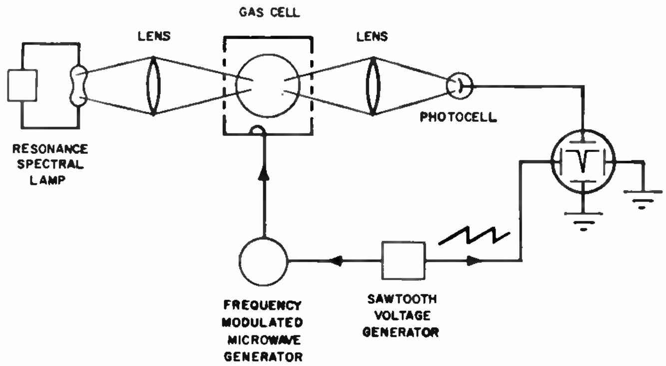


Fig. 4—Block diagram of an optical pumping double-resonance apparatus.

Molecular Beam Magnetic Resonance

Thus far we have discussed techniques in which the interaction between radiation and matter is observed by monitoring the radiation. We examine next the converse scheme in which the interaction between microwave field and matter is examined by observing some change in a property of the latter.

This work grows out of the elegant molecular beam magnetic resonance method of Rabi [13] and collaborators at Columbia University which led to his Nobel award. The molecular beams magnetic resonance method is the earliest of the precision radio frequency spectroscopy techniques and dates back to 1937. We use the generic term "molecular beam" to describe a beam of atoms or molecules.

The technique is as follows: atoms of the material, typically in the condensed phase, are sputtered out of an oven through a series of collimating slits all in a very high vacuum ($\approx 10^{-9}$ mm Hg or better). The resulting beam is composed of atoms so isolated from each other as to behave like independent particles (Rabi once said, "They are more isolated from each other than the stars."). This beam then travels along the axes of a long tube containing two inhomogeneous magnetic fields separated by an RF interaction region in which the electromagnetic fields are introduced. At the far end of the tube are some stops and a particle detector. The sketch in Fig. 5 illustrates a so-called "Flop-in" type of apparatus. Being a magnetic resonance, the property which changes at resonance is the magnetic moment. The first and second inhomogeneous fields deflect unchanged atomic states in the same direction into a stop. The atoms which have flopped are refocused into the detector.

The method of atomic beams requires no relaxation processes to maintain a steady state because the atoms involved in the interaction are constantly being replaced by fresh ones. This is why they may be isolated from each other. If the microwave field is so arranged that its propagation vector is normal to the beam path, Doppler broadening is minimized. Once again as in optical pumping one can conceive of the line width as being inversely proportional to interaction time. The

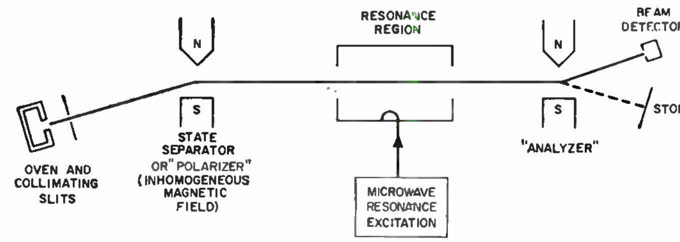


Fig. 5—Atomic beam apparatus.

first use of the conventional magnetic resonance methods at microwave frequencies (*i.e.*, greater than 1 kMc) was made by groups at Columbia [14] and M.I.T. [15] using the so-called weak field $\Delta F=1$ Hyperfine interaction of atomic hydrogen.

In addition to its use as a spectroscopic tool of enormous resolution and sensitivity, the molecular beam with Cs source leads to an X-band microwave resonance line. This has been embodied into a sophisticated frequency standard with a long term stability of the order of 1 part in 10^{11} and sold by one company as an off-the-shelf item!

A particular application of an atomic beam magnetic resonance at microwave frequencies up to 10 kMc was actually used earlier in a brilliant experiment by Lamb [16] and Retherford that led to a Nobel award for Lamb. Technically, one can call this a spectroscopy experiment in that transitions between the metastable $2S_{1/2}$ state and excited P states of atomic hydrogen are observed in a magnetic field. However this is somewhat analogous to calling the development of the nuclear fission bomb an experiment in collision cross sections. For one thing the experiment produces and detects metastable $2^2S_{1/2}$ hydrogen atoms which had not been isolated in a clear cut manner previously, let alone stimulated into magnetic resonances. Space does not permit a recounting of this specialized experiment beyond saying that it had a profound effect on the development of theoretical physics.

Active States and the Maser Principle

At this point we note that both microwave spectroscopy and the molecular beam resonance method deal with passive matter and microwave absorption. However for optical pumping this is not necessarily so. It is apparent from Fig. 3 that if the pumping transition were initiated from ground state No. 1 instead of No. 2 there would be an excess in State No. 2, and the microwave transition would envolve stimulated emission rather than absorption. But the optical transition is still an absorptive one.

If it were possible to obtain a sufficient population excess in an excited state, the fact that stimulated emission is always exactly in phase with the excitation should make it possible to obtain a molecular amplifier, and by feeding back the amplified radiation one should obtain a molecular oscillator.

This was first mentioned in a public meeting by A. H.

Nethercott on behalf of C. H. Townes of Columbia at a Symposium on submillimeter waves at the University of Illinois in May 1951. The first report in print was the proposed design of a beam type device in the Columbia Radiation Laboratory Quarterly Progress Reports of December 31, 1951, by Townes and associates. Successful operation of this device was reported by Gordon, Zeiger and Townes [17]. Independently Weber of the University of Maryland proposed a device at the IRE Electron Tube Confe.ence in Ottawa in June 1952 and published his discussion [18] in 1953, but never published any claims for successful operation of this device. Also independently, Basov and Prokhorov [19] of the Lebedev Institute, Moscow, gave some designs for a beam type device.

The Townes device was called a MASER (Microwave Amplification by Stimulated Emission of Radiation) and made use of an electrostatic state selector acting on a beam of ammonia molecules. To understand its operation, consider a uniform beam of excited ammonia molecules going through a microwave cavity. If the losses are low enough, the time in which the microwave fields and ammonia particles interact long enough, the excited population of the beam dense enough, and the coupling between the microwave field and particles strong enough, the amplification from noise at the resonant frequency will be sufficient to overcome losses, and regenerative oscillation will occur. This can be restated succinctly by the following relation.

Condition for oscillation:

$$(Q)(L/V)(\rho)(\mu^2) = \text{a certain constant},$$

where Q is the cavity Q, L is the cavity length, V is the mean beam velocity, ρ is the mean beam density, and μ is the molecular dipole moment. Since it has received widespread publicity and is probably familiar to most of the scientific community, we will save space by omitting an illustration of the NH_3 maser.

The main reasons that an optically pumped Rb, say, gas cell will not go into maser oscillation is that μ for the magnetic dipole transition of Rb is several orders of magnitude smaller than μ for the electric dipole transition for NH_3, and ρ for the gas cell must be kept low enough to prevent complete "screening" of the volume of gas due to the enormous cross section for optical resonance photons.

The gas maser oscillator possesses the highest short term stability ever obtained at any radio frequency. Because it is an amplifier that does not suffer from the noise due to charged particles, the maser is capable of very low noise when used as an amplifier. The use of gas beams with their sharp resonances and low densities rules out much amplifier bandwidth. The hope for reasonable bandwidths and power has required a solid-state maser.

The first sold-state system which could operate continuously was proposed by Bloembergen [20] of Harvard University. It has a three-level system similar in function to the optical pumping scheme of Fig. 3, and one of his proposed materials was soon made to operate. The use of an external magnetic field to obtain the ground state splitting of the maser materials allowed construction of tunable maser amplifiers. The noise figures were so close to 1 that it became fashionable to use equivalent noise temperatures to describe the noise. Noise temperatures so low that a few microwave photons can be detected have become commonplace, and maser amplifiers have become commercially important in radio astronomy, tracking systems, proposed space communication systems and wherever extremely low-noise amplification is needed.

The Laser

One of the most spectacular achievements of the kind of conceptualizing that produced the maser was ushered in by the appearance of a paper by Schawlow and Townes [21] late in 1958 proposing masers that would operate at optical and infrared frequencies. They included some detailed calculations on general design considerations and operating characteristics as well as on a particular system using potassium vapor.

They proposed the use of a multi-mode cavity with strong mode selection by means of a pair of Fabry-Perot plates. The normally reflecting mode has the highest Q, builds up the quickest, and selectively robs all competing modes of excited population in stimulating emission. Because of the highly directional plane wave emitted the fractional amount of spontaneous emission, which is omnidirectional, is quite small within the highly directional beam output.

A later paper by one of the authors [22] suggested the use of a red ruby rod with optically flat parallel faces, one partially silvered, the other entirely silvered. It would operate at low temperature in a 4-level scheme, in which the pumping is between the ground level and the top level, actually a band of levels that will accept a wide band of pumping illumination in the green region of the spectrum. The stimulated emission would be between the 2nd and 3rd states. In 1961 Schawlow actually got the red ruby scheme to work, although experimental difficulties have at the time of this writing prevented the potassium, or other optically excited alkali-metal-vapor systems, from working.

After the original Schawlow-Townes paper a widespread competition was set off to get the first operational laser as many called it (for Light Amplification by Stimulated Emission of Radiation). The honor for the first successful operation of a laser went to Maiman [23], then at Hughes Aircraft Company, who used a low Cr ion concentration pink ruby in the main fluorescent, R_1 line, for a brute-force 3-level scheme of operation that required the pump to transport more than half the population out of the ground state for laser action. This was because the ground state was the terminal state for the stimulated emission. Maiman was able to get 10 kw bursts of radiation at 6943 Å in the deep red!

He used for his pump an intense source of light, a Xenon flash tube excited by a capacity discharge of over a thousand joules at a rate of several megowatts. (See Fig. 6.)

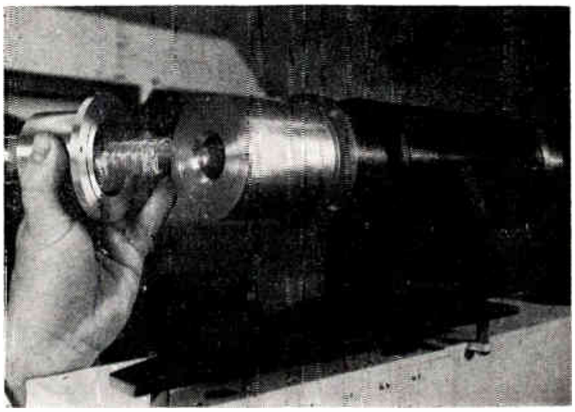


Fig. 6—Ruby laser showing flashtube and ruby.

The chief virtue of the ruby laser is that it is a brightness intensifier. At the typical current state-of-the-art it emits a highly collimated beam whose brightness is approximately 10^6 times that of the flash tube which excites it. If one takes account of the monochromaticity, the brightness per unit wavelength (or frequency) is approximately 10^{11} times that of the exciting flash tube.

The coherence property of lasers is more strikingly demonstrated in a CW gas laser first proposed [24] and later built by A. Javan and collaborators at Bell Laboratories. The technique was to use a mixture of two gases, helium and neon. The first gas is excited by electrical discharge and transfers its energy to the second gas. This device worked in the near IR region at approximately 1 μ and delivered approximately 15 mw power in a beam of a few seconds of arc angular width and with a line width under 1 kc! The equivalent Q here is greater than 3×10^{11}!

The applications which await the successful development of a high-power, monochromatic CW or controllable pulse laser are enormous. Likewise the development of a tunable IR frequency laser, will be of great value to the spectrographic study of chemical reaction parameters.

The Hydrogen Maser

The ammonia beam type maser, too, has evolved to the point where an atomic hydrogen beam maser, using the storage box techniques first proposed by Ramsey [25] of Harvard, has been made at that University. In this technique the resonant cavity contains a coated quartz bulb. The hydrocarbon coating does not quench the interaction between the active atoms and the microwave field in the cavity for a comparatively long period (though it may perturb the frequency slightly). Ultimate absolute precision could be as high as one part in 10^{14}—with even higher short-term stability. And this device could be of great value as an active frequency standard.

Bibliography

[1] C. E. Cleeton and N. A. Williams, "Electromagnetic waves of 1.1 cm wave-length and the absorption spectrum of ammonia," *Phys. Rev.*, vol. 45, pp. 234–237; February 15, 1934.

[2] B. Bleaney and R. P. Penrose, "Ammonia spectrum in the 1 cm wave-length region," *Nature*, vol. 157, pp. 339–340; March 16, 1946.

[3] W. E. Good, "The inversion spectrum of ammonia," *Phys. Rev.*, vol. 70, pp. 213–218; August, 1946.

[4] C. H. Townes, "The ammonia spectrum and line shapes near 1.25-cm wave-length," *Phys. Rev.*, vol. 70, pp. 665–671; November, 1946.

[5] This work was referred to by W. Gordy, W. V. Smith and R. F. Trambarulo, in "Microwave Spectroscopy," John Wiley and Sons, Inc., New York, N. Y.; 1953.

[6] R. L. Cummerow and D. Halliday, "Paramagnetic losses in two Manganous salts," *Phys. Rev.*, vol. 70, p. 433; September, 1946.

[7] D. M. S. Bagguley and H. E. Griffiths, "Paramagnetic resonance and magnetic energy levels in chrome alum," *Nature*, vol. 160, pp. 532–533; October 18, 1947.

[8] G. Feher, "Method of polarizing nuclei in paramagnetic substances," *Phys. Rev.*, vol. 103, pp. 500–501; July 15, 1956.

[9] F. Bitter, "The optical detection of radio frequency resonance," *Phys. Rev.*, vol. 76, pp. 833–835; September 15, 1949.

[10] A. Kastler, "Some suggestions concerning the production and detection by optical means of inequalities in the populations of levels of spatial quantization in atoms. Application to the Stern and Gerlach and magnetic resonance experiments," *J. Phys. Radium*, vol. 11, pp. 255–265; June, 1950. (In French.)

[11] J. Brossel and A. Kastler, "The detection of magnetic resonance of excited levels: depolarization effect of optical resonance and fluorescence radiations," *C. R. Acad. Sci.*, vol. 229, pp. 1213 1215; December 5, 1949. (In French.)

[12] R. H. Dicke, "The effect of collision upon the Doppler width of spectral lines," *Phys. Rev.*, vol. 89, pp. 472–473; January 15, 1953.

[13] I. I. Rabi, "Space quantization in a gyrating magnetic field," *Phys. Rev.*, vol. 51, pp. 652–654; April 15, 1937.

[14] J. E. Nafe and E. B. Nelson, "The hyperfine structure of hydrogen and deuterium," *Phys. Rev.*, vol. 73, pp. 718–728; April 1, 1948.

[15] D. E. Nagle, R. S. Julian and J. R. Zacharias, "The hyperfine structure of atomic hydrogen and deuterium," *Phys. Rev.*, vol. 72, p. 971; November 15, 1947.

[16] W. E. Lamb, Jr. and R. C. Retherford, "Fine structure of the hydrogen atom by a microwave method," *Phys. Rev.*, vol. 72, pp. 241–243; August 1, 1947.

[17] J. P. Gordon, H. J. Zeiger and C. H. Townes, "Molecular microwave oscillator and new hyperfine structure in the microwave spectrum of NH_3," *Phys. Rev.*, vol. 95, pp. 282–284; July 1, 1954. Also J. P. Gordon, Ph.D. dissertation, Columbia Univ.; New York, N. Y.

[18] J. Weber, "Amplification of microwave radiation by substances not in thermal equilibrium," IRE Trans. on Electron Devices, vol. ED-3, pp. 1–4; June, 1953.

[19] N. G. Basov and A. M. Prokhorov, referred to in N. G. Basov and A. M. Prokhorov, "Molecular Oscillator and Amplifier," *Uspekhi. Fizicheskikh Nauk* (*Advances in Physical Science*), pp. 485–501; November, 1955.

[20] N. Bloembergern, "Proposal for a new type solid state maser," *Phys. Rev.*, vol. 104, pp. 324–327; October 15, 1956.

[21] A. L. Schalow and C. H. Townes, "Infrared and optical masers," *Phys. Rev.*, vol. 112, pp. 1940–1949; December 15, 1958.

[22] A. L. Schalow, "Infrared and optical masers," in "Quantum Electronics," C. H. Townes, Ed., Columbia University Press, New York, N. Y.; 1960.

[23] T. H. Maiman, "Stimulated optical radiation in ruby," *Nature*, vol. 187, pp. 493–494; August 6, 1960.

[24] A. Jaran, "Possibility of obtaining negative temperature in atoms by electron impact," in "Quantum Electronics," C. H. Townes, Ed., Columbia University Press, New York, N. Y.; 1960.

[25] N. F. Ramsey, "Resonance experiments in successive oscillatory fields," *Rev. Sci. Instr.*, vol. 28, pp. 57–58; January, 1957.

Section 21

MILITARY ELECTRONICS

Organized with the assistance of the IRE Professional Group on Military Electronics

The Early History of Radar*

R. M. PAGE†, FELLOW, IRE

***Summary*—Five basic ideas are identified, the combination of which consitutes radar. A clear distinction is then made between this combination of ideas, the contemporary technology from which it grew, and the contemporary scientific knowledge on which it was based. The mainstream of the development of radar is traced in a sequence of related events from 1922 to 1941. The technical problems encountered and the solutions employed in the first radar development are outlined in some detail. Two sidestreams of radar development are identified. The relationships among the three streams are discussed.**

TECHNOLOGICAL innovation grows out of contemporary technology, which in turn rests on the research and scientific discoveries of an earlier day. Only when clear distinction is made between innovation, contemporary technology, and contemporary scientific knowledge can lines of interdependence be meaningfully drawn.

The combination of five basic ideas constitutes the innovation which is radar. They are 1) that electromagnetic radiation at high radio frequency be used to detect and locate remote reflecting objects, 2) that the radiation be sent out in pulses of a few microseconds duration, separated by "silent" intervals very many times the pulse duration, 3) that pulses returned from reflecting objects be detected and displayed by receiving equipment located at the point of transmission, 4) that distance be determined by measuring in terms of an independent time standard the time of flight of pulses to "target" and back, and 5) that direction be determined by use of highly directive radio antennas.

We first identify the scientific knowledge underlying these ideas. Faraday and Maxwell had established the theoretical possibility of the electromagnetic field. Hughs had demonstrated its existence at radio frequency. Hertz had demonstrated that radio waves behaved as light waves, obeying the known laws of propagation and reflection. Appleton and Barnet had demonstrated that radio waves could be used in interferometer fashion to determine apparent height of the ionosphere. Their method used phase velocity to measure the difference in length of two propagation paths. Swann and Frayne had suggested and Breit, Tuve, and Taylor had demonstrated that radio waves could be used to measure ionosphere height by observing the relative flight time of pulses of radio transmission. Their method used group velocity to measure the difference in length of two propagation paths. Various experimenters in radio, including Breit and Tuve of the Carnegie Institution, working with ionosphere measurements, engineers of the British Post Office working with short wave radio, and engineers of the Bell Telephone Laboratories working with television, had observed that aircraft flying near their receivers or transmitters created noticeable disturbances in the radio propagation field. This was regarded merely as interference with their experiments, and otherwise ignored. These constitute the elements of scientific knowledge basic to the idea of radar. Since radar was not in any way an objec-

* Received by the IRE, August 8, 1961.
† U. S. Naval Research Laboratory, Washington, D. C.

tive in the discovery of these facts, since no problem was recognized for which radar was the proposed solution, and since radar was only one of many technological innovations dependent on these same scientific discoveries, it is not proper to ascribe to any of the named discoverers any responsibility for the origin of radar by virtue of their discoveries.

Our next step is to identify the contemporary technology out of which radar grew. A catalogue of the state of the art in radio engineering in the 1930's would be both tedious and superfluous. Certain elements have a degree of specificity to radar, however, and require special mention. The cathode-ray tube, devised by the German scientist Braun in 1897 and used on an experimental basis in the early 1920's, became generally available as a laboratory tool in the early 1930's. In 1900 Nikola Tesla suggested the use of electromagnetic waves to determine relative position, speed, and course of a moving object. In 1903 Huelsmeyer applied for a patent in Germany on an anti-collision device for ships, based on directive transmission and reception of continuous waves at very short radio wavelength. Transmitter and receiver were shown on the same ship, but separated as widely as possible. In June 1922 Marconi again suggested the use of radio as an anti-collision device. In 1923 Loewy filed a patent application in the U. S. Patent Office for a radio object detector employing the Fizeau principle. The transmission consisted of chopped CW, with approximately equal intervals on and off. A target would be detected when its reflection coincided with the intervals between transmission. While Loewy's disclosure appears at first to anticipate radar, it fails to meet the requirements for radar, since range indication is ambiguous, and the presence of one target would jam the system for all other targets. Thus it gave no operational advantage over Huelsmeyer, except the possibility of locating transmitter and receiver close together. In 1925 Breit and Tuve proposed a radio pulse method, which they credited originally to Swann and Frayne of the University of Minnesota, for probing the ionosphere. In cooperation with Taylor, Young, and Gebhard of the Naval Research Laboratory the method was used for the first time in that same year. Although this has been said to demonstrate the basic principles of radar, it fails to meet the radar criteria on several counts. The pulses were much too long, being about half a millisecond. This would blank out the first 50 miles of range. The ratio of pulse spacing to pulse duration was too low, being only four or five, therefore subject to saturation by a very few targets. The receiving equipment was not at the point of transmission, so time of flight of radio pulses to target and back could not be measured in terms of an independent time standard. Only the difference in length of two propagation paths was measured, and direction was not indicated. These deficiencies from the radar viewpoint were imposed by the state of the art in 1925. They detracted nothing from the excellence of the method or the apparatus for probing the ionosphere. In 1930 patent applications were filed by Wolf and Hart for a radio pulse altimeter. The disclosures were based on the technology of the ionosphere probe, and were therefore subject to some of the same limitations. No development of radar apparatus resulted from these disclosures. In November, 1933, Hershberger (U. S. Signal Corps) proposed a method essentially similar to that of Loewy and then did some work on microwave generators in a vain attempt to obtain the power required for useful echoes. In 1936 the French liner Normandie was equipped with a microwave anti-collision device similar to that of Huelsmeyer.

It is now obvious that contemporary technology contained much that was suggestive of radar. However, none of the art described contained all five elements necessary to radar, and no radar development resulted from any of it. It is therefore inappropriate to trace the development of radar to any of these proposals or related developments.

The first incident that led ultimately to radar was the accidental observation by Taylor and Young in September 1922 that a ship interrupted some experimental high-frequency radio communication across the Potomac when it intercepted the propagation path between transmitter and receiver. Taylor and Young had for many years been employed by the Navy, and were keenly aware of the problem of screening Naval forces from penetration by other ships in darkness and fog. Though the observation was unrelated to their experiment, the application was obvious to them, and they immediately proposed that high-frequency radio transmitters and receivers be installed on destroyers to detect the passage of other ships between any two destroyers in radio contact. Obviously this was not radar. It did not even involve reflection of radio waves, and was in no way related to Marconi's suggestion, as has sometimes been inferred. It is identified with radar here only because Taylor and Young later originated the first radar development project, and this incident started them thinking in terms of detection of moving objects by radio.

The second incident was another accidental observation, this time by Hyland, a colleague of Taylor and Young. During experiments on high-frequency radio direction finding in June 1930, he detected a severe disturbance of the propagation field by an airplane flying overhead. Hyland was also an experienced Navy employee, and was sensitive to the potential threat of military aircraft and the need for warning devices against them. The observation again was unrelated to his experiment, but the application was obvious, and he immediately proposed that high-frequency radio be used for aircraft warning.

On Taylor's recommendation a project was established at the Naval Research Laboratory in January 1931 for "Detection of Enemy Vessels and Aircraft by Radio." Work on this project continued for several years. The "beat" method was employed, in which transmitter and receiver were widely separated and shielded from each other, transmission was CW and reception observed the fluctuating signals, called "beats," when an airplane flew through the radio propagation field. Detection ranges of 40 miles were obtained in these experiments.

The required wide separation of transmitter and receiver precluded the use of the beat method on ships and limited its usefulness to the protection of large land areas such as cities and military bases. Since this was exclusively the responsibility of the Army, it was proposed in January, 1932, that the Army take over the development for its use in that function. Subsequently, Navy interest in the problem lagged until Young suggested to Taylor that the pulse method be tried. Young's proposal combined for the first time all five elements essential to radar. Ultimately, Taylor accepted the proposal and assigned to the author, working under Young's supervision, the task of developing pulse radar. The author's work on this task was started on March 14, 1934.[1]

The first step was to develop an indicator to display the outputs of transmitter and receiver. A suitable sweep circuit was built for a commercially produced 5-in cathode-ray oscilloscope. The next step was development of a pulse transmitter. The transmitter frequncy of 60 Mc was chosen because that was the frequency then used in the beat method. Pulse length was slightly under 10 μsec, and pulse spacing, 100 μsec, these being chosen as appropriate experimental values. The keyer was an asymmetric multivibrator. The antenna was a single half-wave horizontal doublet with a single resonant reflector. The pulse power was estimated to be between 100 and 200 w. The first question to be resolved was whether echo pulse energy could be detected during the intervals between transmitted pulses, since synchronous detection, characteristic of the beat method, was known to be more sensitive than asynchronous detection, characteristic of the pulse method. Autocorrelation and crosscorrelation were unheard of in those days, and the trade-off between time and bandwidth disclosed by Hartley,[2] as well as the significance of average energy were not too well understood. The only sure recourse was to try it, and skepticism was great. A broad-band high-gain experimental communications receiver was borrowed and connected to a second antenna similar to the transmitting antenna. Coupling between the two antennas was appreciable, and the transmitted pulse caused the receiver to ring for 30 to 40 μsec. However, when a small airplane flew across the beam at a distance of about a mile, the received signal caused the receiver output following the transmitted pulse to fluctuate violently between zero and saturation. This test was completed in December,

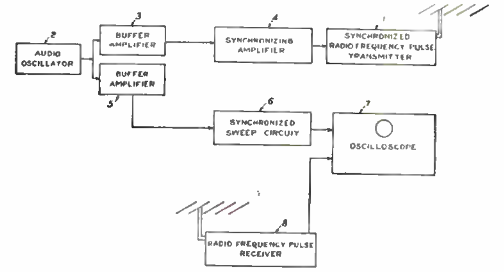


Fig. 1—Block diagram of radars operated in 1936. The block diagram for the radar tested in December, 1934, was similar to this one.

1934. Although synchronous detection prevailed due to the transient ringing of the receiver, the great amplitude of the response left no doubt that an asynchronous detector would also have responded to the reflected pulse. The result was accepted as evidence that echo signals could be detected during the intervals between transmitted pulses, and development of a superior radar receiver was immediately undertaken.

Radar imposed four severe requirements on the receiver which were not encountered in conventional receivers of the time. Close proximity of receiver and transmitter subjected the receiver to paralyzing overload, from which recovery to full sensitivity in the incredibly short time of a few microseconds was mandatory. The first design requirement was to eliminate grid blocking. This was achieved by using a tuned grid circuit with the grid returned to the cathode through the tuning coil. Grid coupling capacitance was then reduced to a minimum by using maximum inductance to capacitance ratio in the tuned circuit, and loading the tuned circuit to the proper Q value with the driving plate resistor.

The second design requirement was to minimize the ring time of tuned circuits from the transmitter-induced high signal level. This was achieved by returning grids to cathodes without bias, thus limiting the level to which the circuits could be driven by the transmitter.

The third requirement was fast response to amplify the short pulse echoes. This meant tailoring the Q values of all tuned circuits so that the composite Q of the receiver would match the pulse length. This was accomplished with the help of the appropriate equation published by Mesny.[3]

The fourth requirement was complete absence of regenerative feed-back in the presence of high gain. A communication receiver of that day was considered stable if it did not oscillate. Equivalent Q, however, is a sensitive function of feed-back, and response characteristics are readily altered by feed-back long before the point of oscillation is reached. This requirement was met by using a superheterodyne receiver, limiting volt-

[1] A. H. Schooley, "Pulse radar history," PROC. IRE (*Correspondence*), vol. 37 p. 405; April, 1949.

[2] R. V. L. Hartley, "Transmission of information," *B.J.T.S.*, vol. 7, p. 535; 1928.

[3] R. Mesny, "Time constants, build-up time and decrements," *l'Onde Elec.*, vol. 13, pp. 237–243; June, 1934.

age gain on any one frequency to one thousand, and changing intermediate frequency as required to accomplish an over-all voltage gain on the order of $10.^7$ In addition, extreme precautions were taken in shielding, filtering, and common point grounding.

The receiver was intended for a 5-μsec pulse. The over-all response was 90 per cent of steady state in 5 μsec. This characteristic was independent of gain up to the point where thermal noise at the input filled the cathode-ray screen.

A new transmitter of the self-quenching or "squegging" type was built to go with the new receiver. The transmitting antenna was a 4×4-wavelength curtain array with resonant reflector. The receiving antenna was a single half-wave doublet with single resonant reflector. The frequency was 28.6 Mc, with pulse length of 5 μsec and pulse recurrence rate of 3720/sec giving a range scale of 25 statute miles. The system went on the air in April, 1936. The receiver recovery to full sensitivity following the transmitted pulse appeared to be instantaneous. Beautifully sharp echoes from aircraft were observed almost at once, and within a few days they appeared all the way to the 25-mile limit of the indicator.

The spectacular success of the experiment was followed by a greatly intensified effort. A primary objective was to reduce the size of the equipment so it could be used on ships. The 28.6-Mc antenna was about 200 ft square. Reduction in directivity of antenna pattern was not desired. A smaller antenna therefore meant higher frequency. On July 22, 1936, a small radar was put in operation on 200 Mc. In that same month the first radar duplexer was successfully tested, also on 200 Mc, enabling both transmitter and receiver to use the same antenna. These two quick developments made it

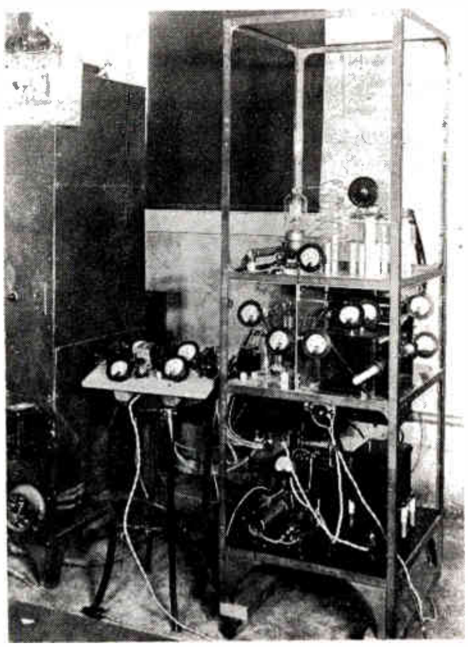


Fig. 2—Original 28-Mc radar transmitter with synchronizing keyer. 17,000 v on exposed wires and condensers on top shelf. April, 1936.

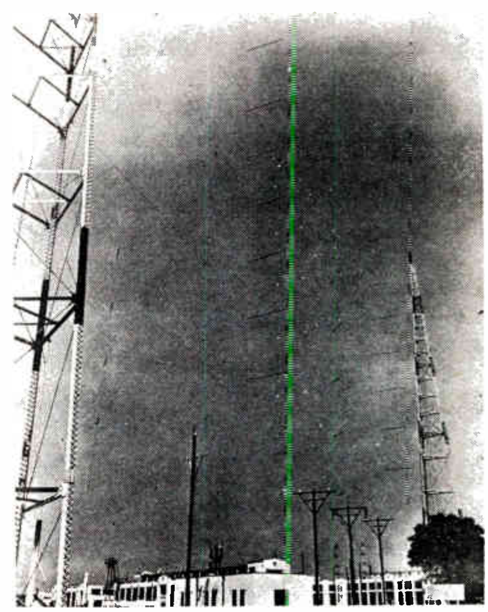


Fig. 3—28-Mc transmitting antenna suspended between 250 ft towers. 1936.

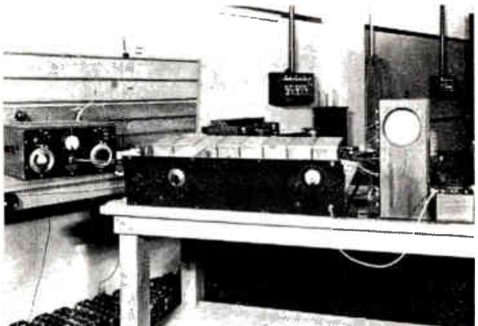


Fig. 4—Original 28-Mc receiver with indicator on test bench showing (left to right) standard signal generator, receiver, indicator, audio output meter. 1936.

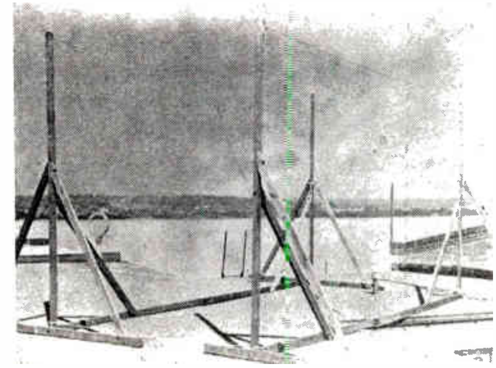


Fig. 5—Original 28-Mc radar receiving antenna; λ/2 dipole with λ/2 reflector. April, 1936.

possible to put radar on a ship for tests at sea. The first seagoing radar tests were made in April, 1937, on the USS *Leary*, an old destroyer of the Atlantic Fleet. The success of these tests led to the development of the model XAF, designed for Naval service at sea. Extensive tests on the USS *New York* in 1939 disclosed operational values beyond all dreams. The XAF was made prototype for the model CXAM, which was in service on 19 ships, the only U. S. Naval radar in service on December 7, 1941. It made an excellent wartime record.

This is a brief outline of the main stream in the early development of radar, resting on a sequence of related events from 1922 to 1941. Up to the summer of 1935 it was a single stream. At that time two other streams started, both remarkably parallel to the main stream. The one in England, sparked by the proposal of Watson-Watt in February, 1935, and conducted under the aegis of the Royal Air Force, was completely independent of the American developments until 1940, at which time the two countries pooled their resources. In the technological trade, America gained the uniquely-British cavity magnetron, and Britain gained the uniquely-American duplexer. The trade was not as one-sided as may have been inferred in some of the postwar literature. The pooled resources formed the technological capital for the newly-formed National Defense Research Committee in the superb development of microwave radar by the Radiation Laboratory of the Massachusetts Institute of Technology. The other stream, in the U. S. Army Signal Corps, sparked by the dynamic leadership of Colonel Roger B. Colton, was independent in part, but received much stimulation, both competitive and cooperative, from the more advanced work of the U. S. Naval Research Laboratory. It is one of the most remarkable coincidences in history that the three streams of radar development, operating more or less independently, issued in three vital but non-overlapping employments, each requiring basically different designs. At the war's beginning the finest mobile ground-based radar came from the U. S. Army Signal Corps, the finest airborne radar came from the Royal Air Force, and the finest Naval radar came from the U. S. Navy.

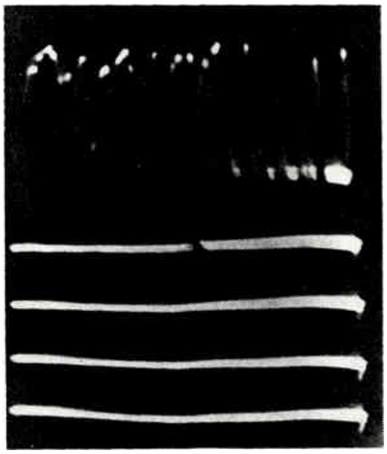


Fig. 6—Echoes of ground clutter (first line, 10 miles) and airplanes (second line, range about 15 miles) with 80-Mc radar installed with duplexer in field house. 5-Line sweep, 10 miles per line—total, 50-mile time base. December, 1936.

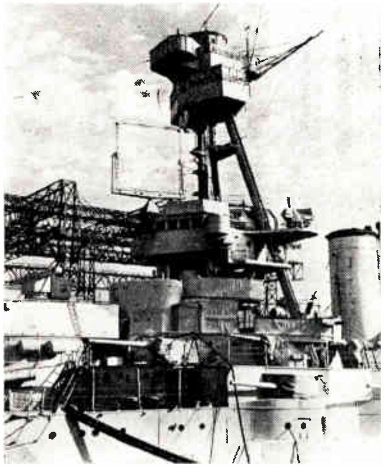


Fig. 7—XAF antenna installed on USS *New York* at the Norfolk Naval Shipyard. December, 1938.

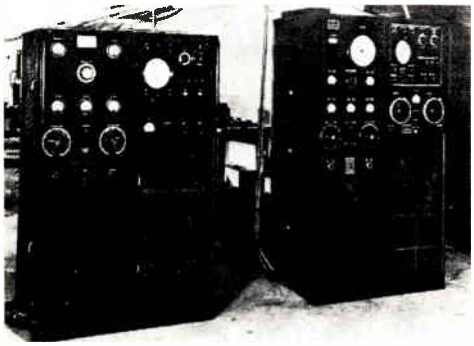


Fig. 8—The XAF radar (left) and the CXAM radar (right). Summer, 1940.

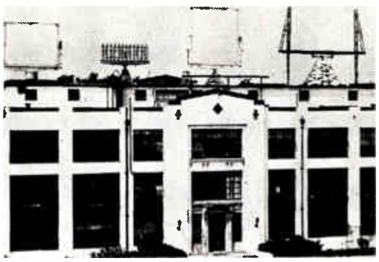


Fig. 9—Penthouse roof, Bldg. 12, showing (right to left), pre-XAF, XAF, 400-mc and CXAM antennas. Pre-XAF is hand driven; all others are motor driven with remote control. Summer, 1940.

The Impact of Missiles and Space on Electronics*

SIMON RAMO†, FELLOW, IRE

Summary—**The ways in which government-sponsored research and development and procurement have influenced electronics in the U.S.A. are discussed. It is concluded that missiles and space have moved to first place in impact on technical advance, industrial organization, and industry-government relations. Major changes have taken place, as a result of these programs, on both technical and administrative fronts, with many unsolved problems remaining in adjusting to the coming expanding missile and space programs.**

In the last two decades, the single most important influence on the electronics profession and industry has been government-sponsored development and production. In more recent years, the fields of guided missiles and space technology have become the most significant areas of military or peacetime government sponsorship. Hardly any organization dealing with the electronics field remains untouched either directly or indirectly by the nation's guided missile and space programs.

This is especially true of everything that is new. The missile and space requirements have pushed the art so rapidly, and technological obsolescence has been so quick, as almost to preclude the attainment of a base of stabilized products. This newness factor has been so dominant, in fact, that a large fraction of the entire electronics industry seems to have been transformed from production activities, for which a good many of the groups were founded, into research and development operations.

This article will call out some of the ways in which the advent of guided missiles and space technology projects have already influenced the electronics field and will indicate something of the likely future trends.

Broad Technological Advances

The missile and space programs have had behind them a sense of urgency previously unknown in the United States in peacetime and an allocation of resources that have made possible parallel attacks on virtually every frontier of electronics research and electronics design engineering. The requirements of missiles and space have been such as to force effort on seemingly every technical front that previously limited what could be done. The research has yielded new answers, and ingenuity has reached new heights. Thus, the requirements, though severe, have been met by the expanding state of the art.

Missiles and space technology have required generation and reception of power over the entire conceivable frequency spectrum of use in communications and control. Power outputs have risen and signal-to-noise ratios have been improved. Decreases have taken place in sizes and weights. Broader bands have been handled. Directivities have been increased. Virtually every determining physical parameter setting specifications for designs has involved higher precisions and higher rates, while all the time the complexity has been growing by leaps and bounds.

With the increase in complexity, accompanied by the new habit of working with devices born yesterday but applied on a crash basis today, reliability problems have assumed a new position. While old lessons were relearned by a wave of new young inventors and designers, an increasing array of devices has nevertheless quickly attained reliabilities that would not have been considered likely a few years back.

The severe requirements of the missile and space program could not have been met by careful or even inspired design if based solely on the physical principles available in the 1940's. There had to be a tremendous effort to achieve a more fundamental knowledge of the physical science underlying electronics engineering requirements. The solid-state field, for example, has expanded largely because of the mass attack on this field made possible directly and indirectly by the accelerated military support for all of electronics. Microwave techniques, given their first great acceleration by the radar work of World War II, have received a vigorous second wave of expansion.

To be sure, an urgent specific requirement cannot and did not trigger fundamental discoveries. But it is usually a long, hard road between such early basic conceptions or experiments and the attainment of practical and useful end results in the numerous forms required to have a major impact on engineering. Parallel and broadly-based attacks on a large scale, once a new approach appears to have promise for solving an important problem, have in the case of the missile and space field provided the accelerated time scale needed.

Electronics research has been expanded from emphasis on the physics that deals with a stream of free electrons in vacuum to the whole range of possible electrical effects in any form of matter right down to the nucleus of atoms—this, as the quest for ability to generate, amplify, and otherwise control electrical effects over a broader range of levels of power and spectrum of wavelengths, and with increasing precision and control, has continued to develop.

* Received by the IRE, August 7, 1961.
† Thompson Ramo Wooldridge, Inc., Los Angeles, Calif.

Similarly, the comparatively narrow role of the electronics engineer in providing a means for communicating intelligence between points, either by wire or unguided electromagnetic waves, has broadened to embrace the full implications of the concepts of control of automated or semiautomated systems. While "communications" still is at the bottom of all that the electronics engineer does in one sense, the notion of what communications is understood to mean has been enlarged. In such examples as operational intercontinental ballistic missile systems, a myriad of men and machines over a wide geographical span are tied together in a complex web of interrelationships and made to work as a harmonious ensemble as a result of the efforts of the electronics engineer.

The match between the urgent needs and the potentialities of recent scientific discoveries has often caused it to be true in missiles and space that new discoveries in physical science are put into almost immediate use, often without time for the digesting of the phenomenon before its application becomes serious. The maser and the broader molecular and atomic resonance phenomena of which it is a part represent an example of this. Electronics engineering always has been closely related to physical research in electronics, and indeed such organizations as the Institute of Radio Engineers always have had a large representation from individuals who think of themselves as "physicists" rather than as "engineers." However, this interrelationship between the engineering and the physics of electronics has become closer.

In electronics engineering today, a larger fraction than ever of the individuals involved were trained as physicists. Most think of themselves as interested in basic underlying physical science rather than in design and application. Actually, of course, in the sense that most of it is so soon applied, most of the group can also be considered as engineers, with the variation only in the degree of formal training and interest in the physical science as against the design of systems using that science. So broadly confused has this interrelationship between science and engineering become that the newspapers tend to refer to all engineers as scientists, even though most of the time the projects to which the public gives its attention are clearly engineering projects, such as lunar vehicles or intercontinental ballistic missiles. Moreover, in almost every university today there is soul searching between the electrical engineering department (now mostly electronics) and the physics department (now at least half electronics, space, solid state, or nuclear resonance) in an effort to understand how they should differ from one another.

Miniaturization and Environmental Engineering

The days are gone when the average electronics engineer can play with circuits alone, disregarding many other aspects of over-all design. Electronics engineering, mechanical engineering, heat transfer engineering, and materials engineering have come together more strongly and have required a more unified design effort as a result of the impact of guided missiles and space. This has often been manifested in the need for major decreases in size and weight of electronic subsystems and components.

While in most guided missile and space vehicle systems the electronics—the "brain" and "nervous system" of the whole beast—constitutes a very small fraction of the total size and weight, the burden for miniaturization nevertheless falls heavily on the electronics engineer. There are several reasons for this.

One is that, in general, the electronics equipment of all airborne or spaceborne systems must go to the "end" while, in contrast, a good deal of the initial weight, notably fuel and oftentimes major portions of the propulsion system, will have been exhausted and purposely dropped along the way to lighten the load. The guidance and control must be present to the last moment where any unnecessary weight exacts a severe penalty.

Another reason for the emphasis on miniaturization of electronic equipment is that it *can* be miniaturized, and any weight that can be saved must be saved. If the art simply will not yield a smaller propulsion system but will yield a smaller guidance and control system, then at least the latter miniaturizing should be done.

In space missions as well as in intercontinental ballistic missiles, there is a fairly direct relationship between the size of the initial booster thrust, which is today a limiting factor, and the size of the final payload, much of which involves electronics gear.

The complex and numerous detailed and sophisticated tasks performed by electronics equipment in guided missiles and space vehicles could not have been seriously contemplated without the major subminiaturization programs and the beneficial results so attained; that is, by the means and components available ten years ago, the equipments that would have been required on board for a typical complex missile or space mission realizable today would have weighed so much more than could have been lifted by a realizable propulsion plant that the whole project could not have been seriously contemplated. The reductions in size and weight attained by the advances in electronics have made possible the project of today.

Computers and Simulators

A large fraction of today's electronics industry is interested in computers, including computers for commercial uses of all kinds as well as for military applications. If one looks carefully at the history of the development of a large-scale computer during the last ten or fifteen years, the impetus given to these developments by military sponsorship becomes evident.

Every missile and space system uses computers. Also, with a large fraction of the industry being in defense, and a large fraction of the defense industry using computers for engineering designs, every missile and space program supports large-scale computer installations. The making of engineering calculations on them and the programming of the computers, in addition to the use of computers directly in the weapons system as ground or vehicle equipment, constitute a large fraction of the total effort.

It would have been out of the question for us to have attained so quickly an intercontinental ballistic missile of proven design such as the Atlas had we not been able to extend the capabilities of the engineering design team by large-scale computers. Actually, only a few dozen full-scale flight tests were required to completely prove out the design in every way. This is, however, because the equivalent of tens of thousands of flights were made by computers in the laboratory. This simulated flight-testing selected, from the infinity of possibilities, a reasonably optimum ensemble of design parameters in a short time. The almost trial-and-error approach that would otherwise have been needed would have had us flying missiles in test after test for many, many years before a design could have been settled on and confirmed.

Of special importance is the relationship between the techniques involved in the handling of information in the computer elements and the process of communication. These two separate fields of electronics engineering have come together in theories that start from basic new concepts of information itself, its quantitative measure, storage, transfer, and processing. To a considerable extent, it is because of the missile effort that the electronics industry is heavily concerned not only with communications, but with every facet of information, including its derivation and its use, for control of a system of equipment and men.

Systems Engineering

Systems engineering—the design of the entire complex of men and machines as against the detailed design of its parts, and the specifying of the individual parts of the over-all so that in combination they will yield the performance desired—is an old facet of engineering. However, in recent years, the term "systems engineering" has taken on broader meaning and greater usage. This is in part because a larger fraction of the engineering body is concerned with the design of the whole as distinct from design of the components, this in turn, as the projects become larger and more complex, with greater interactions.

This expansion of systems engineering has been especially important in guided missile and space phenomena. Moreover, electronics engineers have had a larger hand in systems engineering than have engineers coming from other specialized branches of the whole. This, of course, is natural, because the electronics engineer deals with the intelligence flow throughout the system that ties it together and that has more of a unifying effect on the working of the whole than does, let us say, the propulsion or aerodynamic aspects of a guided missile, or the structural aspects of a space probe. In a sense, the electronic aspects of a project cannot be so easily compartmentalized; electronics is everywhere.

When a large number of people and machines are tied together, we find major problems of over-all stability, unwanted mode possibilities, feedback loops (unavoidable as well as intended), and the accumulation of errors that affect the operation of the entire system. These are, of course, very close to the traditional fields of the electronics engineer in communications, with error accumulation problems very similar to the customary signal-to-noise ratio problems, and with the stability and unwanted mode problems similar to the stability and oscillation problems of typical electronic feedback circuits. Thus, while the over-all problem of systems engineering of a guided missile or a space system involves all fields in addition to electronics, the electronics engineer has a larger fraction of the background and tools needed for the integration engineering.

The electronics engineering fraternity also has a larger fraction of the broader trained physicists who are accustomed to dealing with more difficult conceptual matters and the wider range of science fundamentals applicable in common to all branches of engineering. When one attempts to predict and understand the performance of a large complex system, that is, in effect to try to write its laws and then set up experiments to understand and confirm laws of behavior, the problem is not unlike that of tackling a new branch of nature to bring the level of understanding up to the quantitative point. This, in effect, is the work of the physical scientist.

The analysis of complex ensembles of equipment and people often requires a major emphasis on the use of computers and simulators. Here, again, the electronics engineer is more prepared. The electronics engineer, always closer to physics and mathematics than those in most other areas of engineering, but especially more familiar with the use of electronic tools as represented by the electronic computer, was able to take up the difficult systems engineering problem more rapidly than his fellows.

Large-scale systems analyses often have to be handled on a statistical basis. Random phenomena and statistical techniques are more familiar to the electronics engineer, from his basic noise and modulation experience, and to the physicist, the adopted member of the electronics engineering team, from his basic training in modern physical theory.

For these various reasons, then, the electronics engineer has become the most influential systems engineer, on the average, in the guided missile and space programs. This, in turn, has affected the nature of electronics engineering, the training for it and the way it is practiced, and has influenced the nature of the industry as well as the professional groups. Partially, that is, systems engineering and electronics engineering seem almost to have become synonymous, with the well-trained electronics engineer and the broadly-based electronics engineering organization being one that is capable of handling integration. By contrast, other specialized engineering groups are less inclined to branch out and provide the competence to handle well the systems engineering job.

Industry Organization and Industry-Government Relationships

The electronics industry has become greatly dependent now on government support, as have the universities. It is not only that government support furnishes financial backing for research and development and a large fraction of potential business. The reason is probably more because the fraction of business that is government supported involves the very newest techniques. A commercial venture has to concern itself with the fact that its competitors, if engaged in similar military projects, may develop positions, know-how, and specific products as participants. Competing with this large sponsorship by private funds begins to look less promising than trying to become a part of the family so as to have the same benefits as the competitor.

The large government sponsorship in missiles and space has added a new factor of more rapid technological obsolescence. That is, a private organization engaged in some aspect of electronics may find the field moving past its product line because the whole technology is changing so rapidly, spear-headed especially by the government programs.

The industry has been shifting from a production emphasis, with a relatively small investment in pure research and a good share of earnings plowed back only in new product developments, to dependence on a large research and development laboratory mainly government-supported. This has required major industrial adjustments not yet all complete. Amongst other loose, unfinished ends are needed new concepts in patent protection and, generally, the problem of incentives for private investment.

The total production requirements of missiles and space can be easily handled by the capacity of today's American electronics industry. Also, while there is a shortage of true geniuses, and technological breakthroughs are not everyday affairs, useful research and development (in the sense of making substantive, desirable, and nearly satisfactory advances for the resources allocated) can apparently be performed by a large fraction of organizations involved in electronics today. With supply greater than demand, the entire electronics industry finds itself with more intense competition and the distinct possibility of a poor return on investment. These effects, taken together, tend to reduce private capital available for risk and create in turn a greater interest in assuring government sponsorship to back up private capital risked for new ventures.

Meanwhile, as the technological front widens and deepens and as government control and direction of electronics engineering grow, the government has an increasing problem of planning, directing, and assessing the programs. It becomes increasingly difficult for government to decide how to relate itself to industry and how, as a matter of fact, to organize individual projects and to allocate portions of them to individual competitive industrial organizations.

Since the organization of electronics engineering efforts has a great effect on the quality and nature of that effort, this matter of government-industry relations, as especially manifested by the missile and space programs of the nation, is probably of importance second to none in studying the impact of missiles and space on electronics.

Electronics and the Space and Missile Race

The public in general, understands that the U.S.A., after some years of programs that were much too modest, embarked on a huge crash program to create the intercontinental ballistic missile in record time, and that it succeeded in doing so. Also, it understands that despite this effort we may, in some aspects of ballistic missiles, still be behind the U.S.S.R., whose large program started much earlier. Furthermore, we are behind in some aspects of space technology, notably in size of thrust available for take-off.

It is a special credit to the electronics engineers and to the electronics industry of the nation that in the electronics aspects—guidance, control, communications, accuracy, versatility, and over-all performance—there is no evidence that we are behind the Russians. Complex range instrumentation and precision guidance have been demonstrated on a wide range of guided missiles in the United States, and this not only in the intercontinental ballistic missile field. By having precision guidance, we have been able to count on smaller misses, hence could get by with smaller warheads, and thus could tolerate smaller thrusts.

Turning to space, we note that the American Pioneer V's record for communicating back valuable and high quantity data on many physical phenomena measured and recorded simultaneously at a distance of over twenty million miles still remains unsurpassed today. The best available evidence suggests that U.S.A.'s space flights have brought back much more data than have the Russians', and this can be traced directly to our relative excellence in all aspects of electronics in space.

The Future

It would appear that missiles and space will continue for many years to hold a commanding position in government programs and that government programs, in turn, will continue to hold a dominant position in their influence on the electronics profession and industry.

Early in the guided missile efforts, it was customary to give first priority to getting a substantial mass up to high speed on a trajectory that was more or less the intended one. Soon it was seen, however, that the electronics of the operation was the true, high-priority item because it was the most serious factor in over-all performance and in efficacy of the entire mission. Similarly, it may be assumed that, within a few years, the matter of thrust in space will have been taken care of by proper development programs and space missions will clearly be dependent more on continued development of the payload. The sophisticated electronic passenger will be more important than the vehicle that takes it there, and putting equipment into orbit or out into interplanetary space will be easier than insuring that through this equipment the vehicle will serve a useful purpose.

In particular, we may expect that space technology will have a greater impact on communications than any other new developments in the field of communications. International communications systems, depending upon equipment in space, will expand the flow of information for commercial as well as military purposes.

The desired future missions will, if anything, continue to demand a parade of increasingly advanced electronics developments. Need for smarter circuits with higher reliability, probably in part due to built-in redundancy, will most likely maintain the pressure for reduction in size and weight of electronics apparatus. While there are aspects other than that of missiles and space that influence the effort towards further miniaturization, the requirements from these fields will probably be sufficient to hold the attention of a large body of engineers and physical scientists for years to come.

Electronics engineering will continue to be in close partnership with new physical science, and the systems engineering aspects will continue to press electronics engineering towards over-all system considerations.

The more exotic, newer possibilities in space technology propulsion may well lie in electronic propulsion techniques and in the generation and control of power by direct conversion from heat and chemical energy to electrical energy. These areas may become of increasing interest to electronics engineers, who again will come closest to having the background to move rapidly into these new fields. Thus, we may in the future find that electronics engineering will expand through total information processing, control, and systems engineering to power generation and conversion, with missiles and space as the spearhead.

Very likely space electronics will take over from missile electronics, although these two will always remain related, as the number one sponsorship source for new electronics developments. Because space appears to be a government area as against a private area for all practical purposes, the present trend towards dependence of the electronics industry and the profession on government projects will probably increase rather than decrease.

A Summary of Military Communication in the United States—1860 to 1962*

J. D. O'Connell†, fellow, ire, A. L. Pachynski‡, senior member, ire, and L. S. Howeth||

Summary—**Military progress in communications in the United States really started back in 1860, and both before and since the birth of the IRE, it has been inextricably a part of the development of the communications industry, national communications policy, the IRE and scientific research and development in this art. The U. S. Navy played a particularly important part in the development of maritime and coastal radio systems and in the development of early radio wireless policy. In both major wars and increasing rapidly since World War II, the requirements of the military services for communications to keep pace with successive revolutions in speed, scope, range and nuclear power have increased exponentially, and beyond the capabilities of conventional or commercial systems. The vital significance of adequate communications and the part they play in the success of military operations and weapons systems has been consistently undervalued in the past. The cost of undervaluation and unreadiness in communications has been mounting rapidly in step with orders of magnitude increases in nuclear power and speed of delivery.**

* Received by the IRE, January 29, 1962.
† Lieutenant General, U. S. Army (Ret.); General Telephone and Electronics Laboratories, Menlo Park, Calif.
‡ Major General, U. S. Air Force (Ret.); Lenkurt Electric Company, San Carlos, Calif.
|| Captain, U. S. Navy (Ret.), Box 381, Morris St., Oxford, Md.

THIS YEAR marks half a century of phenomenal progress and growth of the Institute of Radio Engineers. Inextricably involved in this half century of progress of the Institute, of industry and of science have been the Military Services of the United States.

Any attempt to outline the course of U. S. military communications should include another half century of progress because just before and during our Civil War an idea was born which would grow—the first integrated concept of an organized military communication organization.

The record of his achievements and thoughts for the future leaves no doubt that Assistant Surgeon, Dr. Albert J. Myer (later Major Myer, first Army Chief Signal Officer; later Brigadier General Myer, promoter of our National Weather Service and International Weather Reporting) would be at home in the military communication world of today. Aggressively seizing upon every available technical capability of his time, he stimulated the development of the Beardslee portable electromagnetic telegraph, mobile telegraph wagons, observation balloons for combat surveillance, and an over-all communication system with flags and torches. He organized, innovated, trained, administered, promoted and, most important of all, lit a flame of military interest in the art of communication and endowed an organization with the spirit to keep this flame burning through the years.

During this war period the U. S. Navy was also making strides in communication: the two services in an early effort of joint service coordination adopted for the first time the same system of flag and torch signaling and, for a time, instruction in this common system was included in the course at both Annapolis and West Point.

The end of the Civil War brought military reaction and decimation of budgets and personnel—elimination of the concepts of research and study of improved services—"an army should be prepared to fight and not play around in a laboratory." To do both was not considered necessary or practical. But despite apathy and resistance on the part of both military and civilian leadership—typical of the national attitude when not confronted by crises, there was, none the less, continued effort and accomplishment—the linkage of the western army posts with telegraph communication; the initiation of weather reporting from army stations; the formation of the weather service and international weather reporting; hazardous polar exploration as part of the International Polar Year, 1882–1883—significant contributions from military communication beginnings to the progress of a growing nation.

In world communications, the British with great vision and initiative pushed the extension of submarine cable service to most major world ports, thereby linking the British Empire and facilitating their world trade. The U. S. Navy was quick to use this service in keeping naval units informed and coordinated. But in the wars that followed, the U. S. Navy was the first to learn the problems and penalties of dependence on a communication system operated under foreign control.

On inauguration day, March 4, 1897, during the traditional drive from the Capitol to the White House, Mr. Cleveland was prophetic to McKinley: "I am deeply sorry, Mr. President, to pass on to you a war with Spain. It will come within two years. Nothing can stop it."[1]

Sparked by the President's interest, the two military services had an opportunity for advance planning and although still unready in communications when the war came, performed prodigies of recruitment, organization and operations. Receiving orders on April 9, 1898, Captain C. F. Goodrich, U. S. Navy, within 30 days not only organized but had in operation a Coast Signal System of 230 stations connected to telephone and telegraph systems and equipped with telescopes and binoculars.

Thus began a span of years of war emergency in U. S. military communications which educated the Services in broader requirments both for communication to theaters of war beyond the continental limits of the United States and for the establishment of communication within those theaters. Again the vital need was recognized for better communication for joint Navy-Army operations in Cuba, in the Philippines, and in China. Telegraph and telephone sysems were installed by the Army—3500 miles in Cuba, 6434 miles in the Philippines where the Signal Corps entered a new field of submarine cable laying and operation (1326 miles). Both systems were turned over later to commercial use. During this same period, on May 26, 1900, Congress acted to bring the isolated settlements in Alaska into electrical communication with the United States by charging the Army Signal Corps with construction and operation of such a system. Construction started at Valdez within 3 months and was completed by October, 1904, to include 2650 miles of submarine cable and wireless systems and 1400 miles of land lines. Growth and modernization have been continuous to this day, with recent great achievement by the U. S. Air Force and the communications industry in extending communication to the rugged extremities of this great northern state to link the northern radar outposts.

These were important beginnings—important as a backdrop of understanding for what was to come and, doubtless, essential to recognition in the military services of the communication requirements of the future. This recognition, however, was never complete or adequate and resistance to change has imposed an endless succession of obstacles to be overcome.

Coming on the scene with great publicity and fanfare during this period was the fascinating but mysterious radio wireless. It appears certain that the interest, support, publicity and promotion given to this infant by the military services advanced in a major way the application of radio to both seagoing and point-to-point commercial radio-telegraph service.

Both the Navy and the Army had their Service inventors. A brilliant Navy pioneer was Lieutenant Brad-

[1] D. S. Muzzey, "A History of Our Country."

ley A. Fiske who early (1884?) recognized the naval need for a signal means independent of fog, darkness, daylight or distance. His studies led him to Prof. Dolbear's theories, to Dolbear himself in 1888, and thereafter, to devise a system which did succeed in communication from ship to ship over short distances without wires. Had he the means and the time to experiment consistently, the story of naval radio might have been advanced several years. In the Army Signal Corps experimenters were many. Picking up the trail of research blazed between 1892 and 1896 by Preece, Branley, Sir Oliver Lodge, Popoff and Marconi, the experimentation was noted in the New York *Herald*, December 4, 1899: "It is said there can be no infringement of patents as, with the exception of a few parts, the wireless telegraphy system is common property and the Signal Corps of the Army is now experimenting with instruments devised by its own officers." By this time the Signal Corps had assembled equipments and established radio communication circuits between Fire Island and the Fire Island Lightship and between Governors Island, N. Y., and Fort Hamilton, N. Y.[2]

During the turbulent next two decades of radio's growth Army pioneering efforts were continuous and significant but the major military influence during this period came from the U. S. Navy.

From the very outset the U. S. Navy adopted a deliberate and studied approach toward this new invention. With interest and expedition, but refusing to be swayed by newspaper criticism and pressure or by the rival claims of early proponents, in July, 1899, the Navy took note of the Royal Navy's tests of wireless; negotiated in August and September, 1899, with the Marconi Company for tests to be conducted aboard U. S. Naval vessels; observed Marconi's demonstration during the international yacht races in U. S. waters September–October, 1899; produced a set of requirements for service test; and in October–November, 1899, with Marconi presiding, conducted tests off the port of New York. The tests were generally successful but the Marconi Company's uncompromising insistence on a high-priced year-to-year lease to achieve its plan for a monopoly, caused both the Navy and the War Department to reject Marconi's proposal, to seek other sources of supply, and to encourage competitive tests of equipment developed by newly-emerging and highly-competitive groups of U. S. inventors. This decision was in fact a critical one because it gave strong support and encouragement to these small adventurous, imaginative U. S. wireless pioneers who saw visions of future greatness many years before major economic feasibility in the commercial communication field was attained. The Navy's careful study continued; it included the purchase from four European and two U. S. companies of experimental equipment which was ready for test by the end of 1902.

The manifestly sensible effort to explore the whole new field and select the best on a competitive basis immediately placed the Navy in the middle of a hot welter of negotiation, conflicting patent and performance claims, accusations of unfairness to U. S. firms and press criticism. Attack from within the Navy during 1904 on the dangers of relying on wireless in war raised the questions of enemy intercept and interference which were to remain and grow into the major research and operational fields of radio intelligence and radio countermeasures.

Meanwhile, in April, 1902, Admiral Bradford advanced proposals stating the importance of government regulation or control of radio stations established within the territorial limits of the United States. At the first International Radio Telegraphic Conference assembled in Berlin on August 4, 1903, the American delegates were Brig. Gen. A. W. Greely, CSO, U. S. Army; Commander F. M. Barber, U. S. Navy (Ret.); and Mr. John J. Waterbury, Department of Labor and Commerce.[3]

Historically significant was the action of the President in calling into session the first Inter-Departmental Board of Wireless Telegraphy, generally known as the Roosevelt Board, which met July 6, 1904. Despite previous interdepartmental competition, noncooperation and noncoordination, this Board unanimously recognized the Navy's paramount interest in this new field and recommended that

> The Navy be designated to provide efficient coastwise radio communication for the United States Government and when not in competition with commercial stations to receive and transmit all radio messages to and from ships at sea.
>
> The Army be authorized to erect such stations as deemed necessary provided they do not interfere with the coastwise radio system. . . .
>
> Legislation to prevent the control of radio telegraphy by monopolies or trusts should place supervision in the Department of Commerce and Labor.[4]

The story of the rapid development of naval and marine radio under Navy leadership is one of great interest to all.[5] Most of the contentions and problems of later times made their appearance promptly. Through this early period of strife and progress the Navy steered with good judgment, vision and achievement, to create the finest naval radio communication system in service

[2] G. H. Clark, "Radio in War and Peace," unpublished ms.
Note: DeForest (Autobiography), "Father of Radio," Wilcox and Follett Co., Chicago, Ill. p. 150; 1950. DeForest states his belief that the opening of radio telegraph service August 7, 1903, by the Signal Corps between Ft. Safety (near Nome, Alaska) and Ft. Michael, 107 miles south across Norton Sound as part of the Alaska Communications system was the first wireless system in the world operated regularly as part of a telegraph system handling commercial traffic.

[3] This meeting resulted from the 1902 incident when Prince Henry of Prussia, brother of the Kaiser, tried to radio from the Deutschland to President Roosevelt a message of thanks for the hospitality just extended. The Marconi Co. station at Nantucket refused to accept the message because the Deutschland carried competitor's equipment.

[4] "History of Electronics in the U. S. Navy."

[5] It is to be published soon by the Department of the Navy.

during World War I. While accomplishing this primary mission, naval officers left a record of service to the radio industry, to maritime communications, and to the development of national and international radio communication policy.

Many Navy and Army officers representing our government at international radio conferences contributed importantly to the establishment of policies in international radio communication. For example, at the Second International Radio Telegraphic Conference, Berlin, October 3, 1906, the experience of both the U. S. Navy and Army was used to refute the claims and policies of the Marconi Company and others which held that only by restricting intercommunication to message exchanges between equipments of identical manufacture could satisfactory and reliable communications be conducted. The U. S. Delegation seeking to eliminate this barrier to the extension of maritime communication made a declaration which is worth quoting in part

> It has been fully demonstrated by the Government of the United States of America through experiments carried out in climates of every kind, that the different systems of radio telegraphy can be effectively used with the other. The United States Navy is actually using at present eight different systems in its coastal stations and its stations aboard ship; and during the three years in which it has been making these experiments it has reason to be entirely satisfied with the results obtained.

As a result of the determined crusade of the U. S. Delegation and despite strong opposition, the conference finally adopted a provision which stated that every ship station "shall be bound to intercommunicate with every other shipboard station without distinction as to the radio telegraphic system adopted respectively by these stations." One might think that this removed shackles not imposed by nature and freed maritime radio for rapid expansion to serve world needs. But an extended period of contention ensued and the Senate failed to ratify the convention for several years. At the Third International Radio Telegraphic Conference in London, June 4, 1912, the military services were again ably represented. The delegation was headed by Admiral Edwards, USN, and membership included Dr. Louis Austin, Head of the U. S. Naval Radio Laboratory and Majors George O. Squier, Edgar Russell, and Charles Saltzman, U. S. Army Signal Corps.

During this early period of the century both Army and Navy recognized the need for a regularly organized research, development and engineering function. The Army established a Development Division of the Chief Signal Officer's Office in 1902 under Captain E. Russell and the U. S. Navy established, in 1908, the U. S. Navy Radio Laboratory with facilities and help of the U. S. Bureau of Standards and under the direction of a noted Bureau of Standards physicist, Dr. L. W. Austin. This organization and its greatly expanded successors have through the years made achievements well recognized throughout the world. The Navy had started on the course of developing its own detailed specifications for communication equipment and later the Radio Test Shop at the Washington Navy Yard tested commercial equipment and designed naval receiving equipment. Later, 1923, these activities became part of the U. S. Naval Research Laboratory at Anacostia under Dr. A. Hoyt Taylor, whose outstanding leadership of naval research covered more than two decades.

During the period 1912–1945 the naval officer who most continuously and brilliantly contributed to the growth, expanding influence and effectiveness of U. S. naval radio was S. C. Hooper. Even as a midshipman and before his assignment as Fleet Radio Officer in 1912 he had shown great interest and ability in the new field. The influence he attained as a relatively junior officer in the policy councils of the Navy Department was remarkable to say the least and for many years he represented this country's government radio interests in virtually all international conferences. From 1915 to 1928 he was the guiding spirit in developing naval radio from an ambitious experiment to the efficient and reliable system which it became. He has come to be regarded as the father of naval radio and was credited by Owen D. Young as being the motivating force which led to the formation of the Radio Corporation of America.

Just prior to and during World War I there was rapid and manifold expansion of communication activities in both services. The U. S. Navy expanded and perfected its Naval Radio Communication Service and was charged with management of U. S. Government Radio Communications. This included taking over such commercial facilities as were required for the U. S. Naval Communication System. The capacity and reliability of radio communication to Europe was increased by vigorous broad-front development of high-power, long-range systems and the heavy load of war traffic was effectively handled by transatlantic cable and the Naval Radio Service.

Army communication development during World War I required prodigious efforts. The Army Signal Corps had to expand its field signal units about fifty times in personnel strength—build an infant air arm into a major corps—and specify and procure signal equipment for the largest U. S. Army in history.

The primary Army communication problem was to support the AEF—first with the primary means, wire communication, and secondly, with the immature early tactical radio systems.

The Army's Chief Signal Officer sought and received massive support from AT&T, its operating companies, Western Electric, Western Union and many other U. S. companies. Affiliated units from telephone companies and Western Union bulked large in the ranks of units which constructed over 5000 line route miles of pole lines in France.

Tactical radio was pushed with great vigor but never became a major communication factor during World War I. Technical talent in the Army from civilian life was used to man laboratories which did notable work at Fort Monmouth, N. J., and in Paris where Major Edwin H. Armstrong, Fred Lack, William McDonald, Ralph Bowen and many other later famous men made significant contributions to radio—Armstrong's superheterodyne circuit for one.

The years between 1920–1937 became another period of general apathy toward the services and in communications research and development, the result was a loss of momentum in both services. Both were able to maintain research and development laboratories which even grew a little from the point of lowest ebb, but Naval R&D was so reduced around 1930 that the disestablishment of the Naval Research Laboratory was seriously considered. In the midst of and despite these struggles to remain effective both services developed the beginnings of radar in their communication laboratories.

The "cut to the core" budget reductions which followed the economic crash of 1929 almost caused the elimination of research and development in the services. But the Navy's major support heretofore of thermionic-tube development and other projects which had advanced the state of the art was greatly reduced and the Army had virtually no money which could be so employed. Frequently, service research and development was the object of charity from industrial firms who were better able to push forward the state of the art with larger corporate earnings derived from the rapidly expanding broadcast and communication markets. These firms hoped to recover these give aways by getting military production contracts, but few actually did so.

With the establishment of the Federal Radio Commission in 1927—first headed by a retired naval officer and later made permanent and renamed the Federal Communication Commission in 1934, the position of the Navy as government spokesman with the industry terminated after two decades of notable achievement.

Despite the paucity of resources available for the advancement of specialized techniques and the national strife over the role of military air power which characterized the two decades between the World Wars, the Army Air Corps achieved some measure of progress in communications. In 1934 General "Hap" Arnold, then a Lieutenant Colonel, led a flight of ten Martin B-10 bombers from Washington, D. C., to Alaska and back again. Careful advance preparations had been made along the route and the communications essential to navigation and the transmission of weather information were provided. However, a later training mission flight led by Arnold encountered a series of misadventures and inconveniences due to bad weather and unreliable communications. This experience gave impetus to the establishment in November, 1938, of the Army Airways Communications Service, which evolved ultimately into the present globe-girdling Air Force Communications Service.

Meantime the Air Corps had embarked on a program of spanning the continental U. S. with military airways connecting Air Corps fields and installing, in cooperation with the Civil Aeronautics Authority, the required radio ranges and the ground communications facilities required both for exchange of information on aircraft movements and for the air-ground communications. The establishment of the Signal Corps Aircraft Laboratory at Wright Field, Ohio, led to the joint development with the Air Corps of radio communications and navigational equipment designed specifically for military aircraft operations. Under the leadership of Air Corps officers, pioneer work in radio direction finding and instrument landing systems was accomplished.

Except for the radio industry's developments of side band and multiplex equipment and the adaptation of the Baudot code to radio transmissions, the services approached World War II with the major portion of their radio equipment rapidly approaching obsolescence.

Tactical communication techniques and equipment which were standard for the U. S. Army in 1937 were almost without exception inadequate to meet the new requirements of mobility, range and reliability for armor, motorized artillery, infantry and air-ground communication. Nor was development under way which could provide communications comparable to that of the modern, mobile Germany army. Tactical radio sustems of the time were not, in fact, serviceable or reliable communication systems. As late as 1941, wire-system development engineers competing for money against radio projects had a slogan: "Millions for radio but not one cent for communications." Fortunately, but almost too late, the situation was recognized and although beset by equally or more compelling requirements for radar and aircraft communication, a successful fight was waged for the funds to start a new program and a race with time. The technical advances which qualified tactical radio to add deservedly the word "communication" were the wholesale use of crystal control and the adoption of frequency modulation. These fateful decisions were made by Colonel, later Major General, R. B. Colton. Major E. H. Armstrong again threw his personal efforts into the effort to get superior radio communications for the Army, donating his time, his equipment and his patents voluntarily and without compensation. Over 100 U. S. manufacturing companies performed outstandingly in the development of this new equipment for the Army.

In preparing to start the air war against Germany, early dependence was placed on the communications systems of the United Kingdom, developed for the Royal Air Force to meet latest air war concepts. The plan was for U. S. air units to be based in the United Kingdom. Our aircraft were equipped initially with British supplied VHF radio equipment to meet the requirements for operation in a common ground environment. This required U. S. fighters to adapt to the British methods of air defense and the equipment was later manufactured in the United States. These initial adaptations had an impact on our world-wide forces to the extent that at the end of World War II all allied combat aircraft and associated ground units had been converted to VHF for air-to-air and air-ground communications in order to provide compatibility in air operations. The successful adaptation of the VHF to military air operations resulted subsequently in its adoption for commercial air operations.

To meet the requirements for communications electronics equipment in the U. S. Armed Forces in terms of both quantity and variety, the combined scientific, industrial, and military resources of the nation operating in concert were utilized. In the process, miracles were accomplished in the compression of lead times involving the development, manufacture and delivery of new equipments to combat units. Fateful decisions were made to press for delivery of new types of equipment utilizing advanced techniques rather than relying on obsolescent prewar types.

Some understanding may be obtained of the procurement task facing Army, Navy and industry in communication and electronics alone, by considering that March, 1942, estimates showed that 4.3 billion dollars would soon be available to the military services for radio and radar contract awards to an industry whose total annual output was 0.25 billion. The Army Signal Corps preliminary Fiscal Year 1941 budget estimate made in 1939 had been 9 million dollars. The Fiscal Year 1941 funds made available were 256 million. In May, 1941, the Fiscal Year 1942 Signal Corps budget submitted to Congress was 103 million dollars. There were no less than 30 different and enlarged budgets submitted during the year 1941. By the end of Fiscal Year 1942 over 2.6 billion dollars worth of contracts had been placed with U. S. and Canadian firms by the Army Signal Corps alone. Money alone, however, could not produce military equipment on command. Yet as early as June, 1942, the Army Signal Corps was accepting every two weeks as much equipment as it had procured during World War I. The Navy, between November 1, 1942 and September 1, 1945, accepted deliveries of electronics equipment amounting to 2.54 billion dollars.

Radio and wire system development continued during the war. From the Army's Fort Monmouth Laboratories came the development of FM radio relay producing a communications breakthrough which integrated radio into wire systems with equal transmission quality and dependability. From Western Electric Company came a high-speed spiral four-cable and carrier system, and in 1945 the first U. S. time division multiplexed microwave radio relay system which saw service in Europe in the final campaigns.

It is difficult to eliminate completely from this brief summation some mention of the vast field of electronics which developed slowly, then explosively, during the years 1930–1945, because this vast mushroom of electronically-controlled weapons systems grew out of the Service, industry, and university laboratory and engineering groups engaged in advancing the communications art. Radar-groundbased-shipborne-airborne-early warning-gunlaying; sonar; guidance for bombs-torpedos-missiles-drones; proximity V.T. fuse systems; electronic countermeasures—all started from communication beginnings and all were important and some were decisive factors in winning World War II.[6] The incredible progress in all these fields concurrently was made possible only by the establishment of the National Defense Research Committee, its unprecedented mobilization of scientists, the establishment of organizations such as the Radiation Laboratories and the industry-science-military team which resulted.[7]

The urgent requirements for hitherto undreamed of varieties and numbers of highly-skilled specialists created a rapidly increasing crescendo of training demands which swamped the school facilities of all services. Had the multiplying demands resulted solely from the unexpected expansions in communications, they would have seemed impossible. But requirements for training specialists for the needs of radar and other electronic specialties multiplied even faster and fell upon the same communication training establishments. These were required to split and split again and again to establish cadres for training schools throughout the country, some in newly activated military camps and stations, others in universities, in civilian training schools and in schools set up by our British allies within their radar and training establishments.

Of major consequence in filling these insatiable needs for radio- and electronically-trained men was the response of the radio amateurs of the nation. As in World War I, and in even larger numbers, they aided the training operation, and maintenance of electronic equipment in all the services.

The achievement of this literally incredible task is another story which deserves adequate book-length treatment. It resulted after the war in the release to civilian life of the largest influx of technically-trained men ever available in so short a time.

For example, during World War II, the Army Signal Corps alone supervised the training of 432,000 officers and enlisted men. In all, 34,000 officers graduated from 50 courses, most of which had high technical content. If one adds the specialists trained by the Navy, the Air Corps and other branches of the Army, the number of men and women receiving communications or electronics training could be estimated at over a million. Their rapid absorption by a growing electronics industry was a major factor in its postwar growth.

From the brief sidelong glance above at the wartime birth of the electronics industry, a return to communications requires that the growth of military global

[6] J. B. Dow, "Navy radio and electronics during World War II," Proc. IRE, vol. 34, pp. 284–287; May, 1946.

[7] J. P. Baxter, III, "Scientists Against Time," 1945. Other laboratories of outstanding importance and achievement were formed at Harvard, Columbia, Illinois, Chicago, Brooklyn Polytechnic, Stanford and many other universities.

systems be noted as concurrent with the beginnings of U. S. responsibility, first, for war on a global scale and later as the primary trustee of free-world security and peace.

The establishment of U. S. bases, theaters of operation, Army Airways Ferry and Air Transport routes literally all over the world and into combat zones created communication requirements on a scale not hitherto experienced or foreseen. The Army Airways Communications System and the Army Command and Administrative Network owe their growth to these requirements.

The station names stretched out in all directions, with great speed in the perspective of normal actions but with painful slowness from the standpoint of the need.

From Honolulu—the names included New Caledonia, Melbourne, Brisbane, Port Moresby, Noumea, Espiritu Santo, Hollandia, Wake, the Philippines, Okinawa, Japan and Korea. As the U. S. Navy made it possible by control of ocean areas, MacArthur set a record for island hopping and his signal officer, Major General S. B. Akin, developed new concepts for massive use of high-frequency radio to substitute for nonexistent wire communication. The AACS and ACAN followed with long-distance radio circuits.

Over the South Atlantic Ferry Route and beyond, it was Miami, Puerto Rico, Georgetown, Belem, Natal, Ascension, Accra, Dakar, Marrakech, Casablanca, Algiers, Cairo, Asmara, Abaden, Karachi to name only a few.

Over the North Atlantic Ferry Route names like Presque Isle; Goose Bay, Labrador; Gander, Newfoundland; Greenland, Reykjavik, Iceland and Prestwick, Scotland, were most familiar.

Systems development passed rapidly from wide use of hand-keyed and partial use of Boehme speed-tape equipment to an over-all system using 60-wpm radio teletype plus radio telephone with newly developed ciphony equipment between the Pentagon and most important theater headquarters. Both Air Corps and Army personnel had major responsibilities for building these systems.

After the war the Army Command and Communication System was retailored to meet the cold war requirements, U. S. military deployments in 73 different foreign areas and the Korean War, but its speed reliability, flexibility and mobility would soon be inadequate to meet rapidly changing conditions.

The Army Airways Communications System was also retailored progressively to meet the operational needs of a U. S. Air Force deployed on a global scale.

By the end of World War II, the Navy had improved its fleet communications by employment of the latest developments in equipments and techniques which stressed the requirements of time and reduction of distances and the electronic coupling of detection on, above and below the sea. Its shore system had been modernized to share the large volume of joint communications generated by the joint actions of three services.

Communication lessons of the war were major. In addition to the primary realization that equipment rapidly becomes obsolete, it was clear that as the mobility, firepower and complexity of modern war increased, the quantity and quality of communication support must also increase. In World War II the trend was exponential. It was clear also that the advance conceptual planning of the operational nature of the war had been inadequate. Military and industry were unprepared for the large increase in quantity and performance requirements of communication equipment. Organizationally, technologically and production-wise there was surprise; this emergency was met by costly improvisation—some considerable warm-up time had been granted by circumstance. It was made painfully clear at Pearl Harbor that the communication system must be ready to meet the first shock of war and that nothing less than instant readiness of an adequate system in being and in use meets minimum modern defense requirements.

The two most significant features of World War II which dictated the scope and complexity of the communication systems required were:

1) The global nature of the conflict.
2) The advent of Air Power as a decisive influence on the conduct of war.

The operation of the U. S. Armed Forces in far-flung corners of the globe required immediate establishment by the Army and Navy of world-wide communications centering upon Washington, D. C., for strategic direction, intelligence, administration and logistic support and for the considerable load of press traffic generated in the areas of combat operations.

The advent of Air Power resulted in a struggle by each combatant side to achieve air superiority. The inevitable outgrowth of this struggle was the perfection of radar as a basic tool of air defense. The disposition of radars in depth in the combat area required the operation of combat centers, in the ground forces and aboard ship, whose function was to direct the defensive operations of fighter aircraft and anti-aircraft guns. Adequate, accurate and timely information and direction dictated the pattern, quantity, reliability and types of communications installed, an unprecedented requirement. Communications for transmission of weather data and aircraft movements in the area had to be provided. The demands of mobility in a mechanized ground war had to be satisfied by mobilizing heavy radar equipment, communications terminals equipped for a complex of point-to-point and air-ground operations and message centers operating in conjunction with similarly mobilized fighter and anti-aircraft direction centers. In landing operations involving both Air Force and Navy combat aircraft, additional facilities were required for coordinated direction of air operations. The communications systems and facilities thus provided attained an order of accuracy and speed in the handling, assimilation and use of combat data never previously attained. However, in terms of postwar air weapon developments to come, even this complex was soon to become obsolete.

If for the sake of brevity the several phases of the period 1945–1962 must be treated as one, the summarization might be:

In research and experimentation it is the story of a long search by the three services for new capabilities in both strategic global and in tactical communications. As time passed the search was intensified.

In operational achievements it became a period of massive U. S. Air Force, Navy and Army programs to meet the threat of nuclear attack on the North American continent and to extend and increase the effectiveness of our retaliatory delivery. It was also a period of major new system extensions and major advances by all three services.

In appraisal it was a period of progressive realization that the requirements for military communication were changing so drastically and so rapidly as to render obsolete the basic systems and techniques which by 1944 were proving satisfactory for World War II. Part of this recognition was that the new requirements could not be met by adaptation or extension of the current art; that new capabilities must be found.

To outline the search for new technical communication capabilities in such a summary fashion as follows is unsatisfactory to say the least. The major contributions of the National Bureau of Standards, the National Aeronautics and Space Agency, Canadian and British research and our many university and scientific groups should be adequately interwoven into the story and they are not. The role of the military in this search varied in degree and nature: financial supporter, participant, innovater, exploiter and frequently first user.

Communication by earth-moon relay, ionospheric scatter, tropospheric scatter, meteor trail reflection, communication by active or passive relay from satellites, and reflection from orbiting dipole belts—all have been achieved and exploited or are in development.

This 17-year time span has seen the development of weapons which have progressively annihilated the World War II military operational concepts of time and space. The speeds of aircraft have advanced into the supersonic region. The missile age was suddenly achieved. With each new weapon development the effort to provide for more speed, capacity and, most important, reliability in military communications, has been given more impetus.

In this period, the perimeters of U. S. Air Defense were progressively thrust forward to the northernmost reaches of the Arctic, into the North Atlantic and the Pacific and tied to the air defense of NATO in order to cope with the potential air threat. The Soviet missile threat dictated the installation of a Ballistic Missile Early Warning System (BMEWS) capable of surveying the space over the Soviet land mass. The U. S. Strategic Air Command, assigned a deterrent role, acquired a global mission capability from bases in the United States which called for centralized direction and control of SAC units and recently has been engaged in building a capability for operation of intercontinental ballistic missiles.

Each progressive extension of these Air Defense and Strategic Air capabilities has been accompanied by increasingly demanding requirements for communications. The distinction between communications and weapons and radars as separate systems with human operating links began to disappear and each progressively became a subsystem which taken with the others comprised the Air Defense System in which the important objective was to annihilate time as a factor in operations. To accomplish this, a new subsystem was introduced, the electronic computer, which for purposes of compatibility with the communications subsystem resulted in the development of associated data-processing, conversion and transmission equipment. The resulting system was called the Semi-Automatic Ground Environment or SAGE. Gradually the human functions in the system were being reduced to maintenance and decision making by the electron tube and transistor and through a common system language of digital data.

Both in Continental Air Defense and SAC operations reliability in the operation of communications continued to loom large as a problem in the face of the long distances and the difficult Arctic and sub-Arctic environment that had to be spanned. Demands for information handling capacity and versatility increased beyond the capability of techniques available in the immediate post-war period. To satisfy these demands, the U. S. Air Force embarked on a program of installing ionospheric and tropospheric circuits. With the collaboration of the National Bureau of Standards, in 1951, a simplex test circuit (FPIS) of 1000 statute miles was established in 1952 from Goose Bay, Labrador to Sondrestromfjord, Greenland, to determine the feasibility of this technique. Soon after (May, 1952) a decision was made to proceed with a three-circuit multichannel telegraph system known as "Bittersweet" connecting Maine, Goose Bay, Sandestrom and Thule, Greenland, to meet pressing operational requirements. In November, 1953, "Bittersweet" was completed and by 1954 the FPIS system had been extended to the United Kingdom via Iceland.

In this same period the U. S. Air Force turned to tropospheric scatter as another promising solution to its communications problems. Although tests conducted and theoretical data calculated by the National Bureau of Standards, Lincoln Laboratories and Bell Laboratories indicated that the exact degree of reliability and channel capacity to be expected from tropospheric scatter circuits could not be predicted, the consensus among the scientists was that reliable communication with minimum channel capacity could be expected.

Data also indicated a requirement for high power and large antennas. Success in the application of this technique was confirmed with the installation of circuits connecting radar stations in the rugged terrain of the Labrador-Newfoundland Air Defense System which was placed in operation two years after initiation of the Air Force study. Since 1954 tropospheric scatter facilities have been extensively installed to connect the radars of the Distant Early Warning Line (DEWLine), BMEWS and the Alaskan Air Defense System (Project White Alice). This technique has also been extensively applied to military requirements in Europe. Under sub-Arctic and Arctic conditions where other means of radio communication could not be depended on, tropospheric scatter provided a system reliability of better than 99 per cent. Moreover circuits were provided to operate with design capacities of 36 to 72 voice channels depending on antenna size and transmitter power. The effort to date has been toward application of techniques which would increase reliability to the ideal goal of 100 per cent.

During this post-war period the demands of the national economy for space in the VHF portion of the radio spectrum and the needs of the three services for equipment with greater operational flexibility resulted in adoption of the UHF band (225–400 Mc) for short-range military aircraft communications, airborne and ground and land-based radio relay. Equipment was developed which permitted rapid selection of a large number of channels. Use of the UHF band was extended to NATO aircraft and their associated ground facilities so that all Allied military aircraft would be capable of operating in a common communications environment.

In the search for more reliable communications to meet the operational requirements of its global mission, the Strategic Air Command (SAC), in mid-decade 1950–1960, converted its long-range high-frequency equipment from the World War II mode of double-side amplitude modulation to single-sideband modulation and started to increase the power of its ground stations. New records in maintaining long-range communications contacts with Strategic Air Commmand aircraft were achieved as a result. The SAC decision to convert to single-sideband equipment was based in no small part on the distance records achieved by U. S. amateur operators, through the use of this type of equipment. The increasing missile threat was met with a decision to equip SAC with an airborne command capability to offset the vulnerability of ground-based command posts. New techniques to improve this capability in terms of traffic capacity and reliability continue to be applied.

In the Army, Navy and Air Force, recognition of the vanishing time dimension, the need for instantaneous national reaction to military or political perturbations occurring in any part of the globe, the vulnerability of the Army command and communication network and its lack of mobility, led to major action. Programs were carried out for new antenna-transmitter-receiver design and installation and for new modulation techniques. System capacity and speed was increased—the communication system throughout the Pacific and Far East areas was augmented by the most extensive ionospheric scatter system in existence. Mobility was increased by immediate airborne readiness of complete van-mounted stations. During the Lebanon emergency a complete van-mounted station was flown in and communication into what had been called the Army command network was established. Extended and redesigned, the system is now called STARCOM—Strategic Army Command Network.

But even this optimization of current capabilities fell short of meeting requirements for bandwidth, capacity and flexibility. Therefore, Army planning started in 1955 for development of satellite relay systems. Early results were the ARPA-Army Air Force-Industry experimental project "Score," the talking satellite and its experimental successor, "Courier," "Advent," the synchronous active satellite, "Syncom," a NASA-ADVENT experiment and "Westford," a Massachusetts Institute of Technology-Air Force experiment for orbiting a band of reflecting dipoles, are programs which will shortly give birth to new communications capabilities urgently needed to fit the pattern of current and future military needs.

Recognition of the present unsatisfactory variety of analog and digital material now diversely affecting systems design requirements led both the Army and Air Force to contract with industry for new systems concepts. The Army concept is "Unicom," a universal system using one common digital language, accepting all types of signal content and meeting unique military requirements for security and reliability.

The Air Force, in extending its weapons systems management concept to include supporting electronics systems, initiated a series of "L" system projects many of which required associated communications, some on a global scale. Examples of such projects are 413L, extension of the DEWLine; 425L, NORAD Combat Operations Center; 433L, Weather Observation and Reporting; 466L, Electromagnetic Intelligence System and 477L, Nuclear Detection and Reporting System. Experience with such projects as DEWLine and BMEWS led the Air Force to the conclusion that the communications for these and other systems to be implemented might better be engineered within the existing over-all Air Force global communications system and project 480L was established as an over-all framework.

During the mid 1950's there was realization also among the management and membership of the Institute of Radio Engineers of the rapid growth and importance of communication electronics, and the Professional Group on Military Electronics was formed to promote increased IRE assistance, participation and support to the military effort.

During this period all three services pioneered in exploiting electronic computers and computer communications to solve logistical, supply, personnel and administrative problems of these large complex organizations. Punch-card requisitions were now possible direct from overseas commands to depot complexes in the U. S. This revolution is still very much in progress in inventory-stock control—and to provide capabilities for quick reaction to supply-service needs throughout globally deployed military forces. In addition industry developed for the Army van-mounted mobile "Mobidic" computers to accompany and serve units in the field.

The U. S. Navy, faced with all the communication problems of the other two services, had some which were unique to that service. One of these was the problem of communicating to and from submerged submarines scattered throughout the oceans of the world. Besides making communication advances comparable to the other services, the Navy established a system of superpowered (2 Mw) VLF transmitters to cover world ocean areas and reach submerged submarines. The Navy's Marine Corps developed an air-ground support communication system that was unexcelled.

In land warfare, previous maximum artillery ranges of less than a few tens of miles were suddenly expanded by modern missiles to thousands of miles without available observation and communication means to support them. The advent of tactical nuclear weapons suddenly required a revolutionary increase in troop and headquarters dispersions with concurrent increase in range and speed of communication systems. Wholly new concepts for area coverage systems became urgent; also development of new equipment capabilities was required.

As responsible military communication executives have sought to meet the urgent pressures of increasing requirements, the cost of development, implementation and operation of these new systems has increased to constitute a significant percentage of the military budget. Therefore, budgetary and defense management has increasingly questioned the communication programs of the services. Are the requirements realistic? Are new capabilities really needed? Is there unreasonable duplication among the services? Is military communications as a business efficiently managed?

Increasingly the services have examined their requirements more searchingly and in greater detail. Advisory councils of scientific and industry leaders have been constituted in the services to give broader perspective than that of military personnel alone. Contracts have been placed with nonprofit analysis organizations to determine the basis of and reasons for increasing communication requirements. Particular emphasis has been placed on determining organizational and procedural methods for satisfying future needs with maximum efficiency and economy. What can be done to better organize and utilize the systems we now have? Further, the joint chiefs of staff have continually exercised greater supervisory and coordinating influence. The Secretary of Defense has also progressively enlarged his supervisory and decision making role in military communications. To meet the common communication needs of the widely spaced unified and specified commands which are made up of combat elements of all services, there has recently been organized the Defense Communication Agency staffed by communication personnel of all the services. This Agency also has a responsibility for evaluation and coordination of service communication programs. Its scope, authority and effectiveness are growing and should continue to grow.

Looking back over this panorama of some hundred years of U. S. military communication raises questions as to possible recurring patterns or trends. Some appear to be present.

Although given little mention in this fast sweep through time, the most important recurring pattern has to do with people. Fortunately, within the services the challenge of communications has attracted a fraternity of men, with basic military interests, who in addition had strong pioneering interests in providing, maintaining and improving communications. Aboard combat naval vessels, aircraft and in tactical Army and Marine Corps combat units, communicators of whatever rank or branch of service overcame obstacles, shared common hazards, gave dedicated service both as communicators and as fighting men. Their contributions have consistently been a major factor in achieving needed communications throughout this period.[8] Nor can the dedicated service and achievements of civil service men and women in the services go unmentioned.

It seems clear that military communication requirements consistently have preceded and exceeded the communication capabilities available to satisfy them. It appears that primary military emphasis over this period on weapons systems, ships-tanks-guns-aircraft-missiles has not been accompanied by sufficient and concurrent emphasis on the communication systems, equipment and organization to make the weapons employable with adequate effectiveness. Communications have been in the usual position of trying to catch up with needs and the needs were frequently late in being recognized.[9]

Another consistent pattern that emerges is that before each of our major wars of the past there have been periods of warning and warm-up when there was some considerable opportunity to reduce our unreadiness for the

[8] The authors' most serious dissatisfaction with this article is that brevity requires almost complete omission of the names of major outstanding leaders in military communications. The contributions of Army Chief Signal Officers, Chiefs of Naval Communication, Directors of Air Force Communication, Directors of Communication of the Joint Chiefs of Staff and of the Defense Communication Agency—all have been unmentioned—but not overlooked.

[9] "The methods of modern warfare and its wide deployment of forces make effective communications one of the vital elements of victory. . . . Actually war laid far greater demands on Signal Troops and equipment than the War Department had anticipated and the rapid development of electronic devices continued to multiply these demands. Foreword by Major General Albert C. Smith, U.S.A., Chief of Military History to U. S. Army in World War II, *The Signal Corps: The Test* by G. R. Thompson, D. Harris and others in the U. S. Army in World War II historical series.

test of war. Yet each war did in fact find us unready from a communication point of view. The degree of the unreadiness and its cost seems to have become higher as war became more mechanized, complicated, long range and rapid.

Finally, it seems certain that the important relationship of command-control-communications to the effectiveness of all military operations and weapons systems has been consistently misunderstood and undervalued.[9]

Perhaps it may be appropriate in concluding this brief backward look to ask some questions about what lies ahead in military communications. Have the lessons of the past been learned? Is our recognition and evaluation of future needs in military communications now adequate? Do we have realistic answers to the question of what leadership at all levels must do in meeting and after suffering the next shock of whatever type of war may come so that we may plan now and have ready the communications necessary to permit the exercise of this leadership? How can we provide the essential survivability of this system without unnecessary expense? From whom and by what method can the right answers be obtained?

These questions grow more complex with each passing year. The strength of the effort is growing to find better answers for the future than we have had in the past. In the success of these efforts the nation has a vital stake. By what standards can their adequacy be measured? Are there additional ways in which the great technical strengths of this Institute of Radio Engineers can be brought to bear in support of these efforts?

Achievements and Prospects of Artificial Earth Satellites*

S. F. SINGER†

Summary—**In less than half a decade astronautics has developed from crude satellites to systems of unbelievable sophistication. Probes to Mars and Venus and a manned expedition to the moon are well into the planning and design stage. Weather and communication satellites as well as military monitoring satellites will soon be fully operational; navigation satellites are at this stage already, as are manned earth satellites. Elaborate lunar probes and complex astronomical satellites are just around the corner. However, practically all of the scientific achievements so far have been in the field of geophysics. Although some areas still remain to be investigated, many fundamental results have been gathered about the earth and its environment; briefly discussed here are findings about the figure of the earth, the density, temperature and constitution of the upper atmosphere, the nature of the radiation belt and auroral particles, and disturbances of the geomagnetic field.**

THE SCOPE of the present paper is confined mainly to the applications of earth satellites and space probes to space science, *i.e.*, to geophysics and astrophysics. That is not to say that satellites have not been used for other purposes, and certainly in future years and even future months we shall see an increasing application of satellites to important areas of human enterprise, such as communication, navigation, and weather reconnaissance. (The military applications of satellites, embodied in the MIDAS and SAMOS satellites, are also concerned with reconnaissance, but their construction and operation is quite different from a weather satellite.)

* Received by the IRE, February 5, 1962.
† Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Calif. On leave from the University of Maryland, College Park, Md.

Some More Immediate Prospects for Space Technology

We want to speak briefly about these applied satellites. Articles have appeared, of course, which point out the impact that they will have on human affairs and on the way of life of the "man in the street." But it seems impossible to do them complete justice because in their final form, and after their outputs are fully integrated into our economy, they may indeed lead to quite revolutionary (and unpredictable) changes.

Everyone by now is familiar with the importance of a communication satellite. Some pioneer experiments have been made, the initial one by the Army Signal Corps in the SCORE satellite which in 1958 transmitted a message from the President of the United States all over the world.

Communications satellites fall into two major classes: passive satellites and active satellites. A prototype of the *passive* reflector is the ECHO satellite, which is essentially a large sphere of thin balloon material; its purpose is to present as large a cross-sectional area as possible for a minimum weight. Its main advantage is its utmost reliability since it carries no receiving and transmitting equipment. A second advantage is the fact that it can be used to communicate at *any* frequency; its bandwidth is limited only by the opaqueness of the earth's atmosphere and ionosphere. Its only disadvantage, really, is the fact that one requires powerful transmitters and receivers on the earth and high-gain antenna systems in order to communicate with a good signal-to-noise ratio.

A special subclass of passive satellites is represented by the WEST FORD project which consists of a belt of orbiting dipoles. It is effective only for frequencies near the dipole resonant frequency (8000 Mc).

For this reason much attention has been paid to *active* satellites which carry both receivers and transmitters to relay the information sent to them. They may even carry more elaborate equipment, for example, tape recorders, to store and receive information and transmit on command or by preprogramming when they reach a certain location over the earth. In addition, they have the facility, of course, of receiving on one frequency and transmitting on another one. In general they have advantages which are complementary to those of passive satellites. Therefore, it seems likely that both systems will find eventual application.

Two types of active satellites are being pursued: 1) A relatively low altitude satellite, or rather a system of satellites in order to cover the earth; its prototype is the COURIER launched in 1960. 2) In addition, we have the so-called hovering satellite, at an altitude of 22,300 miles, whose orbital period is 24 hours. It, therefore, remains synchronous with the earth, and by suitable minor adjustments can be made to remain almost stationary over a given point on the earth's surface. Three satellites could cover the earth; the system is exemplified by the SYNCOM and ADVENT prototypes now being constructed.

It is too early at this stage to review in detail the manifold activities concerning communication satellites. In particular, their mode of operation has not been conclusively settled, *i.e.*, to what extent individual companies and government agencies participate, how a civilian and military network is to be set up, to what extent foreign countries or associations of foreign countries will participate or compete, etc. It is in the field of the communication satellite where the European community has chosen to make its entry into space technology, and precisely for the reason that it offers hope of a great economic return. One factor which is often not appreciated is the very low cost of the satellite relative to the cost of the ground station facilities which must be built in order to utilize the satellite communication system to the fullest extent.

By contrast, the *navigation satellite* is almost operational as the Navy TRANSIT satellite system. A complete analysis of the Doppler shift of the received radio signal fixes the position of the observer relative to the satellite, and if the latter's orbit is precisely known, then the position of the observer is determined. A review article covering all aspects of the system, from the theoretical point of view to the operational point of view, has recently been prepared.[1] The accuracy of the navigational function depends on how elaborate the receiving equipment is; it can be as good as 100 ft.

Weather satellites are not yet operational, but the time is drawing near when they will be used for weather studies and prediction in a routine way. The TIROS series has been successful as a test of instrumentation, and the forthcoming NIMBUS and AEROS satellites promise to be the prototypes of a complete weather system. The U. S. Weather Bureau has already made immediate operational use of satellite data to forecast severe storms and to check ice conditions in northern shipping lanes.

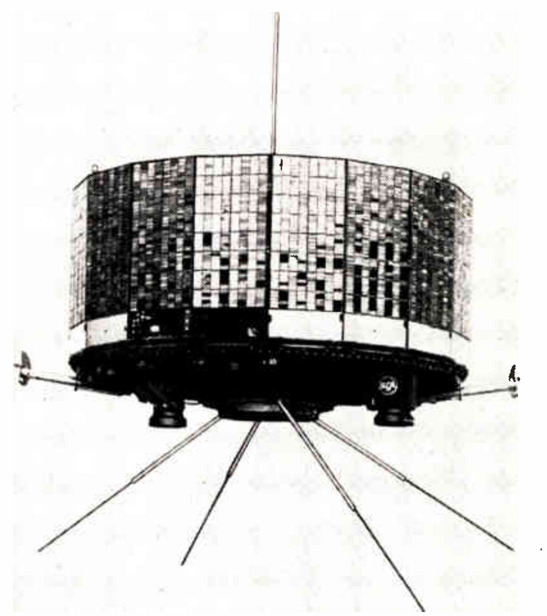


Fig. 1—TIROS III weather satellite, launched July 12, 1961.

Technical Features of Satellites

From the point of view of space science, artificial satellites have turned out to be a tool whose scope probably extends further than any other technique. As we shall see, it has been used to investigate phenomena from the interior core of the earth, extending through the crust and atmosphere all the way into interplanetary space. We shall base our review here, in structure

[1] A. M. Stone and G. C. Weiffenbach, "Radio Doppler method of using satellites for geodesy, navigation and geophysics within the planetary system," in "Progress in Astronautical Sciences," S. F. Singer, Ed., North Holland Publishing Co., Amsterdam, vol. 1; 1962. Wherever possible in this article references will be given to books and review papers in which original publications will be found cited.

at least, on a review published about six years ago,[2] which described the type of research which *might* be done with artificial earth satellites; it gave the theory and the *rationale* for many experiments, but, of course, it gave no results. It may, however, be consulted to illustrate the thinking which prevailed at that time and also as a source for references up through 1956.

Some of the technical features described for satellites in 1956 seem antiquated in modern terms; however, many ideas current at that time have been adopted in actual satellite systems. For example, it was surmised that an inert earth satellite would be scientifically useful since its orbit could be used to measure the figure of the earth as well as atmospheric density. Satellites with telemetering aboard are, of course, much more useful. But in addition to a multichannel telemetering system and telemetering storage which was then anticipated, we have now data-handling systems of great complexity. They automatically select the instrument whose information is to be transmitted, turn off certain instrumentations when the data become less useful and will change the transmitting setup automatically as required by external conditions or by command.

As far as attitude control is concerned, not only has use been made of gravitational torques and of spin stabilization, but also magnetic torques have been found useful for a number of satellites, and automatic attitude-control systems based on horizon scanning have been adopted in many cases.

As anticipated, solar batteries have been found fully useful for at least the present generation of earth satellites. It seems likely, however, that nuclear power supplies will be of increasing importance for many applications where the power demands are great. A nuclear power supply based on the SNAP system, *i.e.*, the use of radioactive isotopes, has already been flown in 1961 on a TRANSIT satellite.

Advantages for Scientific Work Peculiar to Satellites

The chief distinguishing features of a satellite are several: 1) Primarily its location which leads to the possibility of continuous observations above the absorbing effects of the atmosphere. In order to maintain itself the satellite must be located above the appreciable atmosphere which in practice means in the exosphere, about 500 km, well above the ionospheric regions. 2) The uniform coverage of the earth which a satellite can furnish if it is in a suitable orbit.

From these major points follow many minor points in which a satellite may be superior to earth-bound instrumentations:

a) It is capable of nearly simultaneous measurements over the whole earth and, therefore, can give an instantaneous picture of the phenomenon under study.

b) It can make a large number of these simultaneous measurements one after the other, and so average out transient effects.

c) By being at high altitudes it can often de-emphasize disturbances which exist near sea level and are caused by local anomalies.

d) Finally, a satellite measurement may often be less expensive from the point of view of effort and cost than equivalent measurements with other tools such as sea-level expeditions, airplanes, balloons and rockets.

In the following discussions these advantages of the satellite will become evident, but by far the most important characteristic of the satellite is that it is capable of performing some measurements which *cannot be conducted by any other means.* It is therefore a *unique* tool as well as a most convenient one.

The Interior of the Earth

Surprising as it may seem at first glance, satellites are capable of making refined measurements of the interior of the earth which are quite difficult by other means. These refer to the main magnetic field, and therefore to the dynamics of the liquid core, and to the distribution of matter in the earth. While it is possible to made these measurements at lower altitudes, or even at sea level, the advantages of the satellite which have been enumerated above provide an overwhelming argument for its use.

The Earth Magnetic Field and Its Secular Variations

The central problem is to explain the origin of the earth's magnetic field. Sea-level observations over the last decades have shown that the field is complicated, containing in addition to the main dipole also higher-order components, as well as local anomalies; in addition, the field exhibits a slow variation with time, the so-called secular variation, consisting of a westward drift of the nondipole components, as well as a change in the dipole field.[3]

These observations of the secular variation have been handicapped by the fact that the time scale of the observations themselves has a spread on the order of a few years, since simultaneous observations over the earth cannot yet be made. Therefore, it is likely that a satellite survey will reveal quite a different picture concerning the geomagnetic field and its time variation.

[2] S. F. Singer, "Geophysical research with artificial earth satellites," in "Advances in Geophysics," H. Landsberg, Ed., Academic Press, New York, N. Y., vol. 3; 1956.

[3] E. H. Vestine, L. Laporte, C. Cooper, I. Lange and W. C. Hendrix, "Description of the Earth's Main Magnetic Field and Its Secular Change, 1905–1945," Carnegie Inst. Wash. Publ. No. 578; 1947.
E. H. Vestine, L. Laporte, I. Lange and W. E. Scott, "The Geomagnetic Field: Its Description and Analysis," Carnegie Inst. Wash. Publ. No. 580; 1947.

In fact, there are indications today which would show that the sea-level field (as we believe it to be) must contain a large contribution from external current systems, greater than the 1 per cent which is usually assumed.

In spite of the obvious advantages of a satellite for the geomagnetic mapping job, this has not yet been carried out.[4] There are, of course, technical difficulties; for example, satellite magnetometers can be constructed which have an accuracy of 10^{-5} gauss. However, in order to keep the field variation due to altitude changes within this limit, the altitude must be determined to within 20 m for a 500-km orbit.

The Figure of the Earth and Gravitational Anomalies

The determinations of the shape of the earth and of the distribution of matter in its interior are of considerable interest not only in geophysics but also for geodetic applications. The subject had its start with Newton's estimate of the nonsphericity of the earth; this allowed him to give a satisfactory explanation of the phenomenon of the precession of the equinoxes which had been discovered much earlier. Subsequent advances are associated with names of Clairaut whose treatment of the subject proved that in order to satisfy the equilibrium equations, the figure of the earth must be that of an oblate spheroid. However, the earth departs at places by several miles from the shape of a perfect oblate spheroid. It is more closely approximated by the "geoid," an assumed equipotential surface which coincides with the mean sea-level surface over the oceans and is perpendicular at every point to the direction of gravity. The deviations of the geoid from the spheroid normally do not exceed about 100 m. These irregularities are caused by the attraction of land masses, but are partially compensated by isostasy.

The size and shape of the geoid as determined by geodetic data has been reviewed by Heiskanen. Another approach to obtain information on the geoid is from the precession of the moon and from its slow nutation, as shown by Jeffreys.[5]

Earth satellites can be used to measure the geoid by studying changes in their orbits. As is well known the equatorial bulge of the earth leads to two effects on the orbits of satellites:[6,7]

1) A rotation of the orbital plane in space or "regression of the node," *i.e.*, the line on which the orbital plane interesects the equatorial plane of the earth rotates in space at a certain rate, given approximately by 10.0 $(R/a)^{3.5}$ cos i degrees per day, where R is the earth's radius, a the semimajor axis of the orbit, and i the orbit's inclination to the earth's equator. 2) In addition, the line of apsides rotates within the orbital plane, *i.e.*, the perigee moves around the ellipse, at a certain rate given by 5.0 $(R/a)^{3.5}$ $(5 \cos^2 i-1)$ degrees per day. Additional perturbations result from the fact that the earth is not a perfect spheroid.

An analysis of the results has given a more accurate measurement of the equatorial bulge. It has also been found that the northern and southern hemisphere differ slightly in shape, leading to what has been called a "pear-shaped" distribution of the earth. In addition, it seems that the earth is slightly tri-axial, *i.e.*, that it does not possess perfect symmetry about its axis of rotation. This tri-axiality would have an important effect on a 24-hour synchronous satellite.

The results on the shape of the geoid have been discussed in great detail by King-Hele[8] and by O'Keefe.[9] A simplified version of the background is given by Singer[2] and a simplified discussion of the results by Rossi and Jastrow.[10]

The Earth's Surface and Lower Atmosphere

A satellite can be used to study the atmosphere below it primarily by (passive) optical means. It either receives electromagnetic radiations which are *emitted* from the surface and various atmospheric layers, or it observes solar radiation *scattered* from the layers below it. In the radio spectrum, however, the satellite may take an active part in the measurement by emitting radio waves whose modification in traversing the ionosphere gives information about the ionospheric layers. We shall discuss, in order, the earth's surface, troposphere, stratosphere, chemosphere, and ionosphere.

Albedo and Its Relation to Weather

One of the simplest and at the same time one of the most important from a meteorological point of view is the observation of the amount of reflected sunlight. It can be carried out with simple photocells, or when more detail is desired, with television cameras. One can obtain rather detailed information about the distribution of the albedo, and therefore about the distribution of clouds, ice, and snow fields, all of which have a very high value of specific albedo.

Such information is useful from two points of view. The *detailed* distribution of albedo, in particular, therefore, the distribution of clouds, provides one of the

[4] It would seem that the geomagnetic mapping experiment might be a suitable one for the forthcoming POGO (polar orbiting geophysical observatory) series of satellites.

[5] W. A. Heiskanen and F. A. Vening-Meinesz, "The Earth and Its Gravity Field," McGraw-Hill Book Company, Inc., New York, N. Y.; 1958.

H. Jeffreys, "The Earth," Cambridge University Press, Cambridge, Eng., 3rd ed.; 1952.

[6] D. G. King-Hele, "Properties of the atmosphere revealed by satellite orbits," in "Progress in Astronautical Sciences," S. F. Singer, Ed., North Holland Publishing Co., Amsterdam, vol. 1; 1962.

[7] In general these perturbations are not desirable but in some cases they have actually been used; for example, in the WEST FORD project these perturbations were invoked in order to produce a resonance effect of radiation pressure on the orbiting needles. See I. I. Shapiro and H. M. Jones, "Lifetimes of orbiting dipoles," *Science*, vol. 134, pp. 973–979; October, 1961.

[8] D. G. King-Hele, "Satellites and Scientific Research," Routledge, London, Eng.; 1960.

[9] J. A. O'Keefe, "Determination of the earth's gravitational field," in "Space Research," H. Kallmann-Bijl, Ed., North Holland Publishing Co., Amsterdam; 1960.

[10] B. Rossi and R. Jastrow, "Results of experiments in space," in "Science in Space," L. V. Berkner and H. Odishaw, Eds., McGraw-Hill Book Company, Inc., New York, N. Y.; 1961.

most direct means of observing "weather" and can be used for forecasting. On the other hand, the albedo *integrated*, for example, as a function of latitude, is useful for climatological studies and for providing an input to the problem of describing the large scale circulation of the atmosphere.[2]

Finally, there is the problem of measuring the albedo on a long-term basis. The possibility exists that climatic changes are due to long-term changes of the earth's albedo, because of an increase in cloudiness or increased atmospheric turbidity, produced, *e.g.*, by volcanic dust. This point of view has been worked out especially by Wexler; it may be that satellites will provide an answer to this question, but probably not in the near future.

Infrared Emission and Temperature of Atmosphere, Ground, and Oceans

It is well known that the infrared emission from the earth to space can be thought of in terms of two components: 1) the black body emission from the earth's surface with a peak near 10–12 microns; 2) the emission from the upper atmosphere which takes place at a very low temperature of $-50°$ to $-60°$C.[2] Thus, an infrared detector sensitive near 10 microns measures the surface temperature of the earth, while infrared detectors sensitive outside of the atmospheric "window" measure the temperature of the emitting layers at higher altitudes. Both of these possibilities are being followed up in the current meteorological programs.

The measurement of infrared flux leaving the earth is of two-fold importance: 1) it furnishes an important datum for the energy balance calculations of the atmosphere; 2) it allows measurement of ground temperatures. The latter may have an interesting application to oceanography. With sufficient resolution it would allow the tracking of certain ocean currents, *e.g.*, the Gulf Stream, by virtue of their temperature differential with respect to the surrounding ocean surface.

As of this date this application to oceanography has not yet been actively pursued. Neither has the application of satellites for the study of atmospheric electricity and for a measurement of the world-wide distribution of thunderstorms. Again, such applications may be anticipated for the future, perhaps as a part of a full-scale weather satellite program.

The Chemosphere

Solar radiation, particularly in the ultraviolet, can produce excitation and dissociation of atmospheric molecules such as O_2, H_2O, NO, etc., leading to complicated de-excitation and recombination reactions. Often new molecules are formed, such as ozone, having strong absorption characteristics by which they can be detected. Often too, the reaction emits visible radiation which can be observed on the ground as the "airglow."

Ozone Content and Distribution

A method for measuring the vertical distribution of ozone in the atmosphere, as well as the total amount of high-altitude ozone at various places over the earth, has been worked out and operates as follows:[2] ultraviolet light from the sun is partly scattered and partly absorbed in the earth's atmosphere, the absorption in the region of 2600 Å to 3000 Å being principally due to ozone. Therefore, a detector which looks down and measures the spectral distribution of scattered ultraviolet can in this manner determine the altitude distribution of ozone. Interestingly enough, this particular method may be applied to the planet Venus almost before it is applied to the earth.

Light Emission from Airglow and Aurorae

Various emissions can be studied by means of photocells with filters, or similar detectors of restricted bandwidth. The important measurements would be the forbidden oxygen transitions at 5577 and 6300 Å, the OH emissions in the near infrared, and ultraviolet emissions in the airglow which are not observable from the ground. Such measurements would serve to give a synoptic picture of the distribution of the airglow, and therefore some ideas of the state of motion of the upper atmosphere. Measurements by Roach and colleagues have indicated the patchiness of night airglow and its drift, but no world-wide measurements of the phenomenon have as yet been carried out.

Of special interest would be observations of auroral light emissions. They would provide accurate information about the geographic distribution of aurorae and of their time variations. Data obtained from ground measurements appear to be quite exciting. A group at the University of Alaska under the leadership of C. T. Elvey has deduced important information about the morphology of the aurora, while Soviet scientists, principally Lebedinsky and colleagues, have shown the existence of continuous auroral arcs which they suspect may extend not only over the Asian hemisphere but extend into the North American continent as well.

However, no concerted effort has been made to carry on satellite experiments in this field, and no data of any significance are therefore available.

The Ionosphere

Electron Density

Initial observations of the ionosphere were carried out through the Soviet SPUTNIKS, which used a transmitter frequency of the order of 20 Mc, low enough to be affected by the ionosphere. Measurements were made of the appearance and disappearance of the signal, refraction effects, fading, and Faraday rotation, to obtain synoptic information about the ionosphere.[11] However, it was found to be more satisfactory to carry out more specific experiments, and at the present time efforts are being directed towards the development of "topside"

[11] One would like to mention, however, the existence of an ionospheric retardation effect, first described by Berning, which is clearly observable for example in the TRANSIT program.[1] It manifests itself as a frequency shift and is produced by the varying phase shift of the wave as it passes through a varying amount of the ionosphere.

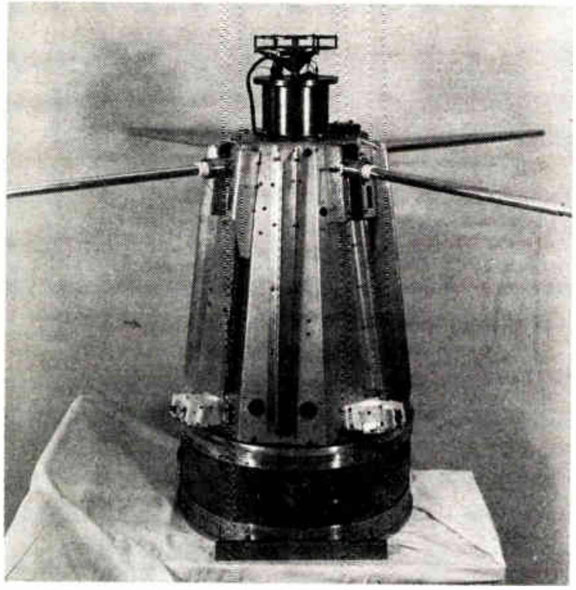


Fig. 2—P-21 Electron Density Profile Probe for upper ionospheric research.

sounders. This device is analogous to the usual ionospheric sounder, except that it is mounted on a satellite and will receive reflections from an electron layer located below the satellite.

A different aspect of ionospheric physics has been stressed in experiments in which LF waves are being received in satellites.[12] The purpose here is two-fold: 1) to measure the transmission of LF waves through the ionosphere which normally reflects the largest fraction of their energy back to the ground; and 2) to study radio whistlers.

No special effort will be made to review these techniques and results in the present survey since many of the papers are published in the PROC. IRE and are therefore available in their original form.[13]

THE EXOSPHERE

Since a satellite spends practically all of its useful life above 200 km, it is well suited to carry out direct physical observations in this region. Above 200 km the atmosphere is assumed to be in diffusive equilibrium, with each component falling off according to its own scale height as determined by its molecular weight. As a result N_2 molecules begin to be unimportant at around 200 km and by the time an altitude of 530 km is reached only oxygen atoms, some nitrogen atoms and a small percentage of helium and hydrogen atoms are present. The altitude of 530 km is generally taken to be the *base* of the exosphere, the region from which gas can escape from the earth's gravitational field without making any further collisions.

From theories of the exosphere it follows that oxygen is the most important constituent just above 500 km, and that hydrogen eventually predominates above about 2000 km. It seems possible from recent evidence that there is an intermediate region, around 1500 km, where helium is most abundant.

The important parameters that a satellite can determine are, in order of increasing difficulty: the density of the atmosphere, time variations in this density and to some extent latitude variations; and finally the composition, state of ionization, and energy distribution of free electrons.

Atmospheric Densities Deduced from Satellite Orbit

Measurement of upper atmosphere densities can be made even with an inert satellite; it is only necessary to observe its orbit. The collisions with atmospheric atoms produce a frictional force which is proportional to the density. The effect of this tangential drag force will be to change the energy of the orbit and therefore its shape and size. A very simplified discussion of the expected changes is given in Singer.[2] The subject has been studied in much greater detail by a number of authors; a recent review has been produced by King-Hele who also gives the currently accepted values of upper atmospheric densities, the day-night variation, seasonal effects, and an indication of a possible solar cycle effect.[6]

So-called "standard atmospheres" have been prepared from time to time. The most recent one is the COSPAR atmosphere of 1961; however, it should be realized that even this compilation is not likely to be the final answer concerning upper atmosphere densities. In fact, if a solar cycle variation does exist (as one would expect from a theoretical viewpoint), then the subject may be much more involved, and a single set of values may not suffice. On the other hand, in many cases it is permissible to deal with an average atmosphere.

More refined methods for measuring atmospheric density have been suggested from time to time but have not been generally used. They are the observations of an angular acceleration produced by aerodynamic torques; a measurement of ram pressure by means of sensitive gauges; and the observation of upper winds, temperature and turbulence by the production of artificial sodium trails.[2] No trails have been produced from a satellite as yet, but in one of the Russian probes sodium vapor was used to pinpoint the position of the probe. In geophysical rockets in the USSR sodium has been used by Shklovsky and Kurt to deduce upper atmospheric densities and temperatures.

Probe Measurements

Among the most active approaches to the study of the ionosphere is the use of probes of various types. We shall discuss the results here in some detail. The subject was pioneered mainly by Gringauz and colleagues, who flew ion traps in SPUTNIK III and more refined instruments in later Russian space probes. In principle, probes

[12] An additional purpose of these studies, not connected with the ionosphere, is the extension of radio astronomy to lower frequencies.

[13] See special issue on "The Nature of the Ionosphere—An IGY Objective," PROC. IRE, vol. 47; February, 1959. See also J. P. Leiphart, *et al.*, "Penetration of the onosphere by very-low-frequency radio signals—Interim results of the LOFTI I experiment," PROC. IRE, vol. 50, pp. 6–17; January, 1962.

can measure the complete properties of the medium in which the satellite or rocket is moving, *i.e.*, the ionic composition and energy spectrum of the electrons. In addition, it is possible to measure such subsidiary quantities as the collision frequency for electrons, and the electric potential of the vehicle with respect to the surrounding plasma.

In practice, however, many of these measurements have been difficult to make. Energetic ions can be studied in a fairly direct way after provisions are made to suppress the photoelectric effect due to solar ultraviolet. For example, Bridge and Rossi have used probes in EXPLORER X to measure some properties of the interplanetary plasma and deduced the existence of "clouds" or "layers" containing very energetic protons (plasma) alternated by layers which do not produce a detectable signal. These observations correlated extremely well with magnetic studies by Heppner in which the magnetic field was found to be turbulent whenever a plasma probe was seeing the hot plasma.

More down-to-earth measurements have been made in EXPLORER VIII and in a large number of research rockets. Bourdeau and colleagues from the Goddard Space Flight Center have measured the energy spectrum of electrons at satellite altitudes and find it to be the same as that of the ions.[14]

The altitude distribution of electrons has been measured through propagation experiments by Berning in the USA and by Alpert in the USSR, but the plasma probes give information which leads to a somewhat different interpretation. From the altitude distribution measured by Hale (as interpreted by Hanson) it is concluded that helium ions form the most abundant species of ions at around 1500 km. Similarly, the results of Gringauz, *et al.*, show that the scale height above 1500 km is extremely long. Before any detailed conclusions are drawn from these data, one must be quite certain that the *potential* of the vehicle is in no way affecting the measurements.[15]

Effects of Satellite Potential

The question of the *potential* of the vehicle is of considerable importance from many points of view. If the vehicle carries an electric charge, this charge will interact with the ions and electrons of the surrounding plasma through which the vehicle is moving, and produce an electric drag. This drag does not affect a large satellite to any considerable extent, but becomes quite important for a very small body, particularly one whose dimension is less than the Debye screening distance, *e.g.*, meteoric dust particles. In fact, the theory of Coulomb drag was first developed in order to deal with meteoric particles in interplanetary space and in the vicinity of the earth. More recently this theory has been applied to calculate the drag of the WEST FORD needles. It was concluded that the needles are likely to have an rms potential of -3.6 v, which in turn limits their lifetime to a few months. If their potential is of the order of 10 v, then the lifetime will be a matter of a few days or a few weeks.[16]

Among the most puzzling of observations have been the unusual radar echoes which are sometimes observed in connection with satellite passes. These were originally described by Kraus, who received bursts of WWV and concluded that somehow plasma clouds were produced in the ionosphere by a passing satellite. The existence of such clouds has not been accepted by all workers in the field, and in fact no detailed theoretical explanation has been produced so far. The writer has been partial to a mechanism which he calls a "plasma compression." It is assumed here that the satellite, because of its electric charge, will affect the plasma; because of the geomagnetic field a condensation of plasma can be produced in the vicinity of the satellite. For further detail concerning the observations by Kraus and others, and concerning possible explanations of these effects the reader is referred to Singer.[16]

Geomagnetically Trapped Radiation and the Earth's Radiation Belts

The motion of charged particles in the earth's magnetic field had been understood for many years, but the existence of particles in trapped orbits was hypothesized only as recently as 1956. Yet the discovery of trapped particles at low altitudes near the equator, in EXPLORERS I and IV by Van Allen and colleagues in 1958, came as a surprise. An interpretation was quickly put forward; the particles were identified as high energy protons even before their nature was experimentally established by Freden and White in 1959. The existence of trapped particles at higher latitudes and altitudes deduced from SPUTNIK III data by Vernov, Lebedinsky, Krasovsky, *et al.*, was similarly interpreted as a component of trapped radiation, but of lower energy and produced by a different mechanism than the equatorial inner belt.

A great deal of experimental effort has been devoted in the past 4 years to the study of trapped particles, unfortunately often with very crude instruments. It turns out that experiments in rockets carrying photographic emulsions have contributed most of the detail concerning the nature of the radiation, while satellite experiments, particularly EXPLORERS VI, VII, and XII, have elucidated the time variations and spatial distribution.

The experimental results may be summarized as follows: There is an inner region of trapped particles beginning at an altitude of about 500 km, at which point the intensity rises very sharply, increasing to a maximum at about 3000 km and reaching a minimum again at about

[14] From theoretical arguments, however, one would expect the electrons to contain a high energy tail since all of them are originally produced by quite high energy photons in a photo-ionization process.

[15] Results of EXPLORER VIII indicate that the potential of the vehicle is very close to zero, within 0.1 v. But this result is at variance with what one would expect from theory. The writer therefore reserves judgment concerning the ultimate interpretation of probe data.

[16] S. F. Singer, Ed., "Interaction of Space Vehicles with an Ionized Atmosphere," Plenum Press, New York, N. Y.; 1962.

6000 km. This inner belt contains high-energy protons with energies up to 700 Mev. The maximum flux is of the order of 3×10^4 per cm^2 per second in the heart of the belt. The falloff beyond 3000 km is thought to be due to the fact that the magnetic field begins to lose its trapping ability for high-energy particles; as a result the radiation becomes progressively softer with increasing altitudes.

The outer belt extends out to about 4 earth radii, and is known to consist of electrons having energies generally below 1 Mev, with perhaps a small fraction above that value. Quite recent information indicates that it also contains protons of the order of 1 Mev. Hence the outer belt is very soft and can easily be screened off with only a small amount of shielding. The fluxes involved are of the order of 10^8 per cm^2 per second and show very large time variations. It is quite clear that the outer belt is closely connected with magnetic storms and aurorae.

As far as interpretation is concerned, the most fruitful approach has been in terms of the so-called neutron albedo theory. Cosmic rays incident on the earth's atmosphere create neutrons, a fraction of which will be emitted in the upward direction. Being uncharged they can travel through the magnetic field without any deviation and will decay along the way. High energy neutrons will pass their kinetic energy to their decay protons, and thus it becomes possible to inject into the earth's magnetic field protons having energies of several hundred Mev. Even very-low-energy albedo neutrons will give an energy of the order of $\frac{1}{2}$ Mev to the decay electron; therefore, both protons and electrons can arise from the neutron albedo. A detailed review has been undertaken[17] and the results may be summarized as follows: Every observed fact about the high-energy protons (>10 Mev) can be explained so far on the basis of the neutron-albedo theory. But in the case of the low-energy particles, electron and proton, it would appear that the neutron albedo only contributes a part of it. It is possible that the largest fraction of high-energy electrons (>200 kev) arise from neutron albedo. But the lower energy electrons are almost certainly due to other causes, most likely a local acceleration mechanism within the earth's magnetic field in which energy contained in solar corpuscular streams is somehow converted to the kinetic energy of trapped particles.

Magnetic Storms and Aurorae

Measurements of the magnetic field were first undertaken in SPUTNIK III and LUNIK rockets, and carried on most fruitfully in PIONEER V and EXPLORERS VI and X. They indicate that the geomagnetic field terminates at about 10 to 15 earth radii, with a "quiet-day" magnetic perturbation existing at altitudes of 7 to 10 earth radii. During a magnetic storm, however, the perturbation is greatly enhanced and also moves inward; its magnitude during the "main phase" decrease of the storm correlates closely with sea-level values. It is generally considered likely that these magnetic perturbations are produced by the magnetic effects of trapped particles. A review article giving the theoretical development of these magnetic effects and a comparison with the observations has recently been published.[18]

Some disagreement exists among the experimenters; for example, the existence of the quiet-day ring current is not generally accepted. Further, a peculiar magnetic variation at 3 to 4 earth radii seen in LUNIK I has not been observed in other vehicles. It is likely, of course, that large fluctuations of the field may exist on a short spatial and temporal scale. In fact, what may be magnetic "shock waves" have been directly observed by Sonett near the boundary of the geomagnetic field.

Cosmic Radiation

Cosmic rays are corpuscular radiations of extremely high energies. The primary cosmic rays consist mainly of protons, but contain also (~15 per cent) helium

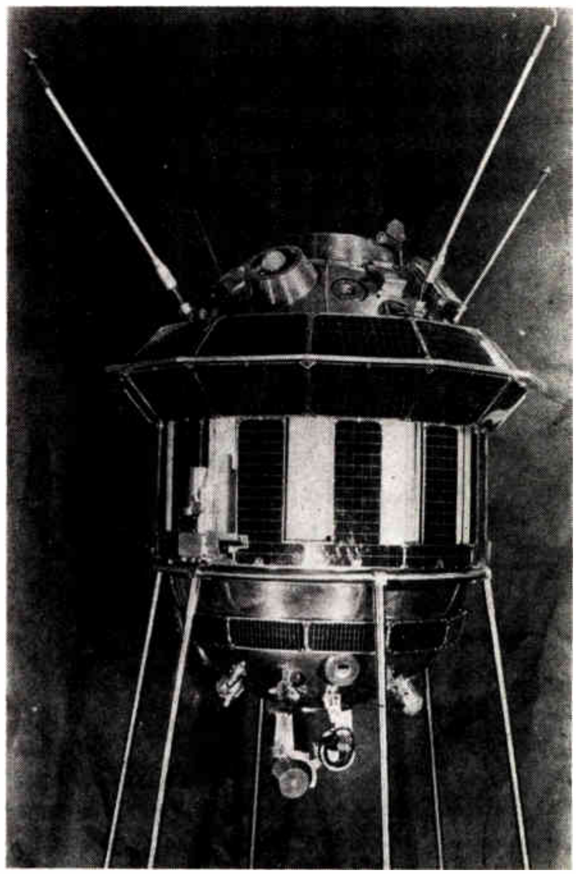


Fig. 3—LUNIK III space probe, launched October 3, 1959.

[17] S. F. Singer and A. M. Lenchek, "Geomagnetically trapped radiation," in "Progress in Cosmic Ray Physics," J. G. Wilson, Ed., North Holland Publishing Co., Amsterdam, vol. 6; 1962.

[18] J. R. Apel, S. F. Singer and R. C. Wentworth, "Effects of trapped particles on the geomagnetic field," in "Advances in Geophysics," H. Landsberg, Ed., Academic Press, New York, N. Y., vol. 9; 1962.

nuclei and to a smaller extent (~1 per cent) the nuclei of heavier elements. They arrive at the top of the atmosphere with almost the speed of light and with energies ranging from a few billion volts (Bev) up to a billion times as much. They constitute the highest energy phenomenon known in nature; but because of the small number of cosmic rays which are received here, the energy they bring is about equal to the energy of starlight. Nevertheless, the effects of cosmic rays on the earth and the earth's atmosphere are important: they account for the production of C^{14}, H^3, and He^3 isotopes, among others, as well as for the ionization below the ionospheric layers.

One of the important means of studying the properties of the cosmic radiation has been through geomagnetic effects. A survey of cosmic ray intensity over the earth is therefore of real interest. Such surveys have been undertaken in the Soviet manned spaceships and published by Vernov and his colleagues. In the United States a lesser emphasis has been placed on cosmic ray measurements in satellites; many of the important results are still being obtained by balloons and high altitude rockets.[19]

Among the most interesting of the investigations are those related to increases of "cosmic" rays associated with solar flares.[20] Satellite data give, generally speaking, the spatial and time distribution of the increase. More detailed studies relating to the energy spectrum and composition of the SHEP radiation, however, have been done in high altitude rockets launched at high latitudes. Worth mentioning is the analysis of recovered satellite material from two DISCOVERER vehicles which were exposed to SHEP radiation. An analysis showed an anomalously large amount of tritium in the skin of the satellite whose origin has not been satisfactorily explained.

Some of the more interesting problems in the cosmic-ray field relate to interplanetary observation outside of the geomagnetic field. A few of these have been made and have given very worthwhile results. For example, in PIONEER V a general decrease of cosmic ray intensity was observed which correlated with a decrease observed at the earth; it established that the mechanism for this decrease was an interplanetary and not a geocentric mechanism.[10]

Meteoric Particles

Satellites are measuring the flux of incoming meteor particles and time variations. These measurements have established so far that the flux is in rough accord with what would be expected from observation of the zodiacal light which gives indirectly a value for the concen-

[19] S. F. Singer, "The primary cosmic radiation and its time variation," in "Progress in Cosmic Ray Physics," J. G. Wilson, Ed., North Holland Publishing Co., Amsterdam, vol. 4; 1958.

[20] The term "solar" "cosmic" rays is anachronistic, and it has therefore been suggested that it be replaced by a more suitable word. SHEP (for "solar high energy particle") seems to be a suitable notation and is currently gaining favor.

Fig. 4—RANGER I lunar probe.

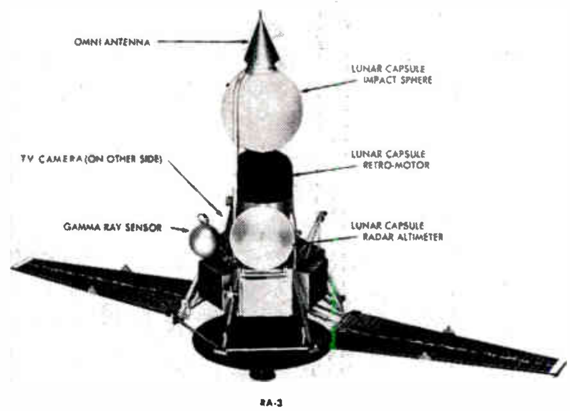


Fig. 5—RANGER III lunar impact vehicle.

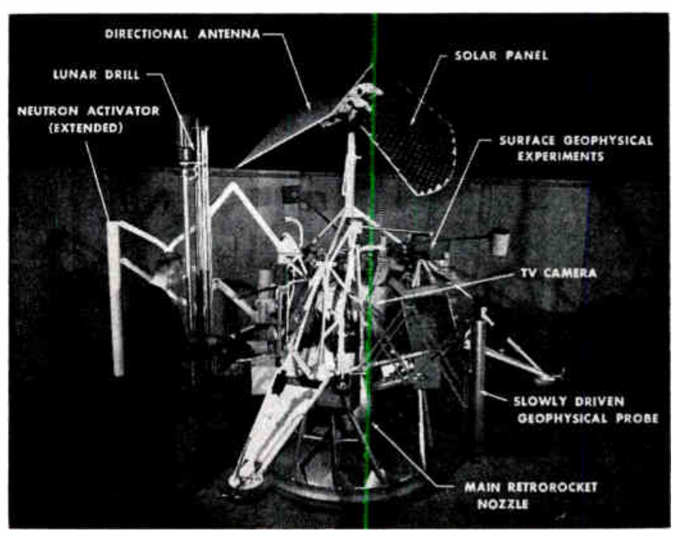


Fig. 6—SURVEYOR lunar surface research vehicle.

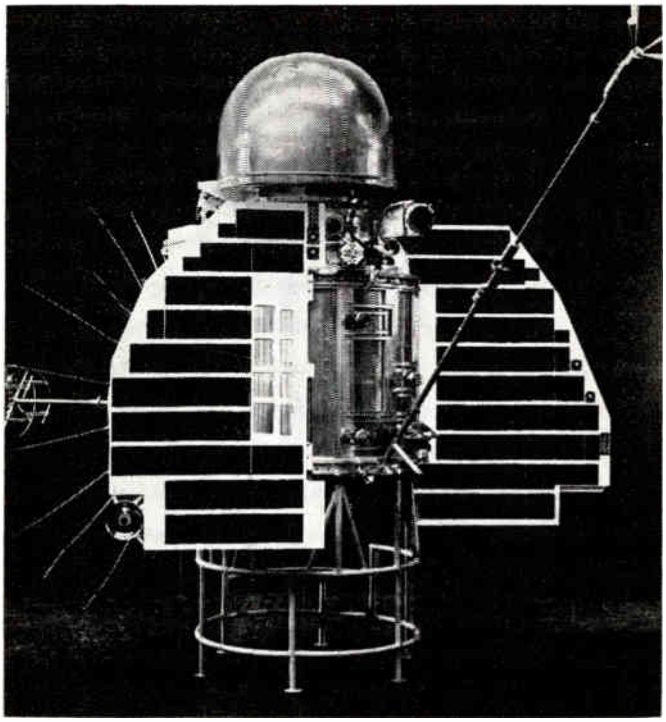


Fig. 7—Russian Venus probe, launched February, 1961.

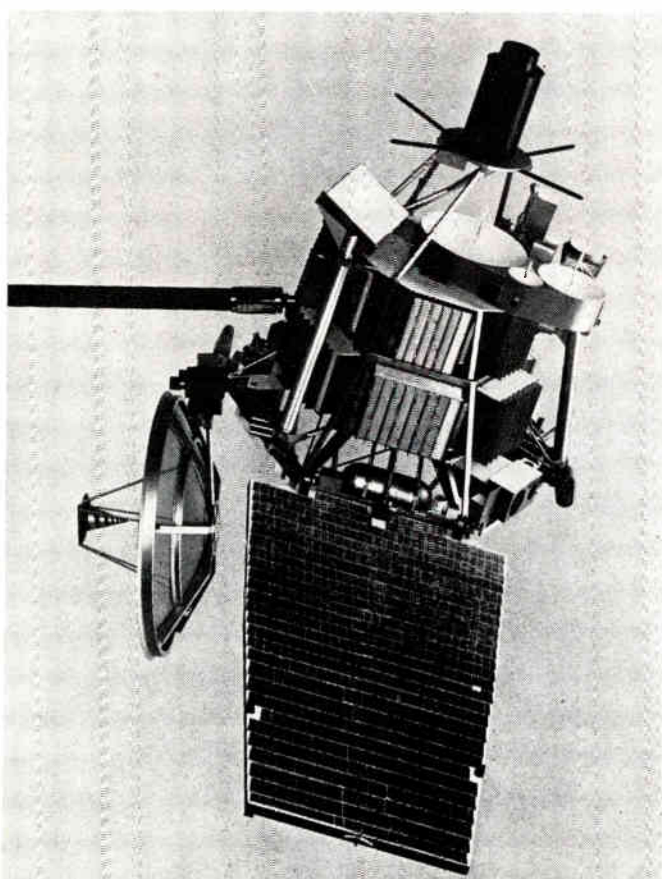


Fig. 8—MARINER A Venus probe.

tration of dust in interplanetary space. The time variations have established that meteor showers enhance the influx of particles, but there still remain many unsolved problems which can only be settled by further measurements, and in some cases by much more refined instrumentation. One problem relates to the spatial distribution of dust around the earth. There are competing theories all of which indicate that there should be a concentration of dust as one approaches the earth, but the expected concentration differs by a factor of 1000. One particular theory predicts a dust *shell* with a maximum at about 2000 km.[21] This could be established by satellites in eccentric orbits, as for example by the proposed EGO satellite.

More detailed problems relate to the size distribution and mass distribution of the dust particles, and their physical properties such as density, nature of the material, shape and so on. These studies are still in their beginning stage and will benefit greatly as satellites become more easily available and as the observations are extended into interplanetary space and into the vicinity of the planet Mars.

Astronomical Observations

It has been mentioned earlier that proper astronomical observations in the classical sense have not yet started, but promise to become one of the main applications of earth satellites. However, a new branch of astronomy, *ultraviolet astronomy*, was started years ago by a research group at the Naval Research Laboratory under H. Friedman. Most of their detailed results have been obtained in high altitude rockets. They have shown in detail that the ultraviolet spectrum of the sun plays an important part in the ionization and photochemistry of the upper atmosphere. They have studied in detail the spectrum as far down as 400 Å, and more roughly down to 1 Å. In pioneer investigations they have also established the incidence of ultraviolet radiation from sources outside of the solar system.[22]

It is in this area where future scientific applications of satellites may well be concentrated. They promise to be one of the most important tools for space exploration and the pre-eminent tool for the new science of ultraviolet astronomy.

The *astronomical satellite* is still in the stage of preparation; its impact will only be felt in the next few years. The planetary probes again will be coming into their own in the next few years, starting with unmanned RANGER and MARINER probes to the moon, Mars, and Venus, leading first to landings of instrument packages on the surface of the planets, and finally to full-scale automatic exploring equipment (SURVEYOR, VOYAGER), and eventually to manned expeditions.

[21] S. F. Singer, "Interplanetary dust near the earth," *Nature*, vol. 192, pp. 321–323; October, 1961.

[22] H. Friedman, "The sun's ionizing radiations," in "Physics of the Upper Atmosphere," J. A. Ratcliffe, Ed., Academic Press, New York, N. Y.; 1960.

Section 22

NUCLEAR SCIENCE

Organized with the assistance of the IRE Professional Group on Nuclear Science

The Instrumentation and Control Circuits of Nuclear Reactors*

E. H. COOKE-YARBOROUGH†, MEMBER, IRE

Summary—**The chain reaction which takes place in a nuclear reactor is a regenerative process and can therefore be unstable, so the fission rate must be closely controlled in order that the nuclear reaction may proceed at a rate which is both useful and safe. Electronic circuits play a vital part in this, and their reliability must be exceedingly high, as many reactors are required to run for months without interruption and for years without an unscheduled shutdown.**

This paper considers in a very elementary way the dynamic behavior of a reactor and the circuit requirements associated with maintaining the power level constant. The additional problems associated with starting-up and the automatic shutting-down of a reactor are briefly surveyed and the use of automatic data-handling is considered. It is concluded that not all instrumentation and control problems presented by existing types of reactor have been satisfactorily solved and that more advanced reactors will present still more difficult problems.

Introduction

THE WORKING of a nuclear reactor depends on a chain reaction in which a neutron entering the nucleus of an atom of fissile material (such as plutonium or uranium 235) causes the fission of this nucleus into two or more fragments, with the release of energy and the emission of further neutrons, some of which in turn cause further fissions.

The chain reaction may be represented in the most elementary way by the idealized regenerative circuit shown in Fig. 1. By inspection we can write down,

$$\frac{dQ}{dt} = i_2(A-1) - i_1 + i_3. \tag{1}$$

If i_3 is small compared with i_2 and i_1, which is true except when the reactor is operating at very low level, then

$$\frac{dQ}{dt} = i_1\left[\frac{R_1}{R_2}(A-1) - 1\right]. \tag{2}$$

The term $R_1(A-1)/R_2$ represents the average number of neutrons which are produced by a fission and which cause further fissions. The symbol used to represent this number is k. To keep the reactor in equilibrium we must keep k very close to unity.

Since $i_1 = Q/CR_1$, we have

$$\frac{dQ}{dt} - \frac{Q}{CR}(k-1) = 0.$$

Solving this differential equation gives

$$Q = Q_0\epsilon^{t(k-1)/CR_1}. \tag{3}$$

If k differs from unity, the neutron flux in the reactor rises or falls exponentially with a period $CR_1/(k-1)$. In a slow-neutron reactor (one in which most neutrons are slowed down by a moderator before causing fissions) CR_1 may be of the order of hundreds of microseconds, but in a fast reactor, which needs no moderator, it is of

* Received by the IRE, July 5, 1961.
† Electronics Division, U. K. Atomic Energy Authority, Atomic Energy Research Establishment, Harwell, England.

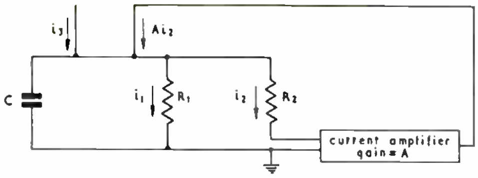


Fig. 1—Simplified equivalent circuit for a nuclear reactor.
Q (charge on capacitor C) represents the number of free neutrons in the reactor, and is thus proportional to the mean neutron flux.
i_1 represents the neutrons lost or captured without producing further neutrons.
i_2 represents the neutrons which cause fissions.
i_3 represents neutrons produced by processes other than neutron-induced fission (*e.g.*, from a neutron source introduced into the reactor or from spontaneous fission).
CR_1 represents the mean time which elapses before a neutron is lost or captured.
CR_2 represents the mean time which elapses before a neutron causes a fission.
A represents the average number of neutrons produced by a fission.

the order of 1 μsec. The expression shows that when k is close to unity the reactor period is much longer than these times.

There are many differences between the regenerative circuit in Fig. 1 and a real reactor, and the most important of these are indicated below:

1) In Fig. 1 it is assumed that there is no delay in the current amplifier, *i.e.*, all neutrons are emitted promptly when fission occurs. In more than 99 per cent of the cases this is true, but in less than 1 per cent of the cases the emission of a neutron is delayed for times ranging up to tens of seconds. When k is within 1 per cent of unity these delayed neutrons make the reactor period much longer than indicated by (3).
2) k may not be independent of fission rate, for it may vary with temperature, or fall as neutron-absorbing nuclei accumulate in the reactor. It is possible, in some types of reactor, for k to have a positive temperature coefficient.
3) The core of the reactor has finite volume, and conditions will not, in general, be uniform throughout this volume.

Nevertheless, the model in Fig. 1 is a useful guide to understanding the problems of reactor control.

To control the reactor we require means of directly or indirectly measuring the fission rate, and of controlling the value of k. The reactor then becomes part of a degenerative feed-back loop.

Reactor Power Measurement

The actual quantity which is stabilized by the control system depends upon the purpose for which the reactor is to be used. A research or materials-testing reactor is required to give a known flux of neutrons, and here neutron-flux measurement is an essential part of the control system. In reactors used for power generation, the power output, and hence the financial return on invested capital, rises rapidly with the coolant outlet temperature. This must therefore be kept as high as possible without allowing the fuel elements or other important parts of the reactor to become too hot. Here thermal measurements are the ultimate factor in the control, though neutron-flux measurements usually also play an important part in the start-up and short-term control of the system.

In slow-neutron reactors the measurement of the fission rate is effected by measuring the flux of slow neutrons, usually by use of ionization chambers containing boron (which emits alpha particles when exposed to slow neutrons) or a fissile material (the ionization being produced by the actual fissions occurring in the chamber). Similar methods can be applied to fast reactors if the ionization chamber is enclosed in a moderator.

The temperature of coolant, moderator and fuel elements are usually measured with thermocouples. These can tolerate the temperatures and neutron fluxes encountered in existing reactors, but in future reactors using higher temperatures and fluxes other methods may have to be found. The use of thermocouples to measure fuel element temperatures presents serious practical difficulties, since the elements must be easily removable, so it is difficult to maintain proper thermal contact between the fuel element and the thermocouple and at the same time to maintain reliable electrical connections to the thermocouple. An over-all measurement of thermal power is usually also required, and is effected by measuring the inlet and outlet temperatures of the coolant and the coolant flow rate.

Control of Reactivity

While measurements of reactor power are almost invariably electrical in nature, the means of controlling reactor power are almost always mechanical. In slow-neutron reactors the usual method is to introduce a neutron-absorbing material, such as cadmium, into the reactor, either in the form of sliding rods or pivoted arms, which is equivalent to reducing the value of R_1 in Fig. 1. In reactors using a heavy-water moderator, control of the heavy-water level can be used to control the reactor power. Precautions in mechanical design must be taken to avoid the possibility of inadvertent or too-rapid withdrawal of absorbers or entry of heavy water. The mechanical design problems are quite difficult, especially as some parts of the control mechanism may be inaccessible after the reactor has first been put into operation.

The Control Loop

The control system of a reactor is essentially an electromechanical one, comprising the reactor itself, some electrical or electronic means of measuring the reactor power and some electromechanical means of controlling the power. In many cases a human operator is interposed between the measuring equipment and the control equipment. With many existing reactors manual control is not only possible but also easy, for the response of the reactor to a small change of reactivity is slow, since it is determined largely by the delayed neutrons. Moreover, these reactors usually have a negative temperature coefficient which, if the control rods are withdrawn slightly, tends to restabilize the power at a slightly higher level. Some more modern types of reactor can, however, show a positive temperature coefficient of reactivity, and this makes manual control much more difficult. Moreover, although the response to very small changes of reactivity is slow, it becomes much faster if the change of reactivity is rather larger, as has already been indicated. A large change in reactivity may therefore be difficult to correct by manual control.

Even if the operator is eliminated, the open-loop transfer function of the complete system is still quite complicated, since it is affected by mechanical lag (and possible lost motion) in the control system, lags in the reactor itself due to delayed fissions, temperature changes and delays in the cooling system, and possible further lags in the measuring system, especially if the measurement is of temperature rather than of neutron flux. To apply effective automatic control it is necessary to know the transfer function of the complete system under all conditions which might arise. Consequently in the design of new reactors, attention is being directed to factors which may affect the transfer function and to methods of calculating the transfer functions of particular reactor designs. Transfer-function measurements on existing reactors are also of considerable importance. They not only facilitate the design of a control system but also throw light on the nuclear behavior of the reactor itself.

Reactor Transfer Function Measurement

One method of measuring the transfer function is to oscillate a neutron absorber in the reactor at known rates and to observe the amplitude and phase of the resultant change in neutron flux at each rate. Each measurement has to be integrated over a period long enough to allow the noise component of the reactor neutron flux to be averaged out. Another approach is to analyze this noise by autocorrelation, but this depends upon the assumptions that the frequency spectrum of the agency causing the noise is known and that the reactor approximates to a minimum phase-shift network. These difficulties can be avoided by programming a control rod to move in a quasi-random manner and performing a cross-correlation between this movement and the resultant reactor flux changes. The first method is at present the most highly developed and accurate, but the noise methods may be more rapid.

The Control of Start-Up

So far we have considered only the problem of keeping the reactor operating at a constant power level. Starting it up and bringing it up to the desired power is often much more difficult, mainly because the range of fission power between shut-down and full power may be as large as 10^{12}. It is necessary to observe the rise of neutron flux continuously throughout the starting-up process to ensure that the reactor period does not become too short at any time. For this purpose logarithmic neutron-flux-indicating instruments have been developed which can cover five or six decades without the need for range switching. Nevertheless, it is difficult to cover the lower end of the range indicated above, and a solution is permanently to install a neutron source in the reactor. Referring again to Fig. 1, the neutron source increases i_3. We have

$$i_3 = i_1 + i_2 - A_{i_2} = \frac{Q}{CR_1} - \frac{Q}{CR_2}(A - 1).$$

Therefore

$$Q = \frac{i_3 CR_1}{(1 - k)}.$$

This increase of i_3 can bring the neutron flux within the range of measurement, and so allows k to be continuously monitored during shut-down. This is valuable, for more than one reactor incident has occurred when the reactor was thought to be shut down.

While the reactor is being started up, its period can be deduced from the rate of change of the logarithm of the neutron flux, but a direct indication of period is valuable and can be obtained by electrically differentiating the output of the logarithmic neutron-flux meter. The design of a satisfactory period meter is far from easy, because of the random fluctuations present in the power-level reading which are accentuated in differentiating the output of the power-level meter. These fluctuations vary with power level and, while they can be reduced by smoothing, smoothing-time constants large enough to deal with fluctuations at low-power level may lead to misleading period indications. This has been the subject of a good deal of theoretical and experimental work which is not yet complete.

Automatic start-up of a low-power reactor can be achieved by linking the period meter with the control system so that the power rises with a predetermined period until the desired power level is reached. In a high-power reactor many other factors, such as temperature gradients and thermal stresses, limit the rate at which power can be changed once the level has been reached at which fission heating is appreciable. In this range the need for automatic control of changes in power level does not seem yet to have been felt, but it may be needed in propulsion reactors or peak-load power stations, and will involve some sort of automatic computation.

Safety Circuits

Whether a reactor is under manual or automatic control there is always the possibility of some failure in the reactor or its associated apparatus which could cause damage or even danger. The function of the safety circuits is to detect a failure as early as possible so that the reactor can be shut down before serious damage occurs. Most important reactor failures result in a local rise in temperature, the detection of which depends on temperature measurements made at very many different parts in the reactor. The cause of the temperature rise is likely to be either a coolant-flow failure or a rise in fission rate, so coolant-flow measurements at every coolant channel together with neutron-flux measurements may prevent fuel element failure by causing a shut-down to be effected earlier than would temperature measurement alone. Fuel element failure may be detected by monitoring for radioactive particles in the exit coolant, though again the monitoring of every coolant channel may be necessary.

Indication of reactor failure may come from very many different sources and so it is difficult to decide whether or not a failure indication from any one source should be made to shut the reactor down automatically. Conservative design philosophy would indicate that any failure should cause a reactor shut-down, including a failure in any of the safety circuits. However, experience indicates that when a reactor has a very large number of such safety circuits the probability of one or more failing is so high that it is difficult to start the reactor or to keep it running for any length of time, even though the reactor itself may be without fault. This can lead to some safety circuits being deliberately put out of action temporarily.

The only real answer to this problem is to improve the reliability of the safety circuits and to make it possible to distinguish between a failure which really does demand a reactor shut-down and a local failure of part of the safety circuits. This can be done by making the circuits continuously self-checking and by introducing redundancy. A typical redundant arrangement will consist of three entirely separate systems monitoring the same point in the reactor. Logic circuits connected to the output of these three systems cause the reactor to shut down if a shut-down is demanded by a majority of the systems. The failure of a single system can be recognized as such, and dealt with, but the circuits must be so designed that while one system is out of action, a failure indication from only one of the remaining systems is able to demand a shut-down. The probability of a spurious shut-down is then increased, so from this point of view a higher order arrangement, say three out of five, is to be preferred. There is a temptation to group together systems covering associated but not identical parts of the reactor into a majority decision arrangement but this is generally bad practice since it might allow a genuine local failure in the reactor to go undetected. On the other hand, the use of majority decision arrangements everywhere in the reactor may involve introducing an inconveniently large number of detectors into the reactor. The difficulties are reduced if the problem is faced early in the design of the reactor.

Reliability can also be increased by using some form of self-checking system. A typical example is the superimposing of an ac signal on the high-voltage supply to an ionization chamber. This will be communicated to the collector electrode through the chamber capacitance and hence through the dc amplifier. The presence of this signal at the output of the dc amplifier is a fair assurance that everything is in working order and connected, though not a complete test of the system. Complete checking can only be effected in a majority decision arrangement, a simulated shut-down signal being applied to each system in turn. Since the system being tested is temporarily out of action, an arrangement of higher order than two out of three is preferable if testing occupies a significant part of the total time.

The mechanical parts of the shut-down system are also subject to failure. Common practice is to use safety rods which can fall into the reactor under gravity, but are normally held out by an electromagnet. It is not unknown for such arrangements to fail due to friction, and the use of multiple shut-down rods is advisable, though sometimes difficult to achieve.

Safety systems which make use of redundancy allow correct operation of the reactor despite failure in part of the shut-down system, but they become ineffective if there is delay in rectifying the faulty parts of the system. This is largely a problem of organization and psychology.

Reactor Poisoning

The fission process generates a large variety of fresh atomic nuclei, some of which, being strong neutron absorbers, reduce the reactivity of the reactor and are referred to as poisons. An important poison is xenon 135 which appears as a result of the beta-radioactive decay of iodine 135 which is in turn a decay product of the fission process. So long as the reactor is running, the xenon 135 is continually being removed, partly by radioactive decay but mainly because the captured neutrons convert it by conversion to xenon 136 which does not significantly absorb neutrons. When the reactor is shut down, the removal of xenon 135 by the conversion process ceases, while its generation by iodine 135 decay continues for some hours, so the xenon 135 concentration first rises and then decays. With high-flux reactors there may be a period after shut-down during which the xenon 135 concentration is so large that it is impossible to restart the reactor. In such cases it is essential to keep track of the xenon 135 concentration and to predict its future behavior. This can, in principle, be computed from a knowledge of the power level of the reactor as a function of time. If the neutron flux and composition of the reactor could be regarded as uniform throughout the reactor, a simple analog computation would suffice, but this is not always adequate, and a more accurate computation involves dealing with many different regions of the reactor concurrently.

Spatial Instabilities

In physically large reactors it is possible for an uneven distribution of reactivity to develop, so that the temperature may rise locally above the safe limit. A normal control system cannot prevent such spatial instabilities though, of course, a local excess of temperature would cause a shut-down. The instabilities can only be damped out by differential insertion and withdrawal of neutron absorbers in different regions of the reactor, but such differential movements can only be effected if differential information on neutron flux or temperature is available. Ionization chambers in the thermal shield of the reactor do not give a good indication of such instabilities and a better method is the monitoring of the temperature at the outlets of the various coolant channels. This, however, only deals with radial instabilities (*i.e.*, those in a plane perpendicular to the coolant ducts); if axial instabilities are likely to occur other means must be found for detecting them. One way is to insert fission chambers directly into the reactor core. This leads to serious problems with regard to depletion of fissile material in the chamber, radiation damage and insulation, but it has been shown that such chambers can be made which will last for more than a year in the core of a gas-cooled graphite power reactor.

Data Handling

It will be seen that the operation of a nuclear reactor involves the accumulation of a very large amount of data on neutron flux, temperature, coolant flow, coolant pressure, etc., at a very large number of points in and around the reactor. In addition, current data needs to be combined with previous data to give such information as the total irradiation of the various fuel elements since they were inserted, the state of poisoning of the reactor, the accumulation of Wigner energy in a graphite moderator and the appearance of spatial instabilities. The need for a small digital data-handling system to do such work is now widely accepted, but there may be a case for going further and making it participate directly in the control of the reactor. For example, instead of using the majority-decision safety circuits considered earlier, all temperature, coolant-flow and neutron-flux readings might be correlated, their space and time derivatives determined and the results used to apply appropriate confidence weighting to the information obtained from each sensing system. This would increase the reliability of decisions regarding control or shut-down and so should permit more economical operation. On the other hand, to rely entirely on a single data-handling system would call for a very high degree of reliability. A mistaken decision to shut down a power reactor might be acceptable if it occurred not more than once a year, but mistakes in the reverse direction would probably be unacceptable even if they occurred on average only once per reactor per thousand years. Whether such reliability can best be obtained from a single suitably designed data-handling system remains to be seen.

Electronic Circuit Design

In the space available little can be said about the details of circuit design. Solid-state devices play an increasingly important part in reactor control and safety circuits, but the dc ionization chamber followed by an electrometer valve amplifier remains widely used for reactor neutron flux measurement. A transistor dc amplifier can very usefully follow the electrometer valve. Replacement of this valve by a semiconductor diode capacity modulator is possible, but the spurious input currents are then higher and the circuits more complex. Modern electrometer valves are extremely reliable, provide an enormous current gain and use voltages and currents which match well with those of transistor circuit. Semiconductor devices are susceptible to damage by fast neutrons and this prevents their sustained use as amplifiers or radiation detectors inside a reactor.

The sensitive slow-neutron measurements required during the early stages of reactor start-up are performed with proportional counters containing boron, which deliver a pulse each time a neutron reacts with a boron atom. These pulses are amplified and their mean rate is measured by a counting ratemeter. The logarithmic characteristic required for period measurement may be obtained by summing the output currents of several diode pump circuits having different suitably chosen constants. With the dc ionization chambers used at higher power levels a nearly logarithmic characteristic may be obtained by exploiting the voltage/current relationship exhibited by an electrometer valve grid, a semiconductor diode, or a transistor emitter.

A disadvantage of pulse counters is that a good signal-to-noise ratio is difficult to achieve if the counter must be connected to the amplifier through a long cable, as may be necessary in a large reactor installation. Associating a head amplifier with the counter may call for high-temperature operation and resistance to radiation damage. A number of ways of dealing with the problem are currently under investigation.

The trend is increasingly towards reactors operating at higher temperatures and this raises severe difficulties in obtaining adequate insulation even in the ionization chamber and its connecting cable. Satisfactory operation of these components at 500°C is now obtainable and 800°C appears possible.

Majority-decision circuits make use of logical techniques not unfamiliar to the designers of digital computers. One method is to use multiple windings on a square-loop magnetic core, the state of magnetization of this core then being determined by the majority decision. Another way is to sample in turn the output of the various circuits and on each cycle to count the number which demand a shut-down. A shut-down is effected when this number exceeds a predetermined value.

Because so many thermocouples are used throughout a reactor it may be necessary to sample their outputs at a rate higher than can be satisfactorily achieved by mechanical contacts. Electronic commutation can be achieved by use of solid-state switches, but is quite difficult because of the small dc voltages which must be switched and the possibility of introducing spurious voltages.

Conclusions

In a short paper it has only been possible to touch on some of the major problems encountered in the control of modern reactors. Space does not allow even a very selective bibliography of the subject. It will be seen that current types of reactor pose control problems which have not all been satisfactorily solved. Future power reactors will be operating under still more stringent condition and, to obtain safe and economic operation, close control of the operating conditions will be needed. The success of these projects depends very largely on the success with which the control and instrumentation problems can be solved.

Acknowledgment

The author wishes to thank R. J. Cox, D. Harrison, and P. K. McPherson of the Atomic Energy Establishment, Winfrith Heath, for comments and advice during preparation of this paper.

Nuclear Radiation Detectors*

G. A. MORTON†, FELLOW, IRE

***Summary*—Nuclear radiation detectors are required in all the major fields of nuclear science and technology. They fall into two principal categories, single element detectors and imaging detectors. Single element detectors can be classified into four types, based upon their physical mode of operation. These are**

1) Scintillation counters,
2) Gas ionization detectors,
   a) Ionization chambers
   b) Proportional counters
   c) Geiger-Mueller counters
3) Semiconductor detectors, and
4) Chemical detectors.

The limitations of these detectors are discussed in terms of sensitivity, energy resolution, speed, and ability to discriminate against background radiations. This gives some indication of the directions along which future developments must proceed.

Imaging detectors include both area imaging, such as radioautography and scintillation counter scanning, and track imaging. The various techniques of track imaging from the Wilson cloud chamber and bubble chamber through scintillation track imaging and spark chambers are discussed.

Introduction

THE FUNDAMENTAL principles of many of the nuclear radiation detectors used today date back to the beginning of the Twentieth Century. The discovery of radioactivity in 1896 by Henri Becquerel was through exposing a black paper-covered photographic plate to a uranium salt. The photographic process is today a very powerful means of detecting and measuring nuclear radiation. A few years after Becquerel's discovery, Rutherford and Geiger,[1] working on the problem of radioactivity and radiation from nuclear processes, used a detector consisting of a chamber filled with a rarified gas and having in it a central wire biased almost to breakdown which gave an electrical pulse each time a particle of radiation passed through the gas. This detector was later developed into what is now known as the Geiger-Mueller counter.[2] Rutherford also found that the arrival of individual particles of radiation could be detected by observing through a microscope the scintillations produced when they impinged upon a fluorescent screen. A logical extension of this principle was to detect these flashes of light with a photomultiplier as was described in 1947 in two nearly simultaneous papers, one by Kallmann working in Berlin, the other by Coltman and Marshall in the United States.[3] This device is now known as the scintillation counter and is one of the most effective radiation detectors in use at present. Other detectors which had their origin in the early years of this century were the ionization chamber,[1] the porportional counter,[1] the Wilson cloud chamber,[4] and the crystal counter.[5]

* Received by the IRE, August 2, 1961; revised manuscript received, October 19, 1961.

† Conversion Devices Laboratory, Electron Tube Division, RCA Laboratories, Princeton, N. J.

[1] E. Rutherford and H. Geiger, "An electrical method of counting the number of α-particles from radioactive substances," *Proc. Roy. Soc.* (*London*) *A.*, vol. 81, pp. 141–161; August, 1908.

[2] H. Geiger and W. Mueller, "Das Elektronenzählrohr," *Physik Z.*, vol. 29, pp. 839–841; November 15, 1928.

[3] H. Kallman, *Natur u Technik*, July, 1947; J. W. Coltman and F. H. Marshall, "A photomultiplier radiation detector," *Phys. Rev.*, vol. 72, p. 582; September, 1947.

[4] C. T. R. Wilson, "On an expansion apparatus for making visible tracks of ionizing particles in gases and some results obtained by its use," *Proc. Roy. Soc.* (*London*) *A*, vol. 87, pp. 277–292; September, 1912.

[5] P. J. van Heerden, "The Crystal Counter," N. V. Noord Hollandsche Uitgevers Maatschappij, Amsterdam, Neth.; 1945.

The development of radiation detectors proceeded rather slowly during the first four decades of the century. Beginning just before 1940, two sets of circumstances combined to greatly accelerate development in the field of nuclear radiation detectors. One was the discovery of nuclear fission and the successful operation of a fission reactor with its implication of practical nuclear power; the other was the great advances that were made in high-energy particle accelerators which expanded the detector requirements from the relatively simple ones for the few naturally radioactive elements to the very complex needs for the host of artificially induced nuclear processes. The stimulus given to the development has continued up to the present and shows every indication of continuing for the foreseeable future. The advances in nuclear radiation detectors and their associated electronics for processing and presenting the information are so rapid that to forecast what the situation will be five years hence is extremely difficult and, beyond that, virtually impossible.

Nuclear Radiation

Nuclear radiation is a generic term referring to the large variety of radiations resulting from nuclear processes. Table I summarizes the more important radiations falling in this category. Of these many types of radiation, gamma rays are the most important since they are produced as the result of almost all nuclear processes. Their accurate detection and measurement is essential in nuclear science and technology. Neutrons rank next in importance because of their role in nuclear power processes and the development of good neutron detectors is imperative.

Radiation Detector Applications

Before considering the detectors themselves, let us first examine the scientific and technological areas where nuclear radiation detectors are required. For convenience, these are classified into four domains, namely: 1) nuclear power, 2) radiation technology, 3) biological and medical, and 4) nuclear physics research.

Nuclear Power

It is expected in the future that the most important contribution of nuclear processes to human society will be in the form of nuclear power generation. Therefore, nuclear detectors for use in this domain must rank high in the general consideration of the problem. At present and in the near future, nuclear power will be obtained from fission reactors. The role of neutrons in the chain of events occurring in a fission reactor has been described in another article in this Journal. It is clear from this discussion that for an efficient control of reactor performance, detectors must be provided which will measure low-level neutron flux and rate of change of flux during reactor start-up and also monitor the high-level neutron flux during operation. In other words, detectors must be provided which will cover a range of ten or twelve orders of magnitude of flux density. As yet, detectors having the requisite sensitivity at low-flux level and at the same time the large capacity, stability, and long life required at the high-level operating flux

TABLE I

Nuclear Radiation

Type	Symbol	Constitution	Occurrence	Remarks
Gamma rays	γ	Electromagnetic waves	Almost all nuclear processes	Corpuscular (photon) behavior when interacting with matter to give up energy
Beta rays	β^-	Electrons	Natural radioactivity Artificial radioactivity High-energy machines	
	β^+	Positrons	Artificial radioactivity High-energy machines Pair production, etc.	
Alpha rays	α	Helium nuclei	Natural and artificial radioactivity	All fission fuels are α-ray emitters
Neutron beam	n	Neutrons (neutral particles of mass of hydrogen nuclei)	Fission reactions Transmutation processes	Important in all nuclear power processes
Proton beam	p	Hydrogen nuclei	High-energy machines Cosmic rays	These radiations are of interest to nuclear physics in the study of elementary particles, high-energy processes, and the basic constitution of matter
Meson beam	$\mu^+\mu^-$ $\pi^+\pi^0\pi^-$ $K^+K^0K^-$	Muons, Pions, Kaons: particles of mass 250 to 1000 that of electrons	↓	↓
Hyperon beam	Λ $\Sigma^+\Sigma^0\Sigma^-$ Ξ ?	Λ hyperons, Σ hyperons, Ξ hyperons: particles of mass greater than proton	↓	↓
Neutrinos	ν	Neutrinos	β-decay, etc.	Of considerable theoretical interest, but of no practical importance

have not been developed. The expedient of several detectors, some of which may be removable, must be resorted to at present.

While the problem of controlled fusion has not yet been solved, it is virtually certain that neutron monitoring will be required in the operation of this type of power source. Thus, it can be expected that neutron detectors will be essential instruments for all nuclear power sources.

In addition to the neutron detectors, the housekeeping of a reactor installation requires gamma-ray detectors and monitors to ensure freedom from radioactive contamination. Finally, alpha-particle detectors will also be required by the health physicist responsible for reactor safety to monitor the air for possible alpha emitter contamination.

Many of the technological processes for preparing reactor fuel whether from the virgin ore or the depleted and poisoned fuel which has spent a period in the reactor will require a number of radiation detectors. These, for the main, will be alpha-particle detectors, since all fission fuels are alpha-particle emitters, and gamma-ray detectors to appriase the amount of radioactive contamination remaining in fuel to be recycled.

Radiation Technology

Under the category of radiation technology is included tracer and activation analysis, food processing, and chemical processing. Tracer and activation analysis is most demanding of the nuclear radiation detector. In general, as has already been mentioned, a gamma-ray spectrum (particularly when supplemented with a knowledge of the half-life of the elements involved) can serve to identify radioactive elements. Therefore, radioactive tracer work is dependent upon gamma ray energy spectrometers. The principle of these detectors will be discussed below. Activation analysis where elements in a specimen being studied are transmuted into radioactive isotopes by neutron bombardment requires the same study of gamma ray energy spectra as does tracer analysis. It should also be mentioned that a very important and fruitful branch of tracer analysis is concerned with organic chemistry. Radioactive carbon and tritium can be used to tag organic molecules so that their chemical reactions can be studied. Both tritium and radioactive carbon emit very low energy nuclear radiation which is quite difficult to detect. Special techniques have been developed for this purpose.

Active experiments are in progress in the use of nuclear radiation for sterilizing food. Many of these experiments have been successful and it may be expected in the future that this technique will be widely used in food processing. The detector problem here is one of monitoring the radiation to ensure that it is of the level required for sterilization but not so high that it will damage the food. Food thus processed must be examined to ensure freedom from residual radioactive contamination. Also, health physicists will be responsible for safe operating conditions in these radio-processing plants. This will require radiation detecting and monitoring equipment similar to that needed for good housekeeping in nuclear power plants.

The use of nuclear radiation to promote chemical reactions which would otherwise not take place or occur too slowly to be of practical use is proving very promising. Chemical plants employing these techniques will need monitoring equipment very similar to that required in food processing plants. Both will demand reliable, long-lived equipment for monitoring the radiation used in the various processes and sensitive, reliable instruments for use by the health physicists.

Medicine and Biology

The application of nuclear radiation and radioactive elements to medicine and biology appears to be giving very promising results. Nuclear processes can not only be used as a tool in research but appear also to be important in therapy. In discussing detector requirements in this field of science, it is necessary to subdivide it into three categories, namely, therapy, medical diagnosis and biological research.

Encouraging results have been obtained from the use of nuclear radiation in the treatment of a number of diseases, particularly the class represented by tumorsand cancers. Successful treatment has been accomplished by using beams of nuclear radiation which include gamma-ray beams, neutron beams, and, less frequently, proton and alpha particle radiation. By properly focusing the radiation beam and orienting the patient, it is possible to make the radiation exposure a maximum in the diseased tissue and to minimize the exposure to the more sensitive portions of the healthy tissue. In order to do this, detection means must be provided which can serve to indicate the intensity and focus of the radiation. In general, high sensitivity is not required for the detectors, but they frequently must have accurate directional properties. In addition, detectors are required to monitor the environment of the radiation equipment for the protection of the physicians and technicians operating the equipment. Therapeutic results have also been obtained by introducing into the patient radioactive elements which are selectively absorbed by the diseased tissue. For example, favorable results have been obtained with thyroid cancer even after metastasis has occurred. Elaborate detecting equipment may be required in order to map the location of the radioactive material in the patient as the treatment progresses.

Techniques similar to radioactive tracer techniques are frequently very valuable in medical diagnosis. The rate at which the patient absorbs and eliminates specific radioactive atoms or tagged molecules will, in many instances, give the physician valuable information as to the state of a patient's health. By employing detectors in mapping systems, the regions where these specific chemicals accumulate can be determined. This application requires rather special detectors. They must be extremely small in size, must be constructed in such a way as to eliminate the possibility of electrical shock to the patient, and must be both sensitive and reliable. Detectors small enough to make internal measurements are sometimes required. Finally, gamma rays, like X rays, may be used to locate and investigate structural damage within the patient; for example, broken bones, malignant growths, and other lesions.

The detector requirements for biological and medical research are extremely varied. They range from simple gamma-ray detectors to the most elaborate gamma-ray spectrometers and from single element detectors to imaging devices capable of giving high definition pictures of the distribution of radioactive material in a specimen. The breadth of detector requirements is second only to the detector requirements for nuclear physics.

Nuclear Physics

Nuclear physics requires by far the broadest diversity of detector types. This area of science employs a large number of conventional detectors in energy spectrometers and radiation monitoring equipment. In addition, detectors are needed which have the maximum possible energy resolution capabilities, the best time resolution that can be achieved, and, in many instances, the very maximum sensitivity. Detectors must be made to respond to a wide variety of radiation including all of those mentioned in Table I. Frequently, the detector or detector system is asked to identify the type of radiation particles passing through it. Extensive use is being made in this field of detectors which present a two- or three-dimensional display of the path of the particle through various media. These imaging detectors are a field in their own right and will be discussed in some detail below.

Radiation Detectors

Nuclear radiation detectors can be divided into two distinct domains, namely, single element detectors and imaging detectors. The former, as is evident from the discussion above, has wide application in all of the fields discussed. Imaging detectors, on the other hand, are primarily of use in the fields of nuclear physics and of medicine and biology. Even in these two fields, single element detectors dominate both numerically and in importance. Because of this, imaging detectors will be discussed only in sufficient detail to bring out the salient features of the principal types.

Single element detectors may be divided into four major classes. These are listed in Table II. The classification here is by the physical mechanism of detection rather than by the type of radiation to which it is sensitive or by its field of application.

Scintillation Counters

The scintillation counter is named first in the table because it is the most sensitive and versatile detector for many types of radiation and, in particular, for gamma rays, that has yet been devised. Furthermore, its speed of response and time resolution capability exceeds that of any other detector.

The scintillation counter consists of a transparent scintillator optically coupled to a photomultiplier whose

TABLE II

Nuclear Radiation Detectors

Type		Application, Remarks
Scintillation Counter		
Scintillator	Example	
Inorganic crystal	NaI:Tl	General survey, energy spectrometry (with photomultiplier having high quantum efficiency, good statistics)
Inorganic powder screen	ZnS:Ag	α-ray detection (with large area photomultiplier)
Organic single crystal	Anthracene	For combined high speed and energy discrimination
Liquid solution	Diphenyl in Xylene	Good time resolution, coincidence measurements
Plastic	Polystyrene+tetraphenyl-butadiene	Large volume, good time resolution
Neutron sensitized phosphor	ZnS:Ag with Li^6	Neutron detection
Cerenkov scintillator	Non-fluorescent plastic	High-energy physics, good time resolution
Gas Ionization Detectors		
Ionization Chambers		Quantitative radiation measurements, personnel monitoring, area monitoring
Proportional Counter		α-particles (with air or inert gas filling), neutrons (with $B^{10}F_3$, etc. filling)
Geiger-Mueller Counter		Survey work, health physics
Semiconductor Detectors		
p-type silicon with diffused surface junction (phosphorus)		α-particles, high-energy particles, low-energy β-rays
n-type silicon with Au barrier layer		α-particles, high-energy particles, low-energy β-rays
Single crystal (diamond, AgCl)		Experimental
Chemical Detectors		
Photographic		Film badge
Glass or crystal radiation darkening (F-centers, etc.)		High-level radiation monitoring, personnel hazard monitoring
Dye reactions		High-level monitoring
Fluorescent activation of glasses		High-level monitoring

output is connected to electronic circuitry for analyzing and presenting the information obtained from the detected radiation.[6] A schematic diagram of the basic arrangement is shown in Fig. 1. When a particle or photon of nuclear radiation gives up its energy in the scintillator, it produces an excitation which generates a flash of light or scintillation. The optical photons thus generated strike the photocathode of the multiplier causing the release of photoelectrons. These photoelectrons are directed into the electrode structure of the multiplier where their number is increased by a very large factor through the cascaded secondary emission process occurring at the successive dynodes. The amplified space current is collected by an anode or collector electrode as a pulse whose amplitude is proportional to the absorbed energy from the particle or photon of nuclear radiation.

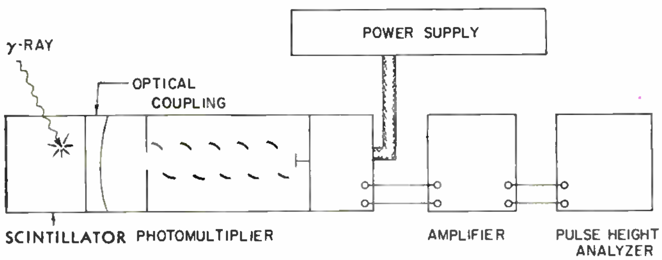


Fig. 1—Scintillation counter.

As indicated in the table, there are a variety of scintillators. The selection depends upon the application for which the counter is designed. For high sensitivity and good energy resolution, thallium activated sodium iodide gives excellent results. For speed and time resolution, a Cerenkov scintillator or a short decay time organic phosphor is employed. Neutron-sensitive scintillation counters are a special problem and will be discussed later. There is also a variety of photomultipliers available designed for different scintillation counter applications.[7]

Gas Ionization Detectors

A large class of nuclear radiation detectors depends for their operation upon the ionization of a gas as a result of the passage of a high-energy charged particle through it. This class may be divided into three types, namely, the ionization chamber, the proportional counter, and the Geiger-Mueller counter.

Basically, the members of this class consist of a chamber filled with an appropriate gas and two electrodes across which voltage is applied. Frequently, the outer wall of the chamber forms one electrode while a small diameter wire through the center serves as the other. The arrangement is shown schematically in Fig. 2(a). The essential difference between the three types in the voltage applied between the two electrodes and the field strength in the filling gas.

The detector functions as an *ionization chamber* if the applied voltage is low.[8] The ions produced by a charged particle are collected on the electrodes and the accumulated charge, measured with an electrometer or other electrical measuring instrument, is a measure of the radiation passing through the chamber.

With an intermediate value of applied voltage, it will be found that the ions released by the passage of the energetic charged particle will be multiplied by secondary ionization. Thus, the collected charge will be proportional to the particle energy, but will be greater than that from an ionization chamber by a factor which may be as large as 100 or even 1000. A detector operating in this mode is called a *proportional counter*.[8,9] The voltage range over which gas amplification occurs is indicated on Fig. 2(b).

If a still higher voltage is applied between the electrodes, the passage of a high-energy charged particle produces a saturated discharge. Provision must be made to quench the discharge either by external circuit means or by the use of a quenching agent (*e.g.*, a higher alcohol or a halogen gas) added to the filling gas. The voltage range over which this type of triggered discharge occurs is known as the Geiger region and a detector operating in this region is termed a *Geiger-Mueller counter*.[9]

The filling gas of a gas ionization detector is, in general, very transparent to gamma radiation. Where these detectors are used for gamma radiation measurements, the detector is designed so that the radiation is

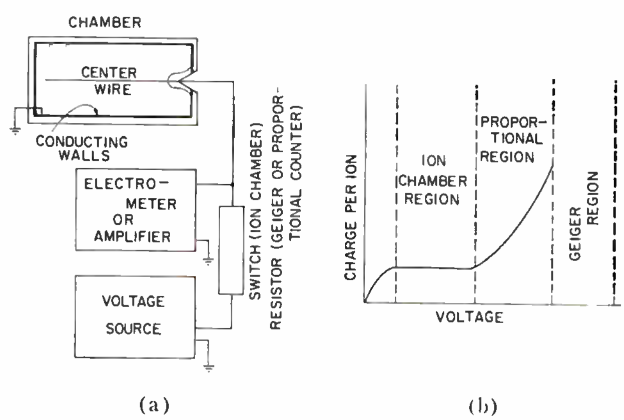


Fig. 2—Gas ionization detector. (a) Schematic diagram. (b) Voltage characteristics.

[6] J. B. Birks, "Scintillation Counters," McGraw-Hill Book Co., Inc., New York, N. Y.; 1953.
G. A. Morton, "Recent Developments in the scintillation counter field," *Proc. Internatl. Conf. on Peaceful Uses of Atomic Energy*, vol. 14, August, 1955. Also in IRE TRANS. ON NUCLEAR SCIENCE, vol. NS-3, pp. 122–137; November, 1956.
[7] Coverage of the field can be obtained from the following: issue on "Proc. of 5th Scintillation Counter Symp.," IRE TRANS. ON NUCLEAR SCIENCE, vol. NS-3, November, 1956; "Proc. of 6th Scintillation Counter Symp.," vol. NS-5, December, 1958, and "Proc. of 7th Scintillation Counter Symp." vol. NS-7, June–September, 1960.

[8] B. B. Rossi and H. H. Staub, "Ionization Chambers and Counters," McGraw-Hill Book Co., Inc., New York, N. Y.; 1949.
[9] S. A. Korff, "Electron and Nuclear Counters," D. Van Nostrand Co., Inc., Princeton, N. J., 2nd ed.; 1955.

absorbed in the walls, causing the release of high-energy electrons into the filling gas of the chamber. Frequently, for example, the wall material will have absorption characteristics resembling that of organic tissue so that the response of the detector is a measure of the biological effect of the radiation. Where high sensitivity to gamma rays is desired in a Geiger-Mueller counter, the walls of the tube are coated with a thin layer of a heavy metal to enhance the useful absorption of radiation.

Ionization chambers are employed for accurate quantitative measurements of radiation. They are also extensively used as monitoring devices to determine the exposure of personnel to radiation. The most frequent application of the proportional counter is for alpha particle detection and for measurement of neutron flux. For this latter, the filling gas is boron trifluoride which has a large capture cross section for low-energy neutrons. The Geiger-Mueller counter is a high sensitivity, simple, easily operated radiation detector. It has been very widely used as a radiation survey instrument to indicate the presence of small amounts of nuclear radiation such as would indicate the presence of radioactive contamination.

Semiconductor Detectors

The operation of a semiconductor detector can best be explained in terms of the solid-state energy diagram given in Fig. 3.[10] The illustration represents a lightly p-type silicon semiconductor with a thin p-n junction close to one surface. The junction is back biased to a point just below that at which appreciable reverse current occurs. An ionizing particle enters the material through the junction and gives up its energy throughout the depletion layer exciting hole-electron pairs. Experimentally, it is found that a hole-electron pair will be excited for each 3.5 ev loss of energy by the particle. The minority carriers (electrons) are drawn through the junction and are collected on the positive electrode. Therefore, a current pulse will be observed every time a particle passes through the material. The amplitude of this pulse (total charge) is proportional to the energy absorbed.

In a practical semiconductor detector, the junction will be a fraction of a micron in thickness while the depletion layer may be a hundred or more microns thick. Semiconductor detectors have been very effective for the detection of heavy charged particles such as alpha particles and are also useful for low-energy beta rays. However, the transparency of the depletion layer is too great for them to be effective as gamma ray detectors. The semiconductor approach to radiation detectors is one of the newest and most promising avenues of development.

[10] K. G. McKay, "A germanium counter," *Phys. Rev.*, vol. 76, pp. 1537–1538; November, 1949. See also issue on "Proc. of 7th Scintillation Counter Symp.," IRE Trans. on Nuclear Science, vol. NS-7, pp. 178–201, June–September, 1960; also issue on "Proc of 7th Annual Natl. Meeting on Solid-State Radiation Detectors," IRE Trans. on Nuclear Science, vol. NS-8, January, 1961.

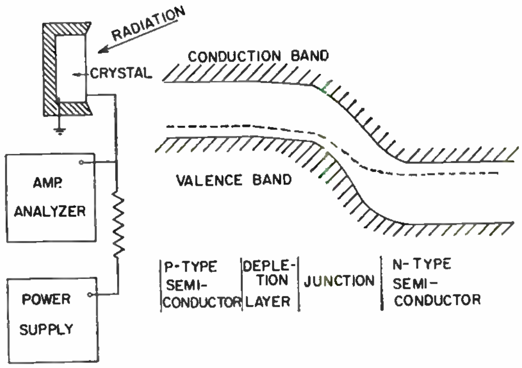


Fig. 3—Semiconductor detector.

Chemical Detectors

As indicated in Table II, there are a variety of chemical detectors. The only one to be mentioned here is the photographic film. Emulsions have been developed which are quite sensitive to gamma rays and whose blackening is proportional to the integrated exposure. Photographic film, in the form of film badges, is universally used as a safety measurement in almost all areas of nuclear work.[11]

Radiation Detector Requirements

A large amount of very important detail is involved in each of the detectors mentioned in Table II. In spite of the fact that this detail is so important that it may determine the operational success or failure of a given detector, it is not the purpose here to discuss this detail, much of which can be found directly or indirectly in the references cited. Instead, the group of radiation detectors as a whole will be considered from the standpoint of some of the more important attributes which they must possess. These are 1) high sensitivity, 2) energy resolution capability, 3) high speed, 4) radiation discrimination, and 5) ability to detect neutrons. The inclusion of neutron detection is, without doubt, a logical misdemeanor; however, the subject is of great importance and should be singled out for emphasis.

Sensitivity

Detectors should be available which will detect every particle or photon of nuclear radiation. The scintillation counter satisfies this requirement for most radiation, including gamma rays, provided that the area over which detection must be accomplished is not too great and that the particle energy is not less than a few kev. With an efficient scintillator, (*e.g.*, NaI:Tl), a good photomultiplier and good optical coupling between the two, each photoelectron released from the multiplier cathode requires 300 ev energy from the nuclear particle. Since this is a statistical quantity, the nuclear particle

[11] G. M. Corney, "Photographic monitoring of radiation," in "Radiation Hygiene Handbook," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 10-53 to 10-75; 1959.

must have several times this energy in order to ensure a reasonable certainty that at least one photoelectron will be released. Therefore, if the photomultiplier generated no spurious "dark" pulses, a scintillation counter could detect each particle of nuclear radiation if the energy were above 1000 to 1500 ev. However, a photomultiplier does generate "dark" pulses due to thermionic emission from the photocathode and other nonfundamental processes. The thermionic emission from a CsSb (S11) photocathode at room temperature is about 1000 electrons/sec/cm². The gain process in a multiplier is, in itself, statistical and the pulse output from single electrons has an exponential pulse height distribution. Because of this spurious pulse background, reasonable certainty of single particle detection requires that the particle energy be 10 to 20 kev. The "dark" pulse rate corresponding to single electrons can be reduced two orders of magnitude by cooling the multiplier to the point where thermionic emission is negligible and the limit is set by other nonfundamental sources of "dark" pulses. As the development of photomultipliers proceeds, it is expected that the nonfundamental "dark" pulses can be materially reduced. However, even if this improvement is made, it is clear that the scintillation counter principle probably can never yield a detector approaching the ideal. The solution more probably lies in the direction of semiconductor detectors, and it may eventually be possible to develop a sensitive element which has a large thermal excitation energy and a nuclear particle excitation energy more nearly equal to the thermal excitation energy.

Alpha-particle detection, because of the low penetrating power of the radiation, presents the problem of an optically opaque window which is not easily solved for a scintillation counter. Here, proportional counters have been very effective. Recent developments in solid-state counters have resulted in outstandingly successful alpha-particle detectors.

Energy Resolution

The ability to measure particle or photon energy is, for many detectors, an important requirement. This is particularly true for gamma-ray detection since the gamma-ray energy spectrum (wavelength spectrum) of radioactive nuclei is quite analogous to the visible light spectrum of excited atoms and is very useful for element identification. Scintillation counters are, at present, the most effective detectors for measuring gamma ray energies. The gamma ray is absorbed in the scintillator by one of three processes, namely, photoelectric absorption, pair production, or Compton collision. This gives rise to one or more charged particles having the energy absorbed from the radiation. These charged particles, in turn, excite fluorescence in the scintillator. Assuming total absorption of the gamma-ray photon, the number of visible light photons and the photoelectrons in the multiplier will, on the average, be proportional to the gamma-ray photon energy. However, the processes involved are statistical and it can be shown that there will be a root-mean-square error ΔE_{rms} in the measured value of gamma-ray energy E given by

$$\Delta E_{rms} = \beta\sqrt{E\epsilon},$$

where ϵ is the energy required to excite the photoelectron and β is the statistical loss in the scintillator crystal and photomultiplier ($\epsilon = 300$ ev, $\beta = 2$ for a scintillation counter using NaI:Tl). With a 1-Mev gamma ray, the rms error will be about 3 per cent.

In order to improve the energy resolution, it would be necessary to devise an excitation process requiring much less energy. Solid-state detectors meet this requirement. It has been found for electrons and alpha particles that a hole-electron pair is excited for each 3.5 ev energy loss of the incident particle. However, the present solid-state detectors are too transparent to be useful for gamma-ray detection. They are very effective for energy measurements of alpha particles and root-mean-square deviations of 0.2 to 0.3 per cent have been observed for alpha particles in the 5- to 7-Mev energy range. The resolution limit is not, in this case, set by the statistics of the excitation process but rather by amplifier noise.

Further work with semiconductor detectors may overcome the two problems indicated, namely, obtaining sufficient absorption to make them effective for gamma-ray energy measurements and the incorporation of a gain process so that the energy resolution limit is set by the fundamental statistics of the conversion process rather than by amplifier noise.

Speed

High-speed capabilities are required of many radiation detectors. Speed here is used in two connotations; one is that the detector be capable of extracting the information (*e.g.*, energy) from nuclear particles when their rate of arrival is very high, the second is that it be capable of resolving the arrival of two nuclear particles separated by a small time interval. Technically, the two place very similar requirements on the detector. At present, scintillation counters with appropriately selected scintillators and photomultipliers have the best time resolution capabilities and speed of any detector known.

The time resolution depends primarily upon the fluorescent decay characteristics of the scintillator and upon the interdynode transit time spread in the early stages of the photomultiplier. With a fast organic scintillator and a commercially obtainable high-speed multiplier, time resolutions of a fraction of a nanosecond can be achieved. Special experimental multipliers have been built with a transit time spread reduced by almost an order of magnitude.

A great deal of effort has been spent on developing electronic circuitry for time measurements. The application of tunnel diodes now plays a major role in these fast circuits. The circuit development includes coincidence, delayed coincidence, and anticoincidence circuits as well as other special circuits designed for time interval determination.

Up to the present, solid-state detectors are not competitive with the scintillation counter for speed and time measurements. However, these detectors are in their infancy and ways may be devised to make them competitive or superior as time measuring detectors.

Radiation Discrimination

A frequent requirement for nuclear radiation detectors is that they be capable of detecting a small amount of radiation of a particular type in a background of other nuclear radiation. One way of achieving discrimination is by making the sensitive elements of the detector responsive to the wanted radiation and relatively insensitive to the background. The second is for the detector to respond to both wanted radiation and background and to obtain sufficient information from each so that the wanted radiation may be selected on the basis of this information. An example of the former is the boron trifluoride proportional counter for neutron detection. The filling gas is insensitive to γ rays and by properly selecting the geometry and materials of the walls, the γ ray response of the detector can be made small. However, the boron (B^{10}) of the filling gas is very neutron sensitive. As an example of the second method, it might be required to detect a high-energy γ ray in a background of low-energy γ rays. A scintillation counter would be employed and its energy resolution capability would be used to separate the wanted radiation from the background. However, the speed of the detector must be sufficient so that the probability of arrival of two photons or more of background radiation during the interval required for the detector to measure particle energy is small. Both of these methods are used either singly or in combination in the problem of discrimination.

The present type of semiconductor detectors with relatively small depletion layers exhibits good selectivity for heavy charged particles against a gamma-ray background. This is because the absorption of the latter in the depletion layer is quite small.

In addition to the two methods mentioned above, coincidence and anticoincidence detectors can frequently be used to aid radiation selectivity. While these and other methods are at present applied with good results, a general solution to the problem has not been achieved as yet and much research remains to be done in this direction.

Neutron Detection

As has been pointed out earlier, neutron detection is fundamental to the development of practical nuclear power. Fairly sensitive neutron detectors have been developed both for fast neutrons and slow or thermal neutrons.[12] The former are based on the impact exchange of energy between a fast neutron and the protons in hydrogenous material. The high-energy protons can excite luminescence if the hydrogenous material is the phosphor of a scintillation counter. Slow neutron detection is efected by using certain isotopes such as B^{10}, Li^6, or U^{235} which transmute after absorbing a neutron with the release of considerable energy. This energy is in the form of kinetic energy of the charged nuclear fragments and can produce ionization or excitation in the medium through which they pass. A europium or samarium activated lithium iodide crystal, for example, makes a fairly satisfactory scintillator for a neutron detector. Similarly, boron trifluoride in a proportional counter is a very effective neutron detector. Finally, Li^6 sandwiched between two semiconductor counters has also given good results.

The principal problem, however, is not that of obtaining sensitivity to neutrons but is due to the fact that most practical neutron detectors must be used in the presence of a very large gamma-ray background. This is particularly true of neutron detectors used in reactor technology. Therefore, the ability of the detector to discriminate against gamma rays is vital. A good deal of progress has been made in achieving adequate discrimination. However, the problem cannot be considered completely solved.

A second problem encountered in the application of neutron detectors to reactor technology is due to the extreme range of flux levels over which the detector must operate. At reactor start-up the neutron flux and rate of change of flux must be measured rapidly and precisely to determine the reactor period. Under these conditions, the flux may be less than 10^5 n/sec/cm². Under full operation, on the other hand, the flux may rise to a value of 10^{14} or more n/sec/cm². An ideal detector must be sensitive enough to make the low level determination but yet be stable and long-lived under the high flux condition. This problem has not been solved and, at present, retractable detectors, multiple detectors and similar expedients are used for reactor control. A possible direction of development of neutron detectors may be toward the use of liquid or gaseous neutron sensitive materials which can be made to flow through the active element of a scintillation counter or semiconductor detector.

[12] C. O. Muehlhause, "Neutron scintillation counters," IRE Trans. on Nuclear Science, vol. NS-3, pp. 77–82; November, 1956.
R. B. Murray, "Use of Li^6I:Eu as a fast-neutron detector and spectrometer," IRE Trans. on Nuclear Science, vol. NS-5, pp. 159–160; December, 1958.
R. B. Owen, "The variation of phosphor decay time with specific ionization and its applications," *Proc. of Symp. on Nuclear Electronics*, Paris, 1958, Internatl. Atomic Energy Agency, Vienna, Austria, vol. 1, pp. 27–35; June, 1959.
T. A. Love and R. B. Murray, "The use of surface-barrier diodes for fast-neutron spectroscopy," IRE Trans. on Nuclear Science, vol. NS-8, pp. 91–97; January, 1961.
W. D. Allen, "Neutron Detection," Philosophical Library, Inc., New York, N. Y.; 1960.

From the above outline of single element nuclear radiation detectors, it will be recognized that while great progress has been made in developing the detectors and their associated electronic equipment to the point where they answer in a fairly staisfactory way many of the problems relating to the measurement of radiation, there is, nevertheless, need for further research and development. Areas in which improvements are urgently needed are in the detection of neutrons, the better measurement of particle energy, particularly for gamma rays, and increased detection efficiency of devices capable of very high time resolution. Of course, there is always need for improvement in reliability, stability, and simplicity of equipment. On the one hand, there is need for very large detectors for work in nuclear biophysics and for survey problems and, on the other, for very small efficient detectors for medical and biological applications. The research paths for achieving some of these objectives are fairly well marked; however, others will depend upon new inventions and discoveries.

Imaging Detectors

The final class of detectors to be reviewed is that of the imaging detectors. It is convenient to divide the imaging detectors into two groups. The first are those which display a two-dimensional pattern representing the distribution of nuclear radiation or radioactive material. This type of imaging finds widest application in the biological and medical field. The second is particle track imaging.

Area Imaging

Photographic: The most frequently used means of area imaging is the photographic film. All ordinary photographic film is sensitive in varying degrees to nuclear radiation. However, special emulsions less sensitive to visible light and more sensitive to nuclear radiation have been developed. At best the sensitivity to gamma rays is not very high, and frequently intensifier screens, in the form of sheets of a heavy metal, or fluorescent screens are used.

A typical example of the use of this form of imaging radiation detector is radioautography.[13] It is used to determine the distribution of radioactive materials absorbed in an object. For example, a fertilizer may be tagged with a radioactive atom and fed to growing plants. Later, the plants are laid against photographic films and allowed to remain long enough to produce an exposure. The degree and location of blackening indicates the distribution of the tagged molecules within the plants. This technique, suitably modified, can be used with small animals and other biological objects.

Scintillation Counter Scanning: The second imaging technique which has been successfully employed, particularly for large objects, is to scan the specimen with a scintillation counter whose angle of sensitivity is limited by a system of apertures. The output of the scintillation counter is recorded as a series of dots by a pen which moves over the recording surface in synchronism with the motion of the scintillation counter. This technique has been used, for example, to trace out the distribution of radioactive iodine in a patient suffering from a metastasized thyroid cancer. Imaging experiments have been carried out with arrays of scintillation crystals which are viewed with an electronic image intensifier. This work is still exploratory, and its usefulness cannot at present be forecast.

Track Imaging

Wilson Cloud Chamber: The second group of imaging detectors displays the paths of nuclear particles through matter. The earliest of this group was the Wilson cloud chamber.[4] This detector consists of a container filled with a gas having a high moisture content arranged so that the gas can be rapidly expanded. This rapid expansion causes the gas to become supersaturated with the vapor. A nuclear particle passing through the gas will produce a track of ions which serve to nucleate droplets of water so that the particle track becomes visible as a thin thread of condensed liquid. By providing suitable windows for illuminating and photographing these mist trails, the paths of particles can be accurately determined. Frequently such chambers are immersed in a strong magnetic field which will bend the particle trajectory. Observation of the density of the mist trail and its curvature under a known magnetic field gives information as to the particle energy, momentum, mass, and charge. Much valuable information in the field of nuclear physics has been obtained with the Wilson cloud chamber.

Bubble Chamber: A second type of track imaging device which had its origin within the last decade is the babble chamber.[14] The sensitive medium is a volume of liquid super-heated by decompression. The passage of a high-energy charged particle through such a volume of liquid leaves a trail of ions which nucleate bubbles of vapor. Sharp, clear tracks are obtained in this way. Of course, these tracks can only exist for a small fraction of a second before the entire volume of the liquid starts to boil. However, by using a high-speed light flash and recording camera, excellent tracks of considerably higher resolution than is possible with the Wilson cloud chamber can be obtained. The high resolution and high density of the sensitive medium gives these bubble chambers a decisive advantage over the earlier track imaging device. Obviously, the equipment required for the bubble chamber is much more elaborate than that for the Wilson cloud chamber. There are at present under construction several very large volume bubble chambers which give promise of contributing very

[13] H. Yagoda, "Radioactive Measurements with Nuclear Emulsions," John Wiley and Sons, Inc., New York, N. Y.; 1949.

[14] H. Bradner, "Bubble chambers," *Ann. Rev. Nuclear Sci.*, vol. 10, pp. 109–160; 1960.

valuable information about the physics of very high-energy particles.

Nuclear Emulsions: A third method of obtaining particle track images is through the use of nuclear emulsions. When an ionizing particle passes through a thick photographic emulsion (especially prepared fine-grained emulsions are used for the purpose), a trail of exposed grains marks its trajectory. These tracks are generally microscopic in size and therefore microphotography is used to record them. Nuclear emulsions have proved particularly useful in showing the interaction of high-energy particles with nuclei of matter. When a nucleus is struck by a high-energy particle, it frequently disintegrates releasing a variety of high-energy particles which leave trails in the emulsion. Analyses of such "stars" have yielded a great deal of information about the creation of new elementary particles as a consequence of energetic collisions.

Spark Chamber: A track imaging system which is presently receiving a good deal of attention and may be the solution to problems requiring low resolution and very large sensitive volume also has the advantage of being simple both in concept and execution. This is the spark chamber.[15] It consists of an array of parallel metal plates with a gas, (*e.g.*, argon, air, etc.) at about one atmosphere pressure between them. If a potential is applied between the plates giving a field strength of a few thousand volts per centimeter, a spark will be formed when a high energy ionizing particle passes through the system of plates. The spark appears to originate at a point where the particle passes through the negative plate and may either follow the particle path or go normal to the surface of the plate (giving a stepped path). The spark paths can readily be photographed from two directions to give a stereo pair.

Scintillation Track Imaging: The final track imaging technique to be considered is scintillation track imaging.[16] When a high-energy charged particle passes through a scintillator, it produces a trail of excitation which, in turn, produces a luminous path or track. The luminosity of the track is so low that it cannot be seen with the unaided eye. However, by the use of electronic image intensifiers, excellent tracks can be obtained.

Two types of scintillators have been used for scintillation track imaging. One is the homogeneous block of scintillator material implied in the above paragraph. The second consists of a bundle of scintillator fibers. As a high-energy particle passes through the bundle, it produces light in each fiber traversed which is then transmitted to the end, the fiber serving as a light pipe. Therefore a projection of the path will appear as a luminous track on the ends of the fibers. Again, electronic intensification is required to bring the brightness of the tracks up to a point where they can be seen or photographed.

Two types of electronic intensifiers have been used for track imaging. One of these, the multistage intensifier image tube, presents the enhanced image on a phosphor screen at the end of the tube where it can be seen and/or photographed. The second is the intensifier orthicon where the reproduced track information is in the form of a video signal which can be reconstructed into a visible image by conventional television techniques, fed directly to computing equipment, or stored on tape for future reference.

Although much research has been done on electronic intensifiers, they are still in their infancy. Basically, the intensifier consists of a primary photocathode and electron current amplifying intensifier screens with electron optical systems between them to image the electrons. The electron image from the intensifier screens is focused onto an orthicon target for video reproduction or on a phosphor screen for direct viewing. Two types of intensifier screens have been used. The first consists of a thin phosphor film in close optical contact with the photocathode. Current gains of 50 to 80 are obtainable. The second is a thin insulator or semiconductor film capable of giving transmission secondary emission. Gains of 4 to 6 are generally obtained with such screens. For track imaging, cascaded intensifier screens are employed since total current gain required is 10^3 or 10^4. Scintillation track imaging appears to offer many advantages but it is too early to predict the full extent of its usefulness.

Conclusion

This paper has attempted to summarize the problem of nuclear radiation detection as it stands today, both in terms of the requirements for radiation detection and the instruments available. It is clear that the development of nuclear radiation detectors has a long road ahead of it before the field is complete. Even at the present rate of development, it will be many years before practical detectors will be able to accomplish all that can be theoretically predicted for them. Many advances in both single element and imaging detectors will come through application of semiconductor and solid-state principles to the detector field. It can be forecast with a great deal of confidence that the development of detectors will continue at its present high rate and that in the future not only will there be major improvements along the lines being investigated at present but also there will be new discoveries which will make radical changes in the field.

[15] "Spark Chamber Symp. Proc." *Rev. Sci. Instr.*, vol. 32, pp. 479–531; May, 1961.

[16] E. K. Zavoisky, *et al.*, "Luminescent camera," *Dokl. Akad. Nauk. SSSR*, vol. 100, pp. 241–242; February, 1955.
G. T. Reynolds and P. E. Condon, "Filament scintillation counter," *Rev. Sci. Instr.*, vol. 28, pp. 1098–1099; December, 1957.
See also IRE Trans. on Nuclear Science, vol. NS-7, pp. 115–158; June–September, 1960.

Nuclear Reactor Plant Kinetics and Control*

M. A. SCHULTZ†, SENIOR MEMBER, IRE, AND J. N. GRACE†, MEMBER, IRE

Summary—**Radio and control engineers have contributed significantly to the understanding of nuclear reactor stability and transient response over the last 15 years. The successful application of frequency response and analog simulation techniques to reactor kinetics problems, together with experimental confirmation, has allowed increased dependence on inherent feedback mechanisms for self-regulation and self-protection. Digital techniques are becoming increasingly important in transient analysis and in automation of reactor systems. Modern large reactors employing boiling coolant and/or natural circulation require thorough study of reactivity feedback, hydrodynamic effects, and spatial stability of the power distribution to assure successful performance. Many important areas exist for future analytical and experimental research, including prompt feedback mechanisms during fast excursions, stochastic fluctuations, reactor noise, and nuclear-thermal-hydraulic interactions.**

Introduction and Scope

THE DEVELOPMENT of reactor kinetics and control had to follow the development of basic nuclear reactor theory. The discovery of fission by Hahn and Strassman[1] in the late 1930's and the intensive exciting work that followed immediately thereafter on chain reactions indicated that a reactor would have an exponential response, and hence was going to be difficult to control. The role of the delayed neutron was appreciated in 1939, and indicated the possibility of setting up a pile or a reactor whose response was based primarily upon the control of the delayed neutrons. This reactor would still be an exponentially responding device, but one whose kinetic response would be slow and more easily restrained.

Early wartime production reactors were built and in some, such as the Hanford reactors, it was discovered that the reactors would run for a while and then shut themselves off. This effect was soon determined by E. P. Wigner and his associates to be caused by the kinetic buildup of xenon 135 acting as a reactor poison. Once the effect was understood, it was possible to design around it.

In this same era, temperature effects in reactors were hypothesized and observed. It was fortuitously noted that many reactors possessed a negative temperature coefficient of reactivity; that is, as the temperature of the reactor rose, the reactor had the tendency to shut itself off.

Up until the late 1940's, the kinetics of the few early reactors were accomplished by the same physicists who did the static criticality calculations. Interest in the kinetics of reactors by engineers did not come about until the art reached the stage where the responsibility of reactor and nuclear plant design began to transfer from the physicist to the engineer. Then the radio engineers who were diverted to the instrumentation and control problems of the plant began to realize that to test the existing kinetic theories on actual reactors could be disastrous. Hence, some form of electronic analog was needed to check reactor response safely in the laboratory and to determine the effects of changing variables.

Once the radio engineers became seriously involved in the reactor control field, they quickly recognized that xenon poisoning effects, temperature coefficients, and other effects could be treated as internal feedback in the reactor, and that the then existing feedback theory could be used to handle reactors as well as amplifiers and other circuits. The success of this approach was so pronounced that as the reactor industry grew up and engineering concepts and methods took over the design and construction of reactors, the transplanted radio engineer became a recognized and acknowledged partner in the reactor design. In this role, he was then able to follow and, in some instances, lead the parallel radio, TV and computer industries and adapt new techniques in the electronics field over into reactor design philosophy and calculations.

This article will present the principal historical developments of the industry from the reactor and plant kinetics point of view, and then elaborate on some of the current and future developmental areas.

Principal Developments in the 1940's

It was recognized that the nuclear reactor was mathematically a very complex device, particularly if one were to treat the kinetic effects on a spatial basis throughout the reactor. Soodak[2] first indicated what the response of the reactor would be if all of the spatial effects were lumped and the reactor was considered as a

* Received by the IRE, July 17, 1961.

† Westinghouse Electric Corporation, Bettis Atomic Power Laboratory, Pittsburgh, Pa.

[1] O. Hahn and F. Strassman, "Uber des Nachweis und das Verharten der bei der Bestrahlung des Urans Mittles Neutronen Entstehendes Erdaltalimetalle," *Naturwiss*, vol. 26, pp. 11–15; January 6, 1939.

[2] H. Soodak and E. C. Campbell, "Elementary Pile Theory," John Wiley and Sons, Inc., New York, N. Y.; 1950.

simple entity, all portions behaving the same way at the same time. The equations for the response of the neutron level in a reactor to any change in reactivity were:

$$\frac{dn}{dt} = \frac{\delta k - \bar{\beta}}{l^*} n + \sum \lambda_i C_i + S \quad (1)$$

$$\frac{dC_i}{dt} = \frac{\bar{\beta}_i n}{l^*} - \lambda_i C_i, \quad (2)$$

when:

n = the neutron level or the power level

l^* = the mean effective lifetime of a neutron in a given reactor design

δk = the reactivity (crudely the amount the multiplication factor of the reactor differed from 1)

S = an internal source of neutrons

$\bar{\beta}$ = the effective fraction of the total neutrons that are delayed

$\bar{\beta}_i$ = the effective fraction (accounting for different birth energies) of the total neutrons that are delayed neutrons in the ith group

C_i = the concentration of the delayed neutron precursors of the ith group

λ_i = the decay constant of the ith group of precursors.

The solution of these equations for the neutron level as a function of time after a step change in reactivity was quickly recognized to be a series of exponentials of the form

$$n(t) = n_0 \sum_{J=1}^{j=i+1} A_j e^{P_j t}, \quad (3)$$

where the first exponent had the same sign as δk, the input disturbance, and where the other exponents were negative.

From this starting point, work went ahead concurrently in a number of laboratories. At the Oak Ridge National Laboratory, Bell and Strauss[3] took these equations and built up the analogs for each term, thus constructing the first reactor kinetics simulator. This crude device, containing an early form of operational amplifier, dispelled a great deal of the fear of the ability to control a reactor with an automatic control. This work turned into the preliminary design of the control system for the LITR and MTR.

Similarly, at the Knolls Atomic Power Labs., Piggott and Owens[4] constructed a simulator, and also began manipulating Soodak's equations toward more familiar engineering forms. A chart was constructed which gave the coefficients and exponents of (3) for input values of step reactivity for reactors having different values of l^*.

In addition to the observation that for any one reactor type there was a definite relationship between A_j, P_j and δk, Piggott and Owens noted that the Laplace transform of (3) was of the form

$$n(s) = n_0 \sum_{j=1}^{j=i+1} \frac{A_j}{s - P_j} \cdot \quad (4)$$

Hence, some type of transfer function or frequency response of the reactor could be obtained. They indeed derived such a transfer function for five groups of delayed neutrons, thus again indicating the parallel-form of development between the reactor and the TV and radar fields. This derivation occurred within a few years of most of the basic work on network theory relating pulse output shapes to frequency and phase response of networks. Franz,[5] using the same approach, in 1949 refined this work to cover the presence of six groups of delayed neutrons. The first experimental measurement of the reactor transfer function was done later at the Argonne National Laboratories.

Once the reactor transfer function was derived, engineering proceeded at an accelerated pace. Pagels at Bettis simplified the reactor simulator by using conventional computing operational amplifier techniques and even built a small portable simulator for demonstration purposes.[6] Concurrent with this work, Harrer and Deshong at Argonne National Laboratories had tied an automatic control system around the reactor simulator and began to accumulate data on the peak and shape of the transient response of the controlled reactor as a function of the external control system parameters.

The reactor by this time had evolved into a relatively simple device. The control engineer approximated Soodak's equations further by ignoring the energy dependence of the delayed neutrons. The internal source was usually ignored as at reactor power levels it was swamped out by the fissioning neutrons. Much of the early intensive kinetic work in the late 1940's involved uranium 235 as the fuel. Therefore, β, β_i, λ, and λ_i were all to be fixed immutable quantities. Consequently the control engineer's concept of a reactor at this time was that of a simple four-terminal network black box in which the only important variable left to the reactor designer was l^*, the mean effective lifetime of the neutrons. The reactor thus had the transfer function block diagram of Fig. 1, and for all practical purposes could be treated as a network component. Pagels'[7] version of the reactor simulator in use in the late 1940's is given in Fig. 2. It will be noted that the only basic peculiarity of these systems was the ever-present nonlinearity in

[3] P. R. Bell and H. Strauss, "Electronic pile simulator," *Rev. Sci. Instr.*, vol. 21, pp. 760–763; August, 1950.

[4] J. Piggott and J. Owens, Knolls Atomic Power Lab., Schenectady, N. Y., Rept. J10-1; 1948.

[5] J. P. Franz, "Pile Transfer Functions," Atomic Energy Commission, Washington, D. C., AECD-3260; 1949.

[6] W. Pagels, "A portable kinetic simulator," *Trans. AIEE*, vol. 70 (*Commun. and Electronics*, no. 2), pp. 1422–1426; June, 1951.

[7] M. A. Schultz, "Control of Nuclear Reactors and Power Plants," 2nd ed., McGraw-Hill Book Co., Inc., New York, N. Y., p. 431; 1961.

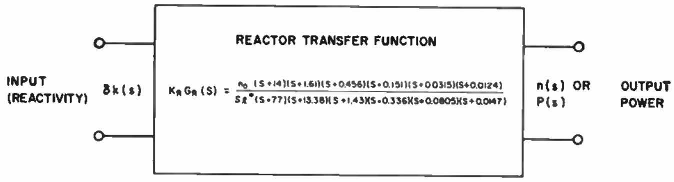


Fig. 1—Reactor transfer function block diagram for early U-235 reactors.

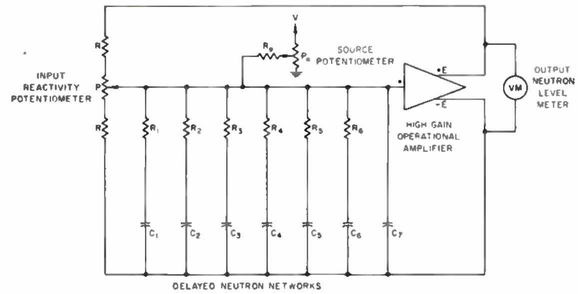


Fig. 2—Reactor kinetic simulator.

which the gain of the reactor depended on the level of operation. In most active networks within the control engineer's experience, linearity might be present over a limited range and then some form of nonlinearity such as saturation would take place. Here the nonlinearity was such that the gain of the reactor, in terms of the ratio of power output to reactivity input, was always a function of the level at which the reactor operated.

Progress in the Early 1950's

After the war, the transfer function concept took hold with the reactor engineers as it did with servo engineers in all other fields. The reactor physicist remained somewhat skeptical of the entire process. It was not until Harrer, *et al.*,[8] obtained the transfer function of an actual reactor that this skepticism disappeared. Their experiment was performed on the CP-2 reactor at the Argonne National Laboratories, and consisted of oscillating a neutron absorbing control rod in and out of the reactor in a nearly sinusoidal manner. They recorded both the position of the control rod and the neutron output of the reactor as a function of frequency and assumed for small excursions that rod position and reactivity were directly related. The transfer function thus obtained compared almost perfectly with the theoretical model and hence with the simulated electronic model.

The faith in the reactor simulator having been vindicated, work could then proceed to refine the automatic reactor control system. A typical block diagram of an automatic reactor control system in the early 1950's is shown in Fig. 3. The only unique feature of this loop is

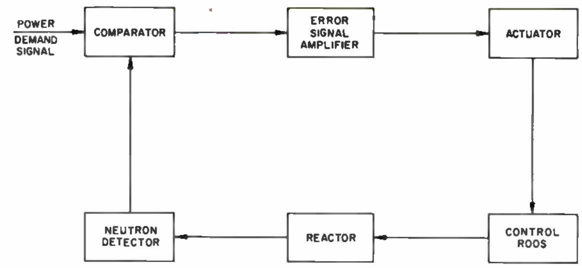


Fig. 3—Reactor automatic control loop.

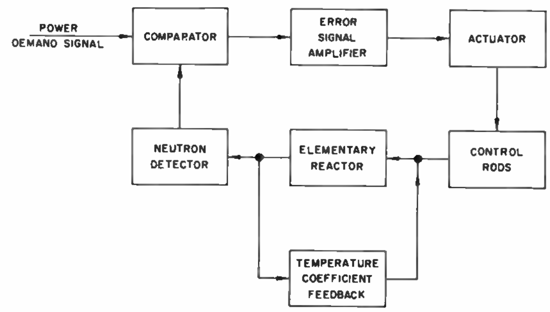


Fig. 4—Reactor control loop with temperature coefficient feedback.

in the comparator. Here, because of the reactor nonlinearity that its gain depended upon its level, a series circuit was introduced in which the gain was inversely proportional to the level. Thus, the nonlinear embarrassment was canceled out. This type of control system was studied intensely in many laboratories, and all of the external parameters were optimized for various reactors.

All of this work was done for plain, simple U-235 reactors with only the internal parameter l^* being somewhat variable. The temperature coefficient effect now came to be recognized as a servo problem, and it was added to the system of Fig. 3 as an internal feedback network around the reactor as indicated in Fig. 4. It was now necessary to optimize the automatic control system again. J. C. Simons and his group[9] at Bettis did extensive simulator work in which they showed that for most practical reactors the temperature coefficient controlled the peak of a fast transient, but the external control system was responsible for the duration and shape of the response. The temperature coefficient in these studies was a lumped coefficient acting with a lag expressed as a simple time constant.

The first attempt to approximate plant feedback effects through the temperature coefficient was done by Henry[10] in 1950. He set down some 13 first-order differential equations describing the thermal and hydraulic behavior of a pressurized water reactor and indicated the means whereby a temperature change affected the reactivity. A transport time delay was then introduced

[8] J. M. Harrer, *et al.*, "Transfer functions of the Argonne CP-2 reactor," *Nucleonics*, vol. 10, pp. 32–36; August, 1952.

[9] J. C. Simons, *et al.*, unpublished work; 1952.

[10] A. F. Henry, unpublished work; October, 1951.

to get this heat to the heat exchanger. Then the equations of the heat exchanger were written for both its primary and secondary sides to indicate how the heat transferred to the working load. The cooled primary coolant was then returned to the reactor via another transport delay in the piping. The average temperature in the reactor was obtained by averaging the returned inlet temperature to the reactor with the outlet temperature leaving the reactor. It was this average temperature that was used to affect the reactivity through the temperature coefficient.

The 13 simultaneous equations describing this system were set up for digital solution. This solution showed that such reactor plant systems were stable.

The reactor control engineers at the time immediately recognized that here was a means of putting more fine structure into the origin of the temperature feedback, in that it now could really have two detailed parts: a local part which acted very quickly, and a delayed part which contained some form of positive feedback for an increase in reactor outlet temperature would result in an increase in inlet temperature. As for stability, servomechanism theory had by now provided a large number of tools which were far more revealing than a single digital computation for the solution of such problems.

The first reactor plant simulator was constructed at Bettis in 1951 to study further the feedback effects on plant transients. This simulator was of the block diagram form of Fig. 5 and did not contain any automatic control loop around the reactor. Stubbs and Single contributed the first all-pass variable transport delay circuit. It was at once apparent that the external plant constituted an inherent control loop and that if the reactor had a negative temperature coefficient, the loop could easily be stable and, in addition, the plant would be self-regulating and follow the power demanded of it.

This situation can be seen as follows. Let the plant of Fig. 5 operate at a constant output load. Now suppose that more output is required by the load caused, for example, by opening the throttle to the turbine. This greater loading causes more heat to be extracted from the heat exchanger, and, for a short period of time, the heat capacity of the heat exchanger and coolant can usually supply the additional load. However, this extra energy extracted from the system requires that the temperature of the coolant into the reactor must drop. If the reactor has a negative temperature coefficient, the reactivity will increase. If the reactor was initially in a critical state, it now temporarily becomes supercritical. The output temperature of the coolant rises, and more energy is then available from the reactor. Finally, in the steady state the reactor returns to its critical condition with the average coolant temperature the same as it was initially. It will be noted that, without any external control mechanism whatever, the reactor system has stabilized itself about a given average temperature and automatically supplies the new power demand placed upon it.

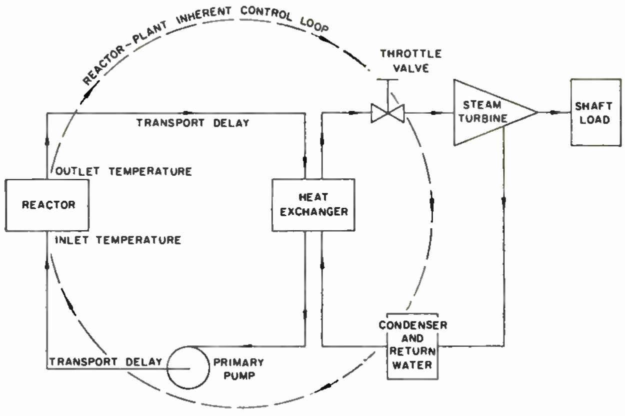


Fig. 5—Block diagram of reactor and associated plant showing inherent control loops.

Analog simulation of reactor plant transients successfully demonstrated the inherent self-regulating and self-protecting characteristics afforded by a negative temperature coefficient coupled with plant feedback. This initiated the exploitation of inherent control as a means of simplifying automatic control of power demand and accident protection. The methods of analysis used were 1) frequency response techniques, for stability analysis and control system synthesis, and 2) analog simulation, for reactor plant transients.

The development of frequency response techniques for application to nuclear reactor plants was the natural inclination of radio engineers familiar with the principles of feedback theory. In 1952, G. S. Stubbs, *et al.*,[9] calculated transfer functions for each reactor plant component using the small perturbation treatment. These transfer functions were then used in both the analysis of inherent feedback loops and the synthesis of automatic control systems, and were supplemented by simulator studies of transients.

The inherent feedback mechanisms which required analysis, in addition to the temperature coefficient, were the effects of 1) system pressure, and 2) fission product poisons (principally xenon) on the reactivity, as indicated in Fig. 6. Understanding of these effects was required before inherent control could be assumed.

Pressure feedback was found to depend on the coupling between the primary loop and the pressurizer. A resonance occurs, illustrated in Fig. 7, which depends on the inertia of water in the surge pipe. Too little damping results in sufficiently strong feedback through a pressure coefficient of reactivity that self-sustaining reactor power oscillations can exist, at a frequency of typically 1 cps. The resonant peak is reduced and the system stabilized simply by orificing the surge pipe.

Xenon poisoning had already been found experimentally to introduce a significant transient control

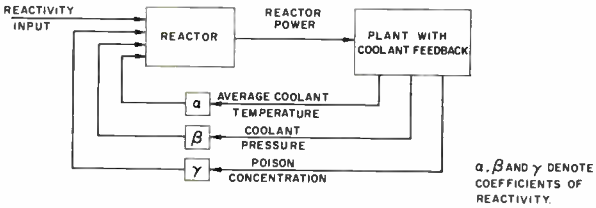


Fig. 6—Inherent reactivity feedback loops.

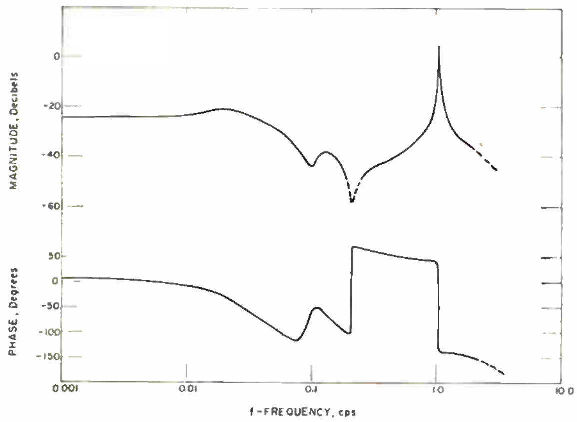


Fig. 7—Loop gain through pressure reactivity feedback (with T_w feedback). Full flow, full power, 4-in surge pipe.

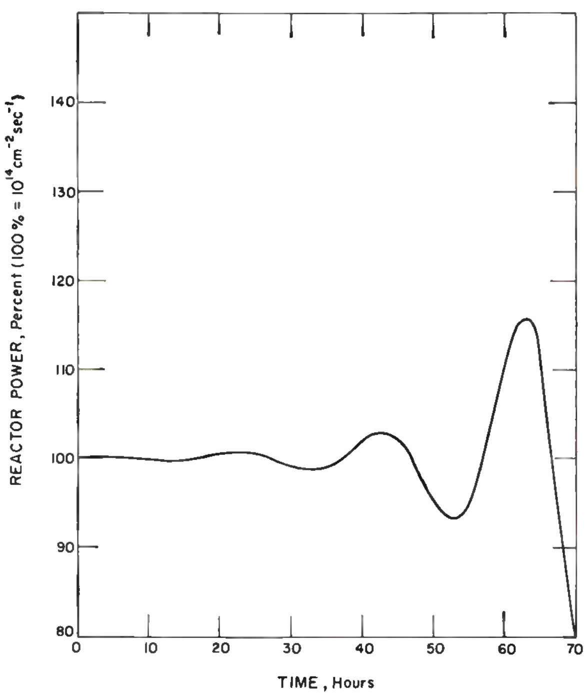


Fig. 8—Reactor power vs time. Average flux $=10^{14}$ cm^{-2} sec^{-1}. Actual gain $=1.1\times$ "critical gain."

problem requiring increased excess reactivity and more control rods. However, it remained to be examined as to whether self-sustaining oscillations could ever occur, although they had never been observed experimentally. A frequency response analysis[11] of the problem revealed that such systems could oscillate when the negative temperature coefficient of reactivity is not large enough (see Fig. 8). However, the frequency of oscillation is typically only 1 cycle per day, which is easily controlled by very slow automatic or manual means. Some reactor plants in operation today would be found unstable if left uncontrolled for several days.

In parallel with the development of the understanding and the means of circumventing inherent instabilities, attention was being given to synthesis of automatic control systems.[12] Two types of control systems were developed. The earliest was a power demand system which did not depend on inherent self-regulation. Control rods were moved in response to a power level error signal, the difference between measured reactor power and a synthesized power demand signal. The power demand signal was the sum of a steam flow signal, representative of the load on the plant, and a temperature error, needed to reset and shim the water temperature to the design level. Rods were moved either continuously or in an on-off manner, based on an early nonlinear systems analysis which today would be called the describing-function method.[7]

The second control system developed depended on inherent feedback for power regulation, and employed only a temperature error signal to move rods—no power error was used.[13] Thus, the principal control loop was that of the inherent temperature coefficient, and the role of the automatic system was reduced to a slow, follow-up, temperature regulating junction.

Both types of systems (Figs. 9 and 10) were analyzed for stability by frequency response techniques and were tested transient-wise by analog simulation. It was found that the simpler system resulted in quite satisfactory stability and transient performance. This, combined with the greater reliability of inherent self-regulation, resulted in the abandonment of power demand-type control systems for reactors having a significant negative temperature coefficient.

The understanding of inherent feedback mechanisms and the methods of stability and transient analysis developed during the early 1950's served as a basis for the design of many reactor plants and control systems.

[11] J. N. Grace, M. A. Schultz, and T. E. Fairey, "Inherent reactor stability," *Proc. 1955 Conf. on Nuclear Engineering*, University of California at Los Angeles, California Book Co., Ltd., Berkeley; 1955.

[12] J. N. Grace, "Synthesis of control systems for nuclear power plants," 1954 IRE National Convention Record, pt. 9; pp. 83–86.

[13] W. H. Hamilton, *et al.*, "Power and temperature control of pressurized water-cooled reactors," *Elec. Engrg.*, vol. 75, pp. 505–510; June, 1956.

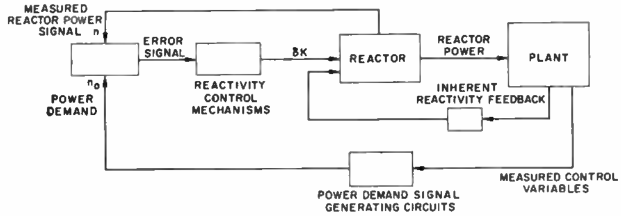


Fig. 9—Control system for a plant with inadequate natural stability.

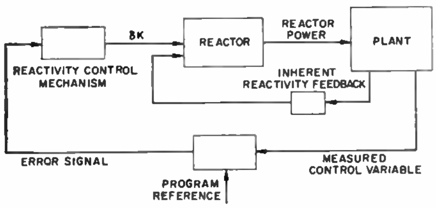


Fig. 10—Control system for a plant which is inherently stable.

The Last Six Years

The last six years has shown a tremendous expansion of reactor kinetics and control studies into many allied fields including boiling reactor kinetics, self-shutdown mechanisms, spatial stability and flux tilt transients, and stochastic effects at low neutron flux levels. Analog simulation techniques have continued to keep pace with the needs, and digital computation has become an important tool in transient analysis.

Pioneering work in boiling reactor kinetics was done by J. M. Harrer, J. A. De Shong, J. A. Thie, *et al.*, at Argonne National Laboratories, using EBWR. Their work included transfer function measurements at various operating conditions and extrapolation to higher power levels.[14] More recently, Margolis, Reihing, *et al.* at Bettis developed analytical methods of predicting the onset of void reactivity instability which were later substantiated experimentally.[15] The frequencies of oscillation predicted analytically are compared with the experimental result in Fig. 11. Zivi and Wright at Ramo-Wooldridge have succeeded in measuring feedback transfer functions experimentally, out-of-pile.[16] Plant feedback on boiling reactor kinetics has been studied by MacPhee[17] and Fleck.[18]

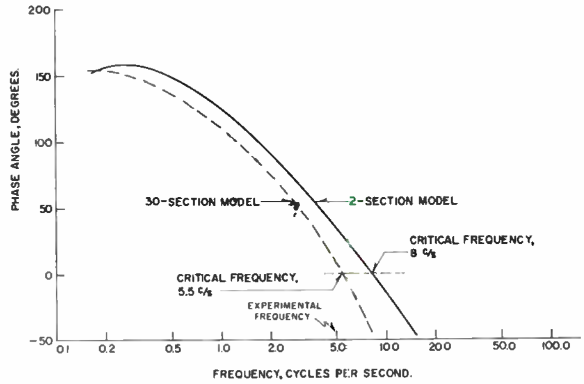


Fig. 11—Phase shift. Boiling reactor reactivity feedback loop.

Many investigators have used the correlation of reactor noise as a means of obtaining the reactor transfer function.[19–23] This method enables a transfer function to be obtained at any power level and is a powerful tool in determining whether a reactor coolant is boiling.

In addition to reactivity instability in boiling reactors, various modes of hydrodynamic instability can occur. Basic bubble generation instabilities and loop instability have been investigated at Ramo-Wooldridge by Zuber, Wright and Zivi and by Beckjord of General Electric.[24] Boiling channel instabilities have been investigated experimentally by Green at Bettis and the results used by Quant, Meyer, Rose *et al.*, as the basis for analytical predictions of oscillations.[25]

Self-shutdown capabilities of reactors were studied experimentally by Argonne in the BORAX reactors, by Phillips in the SPERT reactors, by LASL in the Godiva reactors, and by Atomics International in the KEWB reactor. Considerable success has been realized in the development of empirical correlations for prediction of peak power and energy release of runaway transients down to 5-msec exponential periods and even lower, at atmospheric pressure. More recently, analytical methods for prediction of the entire transient excursion have been under development by Langmuir at Ramo-Wooldridge and by Johnson and Chock at Bettis.[24] In water moderated cores the principal feedback mecha-

[14] J. A. De Shong, Jr., "Power Transfer Functions of the EBWR Obtained Using a Sinusoidal Reactivity Driving Function," Argonne National Lab., Lemont, Ill., Rept. No. ANL-5798; January, 1958.

[15] S. G. Margolis and S. Kaplan, "Transfer functions for boiling reactor stability calculations," *Trans. Am. Nuclear Soc.*, vol. 3, pp. 124–125; June, 1960.

[16] R. W. Wright and S. M. Zivi, "Out-of-pile power-void transfer function measurements and prediction of SPERT IA boiling oscillations," *Trans. Am. Nuclear Soc.*, vol. 3, pp. 114–116; June, 1960.

[17] J. MacPhee, "The relative stability of boiling and pressurized light water moderated reactors," IRE Trans. on Nuclear Science, vol. NS-4, pp. 25–29; March, 1957.

[18] J. A. Fleck, Jr., "The influence of pressure on boiling water reactor dynamic behavior at atmospheric pressure," *Nuclear Sci. and Engrg.*, vol. 9, pp. 271–280; February, 1961.

[19] C. E. Cohn, "A simplified theory of pile noise," *Nuclear Sci. and Engrg.*, vol. 7, pp. 472–475; May, 1960.

[20] G. T. Carrier, "Autocorrelation Methods Applied to the MIT Reactor," M.S. thesis, Mass. Inst. Tech., Cambridge; 1959.

[21] M. N. Moore, "The power noise transfer function of a reactor," *Nuclear Sci. and Engrg.*, vol. 6, pp. 448–452; November, 1959.

[22] R. E. Skinner and D. L. Hetrick, "The transfer function of a water boiler reactor," *Nuclear Sci. and Engrg.*, vol. 3, pp. 573–594; May, 1958.

[23] C. E. Cohn, "Determination of reactor zero power kinetic parameter by noise analysis," *Proc. of the Conf. on Transfer Function Measurement and Reactor Stability Analysis*, Argonne National Lab., Lemont, Ill., May 2–3, 1960, No. ANL6205, pp. 308–313; 1960.

[24] *Proc. 1960 Idaho Conf. on Reactor Kinetics*, Sun Valley, October 12–14, 1960; to be published.

[25] E. R. Quant, "Analysis and Measurement of Flow Oscillations," "Chemical Engineering Progress, Monograph and Symposium Series, Heat Transfer, No. 32-S," pp. 111–126; Buffalo, N. Y.

nisms were found to be fuel element expansion, heating of water by conduction and radiation, Doppler effects, and microbubbles generated by neutrons and protons. The work of Johnson and Chock on the effect of system pressure (Fig. 12) indicates the importance of mechanisms other than steam void formation which predicts self-shutdown even without boiling.

Spatial instabilities of power distribution caused by xenon feedback were observed experimentally in several large reactors, which then led to a number of analytical investigations of the phenomenon.[26–28] Unlike the xenon instability of the total power level predicted earlier, spatial instability is neither easily detected nor easily controlled. Not only is stability in the absolute sense required, but flux tilt transients should be prevented in cores that are subjected to nonuniform perturbations. The newest analytical tools for the study of space-time reactor kinetics employ modal analysis techniques. Stewart at KAPL developed a simple method which Harris at Bettis formalized[29] and applied to experimental data obtained from the Shippingport reactor. This method involves the assumption that the flux and the xenon deviations from equilibrium are everywhere 180° out of phase. Kaplan[30] has developed what have been called the natural modes of the system, which include all pertinent equations in the matrix operator defining the modes, such that the complex eigenvalues obtained indicate directly the frequency and damping factor. Margolis[31] has developed a Nyquist-type stability criterion for overtone spatial modes and successfully applied it to an experiment. Ewen[32] has developed digital programs for calculating the various types of modes. Margolis and Kaplan[33] also have shown the effect of delayed neutrons on flux-tilt transients, as illustrated in Fig. 13. Investigation (by analog simulation) of the effect of nonlinearities on spatial oscillations revealed amplitude limitation and precession, shown in Fig. 14.[34] Space-time kinetics studies become more important as reactors grow larger in size, making them poorly coupled and more sensitive to nonuniform perturbations.

[26] A. G. Ward, "The Problem of Flux Instability in Large Power Reactors," CRRP-657; July, 1956.

[27] A. F. Henry and J. D. Germann, "Oscillations in the power distribution within a reactor," *Nuclear Sci. and Engrg.*, vol. 2, pp. 469–480; 1957.

[28] M. A. Schultz and J. N. Grace, "Diffusion coupled oscillations with xenon reactivity feedback," *Trans. Am. Nuclear Soc.*, vol. 1, p. 164; June, 1958.

[29] D. R. Harris and P. S. Lacy, "A simple approximate test for spatial xenon stability," *Trans. Am. Nuclear Soc.*, vol. 3, pp. 437–438; December, 1960.

[30] S. Kaplan, "The property of finality and the analysis of problems in reactor space-time kinetics by various modal expansions," *Nuclear Sci. and Engrg.*, vol. 9, pp. 357–361; March, 1961.

[31] S. G. Margolis, "A Nyquist criterion for spatial xenon stability," *Trans. Am. Nuclear Soc.*, vol. 3, p. 437; December, 1960.

[32] R. L. Ewen, "MULE—a code for the calculation of two types of overtone modes," *Trans. Am. Nuclear Soc.*, vol. 3, p. 397; December, 1960.

[33] S. Kaplan and S. G. Margolis, "Delayed neutron effects during flux tilt transients," *Nuclear Sci. and Engrg.* (Letter to the Editor), vol. 7, pp. 276–277; June, 1960.

[34] S. G. Margolis and S. Kaplan, "Nonlinear effects on spatial power distribution transients and oscillations with xenon reactivity feedback," *Trans. Am. Nuclear Soc.*, vol. 3, pp. 183–184; June, 1960.

Low-level reactor kinetic behavior differs from conventional power reactor kinetics in that the average neutron population is so low that a statistical treatment rather than a continuous or average treatment is required. Such low levels are actually encountered in some reactors having weak natural sources of neutrons.[35] Transients involving these low levels are not predictable in a deterministic manner and require statistical treatment. Hurwitz[36] assumes that delayed neutron effects dominate the behavior and integrates the stochastic equation by an approximate technique. Harris[37] includes both prompt and delayed neutrons and analyzes the time behavior of the moments of the probability distribution, thus requiring knowledge of the nature of the distribution. Hansen[38] has had some success in appling theoretical methods to transients initiated from low levels. Too little experimental data exists at this time for thorough comparison of the theories.

The analog computer continued to be a useful tool during the late 1950's in the investigation of many of the problems described. General-purpose equipment replaced the homemade special-purpose simulators, and greater flexibility was thereby achieved. The new kinetics problems led to the development of unique circuitry for their solution. Boiling reactors with flow redistribution and void reactivity effects, two-phase coolant loops with condensing heat exchangers, natural circulation with plant motion effects, flux-tilt transients, simultaneous hydrodynamic and reactivity instabilities, fast power excursion transients, and the statistical moments analysis of low-level reactor transients all were simulated on analog computers.

During the last six years, digital computers gradually were coming into use in reactor kinetics problems. Digital programs were written for reactor thermal transients[39] incorporating considerably more spatial sectionalization than was practical on an analog machine. The pattern that predominated was that the analog continued to be the standard tool for 1) reactor plant systems studies involving the simultaneous solution of combinations of the problems cited above, and 2) survey or exploratory studies of new concepts, requiring the flexibility of the analog, while the digital computer was used for more detailed representation of problems of more limited scope. The writing of a digital program was well justified when a problem became a production type, that is, when repeated use of the program was

[35] D. R. Harris, "Natural neutron sources in new uranium fueled reactors," *Trans. Am. Nuclear Soc.*, vol. 3, pp. 534–535; December, 1960.

[36] H. Hurwitz, Jr., *et al.*, "Effects of neutron fluctuations on reactor operation," *Trans. Am. Nuclear Soc.*, vol. 3, p. 477; December, 1960.

[37] D. R. Harris, unpublished work; 1960.

[38] G. E. Hansen, "Assembly of fissionable material in the presence of a weak neutron source," *Nuclear Sci. and Engrg.*, vol. 8, pp. 709–719; December, 1960.

[39] J. E. Meyer, *et al.*, "ART—A Program for the Treatment of Reactor Thermal Transients on the IBM 704," Westinghouse Bettis Atomic Power Lab., Pittsburgh, Pa., WAPD-TM-156; November, 1959.

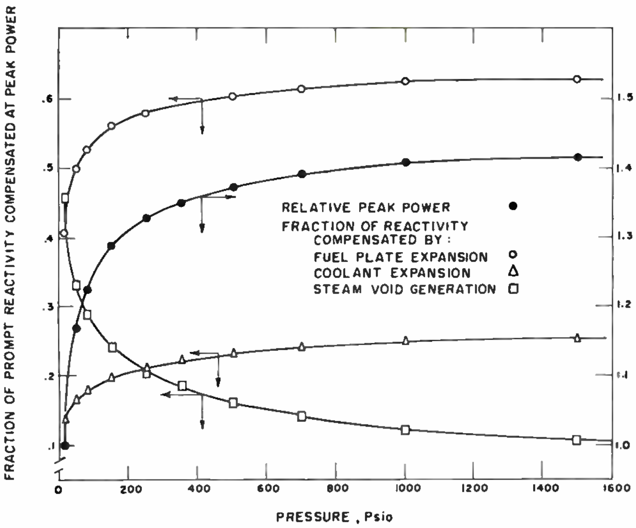


Fig. 12—Relative peak power ●, fraction of reactivity compensated by: fuel plate expansion ○, coolant expansion △, steam void expansion □.

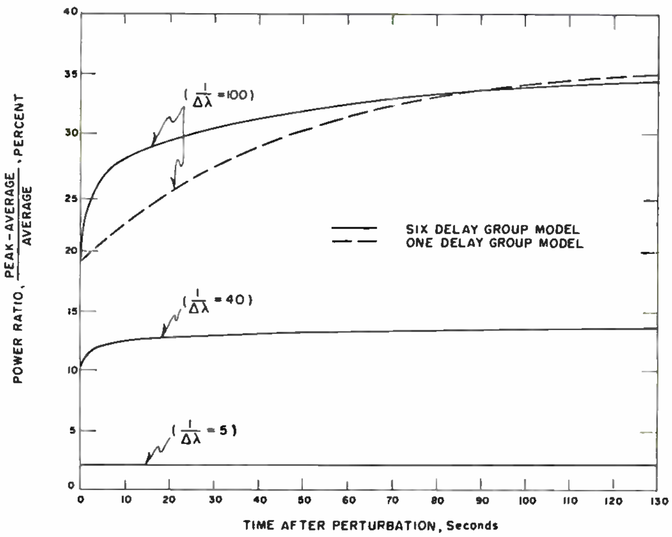


Fig. 13—Transient flux tilt in annular core caused by step temperature unbalance (490°F to 510°F).

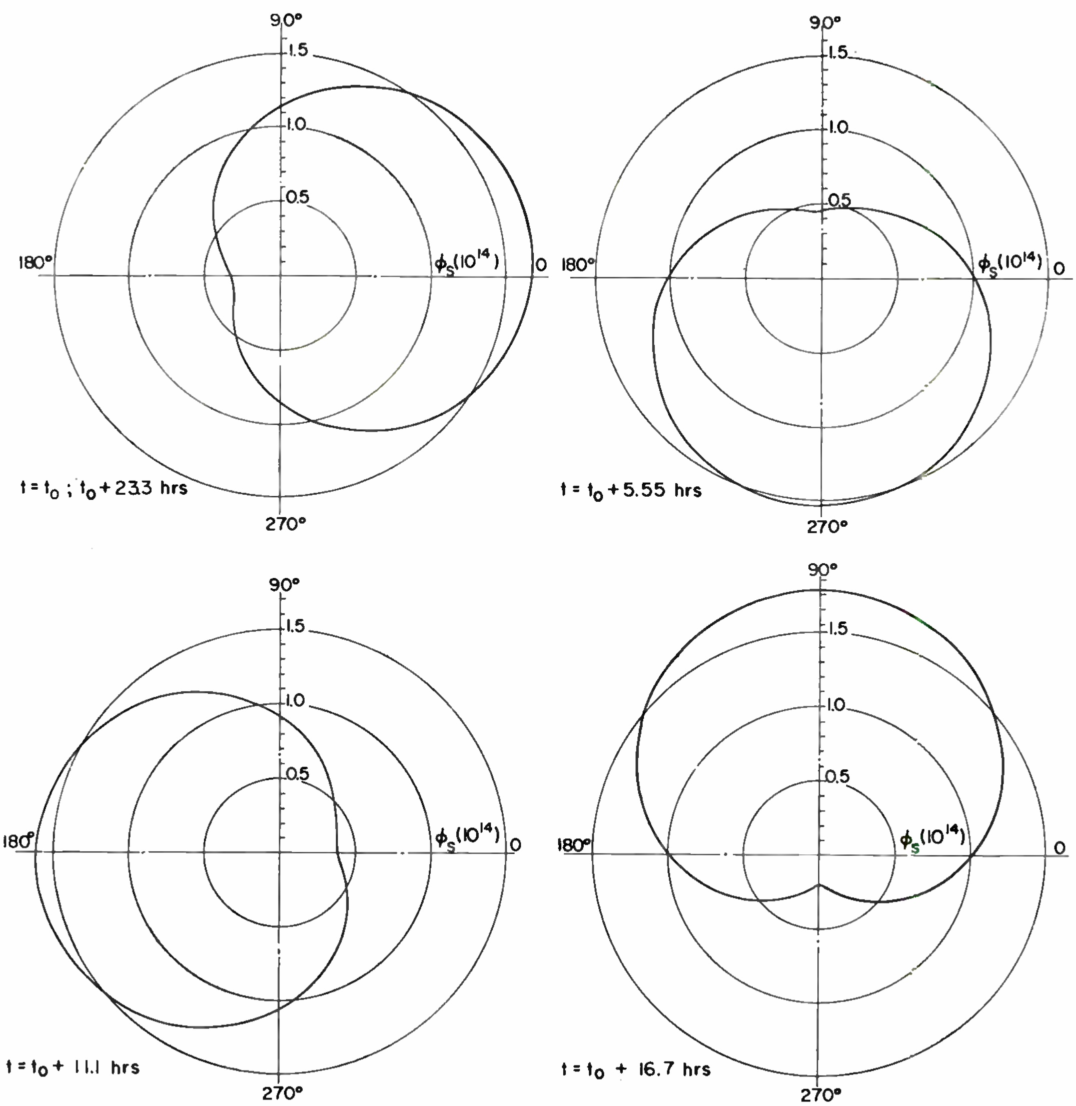


Fig. 14—Polar flux plots at successive times during 23.3-hour cycle. Initial flux 10^{14} cm^{-2} sec^{-1}. Eigenvalue difference $\Delta 2 = 2.7$ per cent.

anticipated. Frederickson and his engineering analysis group at Bettis have developed very useful digital programs to solve hydrodynamic stability, two-dimensional flow transients, heat exchanger oscillations, etc., all of which have found considerable production use.[40,41]

Although reactor plant system studies, even today, are better solved on analog computers, notable progress has been made in finding the right approximations to make digital solutions of reactor plant transients practical. For example, Arker and Greene at KAPL neglect the prompt neutron lifetime in the reactor kinetics equations, allowing precursor delays to dominate the transients. Such a treatment should be adequate and economical of computer time for many of the slowly varying system transients of long duration.

Meanwhile, as external reactor plants become more complex, automation has crept into both reactor startup and secondary plant equipment. The reactor shutdown system has always been highly automated, in that even the first reactor contained many devices for "scramming" or shutting down the reactor quickly. With the addition of a small amount of analog or digital logic, single pushbutton startup has become available in a few installations.

The Future

The continuing trend toward inherent reactor control requires more theoretical work on self-shutdown of fast excursions. Doppler effects, steam generation, and radiation-induced micro-bubbles are not well understood. Prediction of transients initiated from low levels requires an experimental program as a basis for comparison of theoretical treatments of stochastic effects.

Nonlinearities are usually readily taken into account in analog and digital transient analyses. However, more study is desirable on nonlinear instabilities to determine limiting amplitudes, although one would never intentionally design a reactor system to sustain such oscillations.

Digital methods have replaced analog simulators for transient studies of limited scope and short-time duration. There is room for improvement in the digital area to treat reactor plant transients for long duration. One remote possibility is to combine analog integration techniques, on a continuous time basis, with digital techniques, for algebraic manipulations, thus using each machine to its greatest advantage. Another possibility is to exclude automatically the representation of costly short-time-constant effects during those portions of a transient that are slowly varying.

The trend in reactor control and protection in the direction of inherent self-regulation and self-protection will certainly continue in the future. Similarly automatic control of all external functions will proceed at an accelerated pace. Without self-stabilization the inherent instabilities which would require detection and control when added to a power demand system will result in an automatic system of considerably less reliability than a self-controlling reactor. A negative temperature coefficient of reactivity, when sufficiently large, has been found to stabilize pressurizer oscillations, fundamental mode xenon oscillations and, most important in modern large reactors, spatial oscillations.

The control engineer is approaching inherent feedback loops from a synthesis point of view. He is exercising influence on the course of nuclear design and reactor engineering to enhance the desirable self-controlling features. Simultaneously the increasing knowledge of process control speeds the day toward completely automated plants. A marriage of the design of the self-regulating reactor and the completely automatic external plant control obviously will be made.

[40] R. I. Miller and R. S. Pyle, "TITE—A Digital Program for the Prediction of Two-Dimensional Two-Phase Hydrodynamics," Westinghouse Bettis Atomic Power Lab., Pittsburgh, Pa., WAPD-TM-240; to be published.

[41] J. E. Meyer, "Hydrodynamic models for the treatment of reactor thermal transients," *Nuclear Science and Engineering*, vol. 10, pp. 269–277.

Energy Conversion Techniques*

K. G. HERNQVIST†, SENIOR MEMBER, IRE

Summary—The conception and evolution of the basic principles of thermionic energy conversion, the newest of the heat-to-electricity conversion methods, are described. The application of thermionic energy converters to its ideal heat source, the fission reactor, is discussed. It is believed that the state of the art is far enough advanced to make such a thermionic-reactor system practical within the next decade.

The combination of an incandescent electric lamp, a circuit including the vacuous space within the globe of said lamp, and electrical apparatus controlled by the current in such circuit—T. A. Edison (1884).

The method of controlling space charge in an electron discharge device—which consists in generating positive ions in said device independent of ionization by electron impact—I. Langmuir (1929).

Introduction

THE ABOVE claims of Edison and Langmuir describe the basis of operation of the thermionic energy converter, the youngest member of the family of heat-to-electricity conversion devices. There are several reasons for the delay in practical application of this device, paramount of which are the *demand* for a high temperature conversion device and the *technology* necessary for making a practical thermionic converter. Perhaps the most important demand has arisen with the development of the fission reactor. Advances in electron tube technology, primarily the metal-ceramic sealing technique, have made it possible to construct practical converters. Both demand and technology have developed only after the Second World War.

Fundamental Principles of the Thermionic Converter

What Edison had discovered was that an incandescent lamp having two filaments, only one of which was lit, could deliver electrical power, though minute, to an external circuit connecting the two filaments. Electrons emitted from one filament had sufficient thermal energy to overcome the retarding electric field developed between the filaments as a result of the current flowing through the external load. Thus, some of the heat generated in the hot filament was converted into electricity. Later work in the field of physical electronics taught that the conversion efficiency could be improved by making the electron collector of a material with a lower work function than that of the emitter. This difference in work functions, also called the contact difference of potential, can add to or, if large enough, completely override the voltage developed due to the thermal energy of the electrons.

Even with the improvement in performance resulting from the work-function optimization, the vacuum type thermionic converter has an impractically low efficiency. This low efficiency is due to the electronic space charge effects. The negative space charge of the electrons depresses the potential in the interelectrode space and seriously reduces the transmitted electron current. The most practical way of overcoming this effect is to introduce positive ions to neutralize the negative electronic space charge. These ions, continuously lost due to recombination, mainly at the walls, must be generated very efficiently in order to preserve high conversion efficiency. Langmuir's work provides one answer to this problem of efficient ion generation: a hot high-work function electrode (such as tungsten) immersed in cesium vapor provides an excellent ion emitter. Cesium atoms have an ionization potential lower than the tungsten work function and become ionized (resonance ionization) when they touch the hot tungsten electrode and are emitted as positive ions.

Extensive work in establishing the scientific principles of the cesium-vapor type thermionic converter has been performed recently.[1] Two types of converters have emerged, namely, a diode and a triode type.[2] In the diode both electrons and ions are generated at the cathode, which has a relatively high work function. This converter requires heat source temperatures of 1500°C or higher for efficient operation. The triode has a separate high-work function ion emitter and a bias must be applied to overcome its contact difference of potential with respect to the cathode. Operation in the range 1000°C–1500°C appears promising for the triode.

Cesium has other beneficial effects in the converter other than providing space charge neutralization. It provides a cesium coverage of the cold anode, thus automatically keeping its work function low; the other effect has to do with the electron emitter. One of the main requirements for achieving high conversion efficiency with long life is to provide an efficient, durable electron emitter. An excellent approach[3] to this problem is to operate the converter at such a high cesium pressure that a partial cesium coverage on the electron emitter is ensured. Such an emitter, extensively studied by Langmuir, emits both electrons and ions, and it is

* Received by the IRE, August 1, 1961. Part of the work described in this paper was sponsored by Advanced Research Project Agency, Department of Defense.
† RCA Laboratories, Princeton, N. J.

[1] N. D. Morgulis and D. M. Marchuk, "Conversion of thermal energy into electrical energy by means of thermionic emission," *Ukranian Phys. J.*, vol. 2, pp. 370–380; April, 1957.

[2] K. G. Hernqvist, "Plasma synthesis and its application to thermionic power conversion," *RCA Rev.*, vol. 22, pp. 7–20; March, 1961.

[3] V. C. Wilson, "Conversion of heat to electricity by thermionic emission," *J. Appl. Phys.*, vol. 30, pp. 475–481; April, 1959.

self-healing since the cesium is constantly resupplied from the vapor phase. The high cesium pressure (~1 mm of mercury) necessitates the use of a relatively close electrode spacing (of the order of a few thousandths of an inch) to minimize the internal impedance.

Operation of cesium vapor converters, especially of the high pressure type, has led to the discovery that under some conditions improved performance is obtained when the converter is operated in an arc discharge mode.[4] Here all, or at least the major portion of the ions, are generated by impact ionization. The relative merits for the thermionic converter of impact ionization compared to resonance ionization is being investigated extensively at present.

After only about 5 years of development work, the state of the art of thermionic converters is such as to yield conversion efficiencies in the range 10–15 per cent for the 1300°C–1800°-C temperature range and 5–10 per cent in the 1100°C–1200°C range. Further improvements in materials are expected to lead to even better performance.

Construction of Thermionic Converters

In its simplest form the thermionic energy converter is a thermionic diode; thus, the techniques known in electron tube manufacture can be used in converter construction. Special problems are caused by the fact that the heat supplied to the cathode must be conducted through the tube envelope. Also, the presence of the chemically active cesium vapor sets limitations on the choice of materials suitable for tube construction. Work on materials compatibility indicates that the metals normally used in tube construction, such as copper and the refractory metals, are not seriously attacked by cesium. Glass-to-metal seals appear to be unsuitable, but, fortunately, high purity alumina ceramics can be used as insulators in envelope construction. Much work is presently being done on materials for thermionic converter construction.

Fig. 1 shows a typical example of thermionic converter construction. The cathode and the anode are plane-parallel circular disks, the separation between which is accurately set by means of small metal pins resting in insulators fastened to the anode. The thin cathode lead serves as a heat insulator for the cathode and also forms a part of the tube envelope.

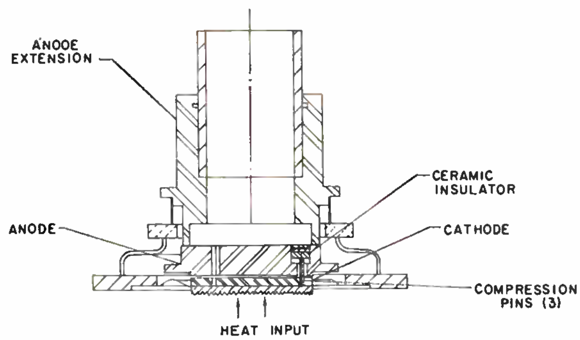


Fig. 1—Diagram showing typical construction of a thermionic converter of the diode type.

When chemical- or fossil-fuel type heat sources are used, one is faced with the severe problem of gas penetration through the hot part of the converter envelope. As will be discussed below, this difficulty can be readily overcome when using nuclear fuels.

Reliability in converter construction, of course, is subject to the same rules as in electron tube manufacture. One hundred per cent reliability is never reached, even in such an inherently simple device as a thermionic diode. Clearly, pretesting of the individual converters must be done before insertion into a power producing system.

In summary, the status of today's tube technology allows manufacture of thermionic converters having long life and good reliability. Future work will widen the choice of materials needed for specialized applications.

The Thermionic Reactor Concept

An idealized form of a thermionic-converter fission-reactor system is shown in Fig. 2. Here the thermionic converter is incorporated in the fuel element. The region occupied by nuclear fuel is bounded by cathode bearing walls. The moderator and anode coolant separate the fuel elements. It is required that the neutron absorbtion be low for the materials making up the thermionic converter and that these materials be undamaged by the nuclear radiation. Although the actual system will look quite different because of the necessity of using a multiplicity of converters in a single fuel element, the system shown in Fig. 2 illustrates the basic thermionic reactor concept. As will be discussed below, there are two important motivations for developing such a system: first *the thermionic converter fills the demand* for a high temperature conversion device *for the fission reactor;* secondly, *the nuclear fuel is the ideal heat source for the thermionic converter.*

[4] L. Malter, E. O. Johnson, and W. M. Webster, "Studies of externally heated hot cathode arcs," *RCA Rev.*, vol. 12, pp. 415–435; September, 1951.

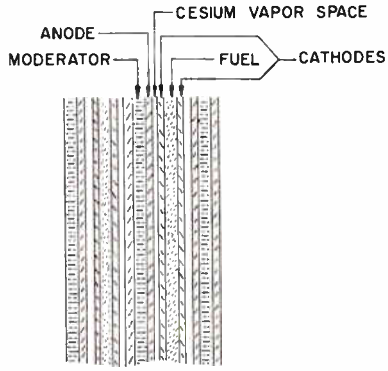


Fig. 2—Schematic diagram showing idealized thermionic reactor. Fuel elements are flat plates.

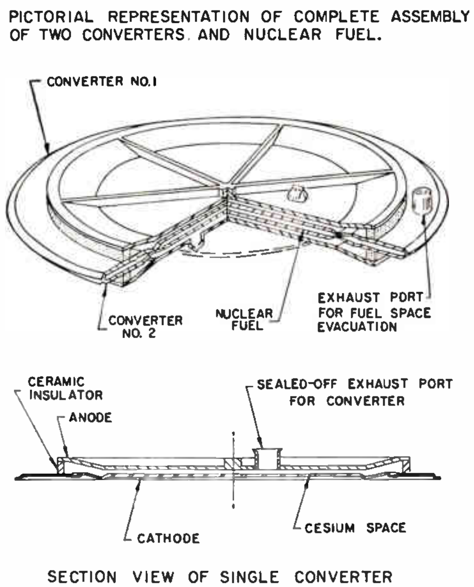


Fig. 3—Sandwich design of a thermionic converter.

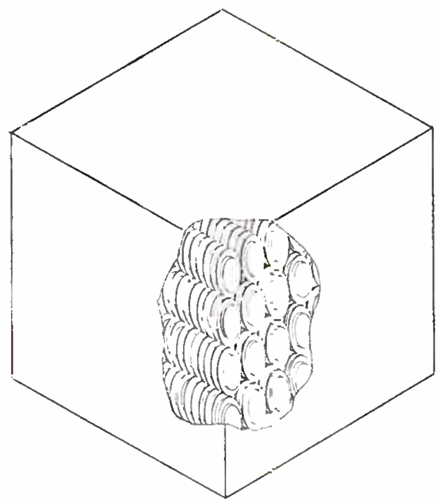


Fig. 4—Diagram showing array of thermionic converters making up a fission reaction.

Nuclear Considerations

The nuclear fuel is naturally capable of delivering heat at high temperatures. The fission fragments have very high energies, and the upper temperature limit is set primarily by the mechanical and thermal properties of the fuel and of the materials confining the fuel. To take full advantage of this property of the nuclear fuel, one must look for a heat engine having a working fluid which can be handled at these high temperatures. The free electron gas of the thermionic energy converter is precisely such a working fluid.

There are two basic requirements on the thermionic converter for its incorporation in a fission reactor: the materials of which the converter is made must not be damaged by the nuclear radiation, and their neutron absorbtion must be small in order to achieve criticality. The first of these requirements is believed to be fulfilled by a metal-ceramic type cesium vapor converter. The requirement on low neutron absorption necessitates the use of a minimum of thickness of converter materials; it may also call for a new type of tube construction materials such as zirconium and niobium.

Fig. 3 shows a proposed[5] thermionic converter assembly for a nuclear fuel heat source. It consists of two converters which can be pretested before insertion of the fuel.

Another factor of importance in criticality considerations is the effect of the inactive volume caused by cathode leads. Since the size of the individual converters is limited, terminals for each converter must be incorporated as shown in Fig. 4. Note in comparison with Fig. 2 that the space occupied by the cathode leads does not contain fuel and that here only neutron absorbtion is taking place. However, calculations show that carefully designed converter units allow criticality.

Converter Requirements

As noted earlier, from the point of view of the thermionic converter the nuclear fuel represents an ideal heat

[5] Design conceived by F. G. Block, RCA, Lancaster, Pa.

source. The neutrons passing through the converter cause fission in the fuel region which can be completely isolated between two converter cathodes as shown in Fig. 3. This assures a minimum of extraneous heat losses, and it allows the hot parts of the system to be completely enclosed in the low-pressure cesium chamber.

The thermionic converter places several requirements on the fission reactor for optimum performance. The high sensitivity of converter performance to emitter temperature sets high demands on power flattening for the reactor. Careful systems engineering is called for to solve this problem.

Another problem is caused by the formation of fission products in the fuel element region. These products must not be allowed to enter the cesium chamber, since some of the elements may condense on the anode and change its electrical properties. The design shown in Fig. 3 fulfills this requirement since the fuel element region is separate from the cesium cell. Also, such a design allows the removal of fission products should they build up intolerably high pressures in the fuel container. It is expected that a system of electromagnetic pumps may provide continuous fission product removal for a considerable period of time.

Conclusions

Although the basic principles of the thermionic energy converter are as old as the electron tube, extensive work on this device has been done only during the last five years. Considerable progress has been made, and this progress has already opened up many promising fields of application for this energy conversion device during the next decade.

The most promising of these applications appears to be the fission reactor which, for many reasons, is an ideal heat source for the thermionic converter. The thermionic reactor has many desirable features such as no moving parts, silent operation, and high coolant temperature, all of which are important for specialized applications. However, the use of the thermionic converter as a topping generator with conventional systems to convert the lower grade heat rejected by the thermionic converter may have wider economic implications.

It appears that the technical know-how for a thermionic reactor is near at hand. Such a power generator will require the use of the best resources and techniques of two highly specialized and separate fields of technical endeavor, namely, electron tube technology and reactor technology.

Section 23

PRODUCT ENGINEERING AND PRODUCTION

Organized with the assistance of the IRE Professional Group on Product Engineering and Production

The Influence of Product Design on Radio Progress*

R. R. BATCHER†, FELLOW, IRE, AND A. R. GRAY‡, FELLOW, IRE

Summary—A survey of the typical construction methods and style concepts that were used during the past seven decades provides an index of progress. From rather haphazard constructions that prevailed during the first quarter of this century, the industry soon reached a period when mass production of radio equipment was possible, whereupon the role of the product engineer, who takes the concepts and ideas of the scientists and devises suitable designs that are effective, reliable, and producible, and of the production engineer, who handles the problems attending the manufacturing of these designs, became important. In this review, the design techniques of several eras will be noted, particularly those that influenced the utility of the equipment. The many developments leading up to new processes, the ever-changing objectives, closer tolerances and increasing complexity taking place during the second quarter of the century, lead us into the space age with the emergence of the transistor, missiles, computers and automation.

Printed-wiring and other assembly techniques are described, from the simple manual operations to the programmed automatic insertion of component parts. The roles of dip soldering and wire-wrap, flexible cabling, and standardization are recounted. The progress of miniaturization is reviewed, from using conventional parts through various methods up to the micromodule and other arrangements with different form factors and integration methods. Finally, more sophisticated microminiaturization approaches—thin films, and fabrication of solid-state circuits by forming, "growing," and electron-microscope techniques. Automatic testing, dynamics, cooling, and audio-visual aids are touched upon.

The roles of the product-design engineer and the production engineer are discussed in relation to current critical needs of the industry, such as the product-development cycle, reliability, and producibility. Trends and predictions indicate that engineers in these fields can look forward to a great future with many challenges and numerous opportunities.

* Received by the IRE, January 8, 1962.
† Electronics Consultant, 240-02 42 Ave., Douglaston, N. Y.
‡ Astrodyne Reliability, Inc., Orlando, Fla.

Introduction

This survey presents several aspects of the design of electronic equipment during the many stages of its development and the work of the designers toward making apparatus that is useful, reliable, less expensive and, above all, producible. It will not just report who did what and when, but rather the constructional philosophies prevailing in each era and how designers handled those problems as they arose. The referenced periods used are not to be considered as being bounded by the precise dates noted, but as the eras where certain distinctive concepts and constructions existed. It is impossible to give here the credit due those who made the notable contributions to this progress in each period.

The equipment styles prevalent during an era provide an index to the advancement of radio communication and its associated arts. Progress was rapid, and a notable piece of equipment in any year usually became just a museum piece before another decade had passed. Equipment construction and styles found during the first four decades of this century will probably be viewed cynically by most of our younger readers. However, in spite of the apparent crudeness of some of the apparatus, judging by present-day standards, many of the current basic techniques were developed during this period.

The Embryonic Period (Prior to 1905)

The final decade of the nineteenth century and the first few years of the twentieth witnessed many independent scientific and technical developments: wireless, X-rays, cathode rays, the automobile and airplane, and the emergence of electricity from the novelty stage into a useful source of power. All of these captured the imagination of the public. In each field usable items of equipment and component parts were not readily available, and improvising was usually necessary. The field of wireless particularly captured the imagination of experimenters. The ease of setting up simple experiments and the desire to know more about the reasons for the effects noted sparked many studies. Advances were reported continually—both real and imaginary—many of which were accompanied by commercial speculation and promotion. The main production problem was that of enhancing the spectacular aspects of the apparatus. The availability of equipment in other fields dictated the type of construction. However, advancements in the theoretical aspects brought about new assembly concepts.

It is probable that Hertz's early experiments were within the range of 50 to 500 Mc, the resulting frequency being incidental to the convenient dimensions for resonant radiators which he used to signal over short distances.

Following the later disclosures of Marconi, the reasons for the switch from the relatively high frequencies of Hertz to low frequencies toward the other end of the spectrum may make little sense now. However, these were two different matters. Few people at the time even realized that Marconi's waves and those of Hertz were even akin. Lack of amplification and effective means for detecting signals over appreciable distances demanded the use of high-powered transmitters. Early transmitters utilized discharges of large capacitor banks charged at high voltages. These large banks dictated relatively long antennas and low transmitting frequencies. The component parts of a transmitter were large and thus were separately placed about the room, sometimes passing through the ceiling into the room above.

Progress in other electrical sciences had usually depended on steady-state concepts, which were inadequate in accounting for wireless phenomena. Measurements using cathode-ray tubes and rotating mirrors trained on the visible sparks in a transmitter provided the early experimenters with proof regarding some of the principles surmised. Tesla, Stone, Fessenden, Fleming, and others had worked out basic theories on some of these. The importance of tuning was well recognized at the turn of the century, when, with more stations in operation, the "other" became saturated with signals. Tesla had stressed the importance of coupled circuits as early as 1891, but the practical realization of any noticeable selectivity was not attained for a decade. It was news when Slaby reported simultaneous operation of two stations at about 240 and 540 m, an indication of the general state of the art at the turn of the century.

Meanwhile a search was on for a better means of detection, and many basic effects were pursued by numerous "inventors." The rectifying properties of many natural mineral ores proved most promising, and hundreds of natural crystalline substances were investigated for use in the "cat-whisker" detector. Incidentally many of these early "solid-state" materials are now being re-examined for application in diodes and transistors. At the end of this period the usual receiver of commerce still used cruder rectifying devices, even coherers.

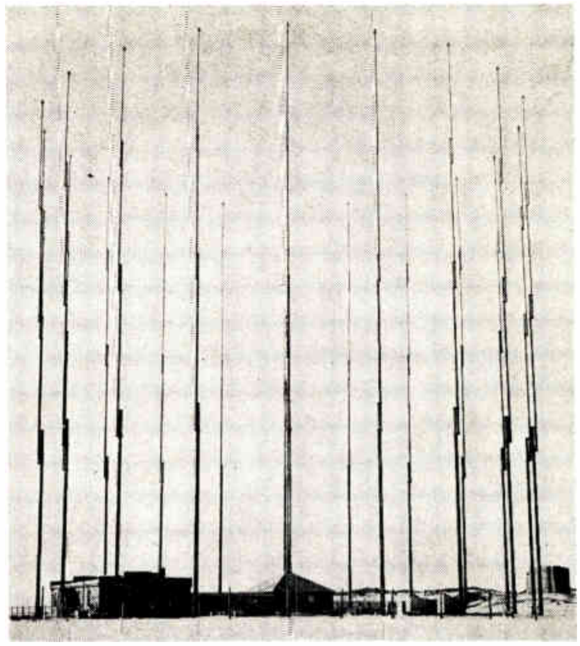


Fig. 1—Marconi station "CC" on Cape Cod in 1901 for trans-Atlantic service by high-powered spark discharges into large antennas. Worked with similar stations at Poldhu. 20 masts, each 210-ft high. They were seen wrecked by a hurricane, but replaced with four-mast array.

The Wireless Era (1905–1917)

By this time more sophisticated equipments were in use. The principles of tuning, quenched-gap discharges and even sustained-wave radiation possibilities were under more intense investigation.

New ideas altered the constructional techniques during the first decade of the century, but many of them were developed by experimenters who were not overly concerned with theories or handicapped by the task of keeping records of results in scientific terms. Whole arts were discovered and lost. Although the newly-discovered tube detectors were produced, they varied over a wide range, so that those having the greatest variations were prized the most. Experimenters spent considerable time playing up these vagaries to obtain the greatest detection "sensitivity," such as using magnetic fields applied across a tube to direct the electron beam on an effective path. Luckily other scientists pursued a principle to a conclusion and laid the foundation for radio electronics. By application of scientific principles and competent measurements a few definite rules for good construction were formulated.

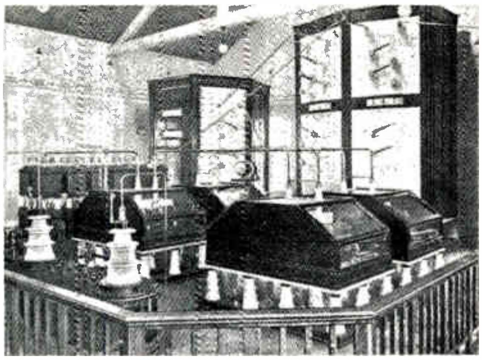


Fig. 2—Transmitting station at Nauen, Germany, in 1912. The transformer (lower left) extends through to floor of room below. 100-kv capacitors in the cabinets in left background, with four cases of quenched gaps in center. Inductances behind marble-panelled cabinets. Wavelengths up to 7000 m.

Fig. 3—Typical receiver and transmitter for shipboard in use around 1910. Receiver (right) consisted mainly of the coupler with its hinged primary and large tuning condenser (below). Detection by crystal rectifiers, usually molybdenum.

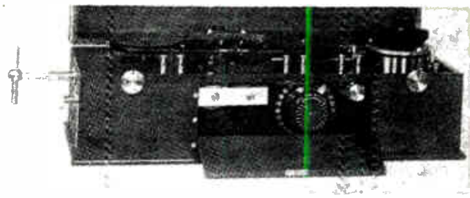


Fig. 4—English Marconi magnetic detector, this assembly being specially designed for receiving time signals. Used until after World War I, when tube receivers became standard. These detectors had to be wound up periodically, a clock-spring motor supplying power to the rotating iron-wire endless cable.

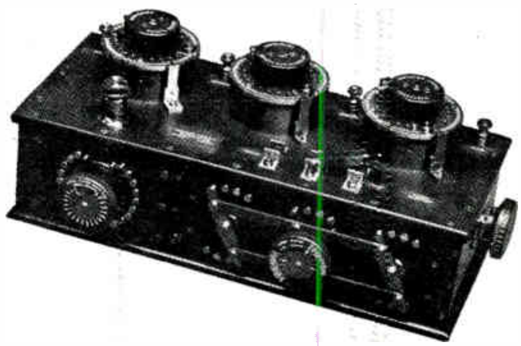


Fig. 5—Standard Marconi tuner also in common use during World War I on British ships. The variable condensers had hard rubber insulation between rotor and stator plates to handle the low carrier frequencies covered.

It is now difficult to imagine radio communication without electron tubes. However, focused and deflectable electron beams were well understood and in common use, in the form of cathode-ray tubes and gaseous path devices, at least a decade before grid-controlled tubes became known. Even oscillation was accomplished early, years before the appearance of triode oscillators.

The production of tubes was advanced during the latter part of this interval by the discovery of the importance of having high-vacuum and precisely located and dimensioned electrodes. However, the cost and fragility of these electrodes led to fantastic attempts to improve other parts of the circuit, so as to obtain the greatest effectiveness from each tube. The era of "low-loss" parts was in full swing. Production efforts were applied to a wide variety of experimental parts.

Continuing the early ideas of Marconi, transmitters used even greater amounts of power with attending lower carrier frequencies. An early "wonder" station was located at Clifden, Ireland, to communicate with Glace Bay, Nova Scotia, powered by a 6000+ cell storage battery delivering some 15,000 v. This battery, charged by steam-driven dynamos (three in series), delivered 300 kw to the capacitor. This capacitor in itself was quite fantastic. It contained dozens of plates spaced about a foot apart, each plate being some 6 by 26 ft in area, requiring an enormous shed to house it. These plates were hung from heavy charging cables. The capacitor was discharged by a rotary gap at a wavelength of around 5900 m. This installation calls attention to some of the unusual difficulties encountered by designers before 1907; a special railroad had to be built to carry peat from the near-by bogs to fire the steam boilers. Operation of this communication circuit continued until after World War I.

The carrier frequencies for some other intercontinental communication circuits extended into the audio-frequency range. Approaches using continuous-wave carriers were becoming popular with some companies, employing vacuum-tube oscillators, oscillating arcs, and high-frequency alternators. Companies were quite patent-minded, and commercial equipment was generally developed around only those methods covered by the organization's patents or its patent licenses. This prevented early utilization of some new developments otherwise useful. To stabilize the operating procedures in marine operations and to simplify the training of operators, a trend began toward the use of rack-and-panel construction to house transmitter parts, and of cabinets for receiver items.

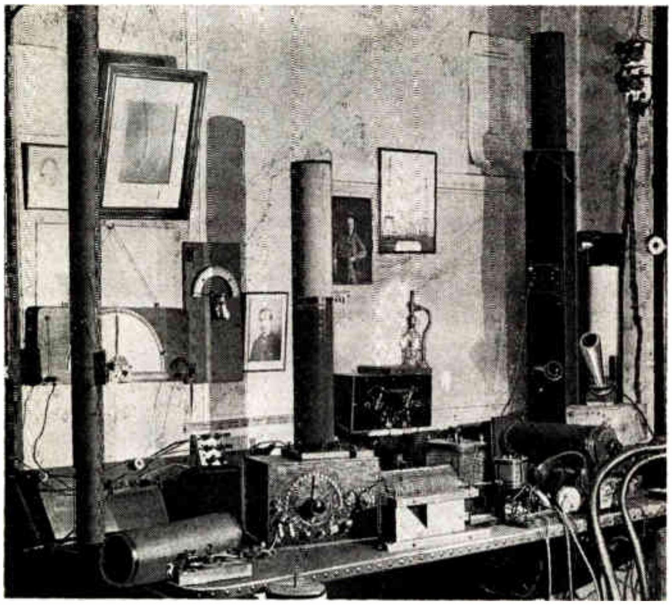
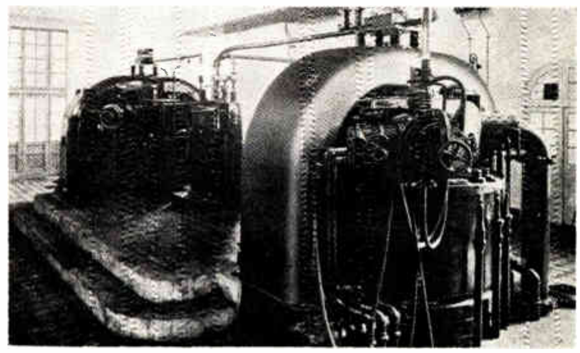


Fig. 6—A typical amateur station found around 1914 to 1917, with long-wave tuning and oscillator inductances hand wound on paper tubes. These circuits were capable of receiving continuous-wave signals from Nauen, Germany, but this particular receiver, owned by C. Apgar of Westwood, N. J., was tuned on the German station at Sayville, N. Y., and recorded (on the phonograph) messages which proved the violations of neutrality and caused this station to be closed down in the autumn of 1915.

Fig. 7—Signal Corps field set developed during World War I. A spark set powered by hand-driven generator.

(a)



(b)

Fig. 8—(a) Assembling 65-ton electromagnets for 1000-kw Poulson arc transmitters during World War I, the most powerful continuous-wave transmitters. (b) Two of these arcs, installed at the historic Lafayette station in France during the war for the U. S. Navy. Several decades later, some of these magnets were utilized in a new role in a cyclotron.

While commercial wireless was concerned mainly with intercontinental and marine services, experimental and amateur applications were countenanced. The use of radio methods for overland uses, where wire circuits were possible, received little attention. The influence of amateurs upon construction processes was marked, and soon the public also became interested when time signals could be picked up each day without having to learn the code, and weather, crop and news reports could also be heard, although in code.

Cascaded amplifiers became commonplace, and packaged receivers with self-contained fully-wired parts became popular. Simplicity of operation was not a factor of design; the greater the number of controls, the greater the impressiveness of the apparatus and the prestige of the operator. Most receivers were of the regenerative type, adhering to the practice of getting the most gain from each tube used. Heterodyne principles, long championed by Fessenden, were then beginning to take hold. More extensive use of radio communication by the military was under way, mostly by the Navy. The Army Signal Corps also started experiments in short-range communication applications.

World War I

At the beginning of the hostilities by the United States, all amateur and most experimental and commercial services were curtailed as a war measure, not as a means of conserving manpower and materials but in order to prevent their use by the enemy for spying. Intensified production schedules were put into operation. A greater amount of information was passed between engineers of competing companies through the media of early IRE meetings. Equipment-design concepts prevalent in other countries also became known here since ships from all countries were used as troop and supply carriers. Even German liners were put to use after being "caught" in United States ports at the sudden start of the hostilities. Many excellent design practices were examined and utilized, such as the use of Litzendraht, and the design of precision variable capacitors for tuning, where the rotor and stator were milled from solid blocks of metal with closely-spaced airgaps.

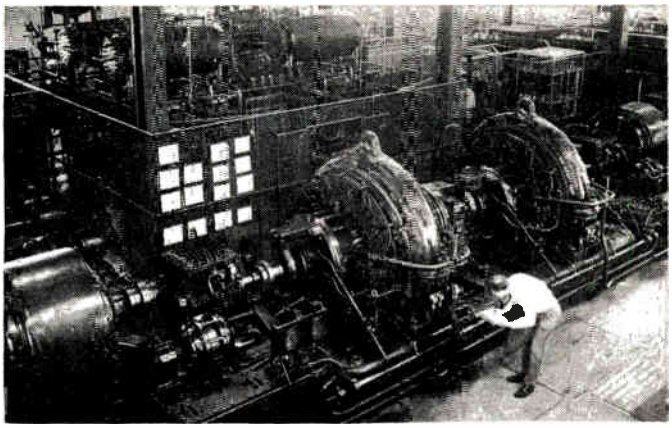


Fig. 9—Another notable attempt to produce high-power radio carriers, the Alexanderson alternator. These two alternators generated 200 kw each and were coupled to separate antennas, properly phased to give directed signals. Frequencies generated were in the range of 12 to 28 kc.

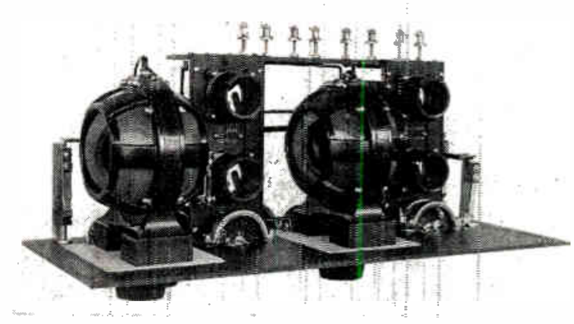


Fig. 10—Amateur receivers in the early twenties made great use of molded phenolics. Inductively-tuned receivers were particularly desired for sensitivity and ease of tuning. This "Grebe" receiver also served for broadcasting reception for those who "had the price."

Wartime equipment designs generally followed the practices of prewar construction, incorporating only such changes that made production easier. The methods of producing tubes with improved and stabilized characteristics was a most important factor, permitting what was then considered their "mass" production. Simple radiophone transmitters using tubes were produced by Western Electric toward the end of the war. These started the trend away from communications by code. A larger radiophone-transmitter design by General Electric used dozens of 50-w tubes in parallel.

Multitube amplifiers were improved for reliability and stability. Also noted were stabilized oscillators, more effective use of heterodyning, and directive communications using either rotatable or crossed loops. The invention of the superheterodyne was disclosed by Armstrong before the close of the war.

The Broadcast Era Starts

The practicability of radio-telephone transmission had been demonstrated before this country entered World War I, and subsequent improvements simplified the equipment. Broadcasting, a new service for home entertainment, started shortly after the termination of the war. Sparked by the great interest in the reopened experimental field, amateurs participated in promoting home construction for broadcasts.

Weekly newspaper supplements disclosed procedures for assembling all sorts of receivers, having to push hard to fill their pages with *new* processes each issue. Hundreds of manufacturing operations were set up, and a great many forms of common component parts were made available. Receiver construction was slanted toward the most impressive appearance. Home constructors often bought only the bare essentials that were difficult to build without extensive tooling (headphones, tubes, variable condensers, etc.), and put together the tuning coils, loops, fixed capacitors and even large storage batteries with arrays of test-tubes and strips of lead pipe.

This rush of construction served to popularize the art, but at the start did little to improve the components. During the early part of this period, factory-constructed receivers were built from designs previously popular for amateur reception. These were often well planned and used many new production processes, particularly the molding of phenolics. For several years these receivers exhibited the general characteristics of the time: they were bulky and spread out, had many controls, boasted many gadgets, and used low-loss elements, the designers striving to get the greatest sensitivity with the fewest number of tubes.

Regenerative effects were encountered and usually encouraged. Fantastic but tricky reception records were expected. Tuned RF receivers became popular, particularly after the disclosure of the neutrodyne principles by Hazeltine, even though more tubes were needed. Reflexing was resorted to in some cases, where the same tubes were used in both the radio-frequency and the audio-amplifier parts of the circuit. Superheterodyne circuits were beginning to be used in deluxe receivers.

Construction Processes

The wiring of the receivers during most of the early broadcast era used bare tinned-copper bus wires carefully arranged in exactly parallel or right-angled paths in an effort to at least stabilize, if not to minimize, regenerative effects. Soon three-dial neutrodyne receivers, containing RF tuning coils that were so positioned (or else wound with toroidal concepts) as to avoid regeneration, were popular. Such receivers, showing far greater operating stability, doomed the use of regeneration as a means of obtaining sensitivity. The growing popularity of the superheterodyne-type circuit completed this change.

Factory-assembled receivers soon replaced, in general, home-built models. The introduction of heater-type vacuum tubes in the midtwenties brought the ac power-line operated receiver into common use. En-

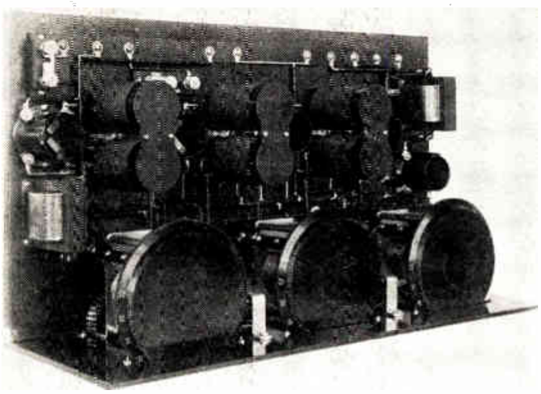
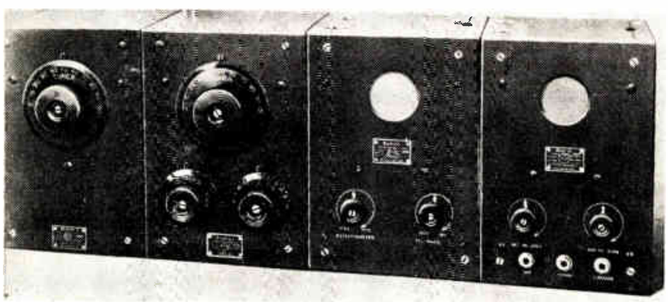
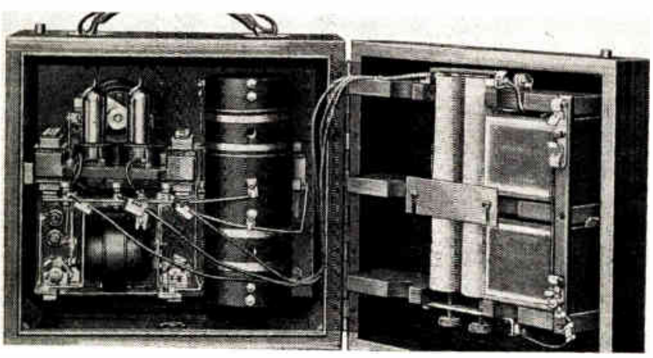


Fig. 11—Broadcast receiver of the neutrodyne type with litz-wound coils that simulated the low-field toroidal coils. This Grebe receiver used the first straight-line frequency shape of tuning capacitor plates. Wired with spaghetti-covered bus wires.



(a)



(b)

Fig. 12—Two early home receivers popular after the start of broadcasting. (a) An RCA-Westinghouse Radiola I, an early "modular" concept, a radio-frequency tuner, and amplifier, detector and audio amplifier. (b) Rear view of a Radiola II—a regenerative battery-operated "portable."

larged production schedules made producibility a most important consideration. During the late twenties, another noteworthy change was introduced when vacuum tubes became cheaper and more reliable. It had become apparent to designers that, if *more* tubes were used in a receiver, not only a *better* set resulted but a *cheaper* one. Two tubes interconnected with simple components could be more effective than a single tube surrounded with highly efficient, precisely adjusted components, which was the previous practice. Thus the era of the "low-loss" circuit was over, and the use of more tubes in roughly-adjusted noncritical circuits was customary. This apparently simple concept, viewed in retrospect, was not easily arrived at. Actually it was the result of several causes: lower cost of tubes, their greater adherence to established rated characteristics, the recent availability of multigrid tubes which permitted greater gains without internal regeneration, and a change in attitude of the public away from the desire for elaborate, mysterious assemblies to simple nontemperamental sets.

Also, strangely enough, it was the financial depression that spelled the end of expensive gadgetry. Perhaps for the first time, manufacturers of radio equipment regarded production as a science, and the role of the production specialist became important, second only to the sales promotion department (who figured that if a few tubes could be added to improve operation, why not add even more?). Purely as a sales feature, it became popular to add tubes to handle minor effects. The greater the number of tubes, the greater the prestige of the set. As a result, a plethora of new tubes were developed with special functions and characteristics. The pages in the tube manuals increased rapidly.

The Electronic Age Expands

New applications of electronics outside of the home-entertainment field appeared: sound pictures, public-address systems, new forms of electronic measurements, electronic phonographs, high-frequency heating, electrocardiography, and others.

From a single long-distance radiophone circuit operating in 1927, this field expanded in a decade to embrace 93 per cent of the world's telephones in 29 countries, incorporating such new developments as *single-sideband suppressed-carrier operation*, *voice-scrambling*, and *voice-operated switching* [1], [2]. Greater economies in space, weight and power consumption increased the use of equipment in aircraft and in flight-control applications. The success of car radio receivers (in 1940 some 2,300,000 were manufactured) spurred activity in related communication fields: two-way radio police channels for taxicabs, trucking fleets, maintenance crews, harbor craft, fishing fleets, and others [3].

The advent of practical frequency-modulation methods spurred some of these applications later. Other innovations included, before 1940, facsimile and television broadcasting. The latter, embracing the first highly complex circuits that had to be mass-produced at economical levels imposed many production problems which were not answered until after the War.

In all of these diverse applications the product designer was the direct link between the abstract discoveries of the scientists and the practical application into equipment that could be sold in a users' market. Such sales, even with the large drop in prices brought about by the market crash at the start of this period, allowed mass production. Home-entertainment receivers then provided the main source of income to back the research and technical development programs. Designing for military applications was again only a minor operation, imminent hostilities not being foreseen at the time. Designing for production simplicity and reliability for special environments gained stature. To improve equipment in several new fields of communications, specialized studies in new production techniques were undertaken along many lines. A brief survey of the trends in constructional methods is of interest.

Conveyor-Line Production

During the thirties, designers had to save every possible penny of manufacturing cost. The individually-placed bus-wire connections of the earlier periods gave way to random-placed flexible-wire connections. With production-line assembly systems, precut and stripped bundles of leads could be prepared in advance for the operators, these being placed and soldered along with other types of components.

As equipments became more complex, another major change took place, to facilitate assembly and maintenance. Here similar parts were placed in aligned rows on subplates which were provided with suitably-placed terminals in advance. Hand-assembled and laced cables tied these subassemblies into the circuit. This method of assembly has long been a standard technique for many types of apparatus. In spite of the introduction of the many more-sophisticated arrangements, it still finds some use in many early conveyor assembly systems, where production chassis were pushed along by hand from operator to operator, as each completed his assigned task. When conveyors were motor-powered, each operator handled from one to six or more items, the assignments being fixed by time requirements. With specialized test equipment, alignment and testing were reduced to simple routine procedures. Conveyor methods are practical only when continuous production over extended periods is possible.

For smaller runs the earliest method assigned a complete assembly to a skilled technician. This method was replaced by the *coach and team* method where this expert and a team of, say, nine others worked simultaneously on ten units. They followed the expert's instructions and demonstrations, and all finished at about the same time, whereupon another group of ten instruments was started. This scheme is actually an antecedent for the visual-and-sound aid programs of operator instruction, to be described later.

The World War II Era (1941–1945)

Near the beginning of 1941, the United States, although still playing the role of a neutral, was committed to be the "arsenal of democracy," and the year marked a gradual transition to a wartime footing. A rapid conversion took place from the low-precision home-entertainment type of radio to high-precision military equipment [4],[5]. Thereafter, in spite of material shortages, production reached many times the normal peacetime level. Complete strangers to radio-construction methods became proficient when the details were laid out by others. Physicists were assigned to communication and electronic problems, and many seemingly fantastic developments were started. The work of making things producible, originated by scientists with so many backgrounds of experience and training, fell to production engineering specialists.

Construction of standard radio-broadcasting transmitters was terminated to conserve critical materials. The production of electron tubes was restricted to a relatively few standard types for military needs. The production of cathode-ray tubes was similarly controlled to the particular types needed for radar and similar uses.

This period marked a transition from independent equipment to the large-system concept, where several or many separated equipments or devices were tied together for coordinated operation. Servomechanisms for remote control or for the transmission of display data and indications to central points brought about new requirements in the matter of precision, by whole orders of magnitude.

Prewar designs of equipment for military applications did not meet requirements in the matter of environmental operation. A change to militarized and ruggedized designs required tying up considerable amounts of engineering manpower. New specifications were quickly established, and soon every major manufacturer had invested heavily in environmental test equipment to cover high shock conditions aboard battleships, vibration in aircraft, the high and low temperatures encountered in some of the fields of action, and the fungus problems in the South Pacific. Other tests were set up for rapid acceleration, humidity, altitude, sand and dust, and salt-spray hazards. Many of these represented entirely new problems for the designers.

In addition to the purely physical and mechanical problems introduced by designing to withstand such environmental conditions, many new classes of components were needed which required development of entirely new manufacturing processes. Naval fire-control problems required new systems of switching and automatic control; submarine uses of Sonar needed new forms of transducers of both the magneto-striction and crystal types. Radar required means for rapidly swinging, large scanning antennas, which introduced mechanical problems of great magnitude. The introduction of pulse techniques into the philosophy of design was a major factor in the expansion of electronic techniques to new applications.

Later the proximity fuze replaced the time fuze in some cases, featuring miniature tubes, the most compact of printed circuits, and the ability to withstand many thousands of g's accelerations [6], [8]. This development, together with the walkie-talkie set [7], represented early steps toward miniaturization by the product engineer. These steps in later years influenced many areas of design.

This entire era of hostilities might be characterized as the period when the most ingenious and seemingly-fantastic ideas of the physicists and scientists were turned over to production-engineering departments and made to work in practical embodiments.

Postwar Trends (1945–1952)

At the end of the hostilities, a period of taking stock of all wartime accomplishments began. It was evident that the use of electronic methods in many fields was an absolute *must*, but also that much of the effectiveness of these methods was missing because of the inherent unreliability [67]–[73]. For one thing most equipments had been worked under all sorts of field conditions by operators who had only meager training.

In general, in the field electronics had suffered a loss in prestige. A lack of feedback of failure information from such field operations usually prevented early corrections. Speed of getting new designs into production for needed apparatus caused design and production responsibility to be divided among many groups. Each group strove for perfection as far as his own work was concerned, but usually without full knowledge of over-all design objectives, and without knowledge of how associated design details were being handled by other groups.

A common fault of the period was over-engineering, starting with the circuit designers. Many relatively unessential parts were added to gain only minor improvements. This added to the over-all complexity and frailty of the product since more parts were present that could go wrong. Another common principle of design of the period developed from an interest in logistics: the *variety* of parts used had to be kept at a minimum. As a result, a particular type and size of component would be used in many places in a given design. This part probably would serve satisfactorily in nine out of ten spots, but in the tenth it might break down consistently in a very few hours, possibly due to some special field condition not anticipated by the designers. How to *use* parts correctly, as well as how to design them, became a problem.

Another part of this stock-taking related to the commercial evaluation of all the applications of electronics resulting from military developments. The carrier frequencies had been extended upward several orders, many interesting applications for pulse techniques were then known, and many forms of automatic control through the use of newly developed transducers, servos, and rapid switching systems, and new ideas in the field of computers became known. These were but a few of the advances that were to be expanded upon.

Computer design problems challenged the product engineer anew, since major mechanical contributions involving many new construction concepts were needed for both the analog and digital types. This field was particularly adapted to the use of subassemblies and modules, and the application of miniaturization, encapsulation and the automation of assembly and testing. The invention of solid-state devices in 1948 was probably the most important event of the era, making possible new approaches for handling circuits of all degrees of complexity.

Meanwhile home television emerged from modest prewar beginnings. In 1946 and thereafter most receivers were completely redesigned so that mass-production methods could be introduced. In 1947 some 175,000 receivers, valued at over $75,000,000 were manufactured. During succeeding years, TV receiver design objectives related mostly to obtaining larger pictures, with even smaller cabinets, and more sensitive receivers, and to increase their range and so reduce the necessity for highly complex antenna arrays which were needed in many areas. Designs included a reduction in circuit complexity and the number of minor controls that had to be readjusted from time to time. By 1951 round and rectangular cathode-ray tubes of both glass and glass-metal construction and a steady parade of smaller improvements were noted. Cobalt-iron magnets replaced electromagnets for focusing. A 3000-mile coast-to-coast microwave relay system augmented the services provided to networks via cables, and by the end of that year 108 TV stations were in operation.

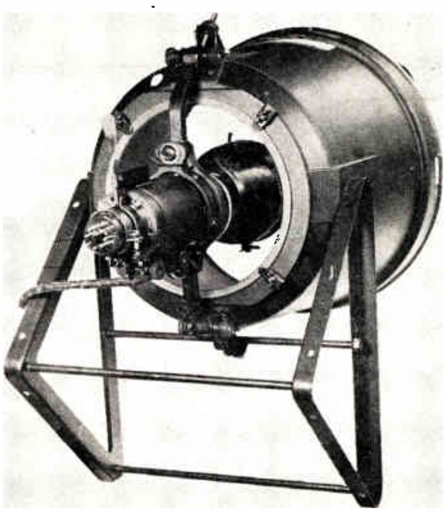


Fig. 13—Television-receiver designers achieved large picture sizes in various ways. The Philco projection receiver (circa 1948) gave the largest picture of any home receiver at the time, using Schmidt optics (shown here) to enlarge the picture from a 4-in diameter tube. Ingenious principles were used to gain picture size equivalent to that from a 24-in tube with equal brilliance at a nominal cost.

Product Engineering and Reliability

A direct result of military applications was a gradual revision of the prevailing idea among many potential users that tubes, *per se*, were fragile and their use should be avoided if possible. The greatly increased "socket" population in the typical equipment challenged the ingenuity of both equipment and component designers. Mass production of many items which previously had been hand-assembled was then possible.

The quest for reliability called for new goals concerning precise tolerances, the maintenance of initial characteristics over long periods, and the handling of a wide range of operating parameters and environmental hazards. A significant step was the establishment of new sets of standards covering many of the common components by joint action of industry and military agencies. To meet these closer limits, intensive programs of the study of all phases of component construction were undertaken independently by many manufacturers. The importance of correct design, the precise control of materials, and better inspection at critical stages of the construction became the rule in many places and were responsible for upgrading the quality of many parts.

Quality control usually caused higher rejection rates, which, combined with greater testing expense brought forth the belief that reliability can be attained only at appreciably higher prices. However, improvement programs in plants which made mass-produced items proved a new maxim: reliability must be *designed* into a product, and no amount of inspection and culling will ever compensate for designs that contain ill-planned dimensions, or an inferior material or uncontrolled construction process at some vital point. In many cases it was found that properly-designed component parts, even complicated ones, could be greatly improved *and* mass-produced at a cost below that previously found with the earlier construction methods which relied mostly on operator training.

Many equipment designers were slow to take advantage of such improvements, believing that the lower costs of mass-produced items were, *per se*, evidence of their inferiority to custom-made components commanding higher prices.

Design-Policy Confusion

Designs were becoming so specialized that engineers could no longer remain individually informed on all details. In the case of military electronics, some companies have been, and in some cases still are, passing through a state of design-policy confusion as a result of these efforts, but new industrial production concepts are emerging.

Military types of electronic systems are presently developed in three basic phases. First, one or more engineering models are designed, built and tested for performance—the so-called breadboard phase. Next, the component parts are extensively tested and a design review is then made, after which any redesign needs are accomplished to rectify environmental weaknesses—the reliability phase. Finally, still another redesign is attempted in the interest of conservation, producibility, automation, etc.—the "industrial" phase. These three phases of development leave much to be desired, especially when it is necessary to perform them in tandem.

In this age of expanding technology, engineering departments use many newly-graduated engineers for the first breadboard phase, because of their knowledge of latest theories and circuits. This reserves the more widely-experienced engineers for the later phases. Sometimes, when poorly supervised, the first stage designs leave much to be desired, requiring extensive changes later. Then too, a system sometimes becomes obsolete because of new discoveries and objectives, even before the breadboard phase is completed. This prompts those concerned to try to compress, time-wise, the over-all development. This, in turn, demands a long lead-time for configuration planning, parts selection and procurement against anticipated future models. The effectiveness of various approaches in this matter are still being evaluated.

The Systems-Engineering Approach

A few years ago certain members of the aircraft industry sensed that the manned military aircraft might be on its way out in favor of the guided missile. It seemed that each guided missile had to be designed as an integrated system, with a single contractor given complete responsibility. Missiles and weapon systems were becoming more and more complex, and most government agencies found it increasingly difficult to supply the detailed technical management required: a complete system design, the assignment of the multiple contracts for parts of the system, and approval of the facilities of the major subcontractors. The systems-engineering concept resulted, which placed over-all responsibility for each major program upon a single organization. Three slightly different solutions emerged. The Army assigned systems-engineering matters to its Ordnance Missile Command at Redstone; the Navy to its prime contractors, and the Air Force to the Space Technology Laboratories. In any case, the efforts of many different organizations were correlated to produce each complete integrated development.

In other prominent areas involving smaller developments, a policy has been considered by Military Agencies which puts more funds into research and development, rather than the actual manufacture of large quantities of apparatus, many of which might become obsolete even before they could be produced, by the emergence of other systems. Instead, an industrial "planning" phase has been substituted in some instances to provide, in theory at least, a well-organized mobilization schedule with all production phases, and in some cases tooling, considered in advance. Such a schedule could be put into immediate action should any national emergency arise. Realistic accomplishments that can be expected from such a plan have not been evaluated or proven as yet.

Integrated Operations Engineering

In regard to general policy problems associated with large systems contracts as they relate to in-plant development programs, it seems that an entirely new approach is being considered. The three phases each contractor must handle—the basic designs for performance, the reliability and value-engineering phase, and a redesign for producibility—should not be done in tandem, as is the usual case at present. On the other hand, it is a gross simplification to infer that these three product-engineering programs can be handled simultaneously.

A new design philosophy seems to be emerging, called by some *integrated operations engineering*—the application of operations research. Here all phases are correlated in an over-all program that provides the right number of engineers with the required talents and experience at the right time during the production schedule of large complicated systems having many development facets. Actually this might be considered as an expansion of a practice long used among smaller limited-product organizations dealing with less sophisticated problems than computers and missiles, where complete coordination of all design and production factors is easily accomplished.

The Space Age (After 1952)

Both the military and industry completed their reexamination of the various impacts made by electronics on their operations, arriving at similar general objectives. Because many of these plans envisioned large arrays of complex equipments, the matter of reliability was uppermost in everybody's mind, and this phase was tackled from all directions. Few doubted the ability of the scientists to devise plans for handling almost every problem put up to them, no matter how fantastic they may have seemed at an earlier age. The problem of devising means for implementing such ideas with equipment that was conveniently sized, producible, trustworthy, simple to handle and economical was then an assignment given to the new expanded field of product engineering.

Standardization Efforts

To increase reliability, many groups, committees and commissions were set in motion, but for the most part, for a long time, the outcome of these groups were warnings and dire predictions as to the results of even minor deviations from absolute reliability in any part. However, other more realistic approaches to the problem were also undertaken, and a number of jointly-operated study groups were set up by, for example: the Joint Electron Tube Engineering Council, RETMA (now EIA) product standardizing committees, the Armed Services Electro-Standards Agency, the Electronic Applications Group (of RETMA) and the Advisory Group on Reliable Electronic Equipment (AGREE), the latter having joint participation by the three Military Services.

Closer studies of ratings, a better insight into the conditions where appreciable derating of component capabilities, means for handling design changes to lessen the extreme operating conditions (by warming, cooling, shock mounting, stabilizing, etc.) were publicized. New specification instructions, procurement policies, test procedures acceptance rules were devised. Above all, a great deal of study was authorized and sponsored relating to the capabilities of current materials and construction processes and to the development of new materials that would better serve the designers' needs in critical applications. A review of early and current progress along these lines follows.

Automation

Printed-wiring and printed circuits had been used before and during World War II [9]–[13] and miniaturization also had been accomplished in some fields: hearing aids, proximity fuzes and walkie-talkies, for example. Neither field had become a serious contender to alter construction concepts, until the last decade [14]–[16], [19], [20]. By 1952 several programs were under way to bring about automatic assembly of standard components using two approaches: automatic insertion of parts into printed-wiring boards and means for integrating small units or modules into large assemblies [17], [18].

The automatic component insertion machine of the United Shoe Machinery Corporation, a conveyor-connected system with 12 insertion heads, can deliver an assembled board every three seconds. By the end of 1954 several manufacturers had trial installations in operation, one of which contained 50 insertion heads, each accounting for a single component. Meanwhile Admiral, Sylvania, and others were using automatic assembly machines of their own designs, all with a relatively high output volume. These also generally used the conveyor-connected "in line" insertion heads.

In 1955 General Mills demonstrated its 24-head "Autofab" in which, in addition to usual features, wrap-

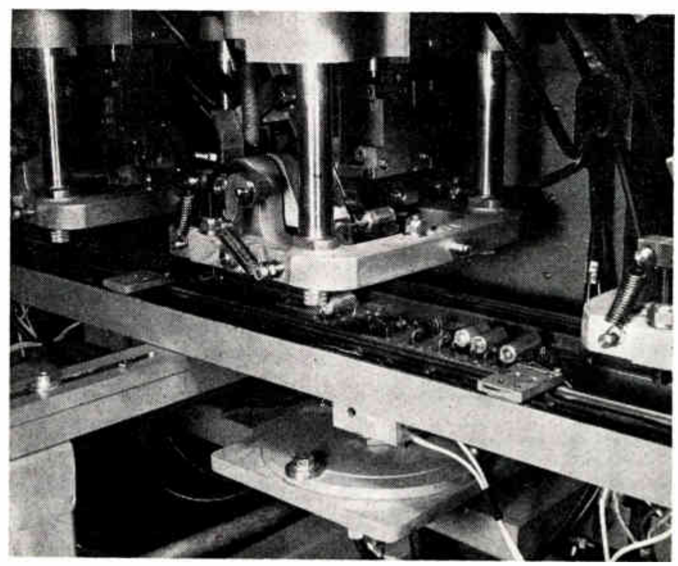


Fig. 14—"Autofab" head inserts pulse transformer. This closeup shows one of the 24 heads about to insert a pulse transformer (dummy parts were used in the early demonstration—1955).

around component lead tips were applied to the formed leads produced by an AMP lead-former. These provided capillary paths for the solder from the dip-soldering operation to simultaneously provide top-side soldering. A General Mills insertion machine had been ordered by IBM for the assembly of printed-wiring cards for the Air Force SAGE computers. Since the equipment in this particular application would be installed in locations practically free from vibration, the mounting of the parts away from the printed-wiring board was feasible, whereas it might not have been acceptable for a subassembly used in high vibration environments, such as in a missile. These wrap-around lead tips also proved useful by increasing soldering reliability even under extreme stresses.

There are now perhaps fifty large conveyorized insertion machines in use throughout the country, about half being home-designed and built by their users. From a long-range viewpoint, at least for many applications, it may be that standard sized components (that now make up the bulk of all construction) are on their way out in favor of integrated circuitry. Therefore these 50 installations may represent a near-saturation point.

Programmed Insertion Systems

If automation is to become generally useful for the assembly of electronic equipment, the set-up time must be greatly reduced in order to permit the system to handle small batches of assemblies, as well as large runs extending over considerable periods. With a programmed insertion machine it would be desirable to be able to assemble a *single* engineering model automatically. To study this idea further, the Navy contracted with Melpar and the Army with General Electric to develop programmed assembly machines. By 1956 Melpar had completed a model which could continuously assemble different circuits on 6- by 6-in boards from tape-controlled programming. GE also completed its program-inventory and programmed-insertion machinery, called ACAS, a much more elaborate development. The insertion speed of both systems was about two or three seconds per head.

In 1958 IBM announced its "punched-card-controlled component insertion machine," featuring a light-weight insertion tool positioned by differential binary air cylinders, the storage of the component parts on reels, and an insertion speed of one-half to one second (with a potential insertion speed that might approach 1/10 second). Other companies have also worked on different approaches to this problem, but to date not much production is being handled by any of them.

Semiautomatic and Methodized Insertion

For standard component parts individual insertion heads have been available from several of the above-mentioned insertion-machine pioneers. In 1955 Minnesota Engineering introduced a head that also drills the holes automatically, thus easing the problems encountered with fibre-glass boards. Many companies use these and similar machines individually or in groups, gaining much of the savings and reliability advantages of the larger conveyor systems, at a reasonable cost. Frequently, four or five such machines

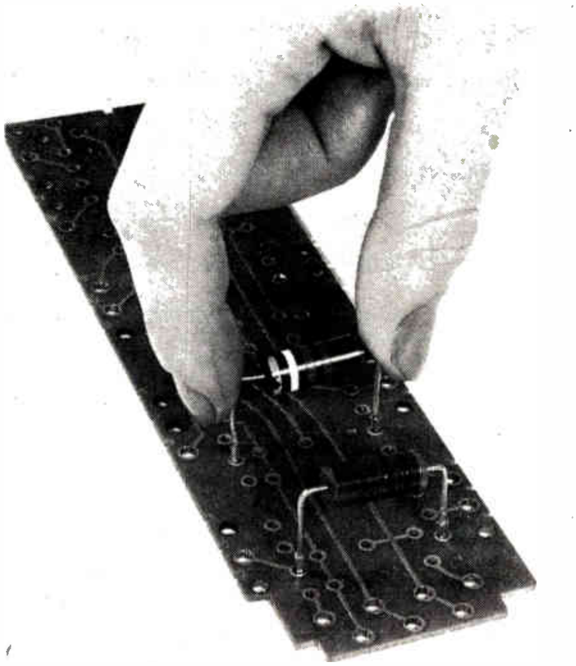


Fig. 15—IBM "Sage" computer board. This board was assembled in 1955 on the General Mills "Autofab" conveyorized automatic assembly machine. Note the wrap-around component tips on the ends of the resistor leads. The board had wiring patterns on both sides and plated-thru holes.

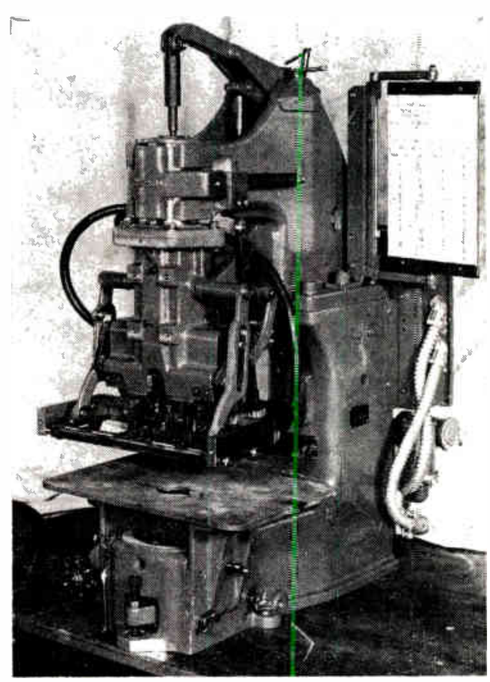


Fig. 16—United Shoe's "Dynasert," an individual component-part machine designed for bench installation. This was in fairly sizable production about 1955.

are arranged in a circle, with one or more operators moving the boards from one machine to another.

Similarly there have been several *methodizing* approaches, which refer to systems that are just short of full automation. Some of these are finding increasing use. Home-entertainment receivers use a considerable amount of automation. Test equipment manufacturers also have taken advantage of such methods, using "walk around," "push-along," and "roller-skate" conveyor systems. RCA made early use of semiautomatic "cut-and-clinch" machines, where the operator inserted four to eight parts in a printed-wiring board, whereupon a foot-controlled air cylinder cuts and clinches the leads of all components simultaneously. Later nearly all manufacturers of missiles and other military electronic equipments used variations of this system. Automation, or at least "methodizing", is being used quite extensively in assembling most forms of military, industrial and home-entertainment equipments [21]-[30].

Fig. 17—"Wheel of fortune" lazy Susan assembly fixture, a *methodizing* approach first used successfully about 1955 to assemble Matador IF strips (and other similar assemblies) at The Martin plant in Baltimore, Md.

Automation WITHOUT Printed Wiring—the Solderless Wrapped Connection

Printed wiring did not have an exclusive monoply on the application of automation. Mallina reported in 1956 that Western Electric had been using a punched-tape-controlled "wire-wrap" machine for several years. This machine wired complicated electrical circuits by means of a six-turn 29,000-psi "solderless wrapped connection," at the rate of 10,000 connections per hour. Three TV manufacturers were also using this method. Clevite-Brush and Keller Tool participated in its commercial development. Assembly for "wire wrap" could be accomplished by either automatic or manual means, provided that the component parts had suitable rectangular terminals. At that time 300,000,000 wrapped connections had been made without a single reject. In 1961 wire-wrap was being used in many assemblies of the Polaris weapon system.

Standardization, the Key to Automation

The success of automation using conventional components depends on *standardization* of the parts themselves, of the engineering layout dimensions and of the machines for their insertion. In recent years the Electronic Industries Association made considerable progress, through its committees, in the standardization of the parts. The United States Signal Corps and others promulgated specifications for the layout of conventional component parts. Other groups set up anvil dimensions and spacing for automatic insertion machines.

These brought us realistic automation with attending lower construction costs, but at a sacrifice of miniaturization goals. Ground-based equipment can afford the latter sacrifice in the interest of reducing assembly costs, but airborne computers, missile guidance systems and satellites can *not* afford to disregard miniaturization [31]-[37].

Miniaturization

Early Miniaturization Programs

Many of the early studies made before 1946 were augmented by new studies which incorporated utilization of automatic methods. Mallory disclosed packaging techniques for integrating five different types of prototype systems at volumetric efficiencies ranging from 2 to 12 per cent. These subassembly units were designed for automatic assembly. Hughes Aircraft has shown a well-integrated airborne digital computer. Other early studies were concerned with conformal coating of printed-wiring assemblies, eyelet failures, improvements in the methods of interconnections, and the use of flexible-printed cabling and flexible-printed assemblies. The studies of eyelet failure brought about the "funnel" eyelet design, and the plated-through hole [38]-[42].

The Miniature Module

In recent years numerous companies have been producing or using miniature modules assembled from conventional but smaller component parts. The new sizes were about as numerous as the companies concerned: Sylvania, Diamond Ordnance Fuze Laboratory, Martin, Ansley, to mention a few. More recently Burroughs, Amphenol (with its micromin connector in a modular package), and Republic have contributed new sizes and shapes.

The Welded Module

Intermediate between the miniature module and the micromodule in size and uniformity, the proponents of the welded module claim high reliability, basing their claims on vacuum-tube lead welding experience, where one failure in 7 million is attained. Francis Associates, Sippican, Raytheon and Litton are heavily engaged in producing welded modules, and others are building development models.

Project Tinkertoy

An attempt to develop realistic modules automatically was undertaken by the National Bureau of Standards for the Navy Bureau of Aeronautics. The NBS process was basically the same as that used by Sprague, Aerovox and Centralab as early as 1942 in building multiple-component plates. Notched steatite wafers, 7/8-in square by 1/16-in thick, were made right in the facility from raw materials. One to four component parts were placed on each wafer. Resistors were applied by an automatic printer. Another machine assembled and soldered four to eight of these wafers on to twelve riser wires to form a module. Other machines tested them automatically. All machines were tied together by an automatic conveyor, geared to produce 1000 tested modules per hour. Sanders Associates later applied the principle of stacked-wafer modules, using printed wiring, to the Navy's Sonobuoy and produced about 1800 units, using semiautomatic assembly methods [43]–[45].

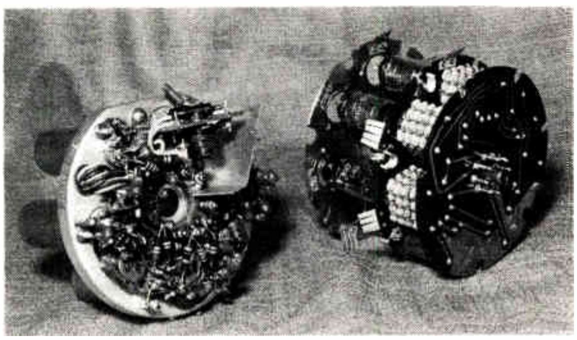


Fig. 18—Modularized Navy sonobuoy. At the left is one of the chassis of the original sonobuoy stacks. At the right is the sonobuoy after conversion to "Tinkertoy" modules by Sanders Associates (about 1955).

By the end of 1954, Project Tinkertoy had served its purpose as a large-scale evaluation program and the whole concept was released to industry to carry on. The construction of a second Tinkertoy facility was started at ACF Electronics, and a complete TV receiver was converted to a stacked-wafer module. ACF also delivered modules in production quantities to other radio and TV manufacturers. But in spite of this progress, it became evident in 1958 that the 7/8-in square stacked-wafer module would not become the standard packaging method of the future.

The Micromodule

In 1958 the RCA completed a study for the Signal Corps on how to microminiaturize functional electronic circuits using current process technology. The Signal Corps then asked industry to carry on, under RCA guidance, most of this followup going to component manufacturers who would change only the size and form factors. Deposition and solid-state approaches beyond the existing state-of-the-art were to be avoided. Size, weight (and cost) factors were required to be a full order of magnitude better than the current *best* techniques, and two orders better than the current *average* techniques [46], [47].

Early in 1961 resistors, applied to a wafer with a form factor of 0.310-in square by 0.01-in thick, were being built by a number of companies. Capacitors, inductors, transformers, diodes, transistors and other items were also being made available by many other companies. As in the earlier Tinkertoy project, twelve vertical risers are used. The integration of these wafers presents the same printed wiring problems, but with a smaller tighter-tolerance scale. Many types of circuit functions were commercially available, but at a cost about double that of more conventional assemblies.

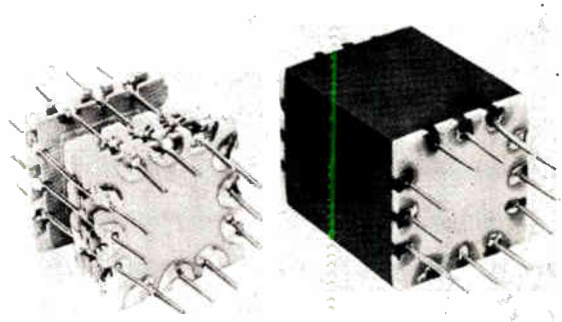


Fig. 19—Signal Corps-RCA micromodules—unencapsulated and encapsulated versions of about 1959. As you read this, there will have been 12 telephone-book size reports circulated to industry on this Signal Corps industry-wide program by RCA, the leading contractor.

Other Microminiature Form Factors

Paralleling this micromodule program of the Signal Corps, other programs were noted in 1961 using other dimensions. General Electric's TIMM program incorporated their ceramic wafer with heaterless tubes and other components to form a cylindrical stacked wafer 0.32 inch in diameter. Part densities reaching one million parts per cubic foot have been attained in functional circuits. An example was shown of a 1.25-in cube system having a density of 2,080,000 component parts per cubic foot after integration. Mallory has combined individual components into pellets of a height equal to the thickness of a printed-wiring board (0.062 in) and of various diameters from 0.100 to 0.25 in [48]–[50].

Integration of Micromodules

The repair of micromodules, when integrated by printed-wiring methods, is usually a difficult procedure, a particularly important problem when used in ground based equipment. Even when used in one-shot missile or satellite equipment, corrections or changes may be necessary. AMP's interconnection system of MECA seems to have answers to many of these problems for ground-equipment maintenance. IBM is using a 7090 computer to establish interconnections, but even without such a facility AMP claims a 25–30 per cent saving in design time [51].

Microminiaturization by Thin-Film Techniques

At present such essentially two-dimensional techniques are mostly in the development phase, but they are of such moment that they are receiving much attention. Development is underway in many places: Bell Laboratories, Sylvania, IBM, CBS, Sprague, Vare, Philco, Servomechanisms, Litton, IRE, Motorola, to name a few. IBM has built up to 17 layers on a 0.300-in insulating substrate. At present thin-film techniques are confined to passive elements (resistors, capacitors, etc.). Although active elements are at present difficult to duplicate, semiconductors grown in thin single-crystalline silicon layers from the vapor state offer promise. Many transistor manufacturers are now working on this phase of the process [52].

RCA and Philco have disclosed progress with evaporation approach. The Diamond Ordnance Fuze Laboratory, Westinghouse, Fairchild and Texas Instruments have completed such items as multivibrator switches, direct-coupled transistor circuits, video amplifiers, oscillators and other functional items.

The formation of laminar layers by vapor growth is another concept. Vare's deposition of vacuum-mixed metals from gaseous atomic forms represents a promising approach. Vitramon suggests a construction where a strata of solid particles of several layers of conducting, nonconducting and semiconducting materials would be fused to form a solid [53], [54].

Microminiaturization by Solid-State-Circuit Techniques

The term *microsystems electronics* has been used to cover many approaches and needs to be clarified for the solid-state-circuit condition. We refer here to either 1) *formed* solid-state circuits, 2) *grown* circuits, or 3) *electron-microscope* forming. The several techniques in microsystems have been given many names with many meanings—microelectronics, molectronics, nanocircuits, microminiature functional electronic blocks, and others.

"Formed" Solid-State Circuits: A number of companies (Texas Instruments, Pacific Semiconductor, and others) have worked independently in this field for several years and have produced operable devices. For example, a typical solid-state method starts with a silicon semiconductor wafer. An integrated single-crystal circuit is formed by a combination of mechanical surface shaping, diamond sawing, sand blasting, mechanical forming or etching, electron-beam machining, oxide and photoresist masking, diffusing, metallic deposition and evaporation alloying. Resistances are formed by the bulk resistance of a certain area in the silicon wafer, sometimes after antimony diffusion. Capacitors are produced by large area biased *p-n* junctions. Semiconductor devices are diffused onto the wafer, and leads are installed by welding, thermocompression bending or with a fusion furnace.

Here, as never before, production engineers work hand-in-hand with physicists in pioneering many radically new processes that may ultimately have wide applications in other fields as well.

"Grown" Solid-State Circuits: Westinghouse has developed for the Air Force functional blocks without using wired junctions or cementing of parts. Their approach used dendritic silicon and germanium providing an excellent surface for epitaxial growth. Motorola is also participating in this approach.

Electron-Microscope Forming: Stanford Research Institute is working on solid-state techniques that have possibilities for obtaining packaging densities more than a million times that of the others by using an electron microscope in combination with mechanical manipulators and molecular-beam etching and deposition equipment. The components have dimensions of about one micron, which makes possible the placement of an entire data-processing system with 10^{11} components in a cubic inch of volume [55], [56]!

Automatic Testing

This phase of the production problem has kept pace with the automatic production advances. Missile and satellite launching facilities are now packed with automatic checkout facilities. Manufacturers of all types of equipment are using forms of automatic testing, often programmed to handle complex tests by punched cards or tapes. Tubes can be tested automatically at the rate of 1800 per hour in one instance, going through twenty static and dynamic tests each [57]–[61].

Dynamics and Cooling

Considerable research on the theoretical aspects of heat-transfer and vibration problems has supplied product engineers with much useful information about the correct placement of parts and other means of alleviating effects of these and other hazards.

The literature will be of interest to those engineers who specialize in design for environment and reliability [62]–[69].

Audio-Visual Aids for Operator Training

During the last few years of the era, the production techniques of the previous decade were not only supplemented by automation for high-production items, but were given an important assist for small-lot production. An audio-visual-aid method is practiced by numerous factories, especially where few-of-a-kind type of production does not permit assembly-line techniques. Here each step of the work is described in a head receiver worn by operator while the particular point on the chassis referred to is projected on a convenient-sized screen. The slides advance as each step is completed. The operator does not have to learn any sequences in the assembly, as the instruction is repeated with every new chassis. Changes in the assembly require only the substitution of one or more new slides and accompanying voice and the replacement of the stock items involved [66].

Product Engineering and Production Techniques

A Challenge and an Opportunity

There is a great future for the product engineer and the production engineer. Since World War II, collegiate curricula for electrical engineering has stressed mathematical and theoretical concepts, physical concepts, and circuits. With such a background, any follow-up activity relating to the conversion of basic ideas and research developments into realistic equipment has been usually left to others. Developing and building the "hardware" has been considered to be unglamorous. The jobs of both the product engineer, with his attention focused on dimensions, tolerances, environments, and the like, and the production engineer, dealing only with machines, schedules, soldering irons, and other mundane matters, have appeared equally unglamorous.

Actually, this work now involves such a variety of specialized knowledge that future product and production phases can be very stimulating and rewarding. The opportunities afforded engineers who operate in these fields are very great, whether they do the initial planning and product engineering, or whether they follow up those designs which provide the articles of commerce that support research. Sales are now at levels reaching many billions of dollars a year, and this can be credited to no small degree to engineers who have created and produced equipment that is desirable to the customer. These engineers are expected to be conversant with a greater variety of skills than anyone else connected with the product, because each year brings forward new approaches to construction.

Whether modules should be repairable or "throwaway," which construction is the most versatile, and other similar matters, have not, and probably never will be, resolved. In any event, the modular-type construction, supplemented by solid-state circuitry, is clearly the main route to future construction. It is in this general area where engineers of our Professional Group will find both the greatest challenges and rewards.

Trends and Predictions

These predictions have been ventured because the panorama resulting from an intimate association with the past provides some vision of the future.

Items like television will undoubtedly use printed wiring and automatic insertion of conventional component parts for many years.

Satellites will use the micromodule and thin-film techniques as fast as they can be adapted. Missiles and giant computers will be only slightly behind satellites in this matter. Later, these techniques will be adopted by airborne electronics.

Printed wiring will get smaller and smaller as registration-capability and machining techniques improve, in keeping step with the availability of ever-smaller component parts.

Conveyorized automatic insertion of conventional component parts has already leveled off. Programmed insertion of conventional component parts may level off within the next 5 to 10 years.

The data-controlled assembly for the future is assured, but directed toward the fabrication of micromodules and other integrated circuit units for subsequent assembly into larger encapsulated units, in an integrated system.

Methodizing, just short of complete automation, will probably always be with us.

For production quantities of large systems, integration harnesses and cables will rapidly narrow down to the use of taper pins, wire wrap, or printed wiring. Flexible printed wiring and flexible printed cables will come into greater usage.

Conventional components may ultimately disappear (20 years?), but they will gradually be reshaped into modular versions.

The micromodule will make great impact on electronic production in a few years, having greater industry support than earlier modular schemes, and having some fifty times the earlier packaging density. However, one or two other promising micromodular concepts will compete with the micromodule during the next ten years.

Integration, cabling, connectors, and reliability will continue to be a major problem—probably indefinitely.

Until the relative reliabilities of the thin-film and the micromodule have been firmly established, it will be hard to predict whether or not the micromodule substrate dimensions will be extended to thin-film techniques. If thin-film reliability turns out to be equal to speculative figures, possibly some multiple of the micromodule dimension may be established, perhaps 2- to $2\frac{1}{4}$-in square. Semiconductor techniques for thin films will anticipate strictly solid-state methods for production units by several years.

Formed solid-state circuits will appear in production quantities by 1968, *grown* solid-state circuits by 1970, and *micron-sized* circuits by 1972, developed using Shoulders' "quantum mechanical tunneling of electrons" principle, or the like.

Automatic testing will continue to expand, and will be supplemented with many computer-controlled functions.

Following another decade of progress, an electronic system will be produced by a truly automatic factory, "engineered" on the typewriter keyboard of a master computer which will create a master memory, representing the desired equipment, from countless standard functional-circuit memories reproduced from its storage. This master memory will then requisition the more common functional devices from an automatic stock-room, and will automatically perform any missing deposition or other necessary operations of integration. The final result will be the desired electronic system which passes all tolerance and functional tests at all levels—from the component-part equipment to the complete equipment.

Bibliography

[1] L. Espenshied, "The origin and development of radiotelephony," Proc. IRE, vol. 25, pp. 1101–1123; September, 1937. (Includes 45 subsidiary references.)

[2] R. Brown, "Transoceanic radiotelephone development," *ibid.*, pp. 1124–1135. (Includes 19 subsidiary references.)

[3] IRE Annual Review Committee, "Radio progress during (the year preceding)," Proc. IRE, vol. 26, p. 279, March, 1938; vol. 27, pp. 161–183, March, 1939; vol. 28, pp. 99–125, March, 1940; and vol. 29, pp. 89–103, March, 1941. (Includes approximately 300 subsidiary references in each issue.)

[4] F. B. Jewett, "The mobilization of science in national defense," Proc. IRE, vol. 30, pp. 113–118; March, 1942.

[5] IRE Annual Review Committee, "Radio progress during (the year preceding)," Proc. IRE, vol. 30, pp. 57–71, February, 1942; vol. 31, pp. 127–132, April, 1943; vol. 32, pp. 125–129, March, 1944; vol. 33, pp. 143–155, March, 1945; vol. 34, pp. 164W–184W, April, 1946; and vol. 35, pp. 399–425, April, 1947. (Includes 70 to 400 subsidiary references in each issue.)

[6] F. Rocket, "Proximity fuze," *Electronics*, vol. 18, pp. 110–111; November, 1945.

[7] D. E. Noble, "Details of the SCR-300 F-M walkie-talkie," *Electronics*, vol. 18, pp. 204, 209, 212, 216; June, 1945.

[8] W. S. Hinman, Jr., and C. Brunetti, "Radio proximity fuse design," *J. Res. NBS*, vol. 37, p. 113; July, 1946.

[9] Staff Report, "Belmont's new pocket radio," *Radio Maintenance*, vol. 2, p. 16; August, 1946.

[10] Z. Benin, "Modern home receiver design," *Electronics*, vol 19, pp. 94–98; August, 1946.

[11] Staff Report, "Printed Circuit Wiring," *Electronic Ind.*, vol. 5, pp. 90–91, 120, 122; April, 1946.

[12] C. Brunetti and A. S. Khouri, "Printed electronics circuits," *Electronics*, vol. 19, pp. 104–108; April, 1946.

[13] C. Brunetti, and R. W. Curtis, "Printed Circuit Techniques," NBS Circular No. 468; November 14, 1957.

[14] H. H. Scott, H. Chrystie, and E. G. Dyett, Jr., "Application of miniature circuit techniques to the sound level meter," *Proc. NEC*, vol. 4, pp. 33–45; 1948.

[15] Seven miniature component parts, potted subassembly and printed-circuit references [(119)–(125)], Proc. IRE, vol. 38, p. 363; April, 1950.

[16] Ten miniaturization and reliability references [(1)–(4) and (19)–24)], Proc. IRE, vol. 40, pp. 388 and 390; April, 1952.

[17] W. C. White, "A look at the past helps to guess at the future in electronics," Proc. IRE, vol. 40, pp. 523–524; May, 1952.

[18] IRE Annual Review Committee, "Radio Progress During (the year preceding)," Proc. IRE, vol. 34, pp. 164W–184W, April, 1946; vol. 35, pp. 399–425; April, 1947; vol. 36, pp. 522–550, April, 1948; vol. 37, pp. 286–322, March, 1949; vol. 38, pp. 358–390, April, 1950; vol. 39, pp. 359–396, April, 1951; vol. 40, pp. 388–439, April, 1952; and vol. 41, pp. 452–507, April., 1953. (Includes 300 to 1100 subsidiary references in each issue.)

[19] W. H. Hannahs, "The shoemaker's child—no longer unshod?" Proc. IRE, vol. 44, p. 1733; December, 1954.

[20] R. W. Melville, "Mechanical design consideration in the ERMA system," IRE Trans. on Production Techniques, vol. PT-2, pp. 115–124; April, 1957.

[21] IRE Annual Review Committee, "Radio Progress During (the year preceding)," Proc. IRE, vol. 41, pp. 452–507, April, 1953; and vol. 42, pp. 705–759, April, 1954. (Includes approximately 1100 subsidiary references in each issue.)

[22] Four printed wiring and dip-soldering references [(472)–(475)], Proc. IRE, vol. 41, p. 471; April, 1953.

[23] Eight automation references [(146)–(153)], Proc. IRE, vol. 41, p. 714; April, 1954.

[24] Eight printed wiring and other production-technique references [(1099)–(1106)], *ibid.*, p. 759.

[25] E. R. Gamson, *et al.*, "Development and Application of Automatic Assembly Techniques for Miniaturized Electronic Equipment," Stanford Research Institute, Stanford, Calif., Contract AF 18-(600)-141; 1952–1954. (Several Reports.)

[26] A. R. Gray, "Guided missile reliability and electronic production techniques," 1955 IRE National Convention Record, pt. 6, pp. 79–91. (Includes 73 subsidiary references on printed wiring, automation, and process technology.)

[27] A. A. Lawson, "The economics of automation—some important considerations," IRE Trans. on Production Techniques, vol. PT-1, pp. 34–37; September, 1956.

[28] C. S. Selby, "Problems of semi-automatic assembly of electronic test equipment," *ibid.*, vol. PT-2, pp. 104–108; April, 1957.

[29] W. A. Schneider, "Automating small-lot electronic production," *ibid.*, vol. PT-3, pp. 50–59; April, 1958.

[30] H. K. Hazel, "A punched-card-controlled component-part insertion machine," *ibid.*, vol. PT-5, pp. 39–42; August, 1959.

[31] E. H. Bosman, "The status of standardization in electronic production and machine tool control," *ibid.*, vol. PT-2, pp. 55–58; April, 1957.

[32] G. D. Gruenwald and A. Attride, "A production design concept for electro-mechanical components," *ibid.*, vol. PT-3, pp. 19–25; April, 1958.

[33] A. W. Rogers, "Development of standards for automation," *ibid.*, pp. 39–44.

[34] J. J. Lamb, "The problem of compatibility between military equipment design and commercial automation techniques," *ibid.*, pp. 45–49.

[35] D. H. Evans, "Design by algorithm: a mathematical method of designing standard assemblies for minimum manufacturing cost," *ibid.*, vol. PT-4, pp. 4–10; June, 1959.

[36] R. A. Gerhold and W. V. Lane, "Modular dimensioning of electronic component parts for mechanized assembly," *ibid.*, vol. PT-5, pp. 12–16; August, 1959.

[37] C. P. Cardani, "Evolution of a system for the production of electronics equipment—mechanization of all lot sizes," *ibid.*, vol. PT-5, pp. 42–49; August, 1959. (Contains some United Shoe anvil standards.)

[38] H. H. Hagens, "Packaging and integration of transistor assemblies," 1958 IRE National Convention Record, pt. 6, pp. 12–19.

[39] P. E. Boran and E. N. King, "An approach to airborne digital computer equipment construction," IRE Trans. on Production Techniques, vol. PT-4, pp. 18–21; June, 1959.

[40] R. A. Martel and L. J. Martin, "Selection of coatings for printed wiring," *ibid.*, vol. PT-2, pp. 125–130; April, 1957. (Includes 10 subsidiary references on epoxy protecting coatings.)

[41] W. J. Hodges, "Eyelet failure in etched wiring," *ibid.*, vol. PT-2, pp. 109–114; April, 1957.

[42] W. B. Wilkens, "Insulated flexible printed circuits," 1958 IRE WESCON Convention Record, pt. 6, pp. 141–156.

[43] R. L. Henry and C. C. Rayburn, "Mechanized production of electronic equipment," *Electronics*, vol. 26, pp. 160–165; December, 1955.

[44] R. L. Henry, "Project tinkertoy," IRE Trans. on Production Techniques, vol. PT-1, pp. 10–11; September, 1956. (Includes 7 references to government reports on modular design fabrication and production.)

[45] A. R. Gray, "Automation for electronics—a 1956 status report," IRE Trans. on Industrial Electronics, vol. IE-4, pp. 29–44; March, 1957. Also, IRE Trans. on Production Techniques, vol. PT-2, pp. 29–44; April, 1957. (Includes 193 additional subsidiary references on printed wiring, automation, and process technology.)

[46] S. F. Nanko and V. J. Kublin, "Micro-modules: component parts and materials requirements," IRE Trans. on Production Techniques, vol. PT-5, pp. 29–38; August, 1959. (See text for source of other micromodule reports.)

[47] M. M. Perugini, "Microminiaturization," *Electronics*, vol. 31, pp. 78–108; November 25, 1960. (Includes 74 subsidiary references on micropackages, integrated circuitry, and molecular electronics.)

[48] S. M. Stuhlbarg and L. P. Sweaney, "Microminiature components and packaging techniques," 1961 IRE International Convention Record, pt. 6, pp. 3–17.

[49] A. E. Hawley, *et al.*, "A dot component packaging system for electronics," *ibid.*, pp. 18–28.

[50] N. Schuster and W. Reimann, "New developments in multilayered etched circuitry," *ibid.*, pp. 29–37.

[51] R. H. Cushman, "Interim packaging," *Electronic Design*, vol. 9, pp. 34–51; October 25, 1961.

[52] J. R. Riggs, "Microelectronics," *Elec. Mfg.*, pp. 121–126; September, 1960. (Includes 22 subsidiary references on micropackages integrated circuitry, and molecular electronics.) See also [47] and [48].

[53] J. M. Carroll, "What's new in semiconductors," *Electronics*, vol. 34, pp. 89–120; September 29, 1961.

[54] J. M. Carroll, "Molecular computer occupies 6.3 cu inches," *Electronics*, vol. 34, pp. 30–31; October 27, 1961.

[55] I. A. Lesk, *et al.*, "A categorization of the solid-state device aspects of microsystems electronics," Proc. IRE, vol. 48, pp. 1833–1861; November, 1960. (Includes 59 subsidiary references on basic, integrated, complex, and array devices.)

[56] J. J. Suran, "Circuit considerations relating to microelectronics," Proc. IRE, vol. 49, pp. 420–426; February, 1961. See also [47], [48], [52] and [53].
[57] H. S. Dordick, "Production testing in the automatic factory," IRE Trans. on Production Techniques, vol. PT-2, pp. 59–67; April, 1957.
[58] J. Tampico and H. B. Rose, "Minimizing production costs through modular automatic test equipment," *ibid.*, vol. PT-6, pp. 19–24; September, 1959.
[59] F. C. Brammer, "The semiautomatic circuit component tester," *ibid.*, vol. PT-6, pp. 12–18; September, 1959.
[60] E. L. Roel, "Trade-off considerations in using automatic test equipment," IRE Trans. on Product Engineering and Production, vol. PEP-5, pp. 39–44; April, 1961.
[61] Staff Report, "Automatic tube tester," and "Self-programmed circuit checkout," *Electronics*, vol. 34, pp. 82–83; October 13, 1961.
[62] D. Ehrenpreis, "Design improvement of an airborne system through mathematical prediction of its vibrations and dynamics," IRE Trans. on Production Techniques, vol. PT-4, pp. 11–17; June, 1959.
[63] W. Stubstad, "The application of thermoelectric spot cooling to electronic equipment," 1961 IRE International Convention Record, pt. 6, pp. 38–46.
[64] M. Nagata and Z. Abe, "Thermoelectric elements for circuit cooling," *Electronics*, vol. 34, pp. 54–66; October 13, 1961.
[65] D. Ehrenpreis, "Analytical dollar tradeoffs for development of electronic equipment," IRE Trans. on Product Engineering and Production, vol. PEP-5, pp. 2–8; April, 1961. (Includes 10 subsidiary references on weight, dynamics, and other tradeoffs.)
[66] A. P. Kromer, "Audio visual instructions for electronics assembly," *ibid.*, pp. 31–32.
[67] R. Lusser, "The Notorious Unreliability of Complex Equipment," Res. and Dev. Div., Ordnance Missile Labs., Redstone Arsenal, revised, November, 1957.
[68] A. R. Gray, "Reliability of missile guidance systems—the statistical cliche vs reality," 1958 IRE National Convention Record, pt. 6, pp. 49–71.
[69] A. R. Matthews, "Reliability and longevity for space technology," *ibid.*, pp. 123–136.

Section 24

RADIO FREQUENCY INTERFERENCE

Organized with the assistance of the IRE Professional Group on Radio Frequency Interference

Evolution of Regulatory Standards of Interference*

E. W. ALLEN, JR.†, FELLOW, IRE, AND H. GARLAN†, MEMBER, IRE

Summary—This paper traces the development of interference regulations from the original requirement that waves be pure and sharp to the present complex formula for attenuation of spurious emissions. It also shows how different philosophies of regulation developed. In broadcasting, the regulations grew up around the concept of a "protected service area." In the communication service, on the other hand, regulations developed in the direction of limiting the bandwidth and confining emissions within the authorized band.

The paper also traces the history of ISM regulation with its philosophy of assigned frequencies on which unlimited radiation is permitted with an associated requirement for severe limitation of radiation on other frequencies. It also discusses the "low-power rules" which establish the conditions under which low-power equipment may be operated without becoming a source of interference. The paper describes briefly the voluntary work in this area by industry, and touches on local regulations concerning interference.

Introduction

THE HISTORY of the regulation of radio interference is essentially the history of the regulation of radio itself. Many of the operating regulations and a substantial part of the technical regulations had as their basic objective the elimination or the control and the prevention of radio interference. The very first international agreement—the Final Protocol signed at the Berlin Conference in 1903—contains a provision in Article 5 that services be organized in such a way as to avoid interference with other stations [1]. This principle was carried over into the Convention drawn up and signed at the Berlin Conference in 1906, and was incorporated in legislation in this country in 1912.

The first radio law passed by the U. S., the Wireless Ship Act of 1910 (36 Stat. 629), required certain ocean-going vessels to be equipped with "an efficient apparatus for radio communication, in good working order, in charge of a person skilled in the use of such apparatus." This act did not impose any technical regulations beyond requiring the apparatus to be capable of transmitting and receiving messages over a distance of a hundred miles, day and night.

The earliest legislation providing in some detail for the technical regulation of radio was "An Act to Regulate Radio Communications" (37 Stat. 302), adopted on August 12, 1912. This act, which implemented the Berlin Convention of 1906, established a licensing procedure for stations and operators and set out specific technical regulations for radio stations. It provided that every license state the wavelength or wavelengths authorized for use by the station for the prevention of interference; it provided for the use of "pure" waves by requiring that, where energy was radiated on two or more wavelengths, "the energy in no one of the lesser waves shall exceed ten per centum of that in the greatest"; it required the use of "sharp" waves by specifying a maximum "logarithmic decrement of two-tenths" except when sending distress signals; the operator was required to listen on the channel before transmitting; he was required to use the minimum power to provide communications.

* Received by the IRE, August 18, 1961.
† Federal Communications Commission, Washington, D. C.

The subsequent approach to the control of interference between radio stations has consisted in a continued emphasis on proper operating procedures which promote time and geographical sharing, in refinements of the initial requirements as to purity and sharpness of waves, and in the development of new requirements as new concepts, new systems, and new techniques have evolved. The old requirement for purity, that no one of the lesser waves exceed 10 per cent, has long been obsolete. Purity of waves is now specified in terms of the required attenuation of spurious, including harmonic, components. Present requirements limit spurious emission at discrete frequencies to fractions of 1 per cent and permit no more than 1 per cent of the total radiated energy to be outside the assigned band. As the earlier use of damped waves (type-B emission) gave way to continuous waves, the sharpness became defined as bandwidth rather than in terms of decrement. The authorized wavelength has changed to assigned frequency upon which continually stricter tolerances for frequency stability are placed.

Although the useful radio spectrum has been expanded many fold during its history, the uses and demands for this spectrum have increased at an even faster rate. Thus there is continued need for ever stricter requirements, both operational and technical, to improve the effectiveness of spectrum use. As new techniques have been devised, they have in general been adopted at national and international levels as rapidly as is economically feasible. The improvement in radio techniques is reflected in the provisions of the Radio Regulations developed at Geneva in 1959, and now being adopted by countries which are signatory to the International Telecommunications Union. Of particular interest are the definitions in Article 1, Section III, and the tables in Appendices 3, 4, and 5 of frequency tolerances, spurious emissions, and bandwidths.

However, the specific figures for frequency tolerance, for spurious radiation, and for bandwidth do not of themselves insure noninterference. While the International Radio Regulations and the FCC Rules require compliance with these limits, the specific figures must be considered basically as design parameters for insuring good practice in the design, manufacture, and use of equipment, in order to minimize the probability of interference. These limitations must in every case be buttressed by a further requirement to take additional measures in the event that harmful interference is caused. The controlling requirement is that the interfering radiations be reduced to the point where well-designed and properly operated equipment can perform satisfactorily in its actual environment.

Improved frequency utilization by the reduction of assigned bandwidth is a concept which has long been emphasized. More recently it has been recognized that this practice does not always make for the optimum spectrum usage. Transmission systems using frequency modulation and pulse-code modulation provide improved systems performance, including discrimination against noise and interference, albeit at the expense of increased bandwidth. Improved discrimination against interference permits greater geographic sharing of frequencies, thus making possible increased frequency utilization. It is necessary, therefore, in each case to weigh the advantages of these improved transmission systems with their need for increased bandwidth against the benefits to be derived from the arbitrary reduction of bandwidth.

A second aspect of frequency usage, which is not apparent on its face and which has been developed over the years, is that increased spectrum usage in some services can be obtained by not providing protection against interference from other stations out to the maximum range at which the signal might otherwise be useful. In domestic broadcasting, duplication of assignments is intensive and neighboring assignments are generally so spaced geographically that service range is more severely limited by station interference than by receiver sensitivity or by other sources of interference. The limits of permissible interference are set in terms of protected contour of field strength and desired/undesired signal ratios, or an estimated protection is given by a table of minimum mileage spacings between stations.

Broadcast Services

The first broadcast rules issued by the Federal Radio Commission during the period 1927–1931 provided for the control of interference between stations by taking into account frequency separation, frequency tolerance, power, and geographic spacing of transmitters. Clear, regional and local channels were established, as well as the 10-kc channel spacing. On the recommendations of the IRE Committee on Broadcasting [2], a 500-cps frequency tolerance was established initially, followed by a reduction to 50 cps in 1932, when a minimum modulation requirement of 75 per cent was also imposed. Although the Committee had recommended a limit of 5 per cent and the need for control of harmonic radiation had been recognized [3], no limit for spurious, including harmonic, radiation was imposed at that time.

The approach to station assignment was changed in 1932 following the development of reliable groundwave and skywave propagation curves, ground conductivity maps, signal ratios, and information on urban and rural noise levels [4]. The field-strength and atmospheric noise-recording program inaugurated for this purpose in 1936 by the FCC was extended in 1938 and has continued until recent years. In 1946–1947 further tests on signal ratios and on urban noise were made by joint FCC-Industry committees in the clear-channel proceeding.

In FM and television broadcasting also, protection against station interference has used both the protected contour method and the mileage separation method at different times. These methods have been supplemented by channel allotment plans, in which specific channels are listed as available for station assignments in each city. These plans are developed in rule-making proceedings, in which interested persons may comment on the rules proposed by the Commission, and they may be supplemented by rules for the relatively simple procedures involving mileage separations or the more complex rules using signal ratios, propagation curves, etc. In general, in these two services, interference protection is afforded by the specification of minimum mileage separations for co-channel and adjacent-channel stations. In the UHF television rules, additional mileage separation requirements are imposed to protect against interference from receiver oscillator radiation and the spurious responses of receivers due to images, intermodulation, etc. Propagation curves and required field strengths and field-strength ratios are specified in the regulations for the purpose of estimating the location and extent of urban and rural service, so as to assist in decision-making processes, rather than for the purpose of providing specific protection to the service so indicated.

Specific limitations on spurious emissions were not included in the broadcast rules until 1953, when a limit of 60 db was imposed on TV stations [5]. Limits for AM and FM stations were imposed in 1959 [6], which provide for suppression of spurious emissions in accordance with a formula (more fully discussed below) which essentially limits the radiation of spurious emission to an absolute level of 50 microwatts.

Communication Services

In contrast to broadcasting, where regulations to minimize interference took the form of developing the concept of a protected service area, an entirely different concept of interference control grew up in the communications services. Since many of these services operate intermittently, it is feasible to assign more than one station to the same frequency for communication over the same, or overlapping, ranges. Accordingly, interference regulation was directed toward improving time-sharing of assigned frequencies. The early regulations required "monitoring before transmitting" (Parts 7 and 8), or "cooperation in the use of frequencies so as to avoid interference" (Parts 10 and 11). Stations were also required to reduce emissions outside the authorized band and to suppress spurious emissions, including audio and radio harmonics, to the "lowest practical level."

After World War II the enormous demand for radio stations in the communications services made it clear that more rigid control over spurious emission would be necessary. Accordingly, in 1947 the FCC proposed [7] that all emissions of land mobile stations outside the authorized band be attenuated at least 60 db. The comments received persuaded the FCC to dismiss this proceeding and to issue a new proposal [8] in June, 1948, which recognized that out-of-band components near the carrier require different treatment than harmonics and other emissions far from the carrier. In this new proposal, which did not apply to sideband components arising from frequency modulation, the required attenuation varied from 40 to 80 db depending on power. This regulation was seeking to provide essentially the same absolute power level for harmonic radiation from all transmitters. These rules became effective on July 1, 1949.

In 1950 spurious limitations were proposed for radiotelephone transmitters above 30 Mc in the maritime radio services [9] in which emission lying between 50 and 100 per cent of the authorized bandwidth removed from the carrier would be attenuated 25 db, and emission falling outside of this band would be attenuated from 40 to 80 db. In this proceeding the FCC also introduced a control technique, type acceptance of transmitters; however, this program was not fully formalized until January 5, 1955, when, in connection with type acceptance for the land mobile services, rules were promulgated [10] specifying in detail tests to be made and the information to be submitted. However, these limitations for spurious emissions were still not satisfactory. The transition from 25 db at 100 per cent of the bandwidth on each side of the carrier to 40–80 db was too abrupt, and the attenuation table included a number of discontinuities.

The first action to eliminate the discontinuities in the attenuation table was taken in 1954, in the rules to add spurious emission limitations for AM and FM broadcast stations [6] which substituted a formula ($50+10 \log P$) for the table. This formula implied an absolute level of harmonic radiation of 10 microwatts. In 1956 the formula was changed to $43+10 \log P$, based on a 50-microwatt level of harmonic radiation, when type acceptance and spurious emission limitations were added to the aviation radio services [11].

In the same year the FCC undertook to review the requirements [12] for out-of-band emissions close to the carrier for all its services. The first order in this proceeding, dealing with ship radio stations, established a three-step requirement by adding a step covering the range 100 to 250 per cent of the authorized bandwidth removed from the carrier. In this range spurious emissions must be attenuated 35 db. This three-step requirement was subsequently applied to maritime stations on land (Part 7) in 1958, to AM and FM broadcast stations in 1959, and to most aviation stations (Part 9) and land mobile stations except domestic public (Parts 10, 11, and 16) in 1960. Action to apply this requirement to the Domestic Public Land Mobile Service (Part 21) is still pending.

Industrial, Scientific and Medical Equipment

The first complaints of interference from equipment which we now group as ISM arose about 1930 when doctors started to use medical diathermy equipment [13]. Since these equipments operated on frequencies between 4 and 25 Mc, interference was particularly troublesome on long-range, point-to-point radio circuits. For the same reason, the interference from diathermy apparatus was experienced at distant points. Reviewing the background of the diathermy interference problem in 1940 [14], E. K. Jett, then Chief Engineer of FCC, reported that a diathermy machine keyed as a transmitter in Boston had been read in Washington, D. C., and at the FCC Monitoring Station at Great Lakes, Ill. Similar observations had been reported by Canada [15] in 1937.

The problem of diathermy interference was considered at the First [16] and Second [17] Inter-American Radio Conferences. The agreement drawn up at the Second Conference urged the participating countries to adopt standards of good engineering for diathermy apparatus.

Although the FCC had been aware of diathermy interference essentially since its inception, and had made a detailed investigation in 1937, it did not come to grips with this problem until 1939 when a conference [18] was held to bring this matter to the attention of the medical profession and the manufacturers of electromedical equipment. A second conference [19] was held in 1940 to implement the Santiago recommendations. Further work on standards for diathermy machines was interrupted by World War II. During the war, diathermy equipment was required by FCC Order No. 96 (May 18, 1942) to be registered to forestall its use for espionage and subversive purposes, since it was known that radiation from diathermy apparatus was capable of being received at great distances.

The question of frequency space for diathermy and industrial heating applications was raised in the allocation proceedings [20] in 1944–1945 when evidence was presented to show the importance of RF heating and that the number of installations and the amount of RF power used was increasing at a startling rate. The importance of diathermy as a therapeutic tool and in surgery was reiterated.

In the face of the demands of the industrial and medical interests for the allocation of large segments of the spectrum and of the conflicting demands of radio users for additional channels for communications and for protection against industrial and diathermy interference, the FCC in 1945 allocated [21] three frequencies—13.66, 27.32, and 40.98 Mc [22]—for medical and industrial purposes, which came to be known as the ISM frequencies. With this action, the FCC took the first step to implement the philosophy that had been suggested at the 1940 Santiago Inter-American Conference, that is, to provide frequencies on which unlimited radiation would be permitted and to prescribe severe limitations on radiation on other frequencies. This philosophy, however, was not specifically enunciated until the FCC adopted Part 18 of its rules in 1947. The same philosophy was accepted internationally in 1947 at the Telecommunications Conference at Atlantic City, N. J. The international frequency allocation adopted at this conference set aside three frequencies—13.56, 27.12, and 40.68 Mc—for industrial scientific, and medical purposes, and required communication circuits operating on these frequencies to accept any ISM interference that may be received.

An engineering conference held in 1946 to discuss standards was followed in 1947 by the adoption of Part 18 of the FCC rules [24]. The rules for medical diathermy equipment were essentially the standards that had been discussed at the 1946 Conference. It is interesting to note that FCC engineers had to build a diathermy machine to prove to industry that it was possible to comply with these rules at a reasonable cost [23]. However, Part 18 also included regulations concerning industrial heating equipment which were new. These rules called for operation on the ISM frequencies allocated previously, and imposed a limitation on radiation on all other frequencies. As a control mechanism the rules provided for type approval of equipment that operated on ISM frequencies and required certification of all other equipment. Although changes have been made from time to time, the basic philosophy and the basic limits—25 μv/m at 1000 feet for medical diathermy equipment and 10 μv/m at one mile for industrial heating equipment—have been retained [25].

Several changes are worth noting. Microwave frequencies were allocated [26] for ISM purposes in 1947, although they were not added to Part 18 until 1956 [27]. In 1948 the ISM rules were extended [28] to encompass other equipment "in which the action of the RF energy is directly on the work load," later changed to "in which the action of the RF energy is applied to materials to produce physical, biological or chemical effects." Separate and more severe provisions were promulgated in 1955 for RF equipment used to generate ultrasonics [29]. The radiation limits for industrial heating equipment were extended [30] into the microwave spectrum (up to 5775 Mc) in 1957. The most recent action (1961) promulgated [31] a standard form (FCC Form 724) to be used in certifying industrial heating equipment as a step toward tightening up the control mechanism over radiation from such equipment.

Low-Power Rules

At the same time that the FCC was struggling with the problem of diathermy interference, a second source of interference came to its attention. Early in 1938 the Philco Radio and Television Corporation placed on the market its Mystery Control, a miniature transmitter remote-control device for its standard broadcast receivers, in the belief that it could be operated without a radio station license. In view of the probability that similar devices would be developed and sold by other manufacturers and that these devices—esentially miniature transmitters—could develop into a serious interference problem unless adequately controlled, the FCC re-examined its position regarding licensing of low-power RF devices.

After a careful study the Commission came up with a proposal [32] to permit the operation of low-power devices without a license on the premise that devices that complied with the proposed standards would not be a source of interference to interstate communication and were therefore not subject to licensing. These rules came to be known as "The Low Power Rules." In addition to requiring the use of minimum power and precautions against interference, these rules required the radiated field to be limited to 15 μv/m at a distance equal to the wavelength divided by two pi ($\lambda/2\pi$). This regulation is often referred to as the "lambda over two pi rule." Much was made of the difference between the induction and radiation fields as justifying the use of $\lambda/2\pi$. The rules were adopted on November 21, 1938, essentially as proposed, and remained in effect for low-power communication devices until 1957. These rules are still in effect for campus radio and other carrier current systems.

The expansion in the use of radio after World War II brought with it an increase in use of low-power devices. But of more significance it brought a demand for use of frequencies above the broadcast band where the $\lambda/2\pi$ formula rapidly became meaningless. Accordingly, on April 13, 1949, the FCC instituted a proceeding [33] looking toward an over-all revision of Part 15, "The Low Power Rules."

The comments received in reply to this proposal indicated that comprehensive studies of the technical and economic aspects of the proposed regulations were necessary before any regulations were put into effect. Such studies were made by joint government-industry committees. On April 14, 1954, a Notice of Further Proposed Rule Making [33] was issued, in which the FCC divided the devices into two categories: incidental radiation devices in which the RF energy is generated as a by-product of normal operation, and restricted radiation devices in which RF energy is generated deliberately. The proposed rules contained a table of field strengths that increased with frequency. These limits were essentially the lowest practical values of radiation that the manufacturers of TV receivers were able to achieve [34]. This second notice also introduced a specific control procedure, certification, which the FCC borrowed from its ISM rules.

The comments received to this proposal were generally unfavorable except with respect to receiver radiation. Since the latter was the most pressing problem, the FCC finalized its receiver rules on December 21, 1955, and established a table of permissible radiation limits for all receivers operating (tuning) between 30 and 890 Mc. These rules provided a conducted interference limit (RF voltage into the power line) on frequencies below 25 Mc and a radiated field strength limit between 25 and 1000 Mc. These limits with two minor changes are currently in effect [35]. As a control mechanism, the rules require certification and permit each certificated receiver to be so labeled.

A second report and order in Docket No. 9288 was issued in 1956 to promulgate limits [35] for radiation from community antenna television systems, a type of carrier current system for the distribution of television programs. Rules for low-power communication devices—essentially miniature transmitters—were adopted in 1957 [36]. The FCC retained its $\lambda/2\pi$ formula recast in a new form as (2400/F) μv/m at 1000 feet for the frequencies 10 to 490 kc, and as (24,000/F) μv/m at 100 feet for 510 to 1600 kc. These rules also permitted use of the ISM band at 27.12 Mc with a power input of 100 milliwatts and a 5-foot antenna. To accommodate the maufacturers of radio controllers for garage door openers, the rules also permitted the use of frequencies above 70 Mc with radiation limited to the value permitted for receivers and subject to a duty cycle limitation of one second on in thirty seconds.

As of the date of this writing, action has been taken only in three categories of devices: radio receivers, community antenna television (CATV) systems, and low-power communication devices. All other restricted radiation devices are still subject to the original "$\lambda/2\pi$ rule" and the various proposals to modify this rule which are contained in FCC Docket No. 9288.

Interference Contributions by Industry

While this paper is written principally from the viewpoint with which the authors are most familiar, government regulation of interference, the extensive contributions of industry [37]–[41] must not be overlooked. These contributions take two forms: 1) Industry has developed a number of voluntary standards for the control of interfering radiation, most prominent of which are the SAE standard for ignition interference [42] and the NEMA standards for electrical equipment [43]. 2) A large part of the information on which industry-adopted limits and governmental regulation are based is derived from work carried on in private laboratories or by industry and joint government-industry committees. Space does not permit a delineation of this work, but this discussion would not be complete without mention of the more important committees such as: the Joint Coordination Committee of the NELA, NEMA, and RMA; SAE Vehicular Radio Interference Subcommittee; ASA Committee C-63; the many IRE committees, such as the committees on Broadcasting and Wave Propagation; the AIEE Subcommittee on Induction and Dielectric Heating; the National Television Systems Committees for monochrome and color television; the Radio Technical Planning Board for the post-World War II expansion of radio uses; the Joint Technical Advisory Committee (JTAC); the Radio Technical Commission for Marine (RTCM); the Radio Technical Commission for Aeronautics (RTCA); the International Radio Consultative Committee (CCIR); and many special committees such as the Ad Hoc Committee for FM and TV, and the Clear Channel Broadcast Committees. The reader is referred to the proceedings of the above committees for details of their contributions to the many facets of interference regulation and control.

Local Governmental Control of RFI

No discussion of interference regulation would be complete without mention of the ordinances that have been enacted by many communities toward the control of radio noise and other interference stemming from the police powers of the community to abate a public nuisance. These ordinances are primarily directed against interference to broadcast reception. The FCC has helped the National Institute of Muncipal Law Officers (NIMLO) to prepare a model ordinance in the interest of promoting uniformity in local regulation. These ordinances basically provide a field-strength limit with penalties for failure to comply. While legally simple, this approach involves many difficult technical problems; one of the most difficult of these is the problem of making reliable measurements. Inability to solve these problems within the means of the local community is undoubtedly one of the major reasons for the lack of success of this approach.

Bibliography and Notes

[1] J. D. Thomlinson, "The International Control of Radiocommunications," J. W. Edwards, Ann Arbor, Mich.; 1945.
[2] "Reports of IRE Committee on Broadcasting," Proc. IRE, vol. 18, pp. 15–37; January, 1930.
[3] S. C. Hooper, "Considerations affecting the licensing of high frequency stations," Proc. IRE, vol. 16, pp. 1240–1251; September, 1928.
[4] A. D. Ring, "Empirical standards for broadcast allocations," Proc. IRE, vol. 20, pp. 611–625; April, 1932.
[5] FCC Docket No. 10353; 1952–1953.
[6] FCC Docket No. 11233; 1954–1959.
[7] FCC Docket No. 8294; 1947–1948.
[8] FCC Dockets No. 9001, No. 9018, No. 9046 (later pt. 21), and No. 9047; 1948–1949.
[9] FCC Docket No. 9797; 1950–1952.
[10] FCC Docket No. 10798; 1953–1955.
[11] FCC Docket No. 11619; 1956.
[12] FCC Docket No. 11654; 1956–1960.
[13] G. E. Sterling, "History of Federal Control of ISM Radio Frequency Equipment," presented at joint meeting of IRE and Industrial Electronics Organization, Cleveland, Ohio; January 25, 1951.
[14] E. K. Jett, "Radio Interference Problems Resulting from the Operation of Electro-Medical Apparatus," speech before the 19th Annual Session of the American Congress of Physical Therapy, Cleveland, Ohio; September 2, 1940.
[15] Canadian Delegation, "Report on Radio Interference from Electro-Therapeutic Apparatus" presented to the 1st Inter-American Regional Radio Conf., Havana, Cuba; November 1, 1937.
[16] "Inter-American Arrangement Concerning Radiocommunications," 1st Inter-American Regional Radio Conf., Havana, Cuba, Section 12; November 1, 1937.
[17] "Inter-American Radio Agreement," 2nd Inter-American Regional Radio Conf., Santiago, Chile, Revision to [16] Annex I, Sections 9–11; January 18–26, 1940.
[18] Conf. on Interference to Radio Reception Caused by Electro-Medical Equipment, New York, N. Y., January 9, 1939.
[19] Informal Engrg. Conf. on the Problem of Interference to Radio Communications Caused by the Operation of High Frequency Electro-Medical Equipment, FCC Docket No. 5973, Washington, D. C.; November 20, 1940.
[20] FCC Docket No. 6651; 1944–1952.
[21] *Ibid.*, "Report of Allocations from 25,000 KC to 30,000,000 KC," May 25, 1945.
[22] Subsequently changed to 13.56, 27.12, and 40.68 Mc to conform to the Atlantic City Internatl. Radio Regulations; 1947.
[23] E. W. Chapin, W. K. Roberts, and M. C. Mobley, "FCC diathermy design for low harmonic radiation and good frequency stability," *Trans. AIEE*, pt. I, vol. 67, pp. 42–46; 1948.
[24] FCC Docket No. 7858; 1946–1948.
[25] W. C. Boese and H. Garlan, "The Federal Communications Commission and the control of radio interference," *Proc. 2nd Conf. on Radio Interference Reduction*, Armour Research Foundation, Chicago, Ill., pp. 161–174; March, 1956.
[26] FCC Docket No. 6651, by Order dated May 15, 1947.
[27] *Ibid.*, by Order (*nunc pro tunc*) dated December 19, 1956.
[28] FCC Docket No. 7858, by Order dated March 10, 1948.
[29] FCC Docket No. 11031; 1954–1955.
[30] FCC Docket No. 11442; 1955–1957.
[31] FCC Docket No. 13511; 1960–1961.
[32] Transcript of informal hearing in the matter of Proposed Rules and Regulations Governing the Operation of Low Power Radio Frequency Devices, Washington, D. C.; September 19, 1938. FCC Docket No. 5335; September 19, 1938.
[33] FCC Docket No. 9288; 1949—still pending.
[34] G. E. West, "Receiver Radiation—A Chronological Digest of FCC—Industry Correspondence with Notations Regarding the Development of Standards and Limits," FCC, Washington, D. C., T.R.R. Rept. No. 6.1.12, June 10, 1953, presents a detailed review of negotiations and correspondence between the FCC and the manufacturers of broadcast receivers regarding the limitation of radiation from FM and TV receivers.
[35] L. G. Whipple and H. Garlan, "Government regulation of unlicensed RF equipment as a means of controlling RFI," *Proc. 3rd Natl. Symp. on Radio Frequency Interference (PGRFI)*, Washington, D. C.; pp. 71–72; June 12–13, 1961.
[36] FCC Docket No. 9288, 3rd rept. and Order; 1957.
[37] E. W. Allen, "Man-Made Radio Noise," presented at Conf. on Radio Noise, Harvard College Observatory, Cambridge, Mass.; April 22, 1958. Also in "The Radio Noise Spectrum," D. H. Menzel, Ed., Harvard University Press, Cambridge, Mass., 1960.
[38] Joint Coordination Committee on Radio Reception (Natl. Elec. Light Assoc., Natl. Elec. Mfgrs. Assoc., Radio Mfgrs. Assoc.), "Radio Interference," NELA Publication No. 33; January, 1933.
[39] P. J. Kent, "The Automotive Industry's Participation in Reduction of Radio and Television Interference," presented at Society of Automotive Engineers Summer Meeting, French Lick, Ind.; June 6–11, 1948.
[40] R. W. George, "Field strength of motor car ignition between 40 and 450 Mc," Proc. IRE, vol. 28, pp. 409–412; September, 1940.
[41] "Recommended Practice for Minimization of Interference from Radio Frequency Heating Equipment," AIEE Publication No. 951; May, 1950.
[42] "SAE Standard for Measurement of Vehicle Radio Interference (30–400 Megacycles," Soc. Automotive Engrs.; January 15, 1958.
[43] Typical of these standards is "NEMA Standards for Small Power Motors," Natl. Elec. Mfgrs. Assoc., Publication No. 47-121, Section MG3, p. 1, MG3-45; October 1, 1948.

References

[44] Annual Reports, Federal Radio Commission, Washington, D. C.
[45] Annual Reports, Federal Communications Commission, Washington, D. C.

Radio Frequency Interference Measurements and Standards*

HAROLD E. DINGER†, FELLOW, IRE

Summary—**Radio frequency interference has been a particularly difficult area for achieving effective standardization, principally because of technical considerations, but also because it is of concern to so many interests and activities. Prior to World War II, standardization efforts were mainly confined to the medium-frequency broadcast band. During the war it became evident that the problem must be considered on a much broader basis. More recently, highly critical systems being used in space research have placed additional requirements on system engineering for interference-free operation. Radio frequency interference can mean the difference between success and failure of a costly space project.**

There are four general categories in which standardization efforts are required: 1) terms and definitions, 2) measurement methods, 3) instrument characteristics, and 4) interference limits. Various professional, industrial, and governmental organizations are working at both national and international levels on standards in these areas.

Introduction

Reaching agreement on any standard is often a long and controversial process, but radio frequency interference (RFI) has proved to be an exceptionally difficult area for effective action. There are a number of reasons why this is so:

1) Most radio noise voltages are of such a nature that they are difficult to specify and measure. They may be periodic or random. They may be narrow band or broad band. The waveform of ignition noise or a radar pulse is greatly different from that generated by a motor or a gas discharge device.
2) The range of levels encountered (the ratio of peak-to-average values of a radar signal or the ratio or the fundamental-to-harmonic power of a transmitter output) requires a large dynamic range, good sensitivity, and good instrument shielding. Direct-reading instruments are usually desired, and this introduces problems concerned with logarithmic vs linear scales and with attenuator design. The problem is further complicated by bandwidth considerations.
3) Specified limits for the interference voltage to be measured are usually quite low, which means that ambient radio noise levels can prevent measurement unless shielded enclosures are used. This is not always possible, as in the measurement of radiation from high-tension lines, for example. The use of shielded enclosures introduces other problems, both technical and economic.
4) Antenna considerations such as size, tuning, directivity, antenna factor, impedance matching, near-field effects, and bandwidth become complicated because of the wide frequency range required, particularly if the antenna is to be used in shielded enclosures.
5) Portability is desired for field measurements. This introduces problems of weight, size, power source, and provision for calibration—all demanding compromises.
6) RFI may be propagated by radiation, induction, or conduction, each of which requires special measurement techniques and specifications.
7) The subjective effect of interference varies with the kind of system affected. An interference that would disrupt television reception might not affect certain radar devices, even in the same part of the spectrum.
8) Many different groups are interested in standardization of radio frequency interference instruments and techniques, and they place different emphasis on certain requirements. Some groups are interested in the complete spectrum; others are interested only in broadcast and television services. Some groups are primarily interested in only one form of interference, others in all forms. Some are interested as manufacturers or users of interference-producing devices, others as manufacturers or users of devices subjected to the interference.
9) Considerable technical skill is required. This is not always readily available.
10) The economic aspects of the problem are often difficult to resolve. If, in order to meet the requirements of several different contractors, a single device must be measured in accordance with several different measurement standards for compliance with several different limit specifications, the cost of such tests can easily be a major factor in determining the selling price. This is, of course, in addition to the cost of any interference-reduction measures required.
11) There has been a tendency in some quarters to downgrade radio interference work, with the implication that this field is rather routine and does not challenge the efforts of the ablest engineers. Fortunately, this attitude is now less prevalent since the complex and many-faceted character of RFI problems is becoming dramatically apparent.

* Received by the IRE, July 5, 1961.
† U. S. Naval Research Laboratory, Washington, D. C.

To attempt a discussion, or even a brief mention, of all the various technical aspects of the radio frequency measurement problems that have been injected into standardization deliberations would require more space than is available. Nor is it practicable to discuss all of the current standards and specifications; what is current at the time of writing may be obsolete by the time of publication. Rather, an attempt will be made to outline the historical development along this line in a general way and to indicate current trends. Since the development of instrument specifications and measurement standards has been interrelated, they will, in general, be treated concurrently. The references will furnish more detailed information in the various areas treated. Frequency allocation questions are not under consideration here, although they constitute one of the major problem areas.

Radio Noise Parameters

One of the first things to be decided in radio noise measurement standardization is what parameter or parameters of the radio noise voltage should be measured. There are several choices: 1) average value, 2) effective value, 3) peak value, 4) a quasi-peak value, or 5) certain statistical distributions. Each of these has been used at one time or another. These parameters are not necessarily uniquely determined since they are influenced by the receiver bandwidth in a manner that varies with the type of noise. Sometimes measurements are desired in terms of power, and frequently the end result will be in the form of a signal-to-noise ratio. The ideal arrangement would be to have one instrument specification, one measurement standard, and one limit specification. Attempts have been made to achieve this, but it turns out that different objectives require the measurement of different parameters. For example, a world-wide program for measuring atmospheric radio noise uses different parameters than those generally used in measuring the interference created by a radar transmitter or an ignition system.

A condensed topical outline of the various sources of radio interference and certain receiver characteristics that must be considered in an over-all RFI measurement or standardization program might appear as follows:

1) Transmitters (including Industrial, Scientific, and Medical)
   a) Co-channel
   b) Adjacent channel
   c) Harmonic
   d) Spurious
   e) Intermodulation
2) Receivers
   a) Selectivity considerations
   b) Cross modulation
   c) Intermodulation
   d) Image response
   e) Hum
   f) Tube and thermal noise
   g) Radiation from receiver circuits
3) Atmospheric radio noise
4) Precipitation static and Corona discharge
5) Extraterrestrial radio noise
6) Man-made (from electrical and electronic devices of many kinds).

It is generally desired to measure these sources in order to determine their interference capabilities to other radio services, but in some cases it is necessary to determine the extent to which they are a hazard to the human body or to the safety of combustibles or explosives. In addition, undesired emissions from receivers and man-made sources can also endanger military security. In general, each of the above kinds of interference requires its own method of specification and measurement.

Prior to World War II the principal RFI standardization efforts were almost entirely confined to the medium-frequency broadcast band of radio frequencies. The first attempts at radio noise measurement were in terms of "nuisance value," and since the primary concern was of broadcast interest, the nuisance value was determined subjectively by listener tests. However, there are now many types of signal display devices, and nuisance value has become rather meaningless. Even in listener tests it was not unique since there was considerable disagreement in its determination. A completely objective system of measurement would require the use of a prohibitively large number of parameters; thus, some form of compromise is essential.

Early History—Some RFI Firsts

The first case of man-made interference that has come to the author's attention was that mentioned in the "Reminiscences" of Dr. A. Hoyt Taylor, a Past President of the IRE, who recalls hearing ignition interference from a two-cylinder automobile driven by R. E. Olds in Lansing, Michigan, in the year 1902. Serious transmitter interference was reported during the press coverage of the International Yacht Races in 1901, and again during the Russo-Japanese War in 1904. The first IRE paper devoted to radio noise elimination was published in 1917,[1] and the first on radio noise measurement in 1921.[2] A paper on electromagnetic shielding appeared in 1925.[3] From 1926 on the general problem of radio interference began to receive increased attention from a limited number of investigators.

The first attempt at standardization of radio noise measurements was made by the Joint Coordination Committee (JCC) on Radio Reception of the Edison

[1] C. J. deGroot, "On the nature and elimination of strays," Proc. IRE, vol. 5, pp. 75–132; April, 1917.

[2] A. M. Curtis, "A system for measuring the amount of static," Proc. IRE, vol. 9, pp. 225–227; June, 1921.

[3] J. H. Morecroft and A. Turner, "The shielding of electric and magnetic fields," Proc. IRE, vol. 13, pp. 477–505; August, 1925.

Electric Institute, the National Electrical Manufacturers Association, and the Radio Manufacturers Association. This committee was organized in 1931 and published a report[4] in January, 1933, in which certain desirable characteristics for a radio-noise meter were specified. The first meter designed expressly for radio noise measurement appeared in the same year,[5] but it was not suitable for impulse-type noise. The first radio noise meter employing a weighted response circuit was probably one built by the Central Electrical Laboratory of Belgium in accordance with specifications published by the International Special Committee on Radio Interference (CISPR).[6] This was followed shortly by a British model. Neither of these instruments were available in quantity, however. In 1938 an American meter incorporating the recommended CISPR time constants was put on the market.[7] In May, 1939,[8] and February, 1940,[9] certain U. S. interests recommended a change in the time-constant values of the weighting circuit and specified other desirable characteristics. In March, 1940,[10] a meter incorporating the new specifications was described. Some of the papers cited above also discussed early attempts at standardization of measurement methods. A more detailed account of the above developments has been treated by Burrill.[11] Several thousand radio noise meters have been manufactured in accordance with the 1940 specifications.

During World War II the frequency range of interference measurements was extended, but otherwise the same instrument specifications and measurement methods were generally used. The military services set up accelerated interference-reduction programs which began to consider some of the problems brought about by the introduction of new and complicated electronic systems, but the available manpower was generally limited to using available instruments and techniques on specific interference-reduction projects. It became apparent, however, that a more objective type of measurement was required. The use of pulse systems and visual display devices called for a reappraisal of the noise parameters to be measured. Measurement accuracy and repeatability left much to be desired, and since meters built to the same specifications often gave different results, the existing specifications were apparently inadequate. In the meantime the American Standards Association War Committee on Methods of Measuring Radio Noise had cooperated with the Armed Forces in the development of a uniform method for the measurement of the conducted interference and the interference field intensity produced by many kinds of electrical equipment being furnished for military purposes. This became known as Joint Army-Navy Specification JAN-I-225, and American War Standard C63.1-1956. It was dated June 14, 1945 and approved August 6, 1946. It became the first widely used standardized RFI measurement technique, but because of rapid advances in radio technology, it was soon supplemented or replaced by more specialized requirements.

Post-War Efforts

During the war some of the factors affecting the accuracy of radio noise meters had been noted, but little effort could be spared for systematic study. Immediately after the war these problems began to receive attention in government laboratories[12] and universities.[13] The frequency range of interference-measuring instruments was extended, both upward and downward, and new measurement specifications were issued by various agencies. The general radio interference situation as it appeared in 1954 was reviewed in a report by staff members of a prominent trade journal.[14] A survey of RFI instrumentation currently in use is given by Haber and Showers.[15]

In the meantime several organizations had been working on standards. There are four general categories in which radio frequency interference standardization is required:

1) terms and definitions,
2) measurement methods,
3) instrument characteristics,
4) interference limits.

The following organizations have issued specifications or standards in one or more of the areas of interest:

1) Military (Army, Navy, Air Force, Department of Defense),
2) FCC,
3) American Standards Association (includes representation from industry, universities, government, and professional organizations),
4) Professional societies (IRE, AIEE),

[4] "Radio Noise Measuring Set," Natl. Elec. Light Assoc., Publ. No. 32; January, 1933.

[5] C. R. Barhydt, "Radio noise meter and its application," *GE Rev.*, vol. 36, pp. 201–205; April, 1933.

[6] International Special Committee on Radio Interference, Rept. R.I. (Belgium) 1, November, 1937, and International Committee on Radio Interference, Rept. R.I. (Belgium) 2; December, 1937.

[7] C. J. Franks, "The measurement of radio noise interference," *R.M.A. Eng.*, vol. 3, pp. 7–10; November, 1938.

[8] C. V. Aggers, D. E. Foster, and C. S. Young, "Instruments and Methods of Measuring Radio Noise," AIEE Tech. Paper 39-144, May, 1939; and *Elec. Engrg.*, vol. 59, pp. 178–192, March, 1940.

[9] "Methods of Measuring Radio Noise," Edison Elect. Inst., Publication No. G9; Natl. Elec. Manufacturers Assoc., Publication No. 107; Radio Manufacturers Assoc., Engrg. Bulletin No. 32, February, 1940.

[10] C. M. Burrill, "New equipment to measure intensity of radio noise," *Broadcast News*, No. 32, p. 6; March, 1940.

[11] C. M. Burrill, "Progress in the development of instruments for measuring radio noise," Proc. IRE, vol. 29, pp. 433–442; August, 1941.

[12] H. E. Dinger and H. G. Paine, "Factors affecting the accuracy of radio noise meters," Proc. IRE, vol. 35, pp. 75–81; January, 1947.

[13] "Investigation of the Measurement of Noise," Progress Repts. Nos. 1 to 58, University of Pennsylvania, Philadelphia, Pa.; March, 1946, through September, 1960.

[14] "How to Suppress Radio Interference," *Elec. Mfg.*, Staff Rept., September, 1954.

[15] F. Haber and R. M. Showers, "Instrumentation for radio interference measurements," *Electronic Industries*, March, 1961.

5) International organizations (CISPR, CCIR, URSI),
6) National organizations in other countries.

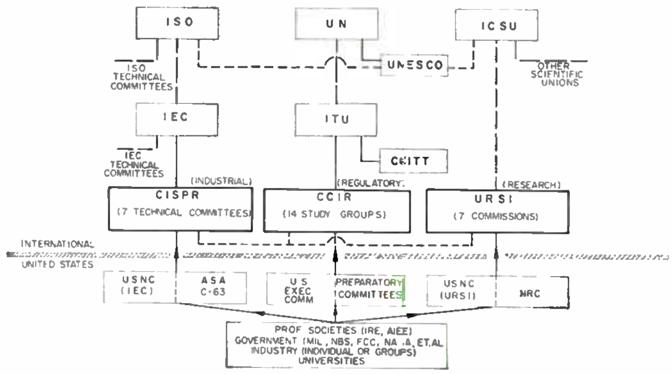
Several agencies that were working in this area during the war or immediately thereafter have ceased to exist, such as the National Defense Research Committee (NDRC) and the Research and Development Board (RDB).

As previously mentioned, the first radio interference standardization was attempted in 1931–1933, by a group of industrial organizations (JCC). The first report issued by the International Special Committee on Radio Interference (CISPR) was that of a meeting held in Paris in June, 1934. This group was formed as a special committee of the International Electrotechnical Commission (IEC) for the express purpose of reaching international agreement on radio interference standards. Except for the war years, the CISPR has been meeting more or less regularly since 1934 and has issued a number of reports[16] containing recommendations for specifications on instruments, methods, and limits.

The American Standards Association (ASA) Committee on Radio-Electrical Coordination (C-63) was organized under the sponsorship of the Radio Manufacturers Association. During the war it was redesignated the War Committee on Methods of Measuring Radio Noise. In addition to working on American Standards, this committee (now sponsored by the National Electrical Manufactuers Association) is in effect the group that, working through the U. S. National Committee of the IEC, represents the United States for cooperation with the CISPR.

Also working in the international radio interference field are the International Radio Consultative Committee (CCIR)[17] and the International Scientific Radio Union (URSI).[18] The CCIR makes recommendations to the International Telecommunications Union (ITU) which is charged with the international control of radio communications. The URSI is a purely scientific organization which enters the picture only from the viewpoint of international cooperation in scientific research programs. A chart showing some of the interrelationships among these organizations is given in Fig. 1. In general, excellent coordination of effort is being achieved among these groups.

The professional societies of the IRE and the AIEE are concerned with standards in their respective fields of interest. The IRE has several committees and subcommittees charged with producing standards in terminology and methods of measurement. The AIEE has been concerned with methods of measuring radiation



AIEE	American Institute of Electrical Engineers
ASA	American Standards Association
CCIR	International Radio Consultative Committee
CCITT	International Telegraph and Telephone Consultative Committee
CISPR	International Special Committee on Radio Interference
FCC	Federal Communications Commission
ICSU	International Council of Scientific Unions
IEC	International Electrotechnical Commission
IRE	Institute of Radio Engineers
ISO	International Standards Organization
ITU	International Telecommunications Union
NASA	National Aeronautics and Space Agency
NBS	National Bureau of Standards
NRC	National Research Council
UN	United Nations
UNESCO	United Nations Educational, Scientific, and Cultural Organization
URSI	International Scientific Radio Union
USNC	U. S. National Committee

Fig. 1—Organizations concerned with international RFI standardization.

from radio frequency heating equipment, or more generally, equipment falling in the industrial, scientific, and medical categories. This category is subject to regulation by the FCC, which is itself concerned with measurements to insure conformity with its rules and regulations. In practice, these various groups work together very closely with some overlapping memberships on the various committees dealing with radio interference.

In 1957 the IRE Professional Group on Radio Interference was formed. This group co-sponsors radio interference conferences such as those held annually at the Armour Institute of Technology, and also holds annual symposia on current RFI topics. In addition, it publishes a TRANSACTIONS and a *Newsletter*, and in general provides an excellent medium for the exchange of ideas and for inciting action toward better measurement and control of radio frequency interference.

It has not been possible to make mention here of all the RFI standardization activities occurring in other countries. Reference to much of this work can be found in the CISPR reports.[16]

It may appear that the RFI problem has been plagued by too many organizations attempting to set up standards in the same field, but, as pointed out in the Introduction, this situation occurs because of the widespread need for standards which are technically difficult to attain, resulting in numerous attempts to achieve a measure of standardization in specific areas. Each organization has had its own field of interest and its own

[16] CISPR Repts. R.I. 1, June, 1934, through Rept. R.I. 14, November, 1958.
[17] J. W. Herbstreit, "The ninth plenary assembly of the CCIR," PROC. IRE, vol. 48, pp. 45–53; January, 1960.
[18] J. H. Dellinger, "International cooperation in radio research—URSI and IRE," PROC. IRE, vol. 44, pp. 866–872; July, 1956.

sphere of action. In general, the various efforts have been well coordinated and progress has been made. Nevertheless, much remains to be done.

Current Trends

The Department of Defense has recently embarked upon an extensive program referred to as Electromagnetic Compatibility. This program entails new specifications and standards as well as some rather complex and sophisticated measurement techniques. A facility called the RFI Analysis Center has been established at Annapolis, Md., which will collect and analyze spectrum signatures of both new and existing electronic equipments and systems. The proper analysis of the spectrum signatures for all equipments to be used in a particular electronic environment should permit the prediction of the degree of compatibility and indicate where further RFI reduction efforts must be exerted.

Initially, considerable research and development is necessary to determine the essential data requirements and suitable test methods. Manufacturers will be required to submit spectrum signature data with preproduction models of equipments and eventually it is hoped that this data can be supplied with design proposals on the basis of prior knowledge of results to be expected for certain design considerations.

This program shows promise of being a definite step forward in achieving an organized engineering approach to the control of radio frequency interference.

In the "space age" costly failures arising from preventable interference cannot be tolerated, whether the interference be due to desired or undesired emissions.

Analysis and Prediction of Interference Factors and Their Impact on the Future*

R. M. SHOWERS†, SENIOR MEMBER, IRE

Summary—**This paper is based on the theme that although there have been, admittedly, many instances of spectrum crowding, and many specific examples can be cited, no one knows just how crowded the available spectrum is. This is because quantitative measures of the use of this important natural resource are not established. The factors contributing to spectrum crowding are outlined briefly and suggestions for the development of methods for determining their importance are made.**

Introduction

PREDICTION of the amount of interference to be expected in any given situation is or should be the ultimate objective of all interference work. The accuracy to which this can be done may be assumed to indicate the degree to which this field has developed from an art to a science. In this paper the present state of the art will be reviewed and some ideas offered as to how this work may develop in the future. There seems to be little doubt that interference considerations will continue to become much more important in the design of electrical and electronic equipment than they are at the present time.

* Received by the IRE, August 9, 1961; revised manuscript received, November 28, 1961.

† The Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.

Interference Factors

The definition of an "interference factor" must necessarily be somewhat vague since almost any "electrical" characteristic of a device or its "radiation" could in some way determine interference in some particular case. All such factors used to describe a communication system or its components should be included in a broad definition. Then interference analysis can be thought of as the process of determining the order of importance of such factors in so far as they contribute to specific interference situations and the methods of utilizing these factors in evaluation of such situations.

The model of Fig. 1 which is used extensively in information theory work is also convenient to use here in modified form.

Interference can be classed as arising from two causes. It may be man-made or natural. In this paper we are primarily concerned with the man-made variety. However, except for lack of control of the source and the usually extremely broad-band nature of natural interference, essentially the same considerations are involved in predicting interference from both.

It is convenient to classify a factor as either a "primary" or "secondary" factor. A primary factor is one which describes a function essential to the transmission purpose, whereas a secondary factor is one which it would be desirable to eliminate in order to get maximum efficiency in spectrum utilization. This distinction is significant for most factors. However, in the case of propagation, a factor may be primary for one circuit and secondary for another depending on the way it is involved. The broadcast service, where ground wave propagation may be used locally and sky wave for longer distances, is an example of this. However, usually both are not used at the same time so that in most instances the distinction is useful.[1]

Allocation of frequencies is usually made on a clear or shared channel basis. A clear channel is defined here as one in which the natural "noise" exceeds man-made noise arising as a consequence of a primary factor. Thus, taking advantage of propagation and geographical factors (describing in part the "channel" above), one can effectively have many clear channel users on the same frequency. The shared channel basis is one in which interference will be experienced if all or some of those assigned to it try to use it simultaneously.

In order to predict interference one must have detailed data on each of the elements shown on Fig. 1, which are outlined in more detail in Table I. It must be recognized that assigning numerical values to the factors must be done statistically. All receivers are not the same, even though they come off the same production line. Likewise, spurious outputs from transmitters will vary from one production unit to the next, and will vary with tube replacement in an unpredictable manner. Thus, a given factor can be specified quantitatively in terms of a mean value and mean square deviation from this mean square value. Design objectives for many of these factors have been stated in some detail in the literature.[2] These objectives have been established on the basis of 1) the apparent importance of each factor and 2) what appears to be technically feasible.

In the last year or so the military services have developed the concept of "spectrum signature."[3] The spectrum signature of any transmitter is essentially a description of the power output as a function of frequency. For a receiver, the spectrum signature, in addition to information on power radiated, contains information on its response as a function of frequency. This includes cross-modulation and desensitization data. In addition to the above, the spectrum signature will include antenna pattern characteristics and data on types of modulation or other descriptions of waveform. In other words, much of the information listed in Table I is included in the spectrum signature.

In some cases one can make a prediction within close tolerance if he takes care to measure all parameters or factors involved in sufficient detail. However, usually the usefulness of a prediction is in direct proportion to its remoteness from such detailed measurements. Furthermore, characteristics of individual devices are subject to variation from time to time, not to mention propagation effects. This is clearly a complex situation. It is obvious that in specific situations many of these factors will be unimportant. On the other hand, any of them may be significant in a particular situation.

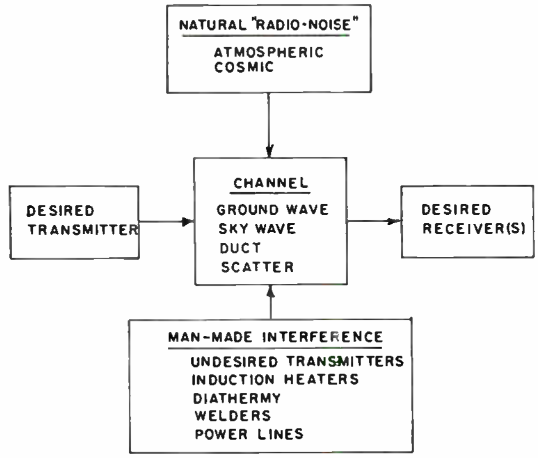


Fig. 1—Model of a communications circuit operating in the presence of noise and interference.

TABLE I
FACTORS IN INTERFERENCE PREDICTION

Transmitter
- Antenna height, pattern, and polarization
- Power spectral output
- Modulation type
- Intermodulation or cross-modulation characteristics
- Frequency stability.

Natural Noise (specified as function of geographical location)
- Spectral and amplitude characteristics (described statistically).

Man-Made Noise
- Where appropriate (for example, ignition and power line corona), spectral and amplitude characteristics (described statistically).
- For Incidental Radiation Devices (Induction heaters, diathermy)
  - Power spectral output
  - Waveform
  - Bandwidth.
- For Undesired Transmitters
  - Same as for (desired) transmitters above.

Channel
- Propagation characteristics (for particular frequencies and geometrical locations of sources and receivers).

Receiver
- Antenna height, patterns, and polarizations
- Bandwidth (of input stages as well as over-all)
- Type of detector
- Nonlinear characteristics of early stages
- Overload
- Desensitization
- Cross- and Intermodulation
- Harmonic and other spurious.

[1] The terms "primary" and "secondary" used in distinguishing factors should not be confused with similar terms used in specifying service areas of broadcast transmitters.

[2] O. D. Perkins, "Design objectives for spectrum economy," *Proc. Symp. on Electromagnetic Interference*, Asbury Park, N. J., November 19–21, 1957. U. S. Army Signal Res. and Dev. Lab., Fort Monmouth, N. J., June 15, 1958. Available from Armed Services Tech. Information Agency.

[3] "Military Collection Plan for Equipment Spectrum Signatures," revised September 1, 1961. Available from the Dept. of Defense, Washington, D. C.

Analytical Techniques

Techniques used in analysis of interfering situations have varied from relatively simple calculations of received power level based upon free space propagation formulas, to extremely complex programs on electronic digital computers. In any case, the objectives may have simply been to determine the extent of interference to be experienced, or to specify an appropriate frequency allocation.

At the more sophisticated end of the techniques spectrum is the computer program. A number of these have been described.[4,5] In general, the usual approach is to feed into the computer data on each element of the system and its electronic environment listed in Table I. Then the computer considers each source and receiver combination in sequence and compiles results. Even with the digital computer, however, it is almost always necessary to resort to approximations in storing the data and in carrying out calculations in order to keep the data storage and computer calculation time within practical bounds. Spectrum signatures and antenna patterns must be "quantized," and parameters subject to statistical variation must be handled so as to produce results that can be correspondingly interpreted. One of the most involved factors in this respect is propagation. Related to this is the question of what value of signal to "noise" (or "interference") ratio should be considered acceptable.[6] Clearly this is a function of various transmitter and receiver characteristics, the frequencies of signal and of interference, their "modulation" characteristics, propagation characteristics of the channel, etc.

Many other methods of interference prediction on smaller scales have been used. Most of these have not been reported in the literature.

In the frequency assignment area a number of specific problems have been solved. These may generally be described as determining the minimum number of frequencies required in a given communications complex under certain stated or assumed restrictions. A formal approach to this type of problem has been described.[7] In one form the restriction is geographical with transmitters and receivers located in a specified way.[8] In another the restriction takes the form of an equation giving a relationship between two frequencies which will produce interference with a third (intermodulation problem).[9]

Another instance of the application of computers is in the case of radar mutual interference. By inserting data on pertinent factors including antenna patterns and geography the average rate at which pulses would be received at various locations can be computed as a function of receiver sensitivity.[7]

Because of the limitations in "brute force" application of computers, consideration must be given to the possibility of "simplifying" the program by various techniques. One should be careful to set up the program in such a manner that unnecessary calculations are avoided. Usually only the strongest received signals will be of interest. To avoid tedious searching of computer files, the data should be stored in some selective way. It is known that such techniques can improve computer efficiency substantially in many types of problems requiring withdrawal of selected data from a large total storage.[10]

Another technique is to adopt the "integrated" approach. In one example of this the point of view was adopted that the only significant measure would be the "probability that interference would occur," it having been granted that it would be difficult, if not impossible, to guarantee that interference would not occur. Thus it was possible to locate transmitters and receivers "statistically" initially, to describe the propagation statistically as well as transmitter and receiver characteristics, and by manual computations to obtain desired results relatively quickly and inexpensively.[11]

Measures of Effectiveness of Spectrum Utilization

When one has calculated the data and made a prediction, how does one evaluate whether the situation is good or bad? A little reflection will reveal that very little is really known about how crowded the spectrum is. Little is available in the form of technical reports which are available generally to those in the technical communications area giving data on such results. Undoubtedly the techniques referred to above have been quite useful in specific areas of interest, and those applying them have definite opinions on the extent of interference in their own provincial environments.

[4] A. L. Fullerton, Jr., "Application of digital computer model to the prediction of radio interference in a large signal complex," *Proc. 4th Conf. on Radio Interference and Electronic Compatibility*, October, 1958, Armour Res. Foundation, Chicago, Ill., p. 579.

[5] D. R. J. White and W. G. James, "Digital computer simulation for prediction and analysis of electromagnetic interference," IRE Trans. on Communications Systems, vol. CS-9, pp. 148–159; June, 1961.

[6] R. F. Schwartz, "Signal acceptability criteria," IRE Trans. on Radio Frequency Interference, vol. RFI-3, pp. 11–18; May, 1961.

[7] W. H. Tetley, "Analytical prediction of electromagnetic environments," IRE Trans. on Communications Systems, vol. CS-9, pp. 175–185; June, 1961.

[8] R. L. Dujack and L. Stambler, "A digital computer program for assigning a set of interference-free frequencies to a communications network," *Proc. 4th Conf. on Radio Interference and Electronic Compatibility*, October, 1958, Armour Res. Foundation, Chicago, Ill., p. 539.

C. L. Becker, "A digital computer technique for determination of permissible co-channel frequency assignment," *Proc. 4th Conf. on Radio Interference and Electronic Compatibility*, October, 1958, Armour Res. Foundation, Chicago, Ill., p. 550.

[9] I. E. Perlin, "Processing of mutual interference charts for maximal lists of noninterfering frequencies," *Proc. 4th Conf. on Radio Interference and Electronic Compatibility*, October, 1958, Armour Res. Foundation, Chicago, Ill., p. 614.

[10] H. J. Gray, *et al.*, "Information Retrieval and the Design of More Intelligent Machines," Moore School of Elec. Engrg., Univ. of Pennsylvania, Philadelphia, ICR Rept. No. AD 59 URL, Contract No. DA 36-039-SC-75047; July 31, 1959.

[11] R. S. Berkowitz, "Computation of probability of interference due to adjacent channel cross-modulation, and intermodulation effects for randomly spaced equipments operating in the VHF and UHF bands," *Proc. 4th Conf. on Radio Interference Reduction and Electronic Compatibility*, October, 1958, Armour Res. Foundation, Chicago, Ill., p. 111.

Most of us use commercial broadcasting as a point of reference. Here very high standards have been set. Relatively little interference is experienced by the majority of listeners, so that excellent reception quality is obtained. Historically, this desirable result has been obtained because of technical and economic feasibility.

The same standards should not be applied to other radio spectrum uses arbitrarily. It seems clear that in many instances the channels are not used to the maximum extent possible, such as where the service involved is operated for only a small percentage of the time, or where the signals transmitted cover only a fraction of the surface of the globe. Evaluating the efficiency of use is, of course, a very complicated problem which should take into account many aspects of the service involved.

The value of the radio spectrum is to transfer information (information used in its broadest sense). One should, therefore, develop a suitable measure for the amount of information transferred. Factors that must be considered clearly include the information rate, number of receivers of the information, and the "value" of the data or information transmitted or received. Costas, in a stimulating article[12] has described a concept of spectrum use which tolerates a substantial amount of "interference" as in radio amateur practice. He also discusses the implication of new methods of modulation in this concept.

Other factors which might be considered are cost (in dollars and men), time delays, and areas served by the transmission system. These we discard, at least initially, as being either irrelevant to the question of spectrum utilization, or incorporated in the other measures.

To apply the measures previously listed, one must attempt to find some means for attaching importance to information. Those familiar with operations research techniques have recognized for a long time that generally dollar value is one of the most useful measures for evaluating different situations. In the case of distress frequencies, it would be necessary to place a value upon human life. This has been done in a number of instances already in other situations. The value of information is frequently a critical function of timeliness (especially in military situations) and consequently there is a basis for comparing clear channel and shared channel operations. The developments of information theory in recent years have provided us with an excellent basis for comparing various types of modulation for efficiency of data transmission.

It is relevant to mention specific examples to which these concepts might be applied.

1) *The Broadcast Band:* The amount of information transmitted is extremely high, although at a relatively low rate, and of relatively small importance when received. However, a relatively high commercial value is placed upon such transmission. The number of receivers is extremely large and the geographical coverage almost complete.

2) *Mobile Services* (for example, taxicabs and police vehicles): Information rate—high, geographical coverage—fair, timeliness—adjustable in accordance with requirements, number of receivers—moderate, over-all efficiency of spectrum use—good.

3) *Radar:* Number of transmitters and receivers variable, depending on the application. For certain functions, such as navigation, geographical coverage is fairly complete; number of transmitters and receivers—large; information rate—relatively low; and commercial value—substantial. For military applications, such as air defense, the geographical coverages are fair, average information rate is extremely low (useful information obtained only in times of emergency), timeliness is extremely important, the value of the information is extremely high.

Other examples could be given. The above would seem to indicate that on the basis of questions that are at the moment only qualitative, the assignment of spectrum space is logical and one begins to get suspicious that when more quantitative means are used we may only justify the arrangements which now exist. This is not surprising. What then is the value of these concepts? Simply, that by attempting more qualitative evaluation of spectrum utilization we shall be able to better determine the relative importance of various possible actions with regard to the secondary factors. For example, consider the seriousness of harmonic radiation which could cause losses of time delays in information transfer. How much harmonic radiation should be tolerated in a given situation? To what level is it economical to use suppression techniques? What will be the economic gain in the application of any given technique for control of a secondary factor? Applications of information theory would give a quantitative basis for changing coding and modulation techniques.

Future Developments

In summary, areas in which one can expect significant developments in the future are as follows:

1) Techniques of prediction.
2) Development of suppression techniques.
3) Frequency allocation.
4) Application of efficient methods of modulation.

A. Techniques of Prediction

The current vogue for using computers to make interference predictions will undoubtedly be extended in the next decade. However, it seems reasonable to expect that not all users of communications facilities will be able to afford such an extensive approach. Furthermore, it stands to reason that requirements for prediction will vary substantially from one user to another so that the

[12] J. P. Costas, "Poisson, Shannon, and the radio amateur," Proc. IRE, vol. 47, pp. 2058–2068; December, 1959.

nature of the problem involved in given cases may be very specialized. Indeed, as the nature of interference becomes better understood, and as the factors contributing to it come under better control, we may expect predictions to become simplified. Until more experience is gained with their use, it may be possible to write computer programs in such a way that specialized problems can be quickly performed. However, the more simple problems will likely be handled by graphical or slide-rule techniques.

B. Development of Techniques for Suppression

The importance of suppressing secondary factors (harmonic and spurious radiation in transmitters and spurious response in receivers) has always been recognized, but relatively little has been done in some cases because of apparent economic factors. Present frequency allocations are based upon avoiding the "worst case" situation. Clearly, if the standard deviation can be reduced through improved methods of manufacturing control, the worst case condition will be alleviated and more use can be made of the frequency spectrum. Reduction of the mean value of secondary factors will probably take more time, require technical advances in the state of the art (such as the development of harmonic-free tubes), and involve substantial expense. Secondary factors can be improved by filtering or by increasing directivity of antennas, as examples, and in many cases the methods used will depend largely upon economic considerations.

Spurious and harmonic suppression is extremely important in some of the modern high-power transmitters, and, indeed, the design of output stages in some of these has been influenced in a major way by the choice of harmonic-and-spurious-free designs. In the case of receivers, the feasibility of new techniques of circuit design in the input stages has been demonstrated,[13] and very likely the additional cost can be justified in many instances. Of course, the importance of suppression techniques, such as the insertion of output filters, has stimulated a great deal of work in the development of new components and it is quite likely that independent of the economic question we can look forward to the development of many new components and techniques in the future.

C. Frequency Allocation

Eventually, a more explicit concept of efficient spectrum utilization with regard to primary factors will evolve. This will include consideration of frequency assignment so that more optimum use can be made of inherent directivity of antenna structures, special modulation techniques, natural propagation conditions and the service requirements. Undoubtedly, present frequency allocations are in many cases sound, and will remain with us for this reason, in addition to the fact that there are large equipment investments involved. On the other hand, as techniques for frequency stabilization improve and new methods of multiplexing information in specific frequency bands are developed, it will become apparent that changes will have to be made in the frequency spectrum. The shortage of spectrum space may force spectrum allocation changes, in spite of the equipment change-over costs involved. This becomes a political problem rather than a technical problem and undoubtedly will be implemented one way or another.

D. Modulation Techniques

In recent years considerable effort has been directed towards examining the relative efficiencies of various methods of putting information on a carrier, from the point of view of bandwidth consumption, signal-to-noise ratio, and reliability (for both data and voice). Implications with regard to spectrum usage have also been considered, but, as previously mentioned, the criteria for efficient spectrum use have not been established. Thus, we may expect to see the development of such criteria take place as this problem is brought into sharper focus.

Conclusions

It is seen from the above discussion that much pioneering work is necessary in the general area of radio interference. What is needed is a truly operational approach to the problem of spectrum utilization. At the present time there is much complaint about shortage of space in the radio spectrum, but as yet no one has come up with a quantitative measure of this shortage. In fact, it is not absolutely certain that a shortage exists. Admittedly, there are numerous instances of interference, many of which could be classified as critical or as of serious consequence.

Undoubtedly, it is essential that the users of the radio spectrum become accustomed to experiencing a certain amount of interference in order that maximum use be made of available channel facilities. We are entering an era in which a better understanding of the relative efficiency of various techniques, including the proper use of such things as the various modulation schemes, directive antennas, the incorporation of coding techniques and assignment of frequencies will take on a new meaning in terms of the required qualities of service.

[13] D. McClenon, "Measurement of communication receiver vulnerability," *Proc. Symp. on Electromagnetic Interference*, Asbury Park, N. J., November 19–21, 1957, U. S. Army Signal Res. and Dev. Lab., Fort Monmouth, N. J., p. 393; June 15, 1958. Available from Armed Services Tech. Information Agency.

Section 25

RELIABILITY AND QUALITY CONTROL

Organized with the assistance of the IRE Professional Group on Reliability and Quality Control

The Reliability and Quality Control Field from Inception to the Present *by Clifford M. Ryerson*

The Reliability and Quality Control Field from Its Inception to the Present*

CLIFFORD M. RYERSON†, SENIOR MEMBER, IRE

Summary—**Starting with a brief discussion of the implications to Reliability of the parallel accelerated growth of modern technology the history of Reliability is traced through four decades. Following the early groundwork of the years before 1940 the decade of the 30's has been described as the Standardization Decade. Here the new emphasis was on specific standards of all kinds. The following decade of the 1940's has been described as the Quality Control Decade. This new field was developed with emphasis on process control and the uniformity of product. This was followed by the Reliability Decade during which the new emphasis was on the inherent reliability of design and the time degradation of performance. The 1960's have ushered in a new decade and new emphasis on coordinated controls in all areas and phases of design, development and production. The new decade has been described as the Product Assurance Decade, wherein the emphasis is being placed on assurance to management and to the customer that all the important product characteristics are being optimized and that all the technical specialties involved in providing the highest value product on schedule are being integrated effectively. This new emphasis is on coordinated programs embracing all the related control specialties sometimes referred to as the Big R approach. These specialties include Reliability Engineering, Maintainability Engineering, Total Quality Assurance, Value Engineering and System Integration among others.**

I. Introduction

A PROMINENT characteristic of this new field of reliability and quality control is the rapid change and growth which has occurred in recent times. In order to clarify the causes and emphasize the pace at which this growth is accelerating, this introduction reviews briefly the parallel growth of modern technology. A brief review of some of the milestones of general progress is presented in a way to emphasize the implications for the reliability field.

* Received by the IRE, July 31, 1961.
† Vice President, EL-TEK Corp., Hawthorne, Calif.

The rate at which technological progress has been made can be illustrated by a brief look at military equipment progress. For example, it was only a little more than 100 years ago that military cannons were cast of iron or bronze and bolted solidly to wooden carriages. A foundryman and a carpenter were the only two specialists needed to design and build a first-quality large-calibre weapon. Since few, if any, drawings or specifications were involved, every model was hand-made and distinctively different. The only design criteria were that the bore be reasonably straight, smooth to the touch, and slightly larger than the balls to be fired. If the artisans decided to make an extra fine firing piece, they might engrave the outside of the barrel or add fancy iron wheels.

Some fifty years ago, the best cannon had become steel rifles with a complicated array of gears and cams for manually cranking the barrel into position. At this stage of development, ballistics became a science, and the mad rush for perfection of power drive and automatic control began. Weapons of this era presented a blossoming array of design criteria ranging from steel specifications to construction and assembly techniques. Not two craftsmen, but hundreds, were required to design, construct and install these advanced weapons. Excellence was no longer a matter of external appearance, but of utility and adherence to strict specifications which had been worked out in accordance with long shop experience and engineering design.

Finally, just a few years ago, the large-calibre automatic weapons used at the end of this last war dwarfed to insignificance those first rifles, by virtue of automatic radar control, long range, high rate of fire, maneuverability, accuracy, and precision. In addition, many uncommon materials were beginning to appear. New processes of construction, radar direction, hydraulic electric drive, automatic interlocking, remote control, and many other factors made military armament into a big business. It became so big that the identity of any one intellect, type of trade or profession, specialist, or personality was lost in hundreds of pages of specifications, thousands of drawings, and hundreds of man-months of all types of technical effort. Now modern missiles with their fire control and self-navigation features are in a different order of magnitude in both functional performance and complexity.

Although this example illustrates the advances made in recent years in the field of military ordnance, similar illustrations can be drawn in other fields. It is so easy to take for granted many of the near-miracle devices of our time that the rate of technological progress they represent is easily overlooked. Consider the airplane, for example. It is hard to realize that man could not fly only some 60 years ago. The first flights of seconds duration were made by the Wright brothers in 1903. Some ten years later, the airplane was still a kite-like affair used for experimental observation and contained no instruments. Several instances are on record during World War I where aerial dogfights occurred using pistols and rifles as armament. Some small explosives were dropped by hand near the end of the war in 1918.

Less than 20 years later, and by the time World War II began, the airplane had become an effective machine with powerful engines and many meters and power controls. In the following four or five years the airplane became a maze of electronic instruments, and today, only a decade or so later, completely automatic, self-guided, self-navigated and computer-controlled world-girdling missiles are commonplace.

Today, we are talking seriously about manned space craft capable of sustained flight for long periods in the vacuum of space. The complexity of functions necessary to support life completely independent of earth environment is a new level of advance in the world of flight.

To emphasize the rate of this advance, let us consider the number of electronic parts employed in a typical equipment at various stages. In 1921, just 41 years ago, no electronic equipment was generally available for use in airplanes. By 1936, simple communication sets containing perhaps 25 to 30 tubes, and from 200 to 300 other parts had become commonplace. By 1956, the all-weather interceptor carried over 600 tubes and the heavy bomber over 2000 tubes. The electronic parts per plane totaled many thousands. Today, although the transistor and other solid-state devices are replacing tubes, the number of electronic parts needed for automatic navigation, communication, and control, is running into the hundreds of thousands for large military aircraft.

We are now in a period in history when a complete revolution from manual to automatic is "booming" in every field. The electronic industry and all its related engineering aspects is perhaps the key element in every phase of this revolution. Born from the merger, not long ago, of the mechanical age to the electrical age, this new electronic offspring has developed to bewildering magnitude in comparison with its still young parents. Friends of the parent fields are frequently ill at ease in the company of this gigantic young progeny. Their attempts to work with it can perhaps be likened to the absent-minded professor trying to put rubber baby pants on an elephant. The trouble is not only with the size of the pants but in the complete inadequacy of the normal techniques.

Coincident with the unprecedented growth of the progeny, new techniques and procedures must be and are being developed to keep pace. The magnitude of this growth can be better understood by comparison of a few interesting statistics. Recent figures indicate that, whereas the population of this country has increased on the average less than 2 per cent each year during the past 15 years, the electronics industry over this same period has averaged better than 15 per cent per year increase. Further, the curve of electronic industry growth has an upward trend which has every indication of continuing for some years. (The 0.4 billion industry in 1940 grew to 2.6 billion in 1950, then to 10 billion by 1960, and is projected to go over 20 billion by 1970.[1] Beyond this, there is still tremendous growth potential.)

The reasons for this enormous growth are many. Furthermore, most of the reasons are of a semipermanent nature which are not apt to reduce in urgency for many years. The progress of aviation, atomic energy, automatic business machines, automatic factories, improved communication, increased standard of living, industrial use of television, complete mechanization of farms, military advances, newly developed materials, and the shortage of technical manpower itself are all factors. Further, these are factors in the true algebraic sense in that advances in one field multiply the problems for electronics in others.

As the trend from "manual" to "automatic" mushrooms in size and the industry growth outstrips the population growth, and as the machine performance requirements outstrip the performance capabilities of the human, it becomes more and more important to replace humans with electronic devices. This occurs in engineering organizations where complex automatic computers do the work of many; it is occurring in the military through the use of automatic navigation, gun-laying and logistic control equipment; it is common in aviation where radio operators are being replaced by compact automatic communication equipment and pilots are being replaced by controls which have a faster reaction time plus automatic computing provisions. There is thus a vicious circle which adds up to a continually increasing requirement for more electronic

[1] EIA figures.

developments which, in turn, create problems which can only be solved by more electronic developments and increased complexity.

An even more powerful factor than the preceding looms on the horizon as a reason for future electronic-industry growth. This involves the process and by-product control of the atomic industry. To date, the atomic industry has not appreciably affected the electronic field. As the atomic industry grows, however, new areas requiring automatic and complex electronic control will become common. In addition to the approximately 500-million-dollar expenditure predicted by 1963 for reactor controls and instrumentation, there is the large field of fuel-fabrication-process control, and the equally large field of waste disposal.

The "ashes" from atomic reactors present a future contamination hazard of staggering magnitude. This problem can only be solved by a new chemical process disposal and by-product manufacturing methods, each with their own special electronic controls. Further, as new liquid-fueled homogeneous reactors replace the present heterogeneous piles, the age of atomic-powered space craft will be ushered in. In all this, the continually increasing need for electronics is readily apparent. It seems reasonable, then, that the present upward trend in the magnitude of the electronic industry will continue unabated for some time.

It seems equally apparent that, as civilization comes to depend more completely on complex automatic equipment, the reliability of this equipment becomes of paramount importance. The higher the dependence on automaticity, the higher becomes the penalty of equipment failure. Indeed, the limiting factor to future advance in nearly all the modern technological fields will be the complexity which is permissible as set by the state-of-the-art in reliability.

The modern quest for higher reliability had its major beginning with the onset of World War II, but the rudiments were established at a much earlier time. The following sections present a chronology of specific events important to this new field.

II. Early Groundwork for Reliability (The Years Before 1940)

At the risk of some controversy, it can be stated that the first step in the new field of reliability can be traced to discussion generated by the AIEE Standard No. 1, published in 1913. The title of this document was "General Principles upon which Temperature Limits are Based in the Rating of Electric Machines and Apparatus." This Standard started the first cycle of appreciation for the chemical composition of electrical materials. During the subsequent development of the electrical and electronic fields, there has been a cyclic recurrence of this emphasis which can be recognized today as "analysis for the basic mechanism of failure in materials used in highly reliable electronic parts." However, in all fairness, it should be stated that the element of time-dependent degradation was not a part of the referenced first document. This element did not become a formally recognized factor until the great upsurge period for reliability in the 1950's.

Perhaps one of the first clear-cut examples of what would now be called a reliability-control program occurred in 1916. The Western Electric Company and the Bell Telephone Laboratories cooperated in a planned program to produce good performing and trouble-free telephone equipment for public use. The following elements of a good reliability program were involved:

1) An ably planned, forward looking R&D program considering the system needs.
2) Development programs leading to mature designs verified by design reviews and engineering model tests.
3) Part improvement projects using test-to-failure methods, consideration of tolerances and performance under stress loading.
4) Standardization and simplification in design.
5) Product evaluation under field application conditions for prototypes and pilot production models.
6) Quality control during manufacture to assure achieving the reliability inherent in the design.
7) Feedback from the field to provide information for the designers.

This same proven approach was later used for several other successful developments in the telephone and communication field. One early application in the military equipment field occurred in 1917 and 1918, when Bell Laboratories and Western Electric cooperated to develop and produce the CW 924 and CW 936 radio equipment. These were installed on submarine chasers at New London, Conn., and were used up and down the East Coast from New London to Galveston, Tex. Daily feedback of failure information was made a part of the program schedule.

Early work for the Navy used similar reliability controls to develop the first vacuum-tube transmitter (1918–1919), the first wide-band superheterodyne receiver (1923), and the special 6-kw tube transmitter installed on the S.S. *Leviathan* (1924).

In 1918 the American Standards Association was founded by five engineering societies to coordinate the development of national standards. It has since served as a national clearing house and coordinating agency for voluntary standards in the United States. In 1948 the ASA was incorporated under the laws of the State of New York as a nonprofit organization. It now is a federation of 122 trade associations and professional societies. In addition, it also has more than 2000 company members. From its earliest beginning, the ASA has had an important influence on the factors contributing to reliability. More specific contributions will be described later.

The 1920's witnessed the birth of our mass-production component-parts industry. In the very early days the processes were crude. For example, grid-leak resistors were even made from lead pencils. The process consisted essentially of screening a supply of "lead" to find those which had the approximate resistance required. Today, the standard molded composition resistors are made automatically on complex and costly machinery. In a similar fashion, many other major lines of parts had come from crude beginnings to a comparatively advanced state of technical development by the end of the 1930's.

The vacuum-tube industry became well established during this growth period, and in the late 1930's came the initial attempts at commercial production of miniature tubes. This development was sparked by the need for small high-gain amplifiers for hearing aids. It was in this service that the first solder-in tubes were used in quantity. Reliability was beginning to be a problem.

The greatest impetus to further miniaturization of electron tubes was felt early in 1941, when rugged tubes were urgently needed for use in the proximity fuse. Early attempts to use hearing aid types of tubes for this purpose were quickly set aside on the basis of unreliability. Not only were the hearing-aid types not reliable enough in ordinary service, but they now had to be rugged enough to withstand the impact of the propellant explosion in the gun. Since that time, the development of reliable subminiature tubes has been pushed forward relentlessly, largely under the sponsorship of military agencies such as the Bureau of Ordnance, the Bureau of Ships, the Air Materiel Command and the Signal Corps.

During the 1930's, a new profession called standards engineering was born. With the quantity production of many types of component parts came the need for standards of all sorts. Standard types came first, then standard methods of specification, and finally, standard methods of measurement and testing. The National Bureau of Standards and the American Standards Association both owe their present strength to the developments of this period.

Test methods have usually been divided into two categories: simple, nondestructive tests for rapid inspection, and the sometimes-simple but sometimes-complicated destructive tests. Throughout the three preceding decades, to date, the same general divisions of testing have been recognized. However, the emphasis on standards during the decade of the 1930's was to define arbitrary limits on the parameters being measured and to specify tests based on a go—no-go basis. This qualitative approach was a real step forward at the time and helped to achieve acceptable reliability for the state of simplicity of the then-current equipment. The emphasis on standards shifted during the 1940's from the go—no-go limits to close regard for the actual variance or "scatter" of the data obtained in performing tests. This change ushered in what might be called the decade of quality control, which is discussed in the next section. This change was brought about by forces resulting from the rapidly increasing technology.

After the 1930's, the end products of mass production began to change complexion. While the considerable volume of home-entertainment receiver production continued to increase (except for the war years), and was finally given tremendous new impetus by the commercialization of black-and-white television, other applications for electronic components began to develop at an even faster rate. These include: radar, sonar, seismic surveys, mobile communication, electronic computers, and various industrial applications of electronics.

Between World War I and World War II, aviation experienced a rapid expansion period, and electronic equipment in military and civil aviation had opportunity to gain service experience. The equipment in use by 1938 had reached a fairly high degree of reliability, and the communication needs of the aviation industry could be satisfied. Suddenly, however, radar appeared and, almost simultaneously, the World War II adventure began. The service that radar could deliver in military applications was so valuable that it was permissible, even mandatory, to produce it without going through the ordinary stages of development and design which are essential to reliability with an entirely new service. Radar and its related equipment was rushed into production and service use without opportunity for adequate attention to important details.

Radar quickly proved to be so effective that, under the urgencies of war, together with all related electronic devices, it came to be handled with short-cut procedures, and with the same minimum attention to reliability and practical details in general. This became regular procedure, and apparatus finally was being chosen almost entirely on scientific merit, with minor attention to operational evaluation of such factors as reliability, complexity, suitability to personnel and training, etc.

Of course, routine attention was given to various standard specifications aimed at reliability, and many were the environmental and functional tests performed. But, frequently, the only results of the tests either damaged the equipment or introduced changes which made the devices even more complex. Often tests were disregarded completely in order to obtain equipment quickly. The atmosphere was such that the ideals of technical performance—more sensitivity, selectivity, range, or intriguing new functions—always won out over the prosaic demands of simplicity, ruggedness, and reliability. There were cases where electronic equipment was installed in planes and carried around unusued, because time was not available to train the crews in its operation and maintenance.

The same frenzy of design—build-and-ship regardless of reliability—appeared in all aspects of industry. For example, there was the case of the magnetic influence torpedo which was manufactured for months and put through a complete series of formal tests long after field evidence of unreliability had caused an order to be issued to use them only for contact firing with the influence feature turned off.

Before the days of radar, common sense and thorough evaluation were familiar ingredients in electronic design, but the recklessness of the times caused a new way of life in industry. Unfortunately, this way of life has left its stamp on many of the current generation of managers engineers, and companies.

What actually happened was that the new creation, "radar," appeared in the category of radio equipment, and, therefore, there was immediate pressure to make it similar and by using the same parts and methods. It could be made small, light, and quickly—as can all electronic gadgets—and the industry responded to these pressures. But one of the costs of this process has been in reliability. The pressures for returning to the former analytical "system-need" approach did not develop until the 1950's, and these pressures have still not achieved their full objectives.

However, some progress was made in the reliability field even as early as 1939. The Ordnance Corps worked out and established statistical methods of surveillance of ammunition in storage. These methods have been standard ever since for controlling the "liveness" of ammunition. In effect, the time deterioration (reliability of ammunition is controlled by a repetitive sampling of typical stock. A quality survey performed on a repetitive schedule was thus demonstrated as a valuable tool to maintain reliability of one-shot devices during storage. This storage phase was recognized as only one area where statistical methods might be helpful, and whetted by this success, leaders in the field worked diligently to usher in the quality-control decade.

III. The Quality-Control Decade (1940–1950)

The emphasis of the standards decade prior to 1940 left many weaknesses in approaches to the manufacture of uniform high-quality products. It was becoming a national scandal that so few products manufactured to the same specifications were similar. Factory inspection had become common, but the basis for acceptance or rejection was frequently a matter of individual choice. Military effectiveness was being severely hampered by poor quality of equipment.

Upon request by the War Department, in December, 1940, the ASA initiated a project on the Application of Statistical Methods to the Quality Control of Materials and Manufactured Products. In the national emergency there was urgent need for the prompt development of standards relating to the control of variability in products and processes. The ASA Emergency Procedure was, therefore, applied to this project and an Emergency Technical Committee was appointed to develop the required standards.

An active committee of six men was chosen which included the following: H. F. Dodge, Bell Telephone Laboratories, *Chairman*, A. G. Ashcroft, Alexander Smith & Sons Carpet Company, W. Edwards Deming, Bureau of the Census, Leslie E. Simon, Ordnance Department, U. S. Army, R. E. Wareham, General Electric Company, John Gaillard, ASA, *Secretary*.

Two American Emergency Standards (since called American War Standards) were published by the ASA through the efforts of this group during 1941, and a third during 1942. The exact titles of these documents follow:

ASA-Z1.1—1941: "Guide for Quality Control"

ASA-Z1.2—1941: "Control Chart Method of Analyzing Data"

ASA-Z1.3—1942: "Control Chart Method of Controlling Quality During Production."

The basic policies and procedures for statistical quality control were established in these three documents. Much of the 1940 decade was expended in an uphill struggle to upgrade inspection procedures and to obtain recognition for the validity of statistical process control as described in these standards.

The Machine Tool Builders Association was one of the first industry groups to accept the challenge of quality control. The spearhead of the supplier group was the Electrical Committee of the National Machine Tool Builders Association. These men worked for "something they could live with." The discussions at their meetings were lively and generally ended in clarification of what the users wanted in quality and what they could reasonably hope to obtain. The first Machine Tool Electrical Standards were introduced by the NMTBA late in 1941, and became the basis of the American War Standard ASA-C74-1942. These dealt with specific quality standards on motors, conduits, cables, etc.

Prior to World War II, military equipment was relatively simple and simple good practices were adequate to produce reliable equipment. With the onset of war, however, the frenzy of produce-and-ship, previously described, set in, and, in addition, came military demands for increasingly complex equipments. The new military demands required developmental efforts for new tubes, new circuitry, new applications of components, new functional components, and an increased ability for equipments to withstand higher levels of environmental stress. One of the major problems in these jumbled times was concerned with tube reliability.

Several organizations, including the M.I.T. Radiation Laboratory, recommended in 1942 that a national coordinating body be established to encourage and direct tube development work as a first step to reliability improvement. This recommendation was approved by the Navy in 1943, and an initial joint Army-Navy Vacuum Tube Development Committee (VTDC) proposal was transferred for implementation to the newly formed National Defense Research Committee (NDRC).

On June 7, 1943, the first meeting of the VTDC was held with Dr. I. I. Rabi as Chairman. This group, under various sponsorships, performed a valuable service throughout the war.

On June 6, 1946, the Joint Research and Development

Board (JRDB) was established. On October 24, 1946, the VTDC was transferred as a specialty panel to the JRDB, and the Panel on Electron Tubes (PET) was formed. Later, a support contract for this panel was established at New York University.

A major technical breakthrough occurred in the spring of 1943, when the noted mathematical statistician, Abraham Wald, devised his now celebrated basic theory of sequential analysis for analyzing U. S. war problems. Initial application of the theory in analyzing combat experience demonstrated its value in obtaining reliable conclusions from a minimum of information, swiftly and economically. This seemed like an ideal tool for use in quality control at this crucial period in our history when speed, precision, reliability, and economy in production were of the essence. To give the new tool a thorough trial, the military released a multiple sampling plan based on Wald's sequential theory. This plan was first released to a limited number of strategic manufacturing firms for use in acceptance sampling. The early success of the plan and the subsequent widespread demands for its use resulted in removal of its "restricted" classification in 1945.

The advantages claimed for the sequential-analysis technique when used for multiple sampling by skilled inspectors included the following:

1) It required a far lower average amount of inspection than did previous methods.
2) It did not require a specific sample size for a given lot.
3) It decided to accept or reject for itself, revealing its decision to the inspector when his last entry of information crossed a line on a graph.
4) It made decisions very quickly and accurately for extremely "good" or "bad" lots.
5) It made equally accurate decisions on "dubious" quality lots when additional information was provided.
6) It greatly reduced the chance of a wrong decision when compared with any other plan, thereby insuring a more consistent product with substantial reductions in costs.

Although the new plan was altogether a remarkable advance, it had weaknesses. Its principle disadvantages were the time-consuming necessity for skilled inspectors to either plot and record after each sample, or to record results of a set of samples, and then compare results to a set of Military Standard Tables. Moreover, if the first sample group yielded no decision, another group of equal size had to be selected and the entire process repeated until a decision was made. This complexity in use and the requirement for skilled inspectors were major factors in causing the plan to be much less of a boon to quality control than the early enthusiastic reception had led many to believe.

In 1944 the Navy began a program to obtain more rugged tubes for ship-board use. The severe shock and vibration accompanying gunfire caused many tube failures in fire-control radars and related equipment. Shock and vibration equipment was developed to aid in the testing, and structural modifications were made in standard tube types to increase their resistance to shock. As a result of this program, the Navy "ruggedized" W-series tubes were evolved, covering some 40 tube types.

In 1944, also, the first plans were established by civilian employees of the Naval Gun Factory in Washington, D.C., and the Naval Research Laboratory in Anacostia, D.C., to establish a National Symposia on Shock and Vibration. Under the able chairmanship of Dr. Elias Klein, a program and format were established which has done much to solidify the thinking on Shock and Vibration in this country. Yearly, symposia are still being held in this series, and this program has proven to be of major national value to reliability.

In 1945 the automotive industry joined the enthusiastic supporters of statistical quality control. Automotive standards and recommended practices were established, following the formation of the forerunner to the Joint Industry Conference (JIC). A basic paragraph numbering system was adopted, which has been retained by many groups and used in JIC and NMTBA standards since 1948.

In 1946 the American Society for Quality Control was formed. Key individuals in this effort to undergird the new field included: W. A. Shewhart, H. F. Dodge, E. G. Olds, H. G. Romig, Leon Bass, and others. During the years in which this society has been active, it has done much to upgrade the professional stature of the practice of quality control. Today statistical quality control of factory operations is an important factor in preventing the reliability inherent in designs from becoming degraded by poor workmanship and uncontrolled processes.

In 1946 Aeronautical Radio, Inc. (Airinc) was established by the Airlines to collect and analyze defective tubes in commercial airline service and return them to the tube manufacturers. As a result of these analyses, certain tube types were improved by exercising extreme care in assembly, giving maximum attention to production details, and checking each part as well as each complete assembly. After manufacture, all improved tubes were tested for shorts, continuity, shock, and transconductance. Samples were subjected to vibration, filament cycling, and life tests.

Special tube specifications were written, and about 10 tube types were improved before the work was discontinued at the start of the Korean hostilities. The Airinc program after 1950 was focused on the military tube problems under a joint services contract. Failed tubes were collected at various military installations, together with reliability data on the failure or cause of removal, hours of use, and use conditions. One of the key results revealed by the Airinc study in 1950 was that only about 33 per cent of the Navy Electronic Equipment on board ship was operating satisfactorily at any one time. This information set the stage for the big reliability push, described in the next section.

Before leaving the chronology of the quality-control decade, some mention should be made of the efforts of the International Electrotechnical Commission during this period. When Technical Committee No. 12 "Radio Communication" resumed its work after World War II, it decided (Stockholm, Sweden, 1948) to establish a special subcommittee to deal with components. After some deliberation, this subcommittee arrived at the opinion that the preparation of a general document for climatic and mechanical robustness tests should be the first consideration. The international efforts on this project resulted in a first draft for discussion at the Paris meeting in 1950, and a second draft for discussion at Estoril and Montreux in 1951. Partial approval was obtained, but the difficulties of obtaining agreement on the required tests and conditions of testing required several years more before final agreement could be reached and the document published in 1954.

Although the IEC Standard "Basic Climatic and Mechanical Robustness Testing Procedure for Components," begun in 1948 and published (1st edition) in 1954, was essentially a type-test environmental condition document, it did contribute to reliability testing in two ways: "Section 4.3 Test C Damp Heat (Long Term Exposure)" called for continued exposed operation for periods up to a maximum of 84 days; and "Section 4.8 Test H Storage" called for degradation in storage tests up to a maximum period of 12 months. These are among the first reliability life tests formally published.

IV. The Reliability Decade (1950–1960)

Early in 1950, the Defense Department, through the Research and Development Board, noted that difficulties with electronic equipments were not governed solely by a focus upon electron tubes. Many other parts were serious causes of unreliability. It became evident that many factors were involved. Consequently, on December 7, 1950, the Research and Development Board formed the Ad Hoc Group on Reliability of Electronic Equipment through its Committee on Electronics.

The Ad Hoc Group contained members from the three military departments, the Joint Chiefs of Staff, the Munitions Board, and various civilian organizations and professions. The Chairman of the group at the start was J. A. Chambers, who later was succeeded by A. F. Murray. The Panel on Electron Tubes was included as an advisor.

During the years in which the Ad Hoc Reliability Group was active (1950–1952), the three military departments undertook increasing numbers of studies in order to add to their knowledge of equipment failures. For example, the Navy contracted for the Vitro Corporation to investigate component failures, for Aeronautical Radio, Inc., to investigate electronic tube failures, and for the Bell Telephone Laboratories to study component-part failures. The Army, through the Signal Corps, entered into a long-term tube analysis program with Cornell University. The Air Force requested the Rand Corporation to investigate the general problem of electronic reliability.

Partially in response to the work of the Ad Hoc Group, and to meet other pressures, General George C. Marshall, then Secretary of Defense, issued on September 12, 1951, the Department of Defense Directive 150 21-1, "Reliability of Electronic Equipment." In this document he reviewed the important work of the Ad Hoc Group and stated that " . . . reliability must be a prime objective in *all* phases of the procurement and use of . . . equipment." He went on to direct that " . . . increased emphasis on reliability of military electronic equipment by all agencies of the Department of Defense is required."

It was in January, 1951, that the American Society for Testing Materials, through Committee E-11, issued its final "Manual on Quality Control of Material." The ASTM, organized with 70 members in 1898, had grown to more than 10,000 individuals and 6,000 company members by 1951. The development of standards had been carried out by more than 80 technical committees of which E-11 had been given the "Quality Control of Materials" assignment in 1933. The early work was done with the ready cooperation of the Joint Committee on the Development of Applications of Statistics in Engineering and Manufacturing, jointly sponsored by the American Society for Testing Materials and the American Society of Mechanical Engineers. W. A. Shewhart was Chairman of this Committee, which later (1946) formed the nucleus of the new American Society for Quality Control.

For a period of some 15 years, the material in the ASTM Manual had undergone a number of modifications and preliminary editions. Its nomenclature and symbolism were adopted in 1941 and 1942 in the American War Standards on Quality Control previously mentioned (ASA-Z1.1–Z1.2 & Z1.3). The manual has since been reprinted seven times, the last edition in July, 1960. The contents are considered by many to be the "bible" on quality control.

This term is also frequently applied to the book entitled, "Economic Control of Quality of Manufactured Product." written by W. A. Shewhart of Bell Telephone Laboratories in 1931. This has been reprinted eight times by D. Van Nostrand Co., Inc., and is still a good seller.

Another development in 1951 illustrates the point that neither the problems of reliability or solutions to the problems are restricted wholly to electronics-type equipment. During 1951, the "roll-pin" was introduced to replace the taper pin as a general-purpose mechanical fastener. Installation was easier, for the roll pin required only a drilled hole and not the additional taper-reaming operation. But a bonus for reliability was realized in high-speed equipment. At 12,000 rpm and above, taper pins, predominantly heavy at one end, tend to disengage and throw out. The roll pin eliminated this trouble.

In 1952 the Ad Hoc Group on Reliability of Electronic Equipment issued a final progress report,[2] making seventeen major recommendations for action by various agencies. The recommendations advised that:

1) Failure data reports be compiled on the basis of field use and be summarized, evaluated, and placed in the hands of designers.
2) Tube, component, and especially system-reliability, programs be continued by appropriate groups.
3) Reliability requirements be added to military characteristics prepared by the Joint Communication Electronics Committee.
4) A study be made of maintenance minimization.
5) A study be made of affects of unreliable equipment.
6) Reliability concepts be involved in procurement, production, and quality control of electronic equipments.
7) The RDB reliability activity be extended from the initiation of military characteristics through its operational use, and that a permanent RDB reliability group be established.
8) Educational activities be expanded and a reliability information center be established to provide reliability data.
9) Testing of equipment, simulating use in the field, be expanded.
10) Analyzing and approving new designs with regard to easy maintenance and reliable performance be established.
11) A reliability section be put into specifications.
12) The training of inspectors be improved.
13) Engineering supervision of installation of equipment by the material agencies and their contractors be improved.
14) Training of operators be improved, with reference to results of operational abuse of equipment.
15) Maintenance problems be investigated to secure better preventive maintenance, training, and simpler test equipment.
16) Reliability organizations be set up in the military department.
17) Requirements of classification as to degree of reliability necessary in specific items be adopted and integrated into military characteristics.

At the IRE National Convention, New York, N. Y., March, 1952, a very significant paper was presented which summarized many facets of the national reliability problem. Such aspects of reliability as the product rule, accelerated testing, designing for safety margins, dependency on the human, and the affects of environment were presented by Capt. A. C. Packard, U.S.N., from the Naval Air Missile Test Center. (The paper was jointly authored by Capt. Packard and Dr. Royal Weller.)

Another significant event for reliability occurred, during 1952, at the suggestion of Dr. W. R. G. Baker. Three planning groups in the Radio and Television Manufacturers Association (RTMA),[3] covering the broad fields of tubes, components, and systems, were merged into one committee entitled "Electronic Applications Committee" (Reliability). The title of this Committee has been variously labeled, but the most common abbreviations are EA(R), and CEAR. Lewis M.Clement was Chairman of EA(R) and M. A. Acheson, Dr.Preston Robinson, and J. A. Chambers were the initial Vice Chairmen.

A major contribution of EA(R) was the distribution of the *Electronic Applications Reliability Review Bulletins* to a list of 4,500, which did much to awaken many people in the military and industry to the problems in reliability.

During 1952, the Navy Bureau of Ships became very active in supplier liaison in order to establish better control of reliability in specifications. The "General Specification for Naval Ship and Shore Electronic Equipment, 16E4 (Ships)" was rewritten and sent to various electronics contractors for comments and recommendations. These comments were reviewed and, where applicable, included in the new "Specification, MIL-E-16400 (Ships)," dated May 1, 1953. This specification stressed greater use of preferred and standard parts, U. S. Navy reliable tubes, and high-quality materials. It also stressed simplicity in design and ease of installation and maintenance. It required the use of check points for fault location, human-engineering analyses during design, and many other specifications aimed at better reliability.

It was also during July, 1952, that the coordinated "Specification MIL-E-5400" was issued to supersede AN-E-19. In addition to many requirements which were beneficial to reliability indirectly, there was one paragraph that was specifically intended to help reliability. Because of the historic nature of this paragraph, it is quoted in full:

> Par. 3.2.9 *Life.*—The equipment shall be capable of 200 hours of operation without servicing. The total operating life shall be as specified in the detail equipment specification. Replacement of vacuum tubes operated within their specified ratings is not considered a failure. The equipment shall furnish specified performance after a life test. The cycle of operation shall be as specified in the detail equipment specification.

[2] Ad Hoc Group Rept. No. EL200/17, 2 vols; February 18, 1952.

[3] Various names for the same organization:

1) RMA—Radio Manufacturers Association, 1924–1950.
2) RTMA—Radio and Television Manufacturers Association, 1950–1953.
3) RETMA—Radio, Electronics, and Television Manufactuers Association, 1953–1957.
4) EIA—Electronic Industries Association, 1957–Present.

This was a wonderful tool for controlling reliability, but, although MIL-E-5400 was specified for years on hundreds of contracts, it remained unenforced for several years. Finally, the Air Force interpreted the first sentence to mean that all production equipment should be put on a special 200-hour life test to prove that the 200-hour requirement had been met. Each batch on test would be accepted if 80 per cent of them passed the 200-hour test. When a total of 22 equipments had been accepted, a reduced sampling of 1 in 12 would be used as the test sample. Failure of any composite sample caused the return to 100 per cent testing.[4]

An interesting document, "Electronics Equipment Reliability," was issued by the Navy (NAVSHIPS-91957) on June 15, 1953, in answer to the Department of Defense Directive Number 150.21-1, issued on September 12, 1951. In the foreword, Rear Admiral H. N. Wallin, U.S.N. Chief of Navy Bureau of Ships, explained:

> This pamphlet contains information showing the broad scope and the many factors involved in the effort to increase reliability of electronic equipment. Also included in this pamphlet is a summary report on the Navy Department's participation in the over-all Defense Department Electronics Reliability Program.

On August 21, 1952, six months following the final report of the Ad Hoc Group, and as one consequence of its recommendations, the Advisory Group on Reliability of Electronic Equipment (AGREE) was formed by the Department of Defense under the RDB. The first step taken by the new AGREE after its establishment was to review all the recommendations of the Ad Hoc Group and place them in priority categories for further study.

The next incident of importance occurred in 1953, when the Rand Corporation issued a research memorandum (RM-1131, dated August 14, 1953) entitled, "A Survey of the Current Status of the Electronic Reliability Problem," written by Dr. R. R. Carhart. This was an excellent survey analysis that set the stage for years of reliability work. One of its immediate uses was to provide information and impetus to the first major reliability-control program in an industrial company established at the Radio Corporation of America, Camden, N. J. Shortly thereafter, the head of this program, C. M. Ryerson, released the first reliability calculator for use in equipment design. It was a later outgrowth of this program that provided the now-well-established reliaability prediction technique known as application stress analysis.

On July 15, 1953, the Bell Telephone Laboratories submitted their final report on the Naval Bureau of Ships component-parts failure analysis program. This report covered the period from July 1, 1951 to June 30, 1953. Many thousands of returned failed parts from the Airinc, Vitro, and other collection programs were analyzed. Specific types and causes of failure were noted and this information made available to the parts manufacturers. Some significant facts were revealed by this analysis. For example, "62 per cent of all returned failed power resistors did not conform to the JAN specifications." Again, "The principle cause of potentiometer failures is overloading due to the failure of another component, primarily vacuum tubes."

An important summary of reliability problems was written by Russell H. Lindsay of the Boeing Airplane Company, in 1954. This was entitled, "Practical Expectations and Limitations of the Reliability Problem," and was printed in the *Aeronautical Engineering Review*.[5]

When the Research and Development Board was abolished in 1953, AGREE was transferred to the Assistant Secretary of Defense (Research and Engineering), and reestablished in 1954 as part of the Office of the Assistant Secretary of Defense (Applications Engineering). The final directive for AGREE, dated March 31, 1954, established the purpose of AGREE as assuring that " . . . the best available scientific, engineering, production and operational talent are applied to the achievement of reliability in the field of military electronics. . . . "

At the same time (March 30, 1954), the Panel on Electron Tubes was redesignated the Advisory Group on Electron Tubes (AGET) with essentially the same charter.

A few months later, (June 8, 1954), the Advisory Group on Electronic Parts (AGEP) was formed. This group's objective was to assist " . . . in achieving a sound, coordinated, and integrated research and development program in the field of electronic parts."

During the period following the formation of AGREE, AGET and AGEP, the question of unreliability in military equipment became a legislative issue. In July, 1954, the House of Representatives' Committee on Government Operations had before it for study a report[6] by its Sub-Committee on Military Operations concerning the development and procurement of a revolutionary new airborne radio transceiver. The sub-committee noted that the stress of the Korean war was a major factor in the need for the new equipment, but felt that the unreliability of the equipment should have been resolved before the Air Force ordered production quantities of the equipment. They stated, "The Sub-Committee is of the opinion that it is not enough to design equipment which meets complex performance requirements. To be of real value, the equipment must be economical in initial cost, operating cost, and maintenance costs, and in providing the flexibility necessary to meet changing operational demands."

In March, 1955, the same sub-committee reviewed the merits of an important piece of air navigation equip-

[4] This requirement of 80 per cent passing a 200-hour test was actually specifying a mean-time-between-failures of 1000 hours according to the exponential law.

[5] Vol. 13, No. 10; 1954.

[6] House Rept. No. 2578, 83rd Congress, 2nd Session; 1954.

ment.[7] Once again, unreliability was discussed and, again, the sub-committee noted that the reliaibility of a piece of gear should be determined before production is authorized.

The National Bureau of Standards', "Preferred Circuits Manual," was published in 1954, containing 17 suggested circuits. These were derived from a study of many equipments, both commercial and military, representing designs that were well stabilized. In all cases, similar circuits had been in use for at least ten years without major improvement. This project was sponsored by the Navy Bureau of Aeronautics and was aimed at reducing drastically the multiplicity of circuits in Navy aeronautical equipment. Later editions of the manual have given more emphasis to modern circuits improved from the reliability standpoint.

An important national event for reliability occurred in New York, N. Y., on November 12–13, 1954, when the first of the National Reliability Symposia was held in the New York Statler Hotel. This symposium was financed by the IRE Professional Group on Reliability and Quality Control, and jointly sponsored by the IRE and the ASQC. Additional sponsors (AIEE, AIA, and EIA) were added for subsequent events which have been held according to the following schedule:

Symposium	City and State	Hotel	Dates
1st	New York, N. Y.	Statler	November 12–13, 1954
2nd	Washington, D. C.	Statler	January 9–10, 1956
3rd	Washington, D. C.	Statler	January 14–16, 1957
4th	Washington, D. C.	Statler	January 6– 8, 1958
5th	Philadelphia, Pa.	Bellevue Stratford	January 12–14, 1959
6th	Washington, D. C.	Statler	January 11–13, 1960
7th	Philadelphia, Pa.	Bellevue Stratford	January 9–11, 1961
8th	Washington, D. C.	Statler	January 9–11, 1962
9th	San Francisco, Calif.	Palace	January — 1963

All the National Reliability Symposia have retained the same general theme, "A Progress Report on Reliability." The concept has been to report to the Nation, at the beginning of each year, concerning the present status and to challenge and spark progress in reliability for the coming year. There have been only minor variations on this theme year by year.

One interesting facet is that the early symposia were restricted almost entirely to reliability in electronics. After the first few symposia, however, it became obvious that the coverage was becoming broader and more general. As the field of reliability control in electronics developed, it became obvious that the principles and techniques involved were applicable to any type of industry. There has been considerable pressure from many groups to drop the limitation entirely, and thus the title of the 8th National Symposium no longer contains the words "in Electronics." Reliability is a universal subject which embraces all fields and technical engineering specialties.

Two important publications were issued in 1954. One was the NEL "Reliability Design Handbook,"[8] published by the Navy Electronics Laboratory, San Diego, Calif., under Capt. Henry E. Bernstein. The other was the companion NEL "Reliability Bibliography."

On March 17, 1955, the Department of Defense issued the directive DOD 4105.10. This directive made possible the award of the first production order to the developer and designer at a price higher than that of the low bidder. This provision is important to reliability for two reasons:

1) The developer has an incentive to invest his own money in doing a good job in developing a mature design.
2) A problem exists in that the developer knows the equipment better than anyone else and also knows the high costs of making it reliable; thus, his bid is not likely to be the lowest.

In April, 1955, the Air Force issued the "Reliability Factors for Ground Electronic Equipment," prepared by the McGraw-Hill technical writing service for the Rome Air Development Center, Griffis Air Force Base, N. Y., under contract NO AF 30(602)-912. Later, McGraw-Hill was given permission to put this book on the market as a commercial book. It was edited by Keith Henney.

In 1955 two important events developed from the progress in reliability which had been made at RCA. First, early in the year the program director, C. M. Ryerson, released the first Industrial Manual on Reliability entitled "The RCA Reliability Program and Long Range Objectives." This book was affectionately referred to as the "Brown Book" by many who used it as a basis for their own programs.

The second related event was a National Conference on "The Reliability of Military Electronic Equipment," held in the Warwick Hotel, Philadelphia, Pa., on August 16–17, 1955. This was jointly sponsored by the Air Force and the Radio Corporation of America. The purpose of the conference can best be described by quoting from the introductory remarks by Major General T. P. Gerrity, Assistant for Production Programming, Deputy Chief of Staff, Materiel, USAF.

> . . . Electronics has come to play a primary role in building modern air power. . . . While the substitution of electronics equipment for human control has given us the high performance we need, it has also greatly increased our dependence on the equipment for the success of our missions. At the same time, it has imposed a heavy responsibility and obligation on the electronics industry to develop and produce equipment of the highest degree of reliability. . . . As time goes on, our dependence on electronic . . . reliability will become even more urgent.

This same plea of urgency and importance has echoed and re-echoed in hundreds of similar meetings since that time. The interesting aspect of this meeting was the amount of valuable technique and information given

[7] House of Representatives Hearings, 84th Congress, 1st Session, 1955.

[8] Dept. of Commerce OTS-PB121839.

out to help bolster reliability efforts in industry. This theme can be explained best by further remarks from General Gerrity:

> ... I am confident that the scientific and production genius of American industry will find a solution to the urgent problem of combining high performance and reliability. . . . Further, I feel that we in the Air Force, jointly with RCA, have gained some very valuable experience over the past few years that will be helpful to industry in the solution to this problem. For this reason we have invited you here to participate in the Air Force-RCA Story on Electronic Reliability.

On October 24, 1955, a significant booklet was released by the Naval Ordnance Laboratory, Corona, Calif., dealing with instrumentation error and other factors in reliability measurement. This was written by Jerry L. Hayes, and was entitled, "Factors Affecting Measurement Reliability."

On December 9, 1955, the Lincoln Laboratory of the Massachusetts Institute of Technology released a helpful booklet for designers, written by N. H. Taylor, entitled, "Designing for Reliability."

In the spring of 1955 the DOD decided that AGREE could contribute more directly to the solution of the problems of reliability if it were reorganized and given specific assignments. Accordingly, a program of nine tasks was established in these areas: numerical reliability requirements, tests, design procedures, component parts, procurement, packaging and transportation, and storage operation and maintenance. A task group of members from the military departments and industry was assigned to each of the tasks early in 1956.

As a result of its investigations, a report of major magnitude was issued by AGREE on June 4, 1957, entitled, "Reliability of Military Electronic Equipment." The core findings of this report serve as a technical basis for the current national approach to military system reliability.

The successful results obtained on many contracts in which the information in this report has been applied and AGREE testing required proves the value of formal reliability control programs. Many applications of the AGREE report could be cited in a more complete history. However, at a minimum, it should be mentioned that the report of Task Group 4 on obtaining reliability in developmental contracts (R&D) was later issued as MIL-Std-441. Also, the test requirements of Task Groups 2 and 3 were included in MIL-R-26667 (USAF).

On October 1, 1956, RCA provided industry with the first breakthrough on the subject of specific part failure rates and derating information. The now famous TR-1100 report was issued as an engineering aid to design and reliability prediction. Later versions of this report were released through OTS Department of Commerce, and now the latest version is available as Section No. 8 of the RADC Reliability Notebook.[9]

The Radio Electronics and Television Manufacturers Association became active in reliability during 1956. At least three major efforts are worthy of mention, as follows:

1) On May 21 and 22, 1956, RETMA sponsored a symposium on "Reliable Applications of Electron Tubes." This was held at the University of Pennsylvania, Philadelphia, and was cosponsored by AGET, the IRE Professional Group on Electron Devices, and the Joint Electron Tube Engineering Council (JETEC).
2) On December 19 and 20, 1956, RETMA also sponsored a symposium on "Applied Reliability." This was held at the University of Southern California, Los Angeles.
3) Also in December, 1956 RETMA published a report prepared by its Systems Reliability Analysis Task Group, under the chairmanship of G. M. Armour of the General Electric Company. This was entitled "A General Guide for Technical Reporting of Electronic Systems Reliability Measurement."

In December, 1956, the Air Force issued the "Design Factors for Aircraft Electronic Equipment,"[10] prepared as a compilation of important work from a group of Air Force consultants.

The Air Force approach for some years, prior to 1956, had been to influence their contractors in every possible way to establish formal reliability programs. A major step in this direction was taken late in 1956 by General Baker, Director of Procurement and Production of AMC. Letters were sent to the major Air Force suppliers, requesting them to describe in detail their reliability-control programs. Then, in order to properly evaluate the answers, an Ad Hoc Committee of AMC-ARDC personnel was established and indoctrinated for this purpose. The Committee consisted of a representative from each of the following: within AMC—The Directorate of Procurement and Production, The Aircraft and Missiles Division, the Aeronautical Equipment Division, the Directorate of Maintenance Engineering, and the Office of Quality Control; from ARDC—The Directorate of Systems Management; and from WADC—The Directorates of Operations, Research, and Development. In all cases, representatives were selected for their knowledge and interest in the reliability area. In addition, representatives participated from the particular Weapons System Project Office or the Buying Office related to the contract program being reviewed.

The first task of the committee was to establish criteria for the evaluations—to create a yardstick against which the submitted programs could be measured. Ten elements for rating were established as follows: Concept and Approach; Organization; Programming; Quality Control; Reliability Requirement Studies; Qualification Testing; Acceptance Criteria;

[9] ASTIA Doc. No. AD-148868.

[10] WADC Tech. Rept. 56-148.

Failure and Deficiency Reporting; Analysis and Correction; Relationship with Vendors and Sub-Contractors; and, Training or Reliability Indoctrination. Eight manned-aircraft contractors, eight missile contractors, and ten subsystem contractors were evaluated. The analysis showed that most of the programs fell short of being good.

Of the ten rated elements, the poorest were Programming and Training. The next poorest were Reliability Requirement Studies, Organization, and Concept and Approach. These five are, of course, closely related and embody the planned, organized and directed approach so essential to good reliability results.

As might be expected, Quality Control was the strongest element, followed closely by Failure and Deficiency Reporting, then Analysis and Correction. Qualification Testing, Acceptance Criteria, and Relationship with Vendors and Sub-Contractors were on the better side of the balance. These five elements have all had considerable attention over a long period of time. They are recognized essentials to any good production program and are important elements in a reliability program. However, they will not by themselves assure *high* reliability.

The preceding is quoted directly from a report of the analysis given by H. G. Spillinger, Colonel, USAF, Chief, Product Engineering Staff Division, Directorate/Procurement and Production Headquarters AMC. This report was presented several times, once at the Air Force-Aircraft Industry Conference sponsored by the Directorate of Flight Safety, The Inspector General, USAF, Santa Barbara, Calif., March 12, 1957. Each time the report was given it included the important statement: "Over-all, our evaluations reveal that too much reliance is being placed on 'fix-it' procedures—insufficient attention to the principle that reliability must be inherent in the basic design." Unfortunately, this is a truism valid to this day.

In January, 1957, the Air Force issued "Techniques for Application of Electronic Component Parts in Military Equipment."[11] This is a very thorough manual on part application which contributed greatly to design reliability.

From 1955 through 1958, the reliability program at the Army Redstone Arsenal was in full swing. Many interesting and thought-provoking booklets were released. Outstanding among these is "The Notorious Unreliability of Complex Equipment," dated September, 1956, and "Reliability Through Safety Margins," dated October, 1958, both by Robert Lusser; and "Optimum Allocation of Funds for Reliability Programs of Guided Missiles," dated January, 1955, revised, April, 1958, by Dr. Erich Pieruschka.

On July 1, 1957, the International Geophysical Year began officially. Actually, a great deal of the research work associated with IGY had been going on for months prior to the starting date and did not finish through several extensions of the IGY. Considerable significance for reliability relates to the IGY effort. Perhaps one of the outstanding achievements was the high-altitude balloon flight of Air Force Scientist Major David G. Simons. This flight reached the height of 100,000 feet and produced evidence of the high-radiation content of the upper atmosphere, later identified as the Van Allen Belts. Many other developments of the IGY have helped take the space satellite out of the realm of science fiction. In this connection, the notorious unreliability of the Vanguard equipments should be placed at the top of the list of examples from which to learn reliability lessons.

In 1957 the Air Force coined the word "air/space" to describe the single operational medium once conceived by many to be two separate mediums: "air" and "space." This term was later modified to "aerospace." Thus, a modern aerospace craft is a vehicle designed to operate both within the earth's atmosphere and in the vacuum of outer space. An example is the Air Force dyna-soar. A true spacecraft is one designed to operate only in the vacuum of space. The implications to reliability of these new operational environments are only beginning to become apparent. Entirely different orders of magnitude of effort, urgency and expense are involved now than were required in the past.

In October, 1957, the Advisory Group on Electronic Parts (AGEP) made a good contribution by issuing a manual entitled "Environmental Requirements Guide for Electronic Parts." This guide was subsequently issued as MIL-Std-446.

Three months prior to the publication of the AGREE report in June, 1957, the Ad Hoc Committee for Guided Missile Reliability (ACGMR) was formed (March, 1957), under the Assistant Secretary of Defense (Research and Engineering). This committee was transferred to the Office of the Director of Guided Missiles on November 15, 1957. The ACGMR was to design a ". . . uniform monitoring program and management procedure that can be used effectively for all types of guided missile projects."

In April, 1958, the ACGMR published a report,[12] frequently referred to as the "Red Book." The management and monitoring program proposed starts when the contract is awarded and continues through all phases of design, development, production, and major product improvement. To ensure compliance with reliability specifications, and to aid the contractor in knowing whether these goals are being reached, eight test points or "milestones" are established in each project, at which time a fully documented report from the contractor, concerning either the predicted or the verified reliability, is required.

These eight reliability monitoring points were based on several conclusions of the committee. First, the committee believed that ". . . reliability is a parameter that can be predicted, assessed, measured and controlled

[11] WADC Tech. Rept. 57-1, vol. 1, ASTIA Doc. No. AD110672.

[12] "Proposed Reliability Monitoring Program for use in the Design, Development, and Production of Guided Missile Systems," DODOA SD-GM; April, 1958.

during the design, development, production, and major product improvement phases of guided-missile weapon systems." Second, they believed that ". . . it is technically feasible and sound to specify and monitor reliability in guided-missile weapon sytems during their growth cycle."

One of the recommendations in the AGREE report of June 4, 1957, was that a permanent group be established at Department of Defense level to include representatives of industry and the three services, and to be charged with the tasks of developing military-component specifications, testing component parts for design capability, and developing inspection methods. This group would serve as a National Reliability Center.

A year later, action was taken to do something about this recommendation. On July 14, 1958, an agreement was made between the Director of Production Policy, OASD (Supply and Logistics) and the Director of Electronics, OASD (Research and Engineering)[13] to establish the Ad Hoc Study Group on Parts Specification Management for Reliability. The assignment to the new Ad Hoc Study Group was to ". . . analyze the recommendations established by the AGREE Task Group 5 in order to advise the Assistant Secretaries of Defense (Research and Engineering and Supply and Logistics) regarding efficient implementation methods and procedures." Expanding upon this objective the group studied specification preparation, requirements for reliability, support from industry and documentation, and also questions of qualified product lists and the need for a management organization for military part specifications at the Department of Defense level.

The report of this Ad Hoc Study Group was issued in two volumes in May, 1960.[14] In addition to establishing three prototype specifications for reliable parts and proposing new procedures for refining the technical documentation of parts characteristics, a proposal was made to establish a new advisory group to report directly to the Assistant Secretary of Defense (Supply and Logistics). The proposed title for this group is AGMEPS, the Advisory Group for the Management of Electronic Parts Specifications. It is understood that negotiations are now underway on methods to implement these latest recommendations.

During the period in which the Ad Hoc groups were active at the DOD level, the military departments were also working on reliability. The Navy (Bureau of Ships) had continued its tube analysis and improvement program, and other Bureaus had pursued various projects for better reliability. A major Navy effort for reliability was initiated by the Chief, Bu Aer, Rear Admiral R. E. Dixon, on October 13, 1958. In Bu Aer Instruction 5420.13, Admiral Dixon announced the establishment of a Bu Aer-Industry Advisory Board on Reliability and Operational Design Requirements of Aeronautical Material (BIMRAB). A brief charter described the function as advisory to the Chief Bu Aer on the reliability, maintainability, and safety requirements for maximum operational integrity. The charter also specified that the Assistant Chief for R&D Bu Aer would be Chairman of the Board, and that a mixed group of Bu Aer and Industry members not to exceed 15 people would be appointed (10 from industry).

An Ad Hoc Group of Bu Aer and industry members met on February 13, 1959, and established the preliminary organization and procedures for the board. Working groups were appointed for Management Procedures, Engineering Procedures, Production Quality Control, Field Maintenance and Support, Information and Education, and Data Processing. Under the initial Chairmanship of Rear Admiral L. D. Coates and others, the Advisory Board performed a valuable service in funneling advice and recommendations from industry and Bu Aer activities to the chief of the bureau.

On February 13, 1959, shortly after the Navy Bureau of Weapons was formed from a coalition of the old Bureau of Ordnance and the Bureau of Aeronautics, a directive (BU WEPS 5420.3 FQQE-21) was released transferring the BIMRAB to the Chief of the Bureau of WEPS, Rear Admiral Paul D. Stroop.

During the years following 1957, a series of Navy-Industry Conferences were held. The later conferences were taken over as a major function of BIMRAB. These can be listed as follows:

Number	Date	Location
1st	October, 1957	Naval Aviation Safety Center, Norfolk, Va.
2nd	November 5–6, 1958	Virginia Beach, Va., Cavalier Hotel
3rd	November 4–5, 1959	Virginia Beach, Va., Cavalier Hotel
4th	November 1–2, 1960	Washington, D. C. Shoreham Hotel
5th	Planned for Fall, 1962	Washington, D. C.

Many events important to reliability occurred during 1958 which were continuations of programs already described, such as the 4th National Reliability Symposium, which was held in Washington, D. C., in January. But an event at this symposium deserving further comment was an evening panel discussion on reliability definitions. This subject is probably one of the major problems in our national reliability picture. Very few conclusions could be drawn from the panel discussion at the symposium, except that nearly every tentative definition was highly controversial, and that the semantics problem was a serious deterrent to reliability achievement. It was obvious that a great deal more work on definitions was needed.

On December 15, 1958, the National Definition Pro-

[13] E. J. Nucci, "Progress report on ad hoc study on parts specifications management for reliability," 1959 IRE NATIONAL CONVENTION RECORD, pt. 6, pp. 120–129.

[14] PSMR-1 Parts Specification Management for Reliability, Supt. of Documents, U. S. Govt. Printing Office, Washington, D. C.; May, 1960.

gram, jointly sponsored by IRE-PGRQC, the ASQC Electronics Division, DOD, and the AIAERP, issued an "Interim Dictionary of Terms and Definitions for Product Assurance and Related Fields." This program was chaired by C. M. Ryerson of RCA, and the interim dictionary contained a compilation of definitions from many sources screened mainly to eliminate major contradiction.

On August 1, 1958, the Air Research and Development Command (ARDC) issued a booklet, "Technical Program Planning Document (TPPD) Release Program." The objective of the TPPD Release Program was to "engage the engineering and scientific talent of the United States under an Air Force-Science-Industry team concept for the purpose of providing a more comprehensive and integrated effort in achieving Air Force goals in research and development."

Quoting from the introduction of the TPPD document, "In the final analysis, national security is founded on our ability to produce quality (reliability) perhaps even more so than quantity,—and it is only common sense that everyone on the team should be kept fully informed of our technical goals as we see them. These technical goals are furnished by qualified people in Science and Industry in the form of documents known as Technical Program Planning Documents (TPPDs)." The efforts of many people to induce the vision of reliability research as a major need to be incorporated in the TPPDs has to-date borne little fruit.

In December, 1958, AGEP[15] issued the first "Capsule Summary of the AGEP Program in the Area of Electronic Parts and Materials," GEP245/1. This document is a compilation of brief summaries on the active and completed research and development projects sponsored by the Army, Navy, and Air Force on parts and materials. Since that date, five other summaries have also been issued at about six month intervals. The last issue (October, 1960) consisted of two volumes; the active projects and the completed ones. Current project sheets on all active projects can also be obtained on a need-to-know basis. The AGET[16] also issues similar sheets and summaries related to projects on tubes.

Just prior to the January, 1959, 5th National Reliability Symposium, the Radio Corporation of America released a "Manual on Product Assurance," prepared by C. M. Ryerson, then newly appointed Staff Manager of Product Assurance. Quoting from the introductory abstract, "Product Assurance is the new management centered program for giving assurance to management and to the customer that all product attributes such as reliability, quality, maintainability, productibility, value, system integration, etc., comply with commitments and will satisfy the customer's needs." The product-assurance program extends into all operations of each division from early contract negotiations to final field follow-up, and includes the responsibility for assuring an optimum effective trade-off among all the controls of product characteristics. Since this manual has been issued, many organizations have adopted similar programs to integrate their various control operations.

In March, 1959, the Inter-Service Data Exchange Program (IDEP) was born. The purpose of this program was to provide for the free and automatic interchange of data among the ballistic missile major contractors, subcontractors, and government agencies of unclassified, nonproprietary reliability data. Three centers for the three services were established: BMD (now BSD), Low Angeles, Calif., for the Air Force; NOL, Corona, Calif., for the Navy; and ABMA, Redstone Arsenal, Ala., for the Army. Data was to have been exchanged on dual aperature microfilm EDP cards. These have since been found to be inadequate and a microfilm strip attached at the top of a printed summary page is now being used. The civilian leader for this project is Martin Barbe, the Aerospace Corporation, Los Angeles, Calif.

Mention should also be made of the Battelle Memorial Institute data exchange program. In March, 1959, Battelle established its Electronic Component Reliability Center, subscribed to at $20,000 per year by a group of industrial firms, all large users of electronic parts. In addition to exchanging test data from its subscribers, the ECRC performs research and data analysis pertaining to the reliability of electronic parts.

On May, 18–20, 1959, the Air Force held a review conference at AMC Headquarters, Wright Patterson Air Force Base, Ohio. This was entitled, "Reliability of Weapon System Equipment," and was jointly sponsored by the Air Materiel Command (AMC) and the Air Research and Development Command (ARDC). Its purpose was to indoctrinate the representatives of the Air Material Areas, Depots and Centers in the philosophy and policies of Reliability. This was a part of the general Air Force reliability training program planned for all divisions and groups. Both military and industry people were invited speakers to present the best known about reliability control.

On July 14, 1959, the Space Technology Laboratories released "STL Reliability Policies and Procedures." Under this system, technical management contract with the AFBMD STL was required to give technical assistance and direction to the various contractors on the subject of reliability. A committee was formed, with H. R. Powell as Chairman, under the technical direction of James H. Allen, to develop a complete guide to reliability. Much of the contents of this document is identical with previous guidance material issued by Allen on the Thor and Atlas programs. It also was closely allied with the contents of AFBM Exhibit 58-10, which is now referenced on nearly all missile contracts.

A significant event occurred during the two weeks

[15] AGEP Secretariat, 200 South 33rd Street, Philadelphia 4, Pa.
[16] AGET Secretariat, 346 Broadway, New York, N. Y.

from October 19–30, 1959. This was the first "Reliability Training Course for Business Management, given under the auspices of a special Reliability Training Committee, jointly sponsored by IRE-PGRQC and ASQC-Electronics Division. This first course was a dual session event given at the Traymore Hotel, Atlantic City, N. J. Sixty-five students from all over the country attended. Twenty-two similar, but one session and one week, courses were given in various parts of the country during subsequent months. C. M. Ryerson was the organizing General Chairman of the joint training committee and the IRE representative. F. M. Gryna, Jr., was general Vice Chairman and the ASQC representative. A training text had been prepared during the preceding year, based on contributions from over 35 authors. An editorial board, under the direction of C. M. Ryerson, performed the final editing.

Many specifications and directives have been written by the military services and it would be futile to try to describe them all here. However, the following paragraphs review some of the more important of these documents as they relate to AGREE and subsequent DOD documents.

On March 23, 1959, Headquarters, USAF, sent a letter to the Commanders of AMC and ARDC stating, in part, that a real ". . . need for increased emphasis on reliability. . ." had arisen as a result of the increased complexity and costs of present weapon systems. As a result of this need, Headquarters, USAF, directed that ". . . all new weapon systems contracts include a meaningful reliability requirement . . .," and it proceeded to define "meaningful" by asserting that ". . . a contract should state clearly what the contractor is to achieve and, of equal importance, what specific tests or means the contractor must use to demonstrate achievement." The letter went on to deal with current contract and contractors, making the point that these ". . . contractors should be required to define present reliability goals and establish suitable tests to demonstrate that the reliability achieved meets operational needs." Although merely a letter to two subordinate commands, this letter is important in that it constitutes the first statement concerning reliability made by one of the military services headquarters.

A few days later, with a primary aim at aiding the quality-control personnel, Headquarters, AMC issued AMC Pamphlet 74-1 (March 31, 1959), "Reliability Evaluation Procedures for Pilot-Production and Production." This Pamphlet is basically ". . . a detailed explanation of the reliability index evaluation procedures recommended by Task Group Number 3 of the Advisory Group on Reliability of Electronic Equipment (AGREE). Specific emphasis is placed on methods for determining and evaluating the reliability index (mean-time-between-failure) of electronic equipment during pilot-production and production." Containing tabled and graphed figures of various functions such as "equipment reliability as a function of mission time and MTBF," "sequential sampling plan," and the like, the pamphlet assists quality control persons to become relatively knowledgeable in those statistical evaluation techniques concerned with sampling for reliability evaluation.

This was followed a month later by Department of the Air Force Letter 84-1, (April 30, 1959), subject, "Reliability Monitoring Program for Weapon Systems."[17] Attached to this letter was a modification prepared by representatives of both AMC and ARDC, of the Department of Defense's Ad Hoc Committee for Guided Missile Reliability (ACGMR) report. In this publication, directive in nature, the Department of the Air Force promulgated three rules: 1) each contract must include a meaningful reliability requirement; 2) the contractor must be prepared to submit to continuous monitoring of his reliability efforts; 3) the contractor must *prove* that the established reliability requirements are satisfied through a formal series of acceptance tests.

Immediately following the issuance of the Department of the Air Force Letter 84-1, USAF Specification Bulletin 506 (May 11, 1959), "Reliability Monitoring Program for Use in the Design, Development and Production of Air Weapon Systems and Support Systems," was published. This specification bulletin, also a modification of the ACGMR report, was a close copy of the attachment to Department of the Air Force Letter 84-1. With the publication of this Bulletin, procuring officers were provided with a document that could be used "externally," in contradistinction to Department of the Air Force Letter 84-1, which was an "internal" document.

During the two-month period following the publication of USAF Specification Bulletin 506 (May 11, 1959), numerous reliability publications were issued most of which drew upon the ACGMR and the AGREE reports. The day after "506" was produced, the first version of MIL-R-26667 (USAF) (May 12, 1959), "Reliability and Longevity Requirements, Electronic Equipment General Specification For," was issued by the Directorate of Laboratories, Wright Air Development Center. On June 2, 1959, it was superseded by MIL-R-26667A (USAF). This specification ". . . covers the analysis and estimation of reliability and longevity, the method of measuring reliability and longevity, the analysis of deficiencies and report of effort to achieve specified reliability and longevity." To assist the contractor, the specification provides for a design analysis to be performed by the contractor in order to establish ". . . the failure rate for each part application and the expected reliability and longevity of the equipment." Repeated analyses are to be performed in order to account for ". . . effects on reliability and longevity which may

[17] An Air Force letter is administrative in nature and general in application. It contains ". . . regulatory material considered to be temporary in duration, or informative matter that may be of either temporary or permanent interest."

occur because of changes in manufacture and supply." If variance does occur and the equipment is below the specified reliability, the contractor has to develop a plan which can raise the expected reliability to the required level. When referenced by a detail specification, the described analyses are added to contractual requirements and are considered part of the acceptance testing.

This general specification also outlines how reliability test environments should be established, how equipment for test should be selected, and how test procedures should be followed. In addition, it requires initial, special, prototype evaluation, progress, and part-failure-rate reports in order to provide the procuring activity with as much information as possible concerning reliability. MIL-R-26667A is an attempt to utilize the test plans of AGREE Task Groups 2 and 3.

On June 18, 1959, the Directorate of System Management, Headquarters, ARDC, issued MIL-R-26674 (USAF), "Reliability Requirements for Weapon Systems," the purpose of which was to cover ". . . the general requirements for the establishment of an organized reliability program by the contractor to assure the attainment of the reliability requirements specified for the weapon system." The specification notes that this program should take into account the efforts of subcontractors and vendors. The remainder of the document discusses those factors which can affect the reliability of equipment, factors such as "trade-offs," derating, redundancy, testing, and quality control. This particular specification, together with USAF Specification Bulletins 506 and 510, presents a complete program for a contractor to follow in establishing his reliability organization and program.

Shortly after this version of MIL-R-26674 was completed, the Air Research and Development Command released its May, 1959, Special Study Group report, "Reliability in Missile Programs." The Study Group, at the direction of the Commander, ARDC, was to investigate the problems associated with current missile reliability programs and develop recommendations for action by the Department of the Air Force to improve its program. When the study was completed, it was sent to forty-four industrial and military organizations for comment and criticism. Thirty-five responded in detail.

Soon after the special Study Group report was released, the Air Force Ballistic Missile Division, ARDC, and Ballistic Missile Center, AMC, issued a new exhibit pertaining to missile reliability. Designed specifically for ballistic missile systems, AFBM Exhibit 58-10 (June 1, 1959), "Reliability Program for Ballistic Missile and Space Systems," provides ". . . optimum requirements and procedures for contractor reliability programs." The exhibit requires that the contractor prepare a description of the reliability program that he intends to follow during the contractual period. His program must include a sequence of monitoring points which provide for time-phased reviews of reliability efforts and the results achieved by that program. These planned reviews will be made by the contractor as well as by the procuring agency. In order to assure that the specification's reliability requirements are met, the program must include provisions for considering reliability during the design stage. The reliability requirements are to be determined and demonstrated by the contractor. Major portions of the document are concerned with the details of the associated reporting system.

In order to provide "internal" direction, as contrasted with "external" direction as represented by MIL-R-26674 (USAF), the Directorate of System Management, ARDC, produced Directorate Office Instruction No. 57-3 (June 23, 1959), "System Reliability." The instruction established ". . . reliability policy, objectives, responsibilities, and procedures for system programs." Applying to all joint ARDC-AMC System Project Offices, the instruction provides the project officers with procedures for system programs and a list of the officers' major responsibilities, including among them, the following:

1) Define and negotiate quantitative reliability requirements in contracts.
2) Require contractors to formulate and implement a comprehensive reliability program.
3) Establish and chair a system Reliability Review Board to monitor the program and results achieved.
4) Designate within the project office a Reliability Monitor.

In addition, a periodic review of each contractor's reliability efforts and results is to be performed by the "Reliability Review Board." This board, organized as a subcommittee of the Weapon System Phasing Group, draws its members from the ARDC System Project Office, the AMC System Project Center, and the ARDC Test Center System Project Office, each supporting ARDC Development Center, the Logistics Support Manager, and the Operating Command. At the end of each review, the board issues a report ". . . summarizing the status of reliability effort, results achieved, and any recommendations on design changes, reorientation of the program, funding, or other actions necessary to insure that the required reliability will be achieved." The board may call upon technical advisors to guide it.

To provide a continuous effort, the instruction further requires that one or more Reliability Monitors be appointed the following responsibilities:

1) Insuring that the provisions of this instruction and the reference documents are applied to programs in a timely fashion.
2) Exercising continuous surveillance of the contractor's reliability effort.
3) Providing the Systems Project Officer with the reliability factors that must be considered in all management decisions which can affect the operational reliability of the system.

4) Insuring that all available and applicable reliability methods, techniques, and criteria are employed by contractors and supporting Air Force activities.
5) Insuring that revisions to existing published procedures and criteria are applied to programs without undue delay including revisions to general specifications applicable to all programs.

Combined with ARDC Regulation 80-21, this instruction completed, for the moment, the ARDC contribution to a formalized internal program. The regulation provided the over-all policy; the instruction told how it would be managed and applied.

With the publication of U. S. Air Force Specification Bulletin No. 510 (June 30, 1959), "Guides for Reliability Organization," the managerial aspect of the Department of the Air Force reliability program was expanded further. This bulletin was designed to aid contractors in establishing their reliability organization. The program required for the reliability organization is complex and varied and, in turn, requires a fine organizational setup. For the administration of the setup, the bulletin presents three alternative means for determining the adequacy of a proposed reliability organization. The first is based upon the criteria developed by ARDC and AMC for their review efforts. The second is based upon the general experience acquired in numerous industrial programs. The third ". . . takes the form of a Master Check List of Reliability Program Practices based on a specific weapon system program and observation of similar programs."[18]

The next publication to appear was the Aeronautical Accessories Laboratory's MIL-R-27173 (USAF), (July 6, 1959), "Reliability Requirements for Electronic Ground Checkout Equipment." This document was designed to detail ". . . the minimum requirements which must be followed by a contractor to assure the design and manufacture of reliable equipment." Like the other specifications, this specification covers the reports and requirements laid upon the contractor. A minimum MTBF is given (300 hours) to be used when the detail specifications do not provide an MTBF. Tests are required, as are reports similar to those required by preceding specifications.

The next major step in reliability occurred on August 22, 1960, when ARDC issued ARDCR 80-1, "Reliability." This new regulation did not rescind ARDCR 80-21, but added, in addition to other things, the very important phrase ". . . reliability will be a major factor in all source selection actions. . . ."

Soon after ARDCR 80-1 had been published, the Air Force issued AFR 375-5, "Reliability Program for Weapon, Support and Command and Control Systems," dated October 17, 1960. This regulation, too, is indicative of the major steps which have been taken in the reliability area since the early days. Because Air Force contracts are so important to the industrial manager, the policy for the Air Force reliability program is given here in its entirety.

a. Reliability is considered a fundamental characteristic of every part, component, module, and complete system.
   (1) Reliability will be stressed during early system studies, source selection, design, development and production.
   (2) Maintenance, storage, transportation, and operation activities should exercise appropriate controls to maintain the reliability of the delivered article.
   (3) Systems in development and production and future systems will be analyzed, and an appropriate reliability program with realistic requirements will be established for each. Each program will include numerical probability values from a minimum acceptable to the desired goal, with such intermediate quantitative values required to measure progression, and a stated minimum acceptable confidence level for each probability value.

b. Specifications, exhibits, and product descriptions and contracts for systems and associated materiel, including GFE for inventory, will include specific minimum acceptable reliability requirements as one of the major engineering factors.

c. Contracts for systems and major sub-systems will include a requirement for a compehensive contractor reliability program.

d. The Air Force will maintain surveillance during all phases of development and production over the contractor's reliability program, his reliability testing, and quality control activities. Collaboration by reliability, production, and quality control personnel with engineering personnel during concurrent development and production will not relieve the engineering agencies of the responsibility for engineering approval.

e. Contractors' reliability capability, both past performance and proposed programs, will be a major factor in all source selection action.

f. If contract reliability requirements are not met, or if the contractors' reliability effort is decreased, the decision to accept or reject the end item or the revised reliability program will be considered with a view toward monetary penalties, unit price decreases, or other considerations deemed equitable.

g. The adequacy of planned reliability efforts, including funds availability, shall be considered when reviewing new programs. Proposals for increasing reliability efforts and funds on programs in existence shall be considered on the basis of the net effect on the over-all Air Force capability and economy, including such factors as spare parts requirements, maintenance workloads, engineering changes, operating costs, and the effect on the system concerned.

h. Reliability monitoring points will be generally established in the following sequence:
   (1) Detailed design study.
   (2) Preprototype.
   (3) Prototype.
   (4) Preproduction demonstration.
   (5) Demonstration of service readiness.
   (6) Service evaluation.
   (7) Full-scale production.
   (8) Demonstration of major product improvement.

   However, this generalization is not intended to delineate the complete or ideal system life cycle, but to emphasize the typical points at which the program should be monitored.

Paragraph F should be particularly noted, since this indicates the first time that monetary or other penalities for failure to meet reliability goals have been officially sanctioned by the Air Force. The various Air Force subordinate commands are assigned the responsibility of actually carrying through these policies.

During this period among the many good publications from DOD which helped achieve better reliability was the "Quality Control and Reliability Handbook," H 108, issued on April 29, 1960. This contains sampling procedures and tables for life and reliability testing based on the exponential distribution. Other handbooks of this series were also helpful.

[18] This "Master Check List" was taken from the "Reliability Handbook, 7-58-2954-9," compiled by W. J. LeVan of Bell Aircraft Corp., Buffalo, N. Y.

Another event for reliability during early 1960 was the establishment of the National Standards Committee for Reliability C-31 at IRE. The chairman chosen for this Committee was C. M. Ryerson. This committee is now organized and places emphasis on top management participation.

On October 30, 1959, the Air Force issued the "RADC Notebook," in loose-leaf form. This is a comprehensive reliability manual intended to be kept up-to-date by subsequent issue of replacement sheets. Various sections treat different subjects. Section 8 now contains the latest revised version of the RCA TR-1100 Stress Analysis and Prediction information.[19] The sponsor of this book at Rome is Joseph J. Naresky, General Engineering Laboratory, Rome Air Development Center.

On December 15, 1959, the Air Force issued a technical report (RADC-TR-59-243), "Reliable Preferred Solid-State Functional Divisions," prepared for Rome Air Development Center by the West Orange Laboratory of Vitro Corporation. This report was an attempt to determine those circuit functions in which revolutionary increases in reliability are achievable by solid-state designs.

On January 8, 1960, the Army Ordnance Corps issued the handbook prepared by H. Walter Price, Chief of the Reliability Branch at the Diamond Ordnance Fuze Laboratories. This is entitled "Reliability Engineering Notes," and contains a collection of application notes obtained by the Fuze Laboratories in recent months.

V. The Product-Assurance Decade

The preceding sections have covered the years before 1940, including the standards decade; the quality-control decade to 1950; and the reliability decade to 1960. A prominent characteristic of all these periods is that, as new control specialities come into the picture and blossom into full technical fields, the old fields continue to grow and bloom. For example, there has never been any secession of the pressing need for standards as quality control and reliability have come into emphasis. Rather, the need for quality and reliability standards has greatly increased the scope and importance of the standards field. Likewise, as reliability came to the front for special emphasis, quality control has also taken on new facets of importance and new duties. The best design has no value unless it is manufactured into hardware with a minimum of degradation to its inherent reliability.

The end of the decade of the 1950's has witnessed the phenomenal growth of many other control specialities. For example, the best design from the reliability standpoint may be very poor from the standpoint of maintainability. Likewise, a system which is very effective functionally, if operated by itself, may not be reliable if operated in close proximity to other systems. System integration is a new specialty which must be given its necessary support. Of course, there are other product abilities, such as producibility, which have long been important design considerations. But now we have new control specialities, such as value engineering. This recital of new specialties in recent history could review a dozen or more, each of which is the most important in the eyes of its devotees.

But no military operation, and surely no company, can long afford to support a dozen or more independent and across-the-board control agencies and programs. Fortunately, if each of these control specialities is analyzed, it is found that they all have many functions in common. They all aim at a form of technical direction by providing information and operational guide lines. They all depend for their effectiveness on some form of auditing or monitoring. They all need data input from essentially the same sources and feed it back to essentially the same people. They all are based on a modern analytical approach which substitutes numbers for intuitive qualitative decisions. They all take advantage of an unbiased and independent second look to "see the woods and not the trees." And, finally, they all must give some form of assurance to management that optimum teamwork in their speciality is being achieved between all organizational groups.

When these similarities are considered from the management standpoint of the necessity of obtaining an optimum trade-off among all the speciality objectives and project restraints, the stage is set as at the present time for the entry of the product-assurance decade.

When good value can be sold at a good profit and on schedule, this is good business for all concerned. But this requires good cost control and good technical management. An all-out effort to achieve this objective is clearly the challenge confronting this nation and the world in the next decade.

By definition, the management program to achieve good value consisting of good quality, reliability, maintainability, producibility, system integration, system effectiveness, etc., at a low sales price and good profit is product assurance. Welcome to the decade of product assurance, and may your specialty prosper.

VI. Addendum

This partial history of reliability has been compiled with the sad realization that many important events probably have been omitted. If those who become aware of an omission or incorrect entry will charitably advise the author of the full correct details, corrections will be made in any subsequent issue.

Although indebted to many too numerous to mention for their suggestions, the author feels obligated to make special mention of Dr. Joseph Spiegel of the Human Factors Department of the Mitre Corporation, and L. M. Clement, for their special contributions.

[19] Copies of this Manual may be obtained as RADC-TR-58-111, ASTIA Doc. No. AD-148868.

Section 26

SPACE ELECTRONICS AND TELEMETRY

Organized with the assistance of the IRE Professional Group on Space Electronics and Telemetry

Milestones in Development of Electronics for Remote Measurement and Control*

CONRAD H. HOEPPNER†, FELLOW, IRE

Summary—**The developments in electronic remote measurement and control as applied to aircraft, rockets and satellites are described. The highlights are considered to be those developments which have contributed most to the technical progress made in the last twenty years. The impact of magnetic tape recorders, automatic digital computers and transistors upon telemetry and remote control is shown. The transition from remote control to automatic control and guidance is noted and the advent of the remote control-computer-telemetry closed loop is discussed. The characteristics of data transmission which differentiate it from ordinary radio communications are introduced and new horizons in the industrial applications of telemetry and remote control are described.**

IT HAS BEEN a rare privilege to witness and participate in progress made in the short interval of the last two decades, progress which, if it had not been for the stimulus of defense and war, might have required a period of several lifetimes. Although the basic research which made the technical developments possible preceded this period, the applications of that research have borne fruit as practical and spectacular new methods of remote measurement (telemetry) and control. In these past two decades there have been dozens of significant advances worthy of mention and a few of a character which may be classified as milestones. In the opinion of this author the significant milestones may be listed as follows:

1) The development of telemetry and remote control for missiles and spacecraft.
2) The development of instrumentation recorders employing magnetic tape.
3) The recognition of data handling as a problem and the subsequent development of automatic data handling equipment.
4) The development of automatic remote control for guidance.
5) The union of telemetry, remote control and automatic computers in a closed loop.
6) Industrial uses of telemetry and remote control.

Telemetry and remote control are old as employed by the electric utilities for remote switching and metering. These systems, while adequate for their purpose, were slow and generally very specialized. The revolution of telemetry and remote control came about as a result of the need for high-speed systems and the requirement for measurement from and control of moving vehicles. The early development of radio telemetry and remote control was centered around the drone and missile programs of the armed forces. It began slightly before World War II and has been growing and expanding to embrace the testing of aircraft and missiles and the transmission of data from and remote control of mis-

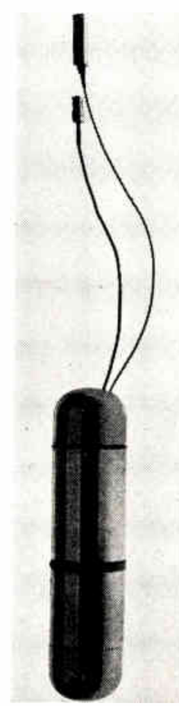
* Received by the IRE, October 9, 1961; revised manuscript received, November 27, 1961.
† Electronetics Corp., Melbourne, Fla.

siles, satellites and spacecraft. Early drones were nothing more than aircraft with the pilots removed and remote radio control substituted. During test phases a pilot was usually carried to perform takeoffs and landings and to observe the performance and deficiencies of control equipment. When the drone was used as a weapon, the pilot was removed and it functioned both automatically and by remote control. It soon became apparent that these missiles could be made better, smaller and more economically if they were originally designed to be pilotless.

One of the early missiles of pilotless design, designated the JB-2, was a copy of the German V-1 buzz bomb. Remote control equipment from the drone programs was adapted to the JB-2. Telemetry was developed to measure the performance of the control equipment and the missile. Meanwhile similar development was carried on to support other programs of strictly "American" design. Remote control developments preceded telemetry by some years, but remote control, as it was initially used, was an "on-off" system, rather than one which provided proportional control. Furthermore, remote control was an intermittent function inasmuch as the vehicles were stabilized by internal automatic equipment. Telemetry, on the other hand, required proportional and linear transmission of measurements on a continuous basis. Consequently, the remote control systems were "adapted" in concept only and telemetering equipment development proceeded quite independently.

Before continuing with the history of remote measurements and control it is well to describe a few of the technical characteristics of typical systems. Remote measurement is generally referred to as telemetry, which may also be considered the inverse of remote control. Telemetry is a special form of communication characterized by 1) the transmission of steady-state (dc) information, 2) flat frequency response from zero to as high as twenty or thirty kilocycles per second, and 3) the multiplexing of numerous different channels of information on a single radio carrier. It is seldom that all of these characteristics are required at one time, but instead, there are many different applications of telemetry with a variety of combinations. Consequently a number of telemetry systems have evolved. A common denominator of telemetry systems and remote control is multiplexing because, with rare exception, data from many sources and commands to many functions are required. Two principlal methods of multiplexing have been utilized; *viz.*, frequency mutiplexing and time multiplexing. In the first, a number of subcarriers are employed, each at a different frequency. All modulate a single radio carrier and then are separated by filters at the receiving station. In time multiplexing a multiplexer or commutator samples each channel successively and applies its voltage to either a subcarrier or the radio carrier. Both methods of multiplexing are used widely. Typical frequency multiplex systems utilize both amplitude-modulated and frequency-modulated subcarriers—usually frequency modulating the radio carrier. Typical time multiplex systems are pulse position modulation, pulse duration modulation, pulse amplitude modulation and pulse code modulation and the radio carrier may be either amplitude or frequency modulated. Characteristics of these systems have been previously published. The frequency multiplex system utilized well-known communication techniques while the time multiplex system developed along lines guided by more recent considerations of information theory. The number of channels, the number of samples per second per channel, and the accuracy of each data sample are the basic considerations of telemetry. The product of these three quantities represents the capacity of the telemetry system and determines the width of radio spectrum which is required to transmit the information.

A simple representative system of FM/FM telemetry is shown in Fig. 1. It was designed to telemeter one or more channels of information over relatively short distances. In this system the output from a resistance bridge tranducer (typically) is transmitted over a radio carrier to a receiving station where a permanent graphic plot is made by a recorder. A resistance bridge controlled oscillator generates the frequency-modulated subcarrier that in turn frequency modulates a radio carrier. An unbalance of the resistance bridge changes the frequency of the subcarrier in proportion. Thus, the frequency of the subcarrier represents the quantity being measured; strain, temperature, pressure, etc. The FM/FM radio carrier is then radiated from an antenna. At the receiving station, an antenna receives the signal and feeds it to a radio receiver which



(*Courtesy of Electronetics Corp.*)
Fig. 1—Two-channel telemetry transmitter.

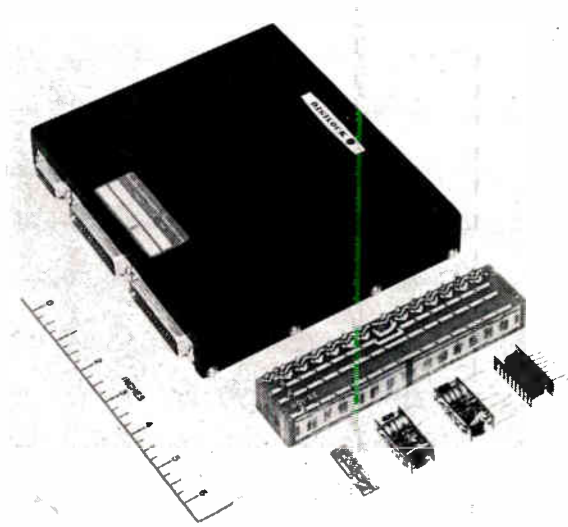
amplifies and demodulates it to reproduce the subcarrier. The subcarrier is fed to a signal converter which provides an electrical output proportional to the unbalance of the resistance bridge. When more channels are to be telemetered, additional subcarriers are added to modulate the same radio frequency carrier. Simple filter networks are usually added at the transmitter to isolate one subcarrier oscillator from another and at the receiver to separate the channels into corresponding discriminators which in turn provide the voltage analog of the information to recorders. In Fig. 2 is shown a two-channel receiving station. For multiplexing a larger number of channels a mechanical commutator such as shown in Fig. 3 is often employed.

Another system of telemetry finding increasing use is pulse-code modulation telemetry. This is a time multiplex telemetry system which applies analog electrical signals sequentially to an analog-to-digital converter. In some cases the sensors themselves may provide digital outputs. If they do, their outputs are multiplexed with the output of the converter to provide a sequence of voltage pulses whose presence and absence represent the encoded data. This "pulse train" controls a radio transmitter to convey the information through space. In Fig. 4 is shown an advanced type of digital telemeter which employs an error-correcting code.

Fig. 5 shows a digital telemetering receiving station. At the receiving station the code may be restored to

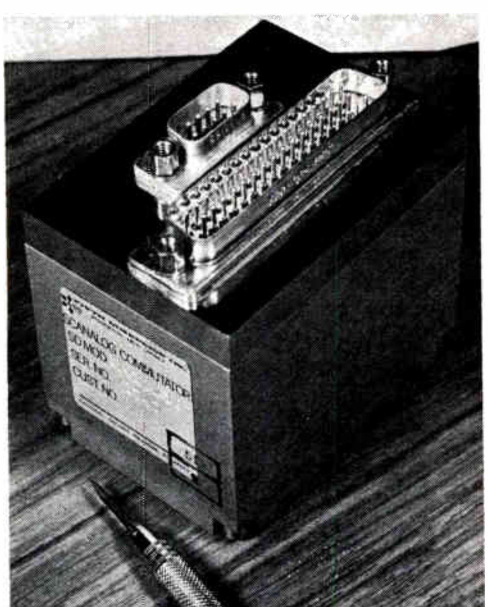
(*Courtesy of Electronetics Corp.*)

Fig. 2—Two-channel industrial telemetry receiving station.



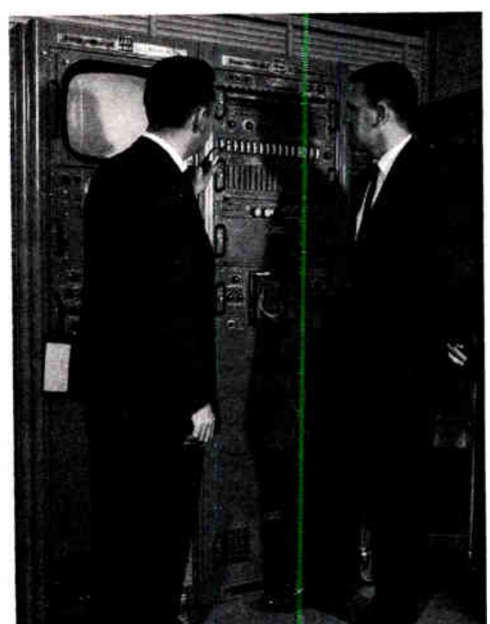
(*Courtesy of Space General Corp.*)

Fig. 4—Digilock telemetry transmitter.



(*Courtesy of Fifth Dimension, Inc.*)

Fig. 3—Modern mechanical commutator.



(*Courtesy of Electromechanical Research Corp.*)

Fig. 5—Pulse-code modulation receiving station.

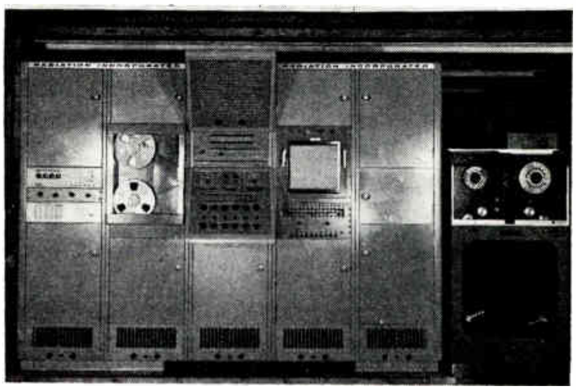
analog form for real time viewing. In most cases, however, it remains in digital form, feeding digital printers, plotters and computers. Then, after it is processed it is viewed as graphs or numbers. Very often the data processing equipment is the largest portion of the "telemetry" system. A typical digital data processing system is shown in Fig. 6.

Remote control systems have also employed frequency-modulated subcarriers to frequency modulate a carrier. The subcarriers are often for the purpose of security rather than for simultaneous controls, particularly in missile range safety equipment. Several subcarriers are utilized concurrently as a code providing the impulse to destroy the vehicle. Remote control systems also utilize pulse-code modulation. In the Telebit system as many as 64 commands to satellite functions are provided by binary digital code transmission. Other remote control systems include radar beam riders, automatic navigation on hyperbolic grids from fixed ground stations and inertial guidance control systems.

The advantages of combining tracking, telemetry and remote control have also been recognized. In the Telebit system these three functions have been combined to provide Doppler tracking in a two-way radio phase locked loop with the earth-to-satellite portion of the loop being utilized to convey command controls and the satellite-to-ground portion for telemetry.

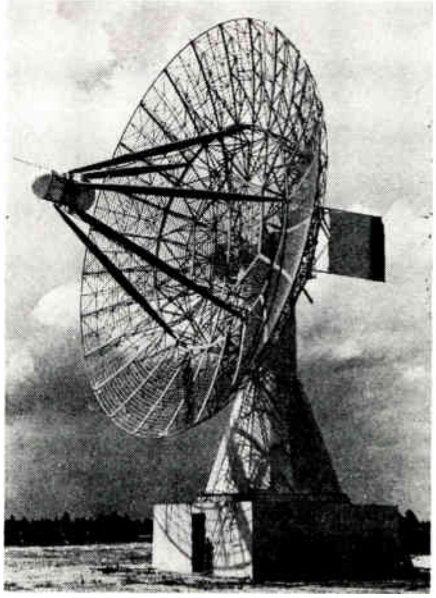
During the development of telemetry and remote control, the impact of generally improving electronic technology was felt. With the advent of transistors and other solid-state electronic components size, weight and power consumption of missile born equipments decreased and reliability increased making it possible to perform many functions of telemetry and remote control that were previously not feasible. The development of large automatic tracking ground antennas (Fig. 7) permitted further reductions in size, weight and power consumption in the vehicle. Consequently the usefulness of telemetry and control were greatly expanded to permit many different measurements to be made in missiles and satellites. These include measurements of the performance of servo systems, remote control and guidance systems, the internal performance of the vehicle, the structure, aerodynamics and propulsion, temperatures, pressures, vibration, external environment such as radiation intensity, magnetic fields, charged particles, micrometeorites and solar light intensities. In order to convert these measurements into electrical analogs capable of being transmitted by radio, it was necessary to develop many new and different transducers. These appeared very quickly and soon vast amounts of data were being assimilated by telemetry receivers and the problem of data handling became severe. Many measurements had to be studied, analyzed and reduced for days after the data was received. Analysis of vibration is an example of such a procedure.

As a consequence, telemetry receiving stations, recording data photographically, oscillographically and by means of pen recorders, grew tremendously in size and complexity as required to record the many concurrent channels of high-frequency data. Certainly then, the appearance of an adequate recorder, the magnetic tape recorder, which replaced the other recorders in the telemetry receiving station, can be hailed as a milestone of progress. With the advent of the instrumentation tape recorder, real time observations and data analysis gradually diminished to be replaced by electrical recording with post-flight playback. A typical telemetering station now consists of an antenna, a receiver and a tape recorder. Recording in the vehicle of electrical data by tape recorder also became feasible. The task of adapting the magnetic tape recorder to telemetry data was not easy, however. Neither its frequency response nor its flutter characteristics were adequate for the task. It was necessary to extend frequency response from 10 kc to 100 kc; tape speed variations (flutter) had to be reduced



(*Courtesy of Radiation, Inc.*)

Fig. 6—Digital data processing system.



(*Courtesy of Radiation, Inc.*)

Fig. 7—Sixty-foot automatic tracking telemetry antenna.

to less than 0.1 per cent and multiple channel capability had to be developed. The fact that these problems have been solved and that satisfactory tape recorders have evolved is a major advance in the telemetry industry. The magnetic tape recorder has also been useful in the remote control field, recording electrical signals from a control system in actual flight and then applying them to a similar vehicle in the laboratory to study its behavior.

Once the ability to assimilate large quantities of data by electrical recording had been developed, the vast possibilities in data processing were soon realized. By far the largest use of telemetry has been in the post-flight analysis of data from internal instruments. Data taken by the telemetry is used for: 1) determining the performance of the vehicle, 2) gathering data for improving the design of the vehicle, 3) determining causes of failure, 4) determining the time occurrence of events and 5) measuring the basic properties of the upper atmosphere and the space beyond the atmosphere. In many cases redundant, unwanted and useless data are taken. It becomes necessary to sort these from the wanted data and the process called "data reduction" evolves. This term is also applied to the correlation of telemetry data with data taken by other means and the computations which may be done to calibrate, convert, smooth, interpolate or extrapolate data. For these computations, since data is already in electrical form, it is possible to enter it directly into digital computers.

In general, measurements made in a test vehicle are recorded as a function of time. Then, on the basis of time comparison, telemetry data may be compared with tracking data and other measurements. In the post-flight computations the parameter time is usually eliminated and the internal measurements are plotted against altitude, velocity or acceleration. Graphic plots are usually acceptable but on occasions numeric prints of the data are required. Special purpose and general purpose computers, both analog and digital, have been employed to perform these functions.

The use of automatic equipment for performing the data processing as described above may be considered another milestone of progress. Automatic equipment is now used to: 1) correlate data, 2) calibrate data, 3) provide numeric printed data, 4) provide graphic plots of data, and 5) perform extensive computation on data. General purpose digital computers with magnetic tape input and output capability are finding extensive use in telemetry programs.

While the art of recording and processing data automatically was developing in the telemetry field, a companion development took place in the field of remote control. This was a transfer of basic concept from remote command control to automatic guidance. The human element and its delay in making decisions has been gradually removed and navigation data transformed directly into guidance and control. A concurrent multivehicle capability soon became practicable and guidance systems employing the principles of: 1) hyperbolic navigation, 2) radio beam riding, 3) target seeking, 4) celestial navigation and 5) inertial navigation have all been outstanding technological developments.

Another ambitious program just now coming into being combines telemetry, remote control and automatic computers for the high speed and automatic testing of aircraft, missiles and spacecraft. A test may be pre-programmed into an automatic computer and the vehicle under test commanded through remote control systems by the computer. Data returned by the telemetry link is used by the computer to verify or modify the test program. In this manner the number of vehicles required to gather sufficient data may be reduced from that required when slower manual methods are employed. Strangely enough it is the general purpose automatic digital computer which fails short in the accomplishment of a complete computer commanded closed telemetry-remote control loop. Its processing speed is not adequate for the programs which are involved. In one current plan, two of the largest computers are multiplexed in order to provide sufficient processing speed.

Still more recent is another trend soon to become a milestone in telemetry and remote control, *viz.*, its application to industry. The techniques perfected for aircraft, missiles and rockets are finding many applications in industry. It is only necessary to apply engineering ingenuity to the old principles to produce long-lived, adjustment free, reliable remote control and measurement devices. Among a few of the applications are: 1) measuring strains on rotating equipment, 2) measuring vibrations, currents and temperatures on high-voltage electric transmission lines, 3) measuring strains in submarine cables, 4) detecting and removing flaws in steel during production, 5) field testing of vehicles such as tractors, cranes, dredges, etc., and 6) biological and medical research. While these applications are only now beginning, they may be expected to outrun the military applications of telemetry and remote control.

Electronics in Planning Space Flights*

ERNST STUHLINGER†

Summary—The phenomenally rapid growth of space flight has placed stringent demands upon the electronics engineer to develop instrumentation of modest size, weight, power consumption and heat dissipation, but of maximum sensitivity, stability, lifetime and reliability. Space flight is creating situations for which no well-developed precedents exist, thus requiring the evolution of new techniques or new basic ideas. Attitude control equipment of space craft will utilize infrared sensors, or the earth-magnetic field, or a radio beam from earth. Cryogenics will be applied to create low operating temperatures. Communications must be maintained by line-of-sight links; the NASA Deep Space Instrumentation Facility will be capable of serving space shots planned for the next few years. Communication systems can be substantially improved by using quantum-electronic devices and light-weight electric power supplies. Telemeter problems include the enormous amount of data which must be transmitted from giant launch vehicles, and the transmission from a space probe of a maximum of data over large distances with a minimum of power. The Saturn C-1 and Mariner telemeter systems are described to illustrate these problems. Space operations such as orbital rendezvous, orbital assembly and launch, approach and descent to the moon and planets will require automated equipment to an extent not previously faced on earth. Electric power supplies with long life, low weight and high reliability are required for the electric and electronic devices on board a space craft; solar and nuclear power appear most attractive to satisfy the requirements. The success of our future space programs depends heavily upon the accomplishments of the electronics engineers today.

Introduction

ROCKETRY BEGAN almost a thousand years ago when Chinese warriors fastened powder-filled pipes to their spears in order to obtain greater throwing ranges. The door to rocket-powered space flight opened only twenty years ago; it began when the first electric instruments were successfully used to force a rocket precisely on a preconceived trajectory. From that time on, evolution assumed a breathtaking pace. Fifteen years after a V-2 accomplished the first fully guided rocket flight, Sputnik I orbited our Earth as an artificial satellite. One year later, in 1958, the Van Allen radiation belts were discovered by Explorer satellites. Another year later, Lunik III photographed the rear side of the moon. In 1960, the Tiros satellite provided more than 22,000 television pictures of the earth's surface from an orbital altitude of 400 miles. During the same year, Pioneer V established a record in communications by transmitting signals from a distance of 22.5 million miles. Our Earth, if seen from that point, would not look much larger than the planet Venus as seen from here. Again one year later, and only 3½ years after the first satellite began to orbit around the globe, Major Gagarin achieved his historic satellite flight. Instrumented probes to Mars and Venus, and deep into interplanetary space, are imminent. Our country is presently at the doorstep of its greatest technical project of all times: to land a man on the moon, and to bring him back safely.

An adequate propulsion system is a mandatory requirement of a space vehicle, but a complex system for control and communications is as indispensable as are the rocket motors. Precise control must be exercised over the performance and the attitude of the vehicle; data must be transmitted to ground stations, and commands must be received and carried out onboard the craft; guidance of the vehicle along the desired trajectory must be accomplished; and finally, each space vehicle has the mission of taking scientific measurements, and of radioing its observations back to earth. Planning for this very elaborate system of guidance, control, and data transmission of a spacecraft is at least as demanding as the planning for its propulsion system. It is of utmost importance that both areas of endeavor find equal emphasis. The National Aeronautics and Space Administration (NASA), being deeply engulfed in the development of spacecraft for an ambitious program of space exploration, therefore pursues very vigorously the principle of concurrency between the development of propulsion and the development of systems for control, guidance, measurements, and communication. Needless to say that the responsibility for these systems rests with the electronics engineer, as the responsibility for propulsion rests with the propulsion engineer.

In planning for space flights, the electronics engineer is confronted with a set of basic requirements which has become almost standard: his systems must be very modest with respect to size, weight, power consumption, and heat dissipation, but they must give a maximum of sensitivity, stability, lifetime, reliability, and resistance against the encumbrances of outer space. The telemeter system which helped to develop the V-2 missile was a 7-channel, amplitude-modulated monster as big as a medium-sized coffin. Two men were required to carry it around. Today, a telemeter system of that capacity, and for that range, is not much bigger than a match box. Thermionic tubes have been supplemented by transistors, ample use is made of printed circuits, and the entire structure is neatly embedded in plastic (Figs. 1 and 2). While the countdown during a V-2 launching became traditionally stalled at "X minus telemeter," the telemeter systems of modern space vehicles are remarkably absent in failure analyses. The solar-powered transmitter of Vanguard I is still transmitting after three years of operation; Explorer VII, after more than two years in orbit, still sends down its

* Received by the IRE, July 31, 1961; revised manuscript received, October 20, 1961.

† Research Projects Division, George C. Marshall Space Flight Center, Huntsville, Ala.

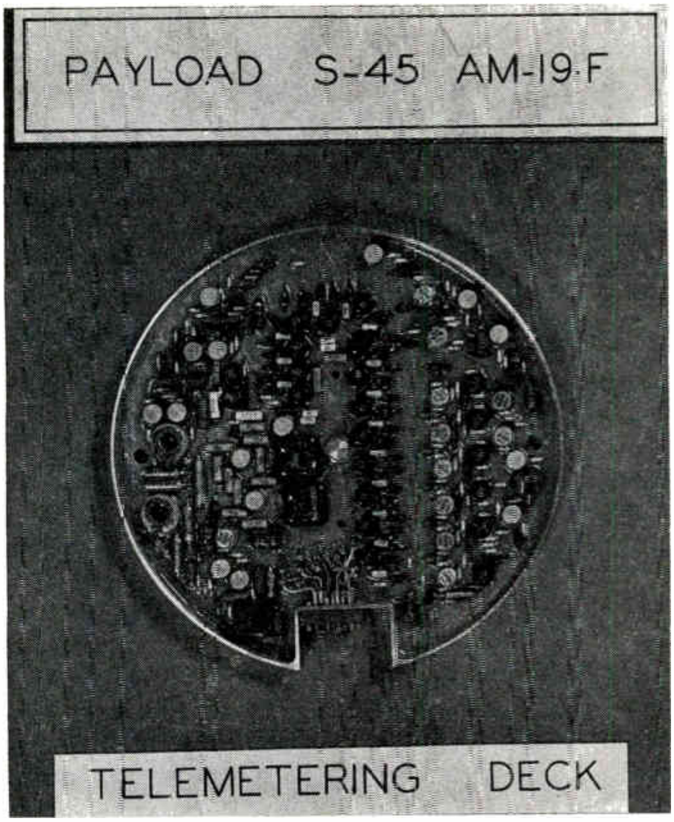


Fig. 1—Telemeter of Juno II satellite (George C. Marshall Space Flight Center, NASA).

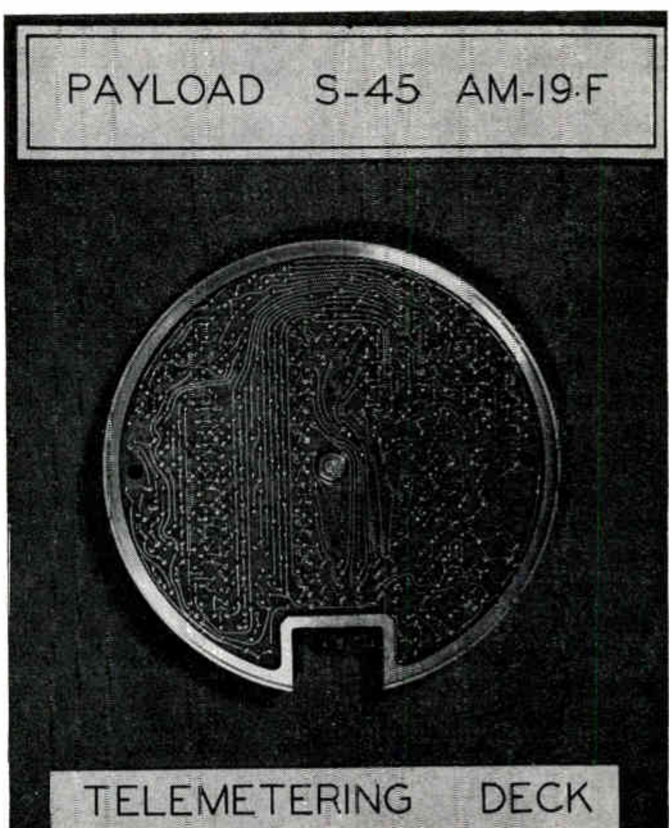


Fig. 2—Printed telemeter circuit of Juno II satellite (George C. Marshall Space Flight Center, NASA).

telemeter signals. Pioneer V was at a distance of almost 23 million miles from the Earth when its batteries were exhausted; up to that time, a 60-ft antenna (28-db gain) received strong signals from the 160-w transmitter, and the 250-ft antenna at Jodrell Bank recorded even signals from the 5-w transmitter over that range. These accomplishments prove that the experimental techniques for these more conventional functions are satisfactorily in hand, and that the limits to our capabilities, at least for present-day projects, are imposed by laws of physics rather than by poor engineering. However, the art of space flight is rapidly approaching situations for which no well-developed precedents exist. They require the evolution of new techniques, and in some instances even the breakthrough of new basic ideas.

Spacecraft Control

Control functions onboard a space probe are required for the propulsion system, for the vehicle attitude, and for the temperature of sensitive equipment. Control of the propulsion system may be achieved by command signals from the ground, or by signals generated by distance or direction sensors. The latter are of particular importance when the craft is to make a slow descent to the surface of the moon or one of the planets. In that case, the distance to the surface will be measured continuously by radar methods, and the propulsion system will be operated according to these measurements. This scheme has been studied in detail for a lunar landing craft in 1959,[1] and more recently as part of the Surveyor Project.[2] The attitude of a space vehicle must be controlled for a number of reasons: to apply thrust in the desired direction, to have narrow-beam antennas or ground observation equipment point in the right direction, to maintain a desired attitude with respect to the sun, or to land on a predetermined point of the lunar or a planetary surface. While close to the surface of a planet or the moon, a space vehicle may use horizon sensors to control its attitude with respect to that surface. Infrared seekers are preferable to seekers using visible light because the difference between day and night signals is less for long-wave infrared than for the visible wavelengths. In typical flight tests, the "black body" temperature of the earth in the 2- to 15-μ spectral region was measured to 250–270°K;[3] the temperature of space may be assumed to be close to 0°K. The most sensitive sensor materials are lead sulfide, lead telluride, and lead selenide. Their response curves are shown in Fig. 3. Present-day seekers employ thermistors of less sensitivity, but

[1] "A Lunar Exploration Program Based Upon Saturn-Boosted Systems," Army Ballistic Missile Agency, AOMC, Redstone Arsenal (now George C. Marshall Space Flight Center), Huntsville, Ala., Rept. No. DV-TR-2-60; February 1, 1960.

[2] "Space Programs Summary," Jet Propulsion Laboratory, Pasadena, Calif., Rept. No. 37–8, vol. 1; April 1, 1961.

[3] R. J. Bordeau, "Horizon Sensor Flight Test Results," George C. Marshall Space Flight Center, NASA, Huntsville, Ala., Rept. No. DG-TN-17-60; May 2, 1960.

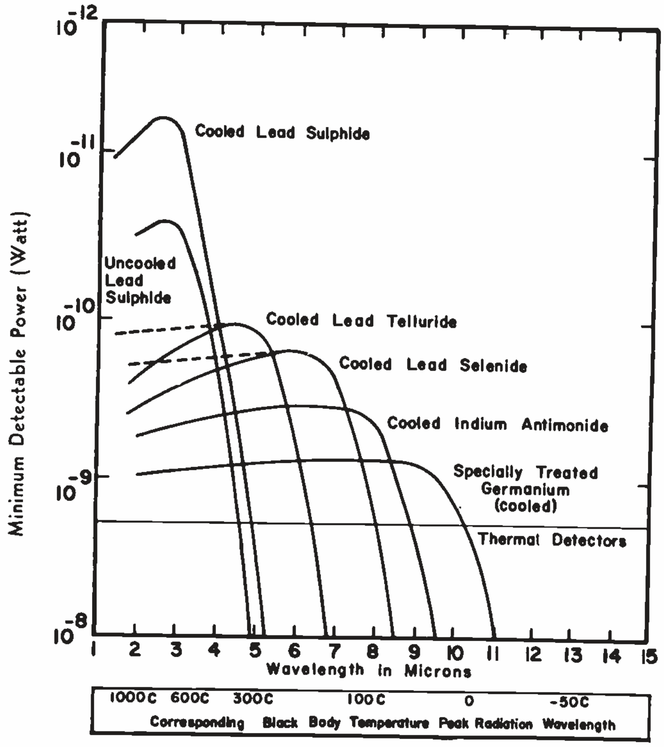


Fig. 3—Response curves of various infrared detectors.[4]

with a flat response curve over the entire spectral range. Germanium, transparent in the 2- to 16-μ region, and with a refractive index of four, is ideally suited for focusing lenses.[4] Attitude-control systems using infrared seekers are installed in Discoverer, Midas, Tiros III, Ranger, and other spacecraft. They maintain the desired earth-oriented attitude within an error angle better than $\pm 1°$. If the infrared-sensitive cells are cooled to the temperature of liquid nitrogen, their inherent noise level decreases considerably, and higher sensitivity is achieved.

An interesting method of controlling the attitude of a satellite evolved from Project Tiros.[5] It is possible to program the current in a satellite-borne coil in such a way that almost any attitude of a specified satellite axis can be obtained through interaction of the current-carrying coil with the local earth-magnetic field. It is also possible to use two permanent magnets instead of the coil. One of the magnets must be pivoted according to a control program.

Earth-oriented attitude control is of particular importance for satellites whose missions are directly related to the Earth.[6] Weather satellites, for example, should orient their cameras with reference to the center of the Earth, and the same is true for communication satellites. A 24-hour equatorial satellite, hovering permanently above the same spot on the ground, will probably take its bearing from a radio beam transmitted from the ground. If solar cell panels are used as power supply, they must remain oriented toward the sun.

Space vehicles traveling beyond satellite orbits on coasting trajectories will keep a specified orientation with respect to the sun in order to keep thermal design problems to a minimum. If thrust maneuvers for midcourse guidance should be necessary, the vehicle would either use an earth-based radio beam, or a star-fixed coordinate system, as a reference. The latter requires two seekers that maintain their directions toward two fixed stars approximately 90° apart.

A star-fixed reference system will be the basis of an inertial space guidance system. The guidance platform may carry gyroscopes to facilitate its control, although they would not be mandatory. It will carry accelerometers to measure thrust forces in three orthogonal directions. Besides the star seekers, several planet seekers will continuously measure the directions toward planets with respect to the star-fixed coordinate system. These directions, if referred to the known positions of the planets within the solar system at any specified time, will immediately give the location of the space vehicle at that time.[7] Strictly speaking, only three planets would be necessary to determine the location in three-dimensional space. However, one or more of the planets may be hidden temporarily behind the sun or the moon, or two of them may be in line with each other; for that reason, four or five planet seekers will be used. The location coordinates will be obtained from the angle measurements with the aid of computers which contain the natural motion of the planets, and the time, as preset programs. A two-dimensional diagram of a vehicle in space, determining its location with a system of star and planet seekers, is shown in Fig. 4.

Once a spacecraft approaches its target—the moon or a planet, or, on a return trip, the Earth—it will use infrared seekers to find the exact direction of the target, and to adjust its attitude. The distance D to a target of diameter $2R$ will be obtained from a simple measurement of the angle α subtended by the target. This measurement, as illustrated in Fig. 5, will give an accuracy below one per cent for distances from the surface between 0.001 and about 80 times the radius of the target, if the angular accuracy of the sensors is ± 1 minute of arc.

An electronics engineer responsible for the control, the guidance, and the communication system of a spacecraft will find that the principal components of these diverse systems share one common feature: they work better at low temperatures. Infrared seekers have a much lower noise figure, and therefore a much greater sensitivity, at cryogenic temperatures. Gyroscopes operate with almost vanishing power requirements if

[4] R. DeWaard, "New developments in thermistor infrared detectors," *Proc. Infrared Information Symp. (IRIS)*, vol. 3, pp. 39–43; March, 1958.

[5] L. H. Grasshoff, "A method for controlling the attitude of a spin-stabilized satellite," *ARS J.*, vol. 31, pp. 646–649; May, 1961.

[6] W. Haeussermann, "An attitude control system for space vehicles," *ARS J.*, vol. 29, pp. 203–207; March, 1959.

[7] E. Stuhlinger, "The flight path of an electrically propelled space ship," *Jet Propulsion*, vol. 27, pp. 410–414; April, 1957.

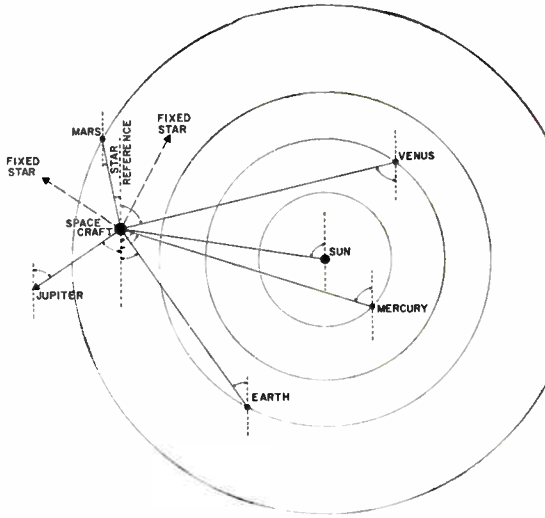


Fig. 4—Two-dimensional schematic of star-fixed guidance system with planet seekers.[7]

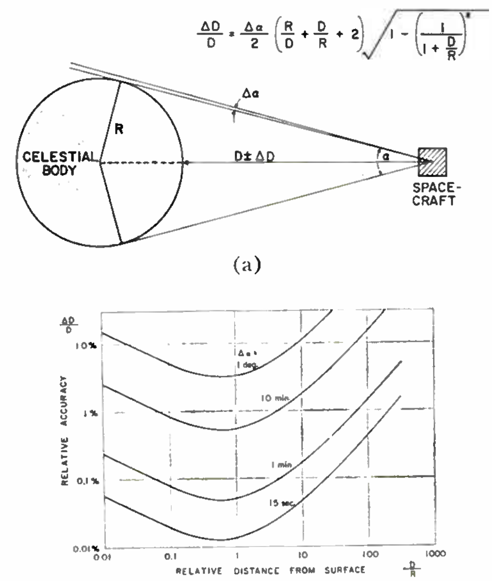


Fig. 5—(a) Distance measurement with angular sensors. (b) Accuracy of distance measurement as a function of distance; angular accuracy of sensors as parameter.

superconductivity techniques can be applied.[8] The same holds for accelerometers and integrators, and also for reaction wheels. Highly efficient binary computer elements have been developed on the superconductivity principle.[9] The noise level of semiconductor amplifiers can be greatly reduced if cooling techniques are introduced. Of particular success is the cooling principle in connection with parametric amplifiers using cavities, and with maser-type circuits. The space vehicle of the future will therefore be equipped with a "cryopackage" inside its guidance and control compartment. The cryopackage will house the gyrostabilized platform with accelerometers, infrared horizon seekers, star seekers, planet seekers, and radar altimeter, and also the low-temperature components of digital computers, programmers, command receivers, radio communication equipment, and television storage and transmitting facilities. The package will be kept at low temperature by means of a cryostat which may be powered by solar energy, or by a nuclear-electric power supply, if this should be carried onboard for other purposes. It is anticipated that the total heat power generated within this cryopackage will not be more than 50 to 100 w, an amount which can be handled easily even by a small cryostat.

Spaceships will keep at least their attitude-control system and their complex communication links active during the entire flight. Spacecraft of the low-acceleration type, like electrically propelled vehicles, will be subject to active guidance throughout their interplanetary trip. In both cases, the cryopackage will be the heart and the brain behind attitude control, launch guidance, mid-course guidance, approach guidance, digital computers, thrust maneuvers, radio communication, data transmission, and even some of the scientific measurements to be made in space and on other celestial bodies. Development, design, and building of such a cryopackage stands out as one of the intriguing challenges that beckon the creative mind of the space electronics engineer today.

Communication Problems

Unmanned space flight would be without value unless means of communication existed that keep the ground team informed of the spacecraft's trajectory, and of technical and scientific data which the vehicle gathers during its flight. Besides, the ground team will desire to control certain functions of the craft by radio control. Even manned space vehicles, which may be completely self-dependent for the duration of their space trip, will maintain their communications with the Earth on a continuous basis.

The electronics engineer entrusted with the planning of all electronic systems of a major space project will leave the details of this gigantic task to specialists. However, he will make a careful assessment of the principal capabilities and limitations of space electronic systems to set the cornerstones in his master plan. First, he must accept the finite velocity of signals transmitted through empty space, 3×10^{10} cm sec^{-1}. Table I lists the one-way times of transmission between the earth and other celestial bodies. When an operator on the earth sends a command to the moon immediately upon receiving a signal from there, the delay between signal generation and command reception on the moon will be almost three seconds. Obviously, the control of lunar operations from the earth therefore can never be as intimate as the control of earthbound operations. In the case of Mars,

[8] T. Buchhold, "The magnetic forces on superconductors and their application for magnetic bearings," *Cryogenics*, vol. 1, pp. 203–211; June, 1961.
W. Haeussermann, "Recent advances in attitude control of space vehicles," *ARS J.*; January, 1962.
[9] R. J. Allen, "Application of cryogenics to electrical and electronic design," *Elec. Engr.*, vol. 79, pp. 1013–1018; December, 1960.

TABLE I

ONE-WAY TRAVEL TIME OF ELECTROMAGNETIC RADIATION BETWEEN EARTH AND OTHER CELESTIAL BODIES

Target	Distance from Sun		Shortest Distance From Earth km	Travel Time		
	10^6 Miles	10^6 km		Seconds	Minutes	Hours
Sun	0	0	1.49×10^8	497	8.3	
Mercury	35.9	57.7	9.1×10^7	303	5.0	
Venus	67.2	108.0	4.1×10^7	137	2.3	
Earth	92.9	149.3	—	—	—	
Mars	141.5	228.0	7.9×10^7	263	4.4	
Jupiter	483.5	776	6.27×10^8	2.09×10^3	34.9	0.6
Saturn	886.2	1424	1.27×10^9	4.23×10^3	71	1.2
Uranus	1783	2870	2.72×10^9	9.07×10^3	151	2.5
Neptune	2793	4500	4.35×10^9	1.45×10^4	242	4.02
Pluto	3676	5905	5.75×10^9	1.92×10^4	320	5.3
Moon			3.5×10^5	1.17		
α-Centauri			4.03×10^{13}	4.3 Years		

the delay time amounts to nine minutes even when Mars and Earth are closest. Direct control of Martian operations by television and command links from Earth would be hopeless. On the other hand, the initiation of programmed mid-course guidance functions, or of a surface operation program, may well be handled by commands from the Earth.

Line-of-sight will be essential in all space communications problems. If it should be interrupted temporarily by intervention of a celestial body, data storage must be applied. A relatively small number of stations on the earth, if well coordinated, will be sufficient to maintain line-of-sight contact with almost any space vehicle. At present, the NASA Deep Space Instrumentation Facility (NASA-DSIF) has a network of three fixed stations with 85-ft parabolas[10] in operation: Goldstone in California, Woomera in Australia, and Johannesburg in South Africa. Each of these stations can measure angles with an accuracy of the order of $\frac{1}{10}$ to $\frac{1}{20}$ of 1°. Since deep space probes, at least within the next few years, will remain relatively close to the plane of the ecliptic, the NASA-DSIF will be capable of serving all space programs which are planned for the foreseeable future. This is true even for electrically propelled spacecraft whose trajectories may form angles of 25 or 30° with the plane of the ecliptic.[11]

[10] R. J. Parks, "The U. S. planetary exploration program," *Astronautics*, vol. 6, pp. 22–25; May, 1961.

[11] J. W. Stearns, Jr., "Applications for Electric Propulsion Systems," Jet Propulsion Lab., Pasadena, Calif., Rept. No. 33–47; April 10, 1961. Committee Report, "Utilization of Electric Propulsion in Spacecraft," Jet Propulsion Lab., Pasadena, Calif., Rept. No. 33–21; October 1, 1960.

The power required to transmit a signal over a given distance, the optimum choice of carrier frequency, the influence of bandwidth and noise, and the capabilities of modern receivers are discussed elsewhere in this volume. The following formula[12] may be useful for a quick determination of the range of a transmitter-receiver link, assuming a signal-to-noise ratio of about ten:

$$D = \tfrac{1}{3}\sqrt{W_T A_T G_R / 4\pi k T \Delta f}. \qquad (1)$$

A range of $D=6\times10^9$ km, equal to the distance between Earth and Pluto, could be obtained with the following figures:

$W_T=200$ w
$A_T=1$ m² (parabola)
$G_R=45$ db (Goldstone, 85-ft parabola)
$k=1.38\times10^{-23}$ w sec deg^{-1}
$T=100$°K (parametric amplifier)
$\Delta f=10$ cps.

Eq. (1) indicates that substantial improvements of space communication systems will come from the use of space-borne masers, and from light-weight electric power supplies. Both subjects present most fruitful research and development areas for space electronics engineers.

TELEMETER DESIGNS

Telemeter technologies have been developed to a remarkably high standard in connection with space

[12] E. Rechtin, "Deep space communications," *Astronautics*, vol. 6, pp. 37–46; April, 1961.

projects of the past years. Two distinct problem areas face the electronics engineer who indulges in present-day projects; both deserve his particular attention and his sincere effort. The first concerns data transmission from such giants as the Saturn vehicle. The problem lies simply in the enormous number of independent data that must be transmitted. The second refers to small, but far-reaching probes with scientific missions. Here, the problem consists of transmitting a maximum of information over a large distance with a minimum of power. Instead of discussing these problem areas in general terms, a short description will be given of the Saturn C-1 telemeter system as used in the single-stage test firings, and of the Mariner system which will transmit telemeter signals over a distance of 6×10^7 km.

The telemeter system which has been developed for the Saturn booster flight tests at the Marshall Space Flight Center (former Army Ballistic Missile Agency) is specifically adapted to the accuracy and response requirements of Saturn booster measurements.[13] Most of these measurements involve engineering data such as flow rates, vibrations, temperatures, and functional procedures. Consequently, some of the telemeter channels must have a frequency response up to about 3000 cps, while others will carry data that vary only at a low rate. Some channels must provide continuous information, others transmit intermittent data at sampling rates not more than a few points per second. The desired accuracy varies with the type of information.

On the basis of these requirements, four different telemeter types were developed and combined into a complex data transmission system. The first uses a single-sideband modulation of AM subcarriers which frequency modulate the carrier (SS-FM); a baseband utilization efficiency of 60 per cent is obtained. The second is a 216-channel PAM-FM-FM system for LF response. Each of 27 submultiplexers has ten channels; eight for one specific measurement on each of the eight engines, and two for identification and voltage reference. The main multiplexer samples each of the 27 submultiplexers in sequence. The third type is an FM-FM system with 15 continuous channels, and also 54 multiplexed channels applied to two subcarriers. The fourth kind consists of the previous type, but modified by seven additional sub-subcarriers that modulate one of the subcarrier channels. This modification is applied to three identical telemeter packages. A schematic sketch of the complete Saturn telemeter system is shown in Fig. 6. The system will weigh 220 pounds. Total input power is about 1600 w, most of it to be consumed by the RF transmitters, and by power amplifiers. The total number of information channels of this system is 546.

[13] J. E. Rorex, "Evolution of the Saturn Booster Telemetry System," George C. Marshall Space Flight Center, NASA, Huntsville, Ala., Rept. No. DG-TN-11-60; March 21, 1960. W. O. Frost and O. B. King, "SS-FM: a frequency division telemetry system with high data capacity," IRE Proc. 1959 National Symp. on Space Electronics and Telemetry, Sect. 7.2, September, 1959.

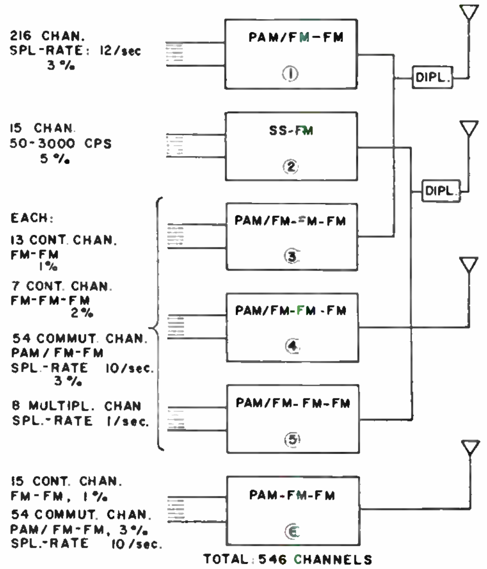


Fig. 6—Schematic diagram of Saturn booster telemetry system[13] (accuracy in per cent).

The Mariner spacecraft, under technical direction of the Jet Propulsion Laboratory,[14] will carry a directed parabola of 1.22-m diameter with 18.8-db gain at a frequency of 960 Mc, and of 26.0-db gain at 2295 Mc. The full cone angle of the antenna beam is 16°; a conical helix provides a circularly polarized beam. The receiving antenna (Goldstone) has a gain of 45 db for the lower, and of 52 db for the higher frequency. The lower frequency will be used until January 1, 1963. After that time, the DSIF stations will operate only at the higher frequency.

The subcarriers transmit the information in digital form. Modulation consists of switching a binary-symmetric channel between two possible states. The two states are represented by two phase conditions of the subcarrier frequency, such as $\phi_0+90°$ and $\phi_0-90°$. This "phaseshift-keying" represents the most efficient digital method of data transmission. A transmitter power of 3 w permits 12.8 bits per second over a distance of 6×10^7 km. The information to be transmitted includes about 90 engineering measurements of the Mariner systems

[14] B. D. Martin, "The Mariner Planetary Communication System Design," Jet Propulsion Lab., Pasadena, Calif., Rept. No. TR 32-85; May 15, 1961.

J. P. Feary, *et al.*, "Radio Tracking Techniques and Performance of the United States Deep Space Instrumentation Facility," presented at the "COSPAR" Symposium, Florence, Italy; April 10–14, 1961.

M. H. Brockman, *et al.*, "Extra-terrestrial radio tracking and communication," Proc. IRE, vol. 48, pp. 643–654; April, 1960.

H. D. Becker and J. G. Lawton, "Theoretical Comparison of Binary Data Transmission Systems," Cornell Aeronautical Lab., Buffalo, N. Y., Rept. No. CA-1172-S-1; May, 1958.

H. Richter, R. Stevens, and W. Sampson, "Microlock: a minimum weight radio instrumentation system for a satellite," *Jet Propulsion*, vol. 28, pp. 532–540; August, 1958.

R. Jaffe and E. Rechtin, "Design and performance of phase-lock circuits capable of near-optimum performance over a wide range of input signal and noise levels," IRE Trans. on Information Theory, vol. IT-1, pp. 66–76; March, 1955.

performance, and about 15 scientific measurements of cosmic radiation, ion densities, cosmic dust, meteorites, magnetic fields, and solar radiations. A maximum of 10^6 bits of information can be stored during any 24-hour period for rapid playback upon command.

Automatic Space Operations

The rapid evolution of space projects confronts the electronics engineer with a requirement that has not existed in the past; it is that of fully automatic equipment for space operations like orbital rendezvous, orbital assembly and adjustment, orbital checkout and countdown, approach and descent to the surfaces of the moon and of planets, take-off from these surfaces, and scientific exploration on other celestial bodies. The number of technical problems associated with these operations is almost endless; their degree of sophistication is limited only by the capability of the engineer, and by the techniques at his disposal. The normal functions which constitute these space operations can be carried out by robot-type devices that act according to a preset program. However, functions such as trouble shooting, failure analysis, locating of faulty parts, repair or exchange of components, or recalibration of subsystems present situations which require judgement, and even decisions, that cannot be preprogrammed easily. Techniques to handle such situations have not been developed thus far. There is no doubt that they will be based at first on a combination of automatic devices, telemeter and television equipment, and man-controlled command signals from the ground. However, the trend will be to eliminate the link to the Earth as far as possible; consequently, the orbital manipulators must be equipped with a complex system of sensors, with a vast store of commands that can be stimulated in many different combinations by sensor signals, and with versatile organs that can carry out these intricate commands.

Problems of this kind involve a degree of automation which has never been even approached in the past. It is not possible to predict how successful our technology will finally be in replacing with electronic circuits the sensory organs, the memory, the capability of intelligent decisions, and the marvelous dexterity of the hands of a human astronaut. Requirements for complex logic procedures under space conditions are by no means restricted to projects of the far future. They are encountered, at least in preliminary form, by the planners of a manned lunar landing project that uses a satellite orbit around the Earth as staging area. The question of the exact role in orbital operations which should be assigned to the human space travelers will not be answered before the capabilities of manipulators, equipped with sensors, memories, computers, and logic circuits, can be assessed. Success in the development of this technology will be one of the most influential factors that determine the rate of progress in future space programs.

Electric Power Generation

The great variety of electric and electronic devices to be built for spacecraft have one feature in common: they need electric power. A supply of electric energy is needed by sensors, by amplifiers, by control and guidance equipment, by communication systems, and, within the near future, for electric propulsion systems. The development of reliable, long-life, light power supplies therefore belongs to the most urgent tasks of the space engineer. Existing batteries with a specific energy of 100 to 150 wh/kg (low discharge rate) are suited for short-time flights up to days or weeks. Solar converters have the advantage of infinite reliability of their prime source, but their effective specific power is not greater than about 0.01 kw/kg. Even thin-film photo-electric converters do not exceed this figure because of reduced conversion efficiency.[15] Nuclear fission energy is presently the most attractive prime source for power supplies in the multikilowatt range. The SNAP-8 project will result in a 60-kw supply with a specific power of 0.04 kw/kg. Larger supplies, also built on the steam turbine and rotating generator principle, may provide 0.3 to 0.5 kw/kg.[16] Much hope concentrates at present on thermionic converters, and particularly on plasma-electric generators in which the kinetic energy of a gas, heated in a fission reactor and seeded with alkali ions to make it conductive, is converted into electric energy by interaction with magnetic fields. Each of these types of converters may eventually generate 1 to 2 kw/kg. Thermonuclear fusion generators promise an even greater specific energy. However, their development is still in a very early state. Considerable progress in our basic knowledge of plasma physics is required, either by breakthrough or by a concentrated effort of many years, before a practical thermonuclear generator can be visualized.[17]

Outlook to the Future

Space flight began a few years ago as an art whose purpose was solely the expansion of our basic knowledge. Space engineers approached their tasks like pure scientists: without looking for immediate recompensation, they were guided by a desire for superior accomplishment. Less than four years after our first steps into space, we are witnessing today how a new element, and a very healthy one, begins to rise in our space effort: it is the practical use of space for weather and communication satellites. Both have been recognized as extremely valuable in the daily needs of modern, civilized nations. Both have passed the state of exploratory research, and both have developed into engineering

[15] H. I. Moss, "Large-area thin-film photovoltaic cells," *RCA Rev.* (*Energy Conversion*), vol. 22; pp. 29–37; March, 1961.

[16] P. J. Valentine, "A Survey of Energy Conversion Systems," Jet Propulsion Lab., Pasadena, Calif., Rept. No. 33–46; June 26, 1961.

[17] G. Warfield, "The present outlook for controlled nuclear fusion," *RCA Rev.* (*Energy Conversion*), vol. 22, pp. 122–130; March, 1961.

enterprises of highest order. The electronics engineers are primarily concerned with bandwidth utilization, with bits-per-second capability, with coding techniques, with read-out speed, with storage volume, with command capability, and with the achievement of longevity. The engineers are joined by lawyers who must concern themselves with problems such as company vs government interest, antitrust laws, proprietary rights, and competitive bidding, because a system of well-functioning communication satellites may easily develop into a multibillion business within a few years. It is evident that the foremost goal of the engineer must be the superior technical quality of his systems. The success of the orbiting communication system will depend almost entirely upon the reliability of its electronic equipment.

While a considerable part of our national space effort will branch out into projects of practical usefulness, a strong endeavor will continue to concentrate around space research. Existing instruments will be refined, and new ones will be invented. Quantum mechanical amplifiers, representing one of the most intriguing areas of modern electronics, have already proven their superior value for space instrumentation. With a maser-controlled transmitter on a spacecraft, and a similar transponder on another spacecraft or on the earth, it will be possible to measure the relative velocity between the two transmitters with an accuracy of centimeters per second.

The greatest challenge is always offered by the unknown. The probing spirit of the research scientist, matched by cleverly designed instruments, will find an unlimited pasture in outer space. The Van Allen belts, showing the strange ways in which our earth catches and stores charged particles thousands of kilometers above its surface, and dumps them toward the poles in beautiful auroras upon a solar disturbance; the hyper-energetic protons which are ejected by the sun from time to time in violent outbursts; the meteoritic dust particles which approach the earth after millions of years of interplanetary travel—these are our first glimpses into outer space. What will we find when we achieve the next step in the exploration of space, the landing on the moon? And what will our first instrumented probes tell us after descending gently to the surface of Mars? The answers will come in the future; they depend on what our electronics engineers accomplish today.

Data Processing and Information Transmission for Space*

ELLIOT L. GRUENBERG†, SENIOR MEMBER, IRE

Summary—**This paper reviews and predicts the roles that data processing and information transmission have played, and will play, in space. The trend toward digital transmission is shown to be a result of increasing requirements for precision, reliability, high information rates, and long paths. Thus, data processing and transmission techniques appear to be converging rapidly, both because of the advantages of digital communication systems and because of the predicted weight and volume reductions for computer equipment. Many improvements in transmission efficiency will result from this merging of techniques, notably in the areas of data compaction and computer-controlled adaptive communications.**

The problem of maintaining a satisfactory SNR in spite of the signal loss from space attenuation is reviewed. General trends in the techniques designed to improve this ratio are reported and several promising developments are emphasized. Among these are the use of signals with large bandwidth-time products, maser amplifiers, and a novel telemetry system which provides high antenna gain with a greatly reduced orientation problem and eliminates spacecraft transmitting equipment by transferring it to earth.

* Received by the IRE, September 27, 1961.
† Communications Systems Center, Federal Systems Division, International Business Machines Corp., Rockville, Md.

Introduction

SPACE ACTIVITIES are, in reality, an enormous array of experiments. Some of these experiments must be done in real time and some must be done separated by great distances. It is inconceivable that these experiments could have been and could continue to be made without reliance on the most modern methods of data processing and the most advanced methods of transmitting information.

The type of communications with which we are concerned here falls under the heading of telemetry, an art which has developed from the scheme for reading "meters" remotely over wired communications links. The term "telemetry" may be rightfully extended to include all communication of measurements; hence, message accuracy is a specific requirement. Because accuracy is a fundamental consideration, it is appropriate that the unit of information developed in information theory, the bit, be used to tie together the diverse areas involved. In a typical space mission the experiment must be planned, the trajectory determined, and the data gathered; and the required *information* must be extracted, transmitted, and reprocessed to present to the experimenter the reading of the meter in the form he requires. If, in the planning stages, the measure of effectiveness for the mission is recognized to be a function of the expected number of bits of information presented to the experimenter, each step can be evaluated according to its contribution to the total. Similarly, the completed mission can be evaluated in terms of the actual number of bits of information obtained, and the sources of inefficiency can be established.

Thus information theory has effected a marriage of the sciences of instrumentation, communication, and data processing. In addition, recognition of the bit as a common unit has led to the merging of the techniques of data processing and information transmission, and to a growing emphasis on digital methods of transmission.

The purpose of this paper will be to review briefly the important roles that data processing and information transmission play in man's entry into space, to predict future trends, and to suggest several ideas that might provide significant advances.

The Role of Data Processing in Space

Man's entry into space has depended upon a great deal of preparation and experimentation, much of which has required the use of high-speed data-processing equipment. Among the more prominent uses of ground-based equipment have been:

1) Design calculations for engine, structure, etc.
2) Static tests of rocket engines.
3) Stress analysis.
4) Flight simulation.
5) Determination of optimum trajectories.
6) Track prediction and orbital calculations.
7) Launch control and range safety.
8) Smoothing and reduction of remotely-sensed data.

Since these tasks have not imposed very specialized requirements on the processors, general-purpose computers have usually been employed. The basic trend in these activities seems to be toward higher speeds, with interest in real-time operation increasing. Examples of these ground-based activities abound. A few will now be mentioned.

In the Vanguard program, observations from the many Minitrack stations arranged as a "fence" around the globe were transmitted by teletype to a computing center in Washington, D. C. Orbits were computed for the satellites and ephemerides were established. One important result of these computations was the revelation of the earth as being "pear-shaped." This example illustrates the marriage of data processing and information transmission because continuous tracking would have been impossible without data transmitted from remote points on the globe. It also illustrates how data-processing techniques remove redundancy and noise from the enormous amount of data and provide the experimenter with the *information* he requires.

A significant example of the necessity for real-time data transmission and processing is Project Mercury. To achieve maximum safety the exact position and velocity of the Mercury capsule must be known at all times. To this end, a world-wide network comprising radar (FPS-16 and Vercort) and telemetry sites has been installed. To process the data into and out of the central computing facility, two types of communication paths have evolved. High-speed (1000 bps) communications lines are used for bringing impact prediction information from Cape Canaveral into the computers and for returning digital display and plotting information back to the Mercury Control Center to enable monitoring and control action to be taken. Low-speed inputs bring into the computers position information acquired at the radar sites. At the computing center two 7090 computers are used in parallel and are equipped with special data communication channels which enable real-time access by new data, as well as the processing of the necessary display and spacecraft acquisition information.

An example of an existing facility for reduction of space data is that located in Philadelphia Pa., at the Missile and Space Vehicle Department of General Electric Company. Programs have been written for this computing center to demultiplex data recorded from missile tests, to correct it for calibration, and to smooth it by digital filtering techniques.

It seems that the area of data reduction will probably see the largest increase in ground-based computer usage in the near future. Enormous amounts of data are accumulating daily from space vehicles that have already been launched. Consequently, a real need exists for computer programs that will not merely process these data by "brute force" techniques, but will minimize processing time by automatic redundancy removal and sophisticated statistical analyses. Such programs are the best way to provide the experimenters with *current* information in the most useful form.

With regard to spaceborne computers, the major role they have played in the past has been in guidance and flight control. One of the major requirements for this application has always been resistance to the considerable shock and vibration occurring during launch, the time of greatest activity of the computers. The trend has been toward digital processors for the increased accuracy they provide. The digital, stored-program guidance computer for Titan is shown in Fig. 1.

In the future one of the primary applications for spaceborne processors will be data compaction (or compression). There is an ever-increasing demand for reduction of the amount of data that must be telemetered to earth. Techniques for accomplishing this will now be discussed.

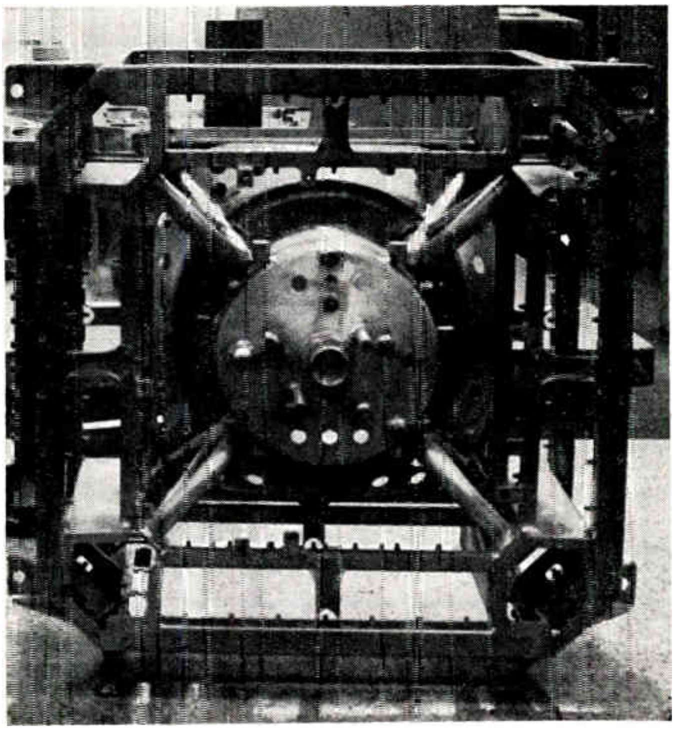


Fig. 1—The Titan guidance computer (IBM).

Data Compaction

Data compaction, or the reduction of the time or bandwidth required to transmit a message, can be conveniently described as either information-preserving or information-destroying. By information-preserving compaction we mean that the (compacted) data received by the experimenter on the ground must be sufficient to permit him to reconstruct exactly the sequence which caused these data.

Although both of these methods hold great promise for increasing the amount of information that can be obtained from a mission, most of the compaction techniques that have been implemented to date have been information-destroying. Examples of these are particle counters, peak detectors, and the spectrum analyzer proposed by Ratz [1]. In general, these devices can be said to extract those statistics of the input data that are of interest to the experimenter.

A generalized, information-preserving data compactor for use with a variety of input signals was recently proposed [2], [22]. One version of this device, which appears to be quite feasible for spaceborne telemetry, is shown in Fig. 2. This compactor exploits the fact that the probability of a change in a bit decreases as the "significance" of the bit increases (in a straight binary code). Run length coding is used because it requires only simple equipment to implement and because it has been shown by Elias [3] to approach the theoretical compaction limit closely.

One area that is now starting to receive more attention is the application to spaceborne equipment of some of the vast amount of research that has been done in

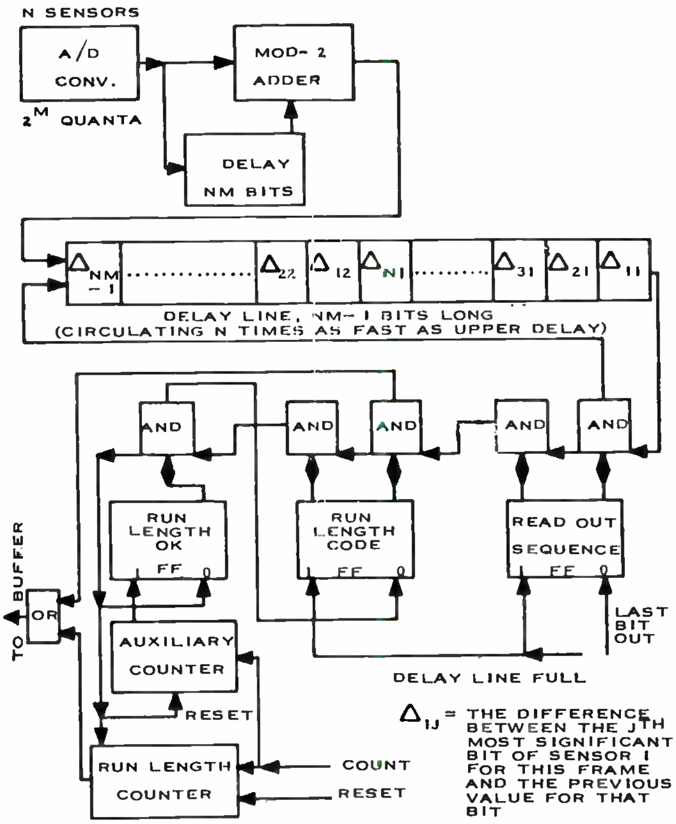


Fig. 2—A compactor for time-multiplexed sensors.

the compaction of television data. Information-preserving techniques appear likely to yield compaction ratios of 2 or 3. These techniques can probably be supplemented by information-destroying methods, but careful selection of the "information" is necessary. For instance, the "brute force" approach of degrading the resolution of the picture is probably inferior to the contour determination obtainable from the use of the gradient and the Laplacian operator as suggested by Kovasznay and Joseph [4].

As experiments and missions become more complex, it is probable that a general-purpose data processor will be included aboard advanced vehicles, having data compaction as one of its primary functions.

The Role of Information Transmission in Space

The second major topic of this paper is information transmission, which takes two forms for space: transmission between ground-based stations; and transmission from, to, and between space vehicles. Although we will concentrate on the latter form, the former deserves a few words.

An increasingly larger percentage of the capacity of ground-based transmission facilities is being used for data with extreme accuracy requirements (*e.g.*, telemetry, record transmission, data transmission between computers). To accommodate this traffic many transmission methods are being employed, special modems are being built, and special facilities are being supplied by the common carriers. All of these activities have a common goal: to increase the expected number of bits of *information* presented to the experimenter.

Most surface transmissions of space data result from the need to concentrate data at one computing point. This centralization allows the use of specialized programs and the correlation of data from several remote stations. The Mercury and Vanguard programs are excellent examples.

Data transmission from, to, and between the space vehicles affords the greatest challenge for communications. The major obstacle arises from the enormous space attenuation. The power received from a space vehicle is

$$S_R = \frac{P_T G_T A_R}{4\pi R^2}, \tag{1}$$

where

P_T = transmitted power,

A_R = effective area of receiving antenna,

R = range between transmitter and receiver,

$$G_T = \text{gain of transmitting antenna} = \frac{4\pi A_T}{\lambda^2}, \tag{2}$$

λ = wavelength,

A_T = effective area of transmitting antenna.

A_T, A_R, λ, and R must be in consistent units.

The power received must be maintained at a sufficient level, usually 10 to 100 times, above the competing noise N_R:

$$N_R = kTB_w, \tag{3}$$

where

B_w = effective receiving bandwidth,
T = effective receiver temperature in °K,
k = Boltzmann's constant.

The term least under the system designer's control is $1/4\pi R^2$, the space attenuation. There are three approaches to maintaining communications in the presence of this attenuation:

1) To reduce the required ratio of signal to noise (S_R/N_R).
2) To reduce the noise at the receiver input.
3) To increase the received power S_R.

Each of these approaches is being actively investigated and will be discussed in the subsequent sections.

Modulation Techniques for Reducing the Required SNR

Modulation techniques have probably been investigated more thoroughly than any other method for improving transmission efficiency. Since Shannon demonstrated that channel capacity could be approached more closely by the use of a large bandwidth-time product, modulation methods have been evaluated as a function of this product. Sanders [5] evaluates efficiency in terms of β, the received signal energy required per information bit transmitted in the presence of white Gaussian noise:

$$\beta = \frac{S_{R_{\min}}}{\epsilon^2 R}, \tag{4}$$

where

$S_{R_{\min}}$ = minimum received power required,
R = transmission rate (bps),
ϵ^2 = noise spectral power density (positive frequency).

Sanders' comparison of various modulation techniques is shown in Fig. 3. The Shannon limit, obtained when transmitting at channel capacity (for a given error rate), is

$$\beta_0 = (2^{1/\alpha} - 1), \tag{5}$$

where

$$\frac{B_w}{R} = \alpha.$$

If unlimited bandwidth is available, $\alpha \to \infty$ and $\beta_0 \to 0.693$.

The decrease in the theoretical limit of the required SNR with increasing bandwidth-time product (per bit) is especially interesting for space applications, where bandwidth is often readily available. It should be noted, however, that smaller and smaller gains result from increases in bandwidth as B_w/R increases past the knee of the curve.

Some systems, such as FM, do not even display this characteristic at minimum SNR because of the threshold effect. To reduce this effect the phase-lock loop [6] and the FM feedback system [21] have been proposed. In an analysis of the latter system Enloe [7] shows that reduction of the threshold is limited to 5 or 6 db because of cross-multiplication of in-phase and quadrature noise components.

The Shannon limit can be approached most closely by orthogonal coding techniques such as those investigated by Viterbi [8]. One example of these is the 5-bit-per-word Digilock system shown in Fig. 3. Further progress in improving communication system efficiency by going to higher bandwidth-time products and hence longer decoding words is limited in payoff. However, there are other uses for this technique such as the asynchronous communications system discussed later.

Many factors other than SNR improvement enter into the selection of a modulation system. Because precision is a fundamental requirement in telemetry, the modulation technique must guarantee the necessary accuracy at the required data rate. This must, of course,

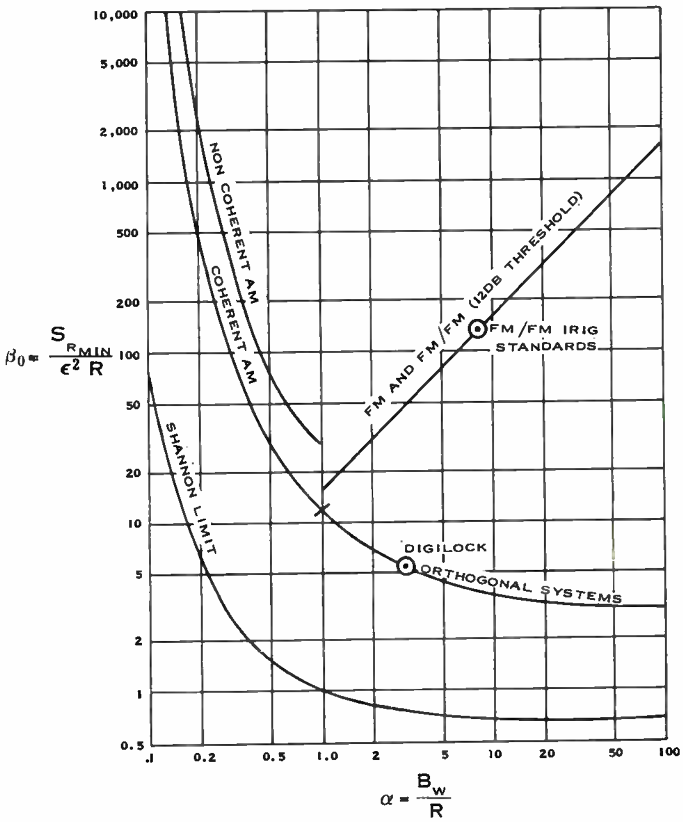


Fig. 3—β efficiency of various systems as a function of $\alpha = B_w/R$ for digital equivalent systems where the bit error rate $P_e = 10^{-6}$ (after: R. W. Sanders, "Communication efficiency comparison of several communication systems," Proc. IRE, vol. 48, pp. 575–588; April, 1960. Fig. 3, p. 578).

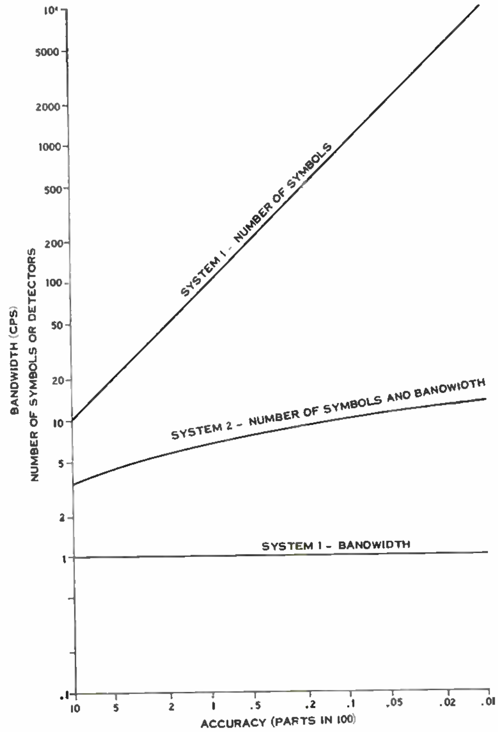


Fig. 4—Trade-off between bandwidth and number of symbols. (Sampling frequency = 1 cps.)

be accomplished within the weight limitations and with the desired reliability.

The fundamental factors determining the required data rate in a transmission system are the number of inputs (variables), the measurement accuracy of the variables, and their rate of change with time. The accuracy, expressed as the smallest recognizable difference in measurement in the range of possible values, establishes the number of messages that might be sent about a given variable at a given time. The maximum rate of change determines the sampling frequency.

Considering a single variable for the moment, one method of providing the N possible values for this variable would be to utilize N signal generators to obtain a set of N orthogonal symbols. This method achieves the minimum system bandwidth (and hence minimum transmitting power), requiring each filter in the receiver to have a bandwidth equal to the sampling frequency.

Such a minimum bandwidth system is shown as System 1 in Fig. 4. Here the sampling frequency was assumed to be one sample per second. Although the filter bandwidth can be seen to remain fixed as the accuracy increases, the number of symbols generated must increase. Thus 50 symbols are required for 2 per cent accuracy and 10,000 symbols for 0.01 per cent.

The number of signal generators could be reduced markedly by representing each possible value for the variable by a combination of several orthogonal signals.[1] One possible implementation would construct the N symbols from combinations of $\log_2 N$ orthogonal signals, each signal being either present or absent. This system is shown as System 2 in Fig. 4. The number of signals required has been reduced by an exponential factor; only 14 signals are required for 0.01 per cent accuracy. The price paid for this reduction is the increase in the effective bandwidth which is now $\log_2 N$ times the previous bandwidth.

The designer of modulation and demodulation systems must compromise between the considerations of symbol generation and bandwidth. Symbol generation really means circuit complexity since each symbol might be represented by a coded stream [8] or an orthogonal frequency. The choice depends upon the cost of the modulating equipment vs the cost and availability of transmitting energy, and will vary with the application. Some present systems might be classified:

System 1	System 2
QPAM/FM	PCM
PFM	
Digilock	

An attempt to compromise these conflicting trends is the PACM system [10]. Here pulse amplitude modulation is used for low accuracy and PCM for high accuracy. An extensive study has been made of the com-

[1] A method of combining the signals has been suggested by Franco and Lachs [9].

patibility of these systems [11]. Although PCM systems minimize the complexity of the space-craft modulator, bandwidth reduction is of such importance in space that attempts to employ System 1 types will continue by minimizing the size and weight of the generators.

It might be noted that sequential-type PCM systems have been the rule in the past. Some attention might be given to parallel symbol generation as a means for reducing peak power requirements aboard the vehicle.

Techniques for Reducing Receiver Noise

The noise with which the signal competes may be minimized by reducing either the information bandwidth or the effective temperature at the receiving antenna. This section discusses the advances anticipated in temperature reduction.

The effective temperature[2] depends upon a number of factors, as shown by de Rosa and Keller [12] in Fig. 5. This figure shows some of the unexpected noise sources which may be quite important in a poorly designed system. Careful attention to design will result in systems with effective temperatures determined primarily by the active receiver noise temperature (T_{RA}) and the space noise sources.

As an example, de Rosa and Keller consider the following system:

$f = 5.65$ kMcps,
$B_w = 25$ Mcps,
$G_R = 35$ db,
$T_{RA} = 10.5°K$ (TW maser).

If this system were well designed to minimize sidelobes and were operating in a quiet sky, the contribution from space noise would be approximately 6.0°K. The contribution from all other noise sources in this hypothetical system could be reduced to 2.0°K, resulting in an effective temperature at the output of the active receiver of 18.5°K.

This prediction coincides quite well with the performance forecast for the Goldstone antenna of Jet Propulsion Laboratory. Martin [13] indicates that the effective temperature for this receiving system will be reduced from 1430°K in 1961 to 20°K in 1964 by such measures as matching polarization, improving the antenna design, and utilizing a maser receiver.

Fig. 6 (from de Rosa and Keller [12]) indicates the limits of performance in terms of the space noise contributions from galactic, corona, ionospheric, and atmospheric noise. It must be pointed out, however, that this noise refers to open areas. If the antenna must be pointed at the sun or any other discrete radiator another component must be added. The most troublesome bodies for earth-based receivers will probably be the sun (up to 100,000°K) and the moon (200–300°K). The planets will subtend only a small portion of the beam and should not contribute more than 3°K. Thus, unless direct communication with the moon or within about 5° of the sun is required, the window in the 1–10-kMcps region appears to present a promising application for earth-based low-noise amplifiers (see Fig. 7).

For the receiver in the spacecraft, Fig. 7 indicates the considerable superiority of the maser throughout the useful frequency range. The maser in conjunction with a cryogenic computer appears attractive on space vehicles as pointed out by Stuhlinger [14]. However, such a receiver would drop in efficiency whenever its antenna points to a planet or the sun to receive from a particular direction, or if the earth must subtend an appreciable part of its beam. In such cases the simplicity of the Esaki diode amplifier might suffice.

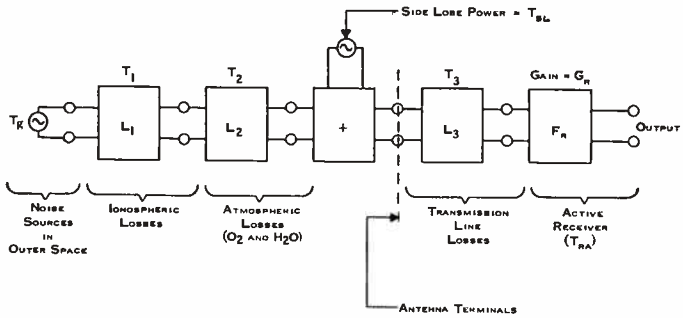


Fig. 5—Noise sources in a receiving antenna (after [12]).

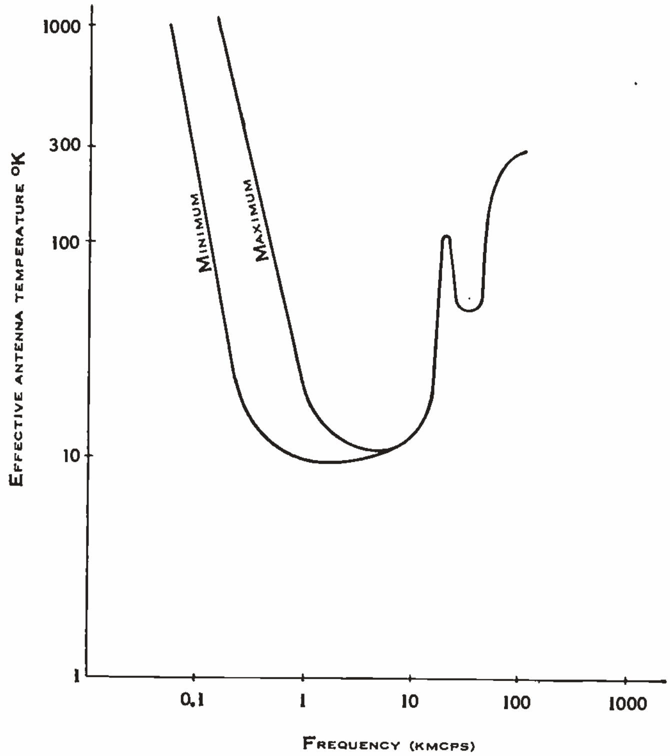


Fig. 6—Space noise from galactic, corona and atmospheric contributions (after [12]).

[2] Noise figures are not used here because noise temperatures can be measured readily, and added directly when referred to the same point, and because small errors in noise figure can make a very large difference in high performance systems. Actually, temperature can be considered to be noise power per unit bandwidth normalized by the Boltzmann constant.

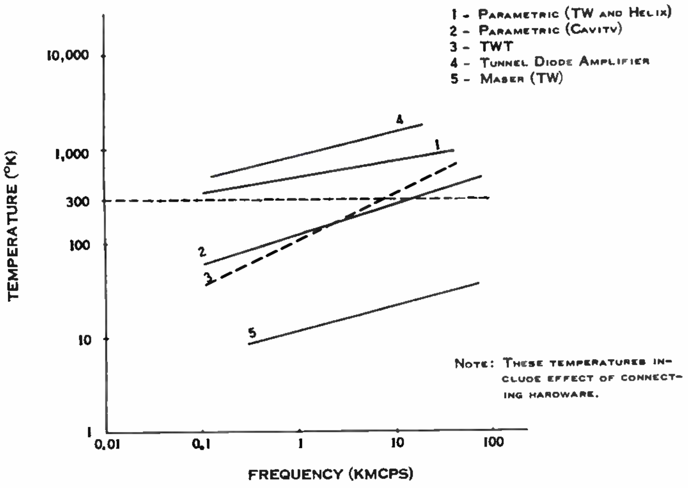


Fig. 7—Effective input noise temperatures of low-noise amplifiers (after [12]).

We conclude that we may expect to improve communications to space vehicles by reducing inherent noise by a factor of 10–70 under ideal conditions.

Augmenting the Received Power

Several methods exist for augmenting the received power:

1) Increasing the power transmitted.
2) Choosing the frequency judiciously.
3) Increasing the gain of the transmitting and/or receiving antenna.

These will now be considered in turn.

The power available in space is quite limited at present, and probably will continue to be in future missions. For the next few years solar power will be relied upon, with eventual transition to nuclear sources. Neither of these sources is particularly lightweight or efficient (typically, 10 per cent for solar cells). There exists the additional problem of converting the dc output of these sources to usable RF frequencies, usually at efficiencies of about 10 per cent. New transmitting components are becoming available to accomplish this conversion, particularly in the 1–10-kMcps region. Of particular note is the development of traveling-wave tube sources having high reliability and power in the 3–10-watt region. Solid-state components such as Esaki diodes are available up to SHF but are limited in power to milliwatts.

The standard approaches to increasing antenna gain are enlarging the antenna and increasing the operating frequency. By substituting (2) into (1),

$$S_R = \frac{P_T A_T A_R}{R^2 \lambda^2}, \tag{6}$$

it becomes apparent that the received signal is enhanced by the square of the operating frequency and directly as the aperture area of the antennas. There are these further considerations:

1) Atmospheric attenuation and noise increase sharply above 10 kMcps because of water vapor.
2) A window in the space noise exists between 1 and 10 kMcps.
3) The availability of components generally favors lower frequencies.
4) At high frequencies tracking problems are more difficult.

After careful consideration of these factors Rechtin [18] recommends operation in the 1–10-kMcps region.

For earth-based antennas the trend to large sizes is typified by the Jodrell Bank radio telescope, the Sugar Grove, W. Va., Project, the 250-foot dish planned for Goldstone, and the 1000-foot reflector to be embedded in a Puerto Rican hilltop.[3] A possibility for the future is the use of electronically steerable phased-array antennas. Though expensive, these antennas are a convenient way of phasing multiple transmitter power sources and augmenting the transmitted power greatly.

For spacecraft antennas, increases in size bring major structural and orientation problems. The structural problem is being alleviated by the advent of unfurlable antennas, feasibility models of which have been described by Kennedy [16].

A major step toward the reduction of the orientation problem has been devised.[4] A data-transmission system called TELARRAY uses a spaceborne Van Atta array [17] which modulates the ground-supplied power to return information from the spacecraft to the original transmitter, as shown in Fig. 8.

The unmodulated power from the earth-based transmitter is modulated by the microwave diodes in the interconnecting lines, providing identical modulation to the energy from each array element. Because the delay between a pair of interconnected elements is the same for any pair, the returned energy is reflected back in the same direction as the incident energy.[5] Thus, as long as the earth-based transmitter is within 60° of the center line of the array (*i.e.*, $\theta < 60°$), the antenna will provide the gain specified by (2).

The advantages of this system are then:

1) High gain with a much reduced orientation problem.
2) No transmitting power is required in the spacecraft.
3) Greatly augmented *reliability* because of the natural redundancy of the antenna configuration.

The high gain partially compensates for the space attenuation. The power is derived from the ground transmitter, where it can be economically supplied.

[3] There are some indications, however, that this trend may be reversing itself because of the cost of the larger antennas [15].

[4] C. M. Johnson and E. L. Gruenberg, "A Semi Active Telemetry and Communications System for Space Applications," to be presented at the National Telemetry Conference, Washington, D. C.; May, 1962.

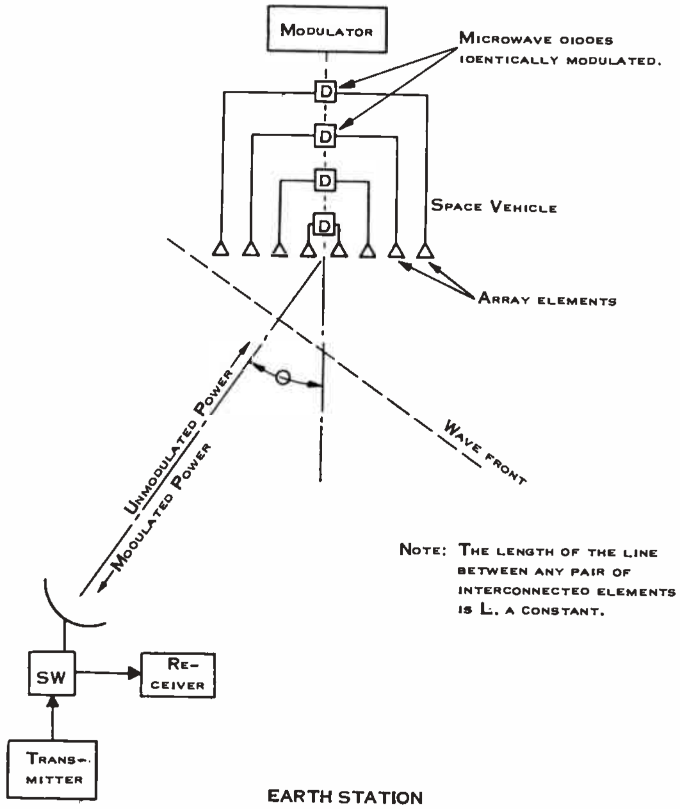


Fig. 8—TELARRAY data-transmission system.

The expression for S_R in this case is

$$S_R = P_T \frac{A_T^2 A_R^2}{R^4 \lambda^4}. \tag{7}$$

To illustrate the performance of this system consider the following example:

$P_T = 20$ kw
$f = 10$ kMcps (corresponding to $\lambda = 3$ cm)
$R = 4000$ miles
Ground antenna, 60 feet in diameter
Space antenna, 4×4 feet
Noise Temperature, 500°K.

If the bandwidth is 1 Mc, S_R will be 15 db above the noise. Thus nearly 10^6 bps can be transmitted.

It is possible to make use of this technique to provide a highly reliable communication satellite. For communication between spacecraft, the utilization of masers and optical frequencies appears most promising. Methods of modulation and efficiency of sources must be improved. Techniques of orientation stabilization must be found so that extremely narrow beamwidths may be employed. The quantization of energy in blocks[6] of $h\nu$ becomes a serious consideration at high frequencies, and extremely directive beams must be employed to concentrate sufficient energy at the receivers.

[6] Planck's constant times frequency.

Asynchronous Communications

The transmission of command and control information to a spacecraft is a special problem in itself. The time lag for communications[7] makes the acknowledgment of commands very difficult, and requires rapid and reliable synchronization of the spacecraft command decoder with the incoming message.

A promising solution to this problem is the use of an asynchronous communication system with a large "codebook" of stored messages (see Fig. 9). Any message would be decoded and recognized immediately with high reliability, regardless of the time of reception of the sequence. The large codebook would permit the receiver to operate at high detection efficiency and approach Shannon's ideal communication system. This efficiency increases with the number of symbols in the codebook, but the increase is relatively slow because of the required orthogonality of the symbols [19].

The requirement of this system for considerable message storage will probably be met by the increasing miniaturization of memory devices. An optical system using photoelastic delay lines appears to be the most promising system for implementing the configuration shown in Fig. 9.

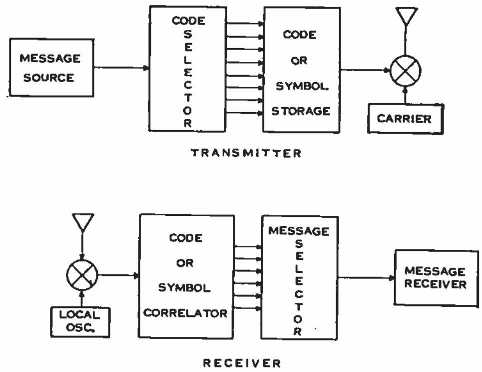


Fig. 9—A transmission system with a large codebook.

Component Development

The increasing complexity that has been discussed demands major reductions in the size of the components for future spaceborne data-processing and transmission equipment. That this matter is receiving increasing attention from many organizations is evidenced by Gerhold's volume and weight predictions [20], shown in Tables I and II. These tables compare a "conventional" printed wiring design, improved packaging techniques, an integrated micromodule design using advances in thin-film and solid-state circuits, and a future hybrid design incorporating the promising molectronics advances.

[7] For a round trip to the moon, 2.6 seconds; for a round trip to Mars, about 9 minutes minimum.

TABLE I*

SIZE COMPARISONS FOR A WEAPONS SYSTEM†

Item	1959–1960 Conventional		1960–1963 Micromodule		1963–1967 Micromodule and Solid State		1967 Micromodule, Solid State, and Molectronics	
	Size (cubic feet)	Size Reference	Size (cubic feet)	Size Reduction	Size (cubic feet)	Size Reduction	Size (cubic feet)	Size Reduction
Computer	3.5	1	0.2	18/1	0.05	70/1	0.01	350/1
Power supply	0.25	1	0.12	2/1	0.08	3/1	0.056	4.5/1
Communications	2.62	1	1.37	1.8/1	1.02	2.5/1	0.73	3.6/1
Guidance system	4.65	1	0.93	5/1	0.74	6.3/1	0.67	7/1
Complete system	11.02	1	2.62	4/1	1.89	6/1	1.47	7.5/1

* After: R. A. Gerhold, "Integration of microcircuitry into microassemblies," IRE TRANS. ON MILITARY ELECTRONICS, vol. MIL-5, pp. 227–233; July, 1961. Table II, p. 229.
† Based on data furnished by the Missile Electronics and Controls Div., Radio Corp. of America, Burlington, Mass.

TABLE II*

WEIGHT COMPARISONS FOR A WEAPONS SYSTEM†

Item	1959–1960 Conventional		1961–1963 Micromodules		1963–1967 Micromodule and Solid State		1967 Micromodule, Solid State, and Molectronics	
	Weight (pounds)	Weight Reference	Weight (pounds)	Weight Reduction	Weight (pounds)	Weight Reduction	Weight (pounds)	Weight Reduction
Computer	125.0	1	12.0	10/1	3.0	40/1	0.6	210/1
Power supply	19.0	1	9.5	2/1	6.5	3/1	5.0	3.8/1
Communications	212.5	1	108.0	2/1	65.0	3/1	40.0	5/1
Guidance system	181.5	1	47.5	3.8/1	26.7	6.8/1	21.1	8.5/1
Complete system	438	1	177	2.5/1	101	4.3/1	66.7	6.5/1

* After: Gerhold, *op. cit.*, Table III, p. 229.
† Based on data furnished by the Missile Electronics and Controls Div., Radio Corp. of America, Burlington, Mass.

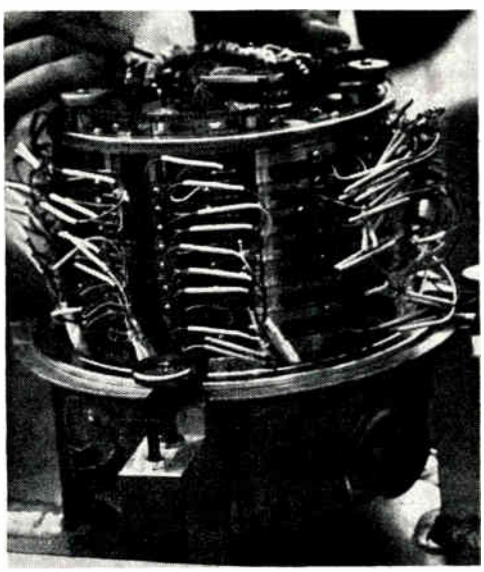


Fig. 10—Magnetic drum memory for the Titan guidance computer (IBM).

Fig. 11—Ferrite cores for random-access memory for NASA's Orbiting Astronomical Observatory (IBM).

A striking situation pointed up by Tables I and II is the small amount of reduction anticipated in the communications and power equipment compared with that predicted for the computer equipment. This situation emphasizes the desirability of increased utilization of those logical functions best performed by digital equipment. Examples of these are data compaction, "adapting" the transmission to the medium, and the asynchronous communication system of the preceding section.

In addition, more attention will probably be devoted in the future to those transmission systems that minimize the amount of power and communication equipment required on the spacecraft. The TELARRAY system described earlier is an excellent example of such a system. It reduces the spaceborne power and transmitting equipment markedly by increasing the ground-supplied power and the antenna gain.

Representative of the improvements in data-processing equipment expected in the near future is the comparison between Figs. 10 and 11. Fig. 10 shows the state-of-the-art magnetic drum memory utilized in the Titan inertial guidance system that was successfully demonstrated in a recent test firing. Fig. 11 depicts the random access (nondestructive readout) memory elements that will be incorporated in NASA's Orbiting Astronomical Observatory. It seems likely that densities of 10^6–10^7 bits per cubic foot will be achievable with ferrite elements similar to these in the near future.

The Future

An estimate of information transmission capacity in the near future for deep space vehicles has been made by Martin [13] and is shown in Fig. 12. The rising bit rate is expected to be achieved by such techniques as increasing spacecraft power, increasing ground antenna size, raising the operating frequency, and using masers, steerable directional antennas on the spacecraft, orthogonal coding, and coherent demodulation.

Further gains may be expected if techniques suggested in this article prove themselves. The compacting of data will result in the transmission of necessary information only, effectively increasing the channel capacity by a factor of perhaps 3 to 10. Other benefits, such as reduction of the time required to transmit the real message, may be anticipated from compaction.

Several orders of magnitude improvement in channel capacity could well result from the exploitation of very large antenna apertures in space, beyond those estimated by Martin. Orientation instabilities have been the limitation in the past, but the passive modulated antenna (TELARRAY) shows one way of alleviating this problem.

The startling difference in size reduction potential between computer components and transmitter and other components associated with the communications art suggests that computer components will take over more

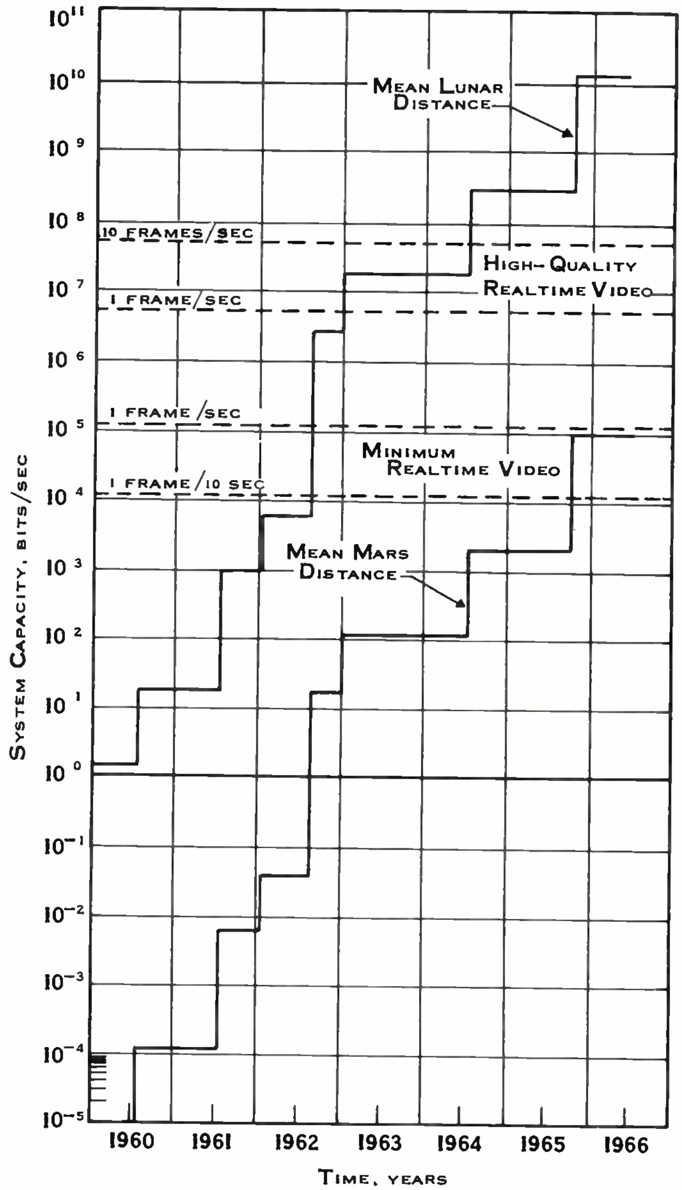


Fig. 12—Projected deep-space telemetry system capacity (after [13]).

of the data transmission job. Such functions as message programming, multiplexing, coding, and compacting are fair game for the computer. Others will be found.

Improvements in communication systems are appearing which will make possible, within the inner solar system, data rates high enough to permit real-time remote observations of the most intricate nature. The rapid and timely evaluation of such data will be a challenge requiring the most advanced data-processing equipment. We may then look forward to the performance of remote experiments by remote handling equipment under earth-based control.

Acknowledgment

The author wishes to thank F. W. Ellersick for his assistance in the preparation of this paper.

References

[1] A. G. Ratz, "Telemetry bandwidth compression using airborne spectrum analyzers," Proc. IRE, vol. 48, pp. 694–702; April, 1960.
[2] F. W. Ellersick, "Data compactors for space vehicles," *Record IRE East Coast Conf. on Aeronautical and Navigational Electronics*, pp. 6.2.6-1–6.2.6-5; 1961.
[3] P. Elias, "Predictive coding," IRE Trans. on Information Theory, vol. IT-1, pp. 16–33; March, 1955.
[4] L. S. G. Kovasznay and H. M. Joseph, "Image processing," Proc. IRE, vol. 43, pp. 560–573; May, 1955.
[5] R. W. Sanders, "Communication efficiency comparison of several communication systems," Proc. IRE, vol. 48, pp. 575–588; April, 1960.
[6] R. M. Jaffe, and E. Rechtin, "Design and performance of phase-lock circuits capable of near-optimum performance over a wide range of input signal and noise levels," IRE Trans. on Information Theory, vol. IT-1, pp. 66–76; March, 1955.
[7] L. H. Enloe, "Decreasing the threshold in FM by frequency feedback," Proc. IRE, vol. 50, pp. 18–30; January, 1962.
[8] A. J. Viterbi, "On Coded Phase Coherent Communications," Jet Propulsion Lab., Pasadena, Calif., Tech. Rept. No. 32–25; August 15, 1960.
[9] G. A. Franco and G. Lachs, "An orthogonal coding technique for communications," 1961 IRE International Convention Record, pt. 8, pp. 126–133.
[10] P. J. Klass, "Hybrid telemetry proposed for future use," *Aviation Week*, vol. 72, pp. 115–116; April 4, 1960.
[11] "Telemetry System Study, Final Report," vol. I, Aeronutronic, Newport Beach, Calif.; Publ. No. U-743, ASTIA Doc. No. AD234959; December 18, 1959.
[12] L. A. de Rosa and E. W. Keller, "Potential Application of Recent Advancements in Communications Technology," IT&T Communication Systems, Paramus, N. J., prepared for 5th Natl. Symp. on Global Communications, Chicago, Ill.; May 22–24, 1961.
[13] B. D. Martin, "The Mariner Planetary Communication System Design," Jet Propulsion Lab., Pasadena, Calif., Tech. Rept. No. 32-85 (Revision No. 1); May 15, 1961.
[14] E. Stuhlinger, "Electronics in planning space flights," this issue, pp. 1344–1351.
[15] D. Winston, "New antenna patterns improve radiotelescope range and sensitivity," *Electronics*, vol. 34, pp. 28–29; September 8, 1961.
[16] P. D. Kennedy, "Unfurlable antennas for space communication," *5th Natl. Symp. on Global Communications Convention Record*, Chicago, Ill.; May 22–24, 1961.
[17] L. C. Van Atta, "Electromagnetic Reflector," U. S. Patent No. 2,908,002; October 6, 1959.
[18] E. Rechtin, "Space Communications," Jet Propulsion Lab., Pasadena, Calif., Tech. Release No. 34-68; May 1, 1960.
[19] R. M. Fano, "Transmission of Information," John Wiley and Sons, Inc., New York, N. Y.; 1961.
[20] R. A. Gerhold, "Integration of microcircuitry into microassemblies," IRE Trans. on Military Electronics, vol. MIL-5, pp. 227–233; July, 1961.
[21] J. G. Chaffee, "The application of negative feedback to frequency modulation system," Proc. IRE, vol. 27, pp. 317–331; May, 1939.
[22] R. Van Blerkom and H. Blasbalg, "Adaptive Redundancy Removal," IBM Communications Systems Center, Rockville, Md., prepared for NEREM Conf., Boston, Mass.; November 14–16; 1961.

Trends in Space Navigation*

K. N. SATYENDRA†, SENIOR MEMBER, IRE

Summary—The theme of this paper is based on the growing use of self-contained navigational systems in space navigation. A figure of merit, called the self-containment index (SCI), is defined in order to define quantitatively the degree of autonomy for which a system design is directed.

The navigational problem of spacecraft is divided into problems of point-dynamics and rigid-body dynamics and the major factors are defined which deserve consideration in the understanding of the space navigational problem. Consistent with the requirements for low size, weight and power allocations and for very high accuracy and reliability tolerances, optical techniques are found ideally suited for space navigation. Open-loop and closed-loop methods of earth satellite navigation are presented and extended to the lunar navigation problem.

Some advanced navigational concepts are given using nonrotating and track-while-scan instruments. The importance is emphasized of the increase in data-processing requirements and their part in accuracy enhancement through statistical techniques. A review of optical Doppler and lasers for measurement of distance and velocity is given. The concept of guidance management is proposed for consolidation of all the launch, orbital, midcourse, return and other guidance phases of a future multistage spacecraft into a single, integrated, self-contained navigation system within the final upper stage.

Introduction

THE PURPOSE of this paper is to point out some trends in space navigation and some advance navigational concepts which appear to be feasible and practicable, especially in the light of recent improvements in signal-sensing and signal-processing techniques. The paper is therefore not an original and detailed discussion of any specific navigational method; however, heavy emphasis is placed on optical celestial sensing techniques since these techniques appear to be an essential aspect of nearly all future space systems.

With increased skills in astronautics, the space traffic consisting of scientific and military payloads is rapidly increasing. An educated guess of five to six thousand objects in orbit during the next decade is not unreasonable. Some of these objects may be useful payloads, which may be manned or unmanned scientific or military payloads, while the others may be inactive payloads and space debris—booster tankages—and the like. The tracking and cataloging by ephemeris prediction of such a vast number of space objects creates a data-processing problem at a ground station resulting in possible denial of navigational data to useful orbital payloads which depend on tracking and computations made on the ground. Unlike the air traffic problem over modern airports where aircraft can be made to fly a pattern while awaiting landing instructions, orbital vehicles have a limited time for communication over the ground station before they get lost over the horizon.

The manned spacecraft, very much like the manned aircraft, demand on-board guidance and control capability. The term "self-contained navigation" has become a familiar term in the aerospace industry and an ever-increasing demand is made on the electronics industry to provide all navigational measuring and computational instrumentation on board the spacecraft itself with very little dependence on the ground complex. The extent of "self-containment" or the degree to which the spacecraft can navigate without receiving inputs from outside depends on the mission requirements, payload and equipment limitations, cost and reliability considerations. No clear definition of the term "self-contained navigation" appears to be present and therefore the following definition, which is especially illustrative for earth satellites, is offered:

Definition: A self-contained navigation system is defined as one which, for a given portion of the mission flight path, functions independently of concurrent man-made friendly signals originating from points external to the spacecraft.

For a long time to come, the earth will still be the main source of friendly signals which will be generated in ground-based communication and control stations. With the above definition of self-containment, three classes of self-contained or autonomous navigation systems may be studied:

Class I: Autonomous outside the horizon of ground communication and control center. This class covers those space vehicles which are independent of communication with the earth when outside the line-of-sight of a set of predefined ground tracking and communication control centers. During this time any navigation performed is accomplished by the on-board systems.

Class II: Autonomous for more than one pass over the ground communication and control center. This degree of autonomy may be necessary to accommodate those instances when the ground-based communication stations are saturated because of too many space vehicles overhead and cannot transmit updated navigational data to an individual spacecraft.

Class III: Indefinitely autonomous or continuous systems. This class covers those systems which are completely autonomous and do not depend upon earth-based tracking and communication networks to navigate. The degree of self-containment for this class actually depends on the reliability of the system.

* Received by the IRE, March 1, 1962.
† Nortronics, a Division of Northrop Corporation, Palos Verdes Estates, Calif.

A self-contained navigation system for Class III, requiring more advanced and sophisticated components, will chronologically follow the development of simple systems to cover Classes I and II. A fully self-contained navigation system is therefore the ultimate one can hope to accomplish and suffers from all the headaches such aspirations usually involve.

Quantitatively, one can define a figure of merit to distinguish between the classes of self-containment in terms of the ratio T_Q/T_o, where T_Q is the quiescent period during which there is no communication between the space vehicle and the earth, and T_o is the orbital period or time for a round trip. This ratio, rounded off to the nearest integer, may be defined as the self-containment index, and abbreviated as SCI. For example, if an earth satellite were autonomous for 260 minutes and its orbital time period were 90 minutes, then its SCI = 3. With this definition, the SCI of Class I–III systems would be respectively less than unity, greater than unity and approach infinity.

The methods involved in solving the three classes of problems differ, not in principle, but in technique and detail because of the differences in accuracies, component life times, duty cycles and degree of sophistication permissible. Present day technology is adequate to achieve values of SCI between 1 and 3.

Higher degrees of self-containment await improved techniques for developing components with reduced size, weight and power allocations and increased reliability in a hostile space environment. Almost all aerospace systems of the future will demand a fully self-contained navigation ($SCI \geqq 3$) as the *primary system* and a ground-based or earth-dependent navigation system as a *back-up*. Some illustrative examples of systems with different requirements for the SCI will be presented here.

The Navigation Problem and Its Solution

In its simplest form, the navigational problem of a spacecraft consists of two basic parts:

1) *Point dynamics*, defining the motion around the geocenter of the spacecraft treated as a point-mass.
2) *Rigid-body dynamics*, defining the interrelation between the body axes of the spacecraft and certain directional space references. This aspect of navigation includes the definition of a system of coordinate axes within the spacecraft, and their relationship to one or more systems of directional reference in space.

Before proceeding with the detailed exposition of the two problems, it is desirable to distinguish between the terms, *navigation*, *guidance*, and *control*, at least to classify the instrumentation associated with each of these terms.

The origin of the term *navigation* lies in the French words *navis* meaning *ship*, and *agere*, meaning *to move or direct*. Space navigation may be defined by analogy with sea navigation, by a modification of Webster, as "the art or science of conducting a spacecraft from one place to another, including the method of determining position, course, distance passed over, etc., by the principles of geometry and astronomy. . . ." Revising Webster's order, one may single out as navigation functions,

1) the determination of position, present course, distance covered, etc.,
2) the action of *conducting*.

Based on the line of reasoning above, Hoelker [1] defines *guidance* as follows: Guidance is the technological field concerned with the following sequence of action:

1) Measuring the data for the purpose of the determination of the current state of flight,
2) Determining the current state of flight from the measured data,
3) Determining the path ahead that joins the current state of flight with the desired terminal state of flight,
4) Providing to the control mechanism an immediate reference that serves for implementation of the determined path ahead.

The action of *conducting* may be interpreted as the enforcement of the course of action through the use of the *control* system. The above definition separates the functions of guidance and control, and guidance terminates at providing references, computations of maneuver ballistics and inputs to the control loop or loops.

With the concepts stated above, it may be desirable to relate the two broad problems of navigation with guidance and control as follows:

Guidance implies preparation of the *inputs* for solving the point-dynamics and rigid-body dynamics of a spacecraft. Control implies enforcement, through the use of control loops and additional force, of the point-dynamics and rigid-body dynamics of the spacecraft to a prescribed course of action.

The solution of the *guidance* problems and the preparation of the necessary *inputs* requires the use of instrumentation involving two basic functions, namely, *sensing* and *signal processing*. The sensing function involves establishment of directional and positional references through gyroscopic or celestial (planetary or stellar) means, or man-made optical and/or radio-frequency means. The signal-processing function involves the computation or extraction, usually by means of a computer, of the navigational parameters of the spacecraft from the sensed signals. The most demanding problem is the establishment of the orbital elements of a spacecraft to a high degree of accuracy in the shortest possible time. The accuracy in sensing direction references or velocity or position of a spacecraft has a physical limit set by the equipment and laws of physics. Space missions, especially those requiring self-contained navigation techniques, demand refinements in accuracy which, therefore, must be obtained by statistical techniques, yielding what is usually referred to as "smoothed" data. Such refinements are achieved mainly through high-speed computers and the trend in space navigation is therefore an optimization of the signal-sensing and signal-processing techniques, usually resulting in a higher burden being placed on the latter.

In summary, the following major factors deserve consideration in the understanding of the space navigation problems and their solution:

- I. Directional References
  - A. Inertial
    - 1) celestial (stellar)
    - 2) gyroscopic
    - 3) combination or astronertial
  - B. Local Vertical
- II. Position Determination
  - A. Static
    - 1) triangulation
    - 2) stadiametry
  - B. Dynamic
    - 1) active ranging
    - 2) explicit solution of point-dynamics
- III. Types of Measurement
  - A. Angular—angle, angle-rate, angular acceleration, etc.
  - B. Range—range, range-rate, range acceleration, etc.
- IV. Accuracy Refinements
  - A. Systematic Error Isolation—intrinsic and external to instruments
  - B. Statistical Techniques—time smoothing and space smoothing
- V. Forms of Solution
  - A. Time
    - 1) real-time solution
    - 2) delayed-time solutions
  - B. Location
    - 1) earth based
    - 2) on-board or self-contained
- VI. Signal Sources
  - A. Natural—stars, planets, fields of force
  - B. Man-made—radio or optical, both passive and active
- VII. Physical Description
  - A. Size, Weight, Area, Power, Volume
  - B. Reliability, Cost, etc.

Self-Contained Navigation of Earth Satellites and Lunar Spacecraft

Earth Satellites

The most difficult problem of navigation is the solution of the point-mass dynamics. In general, the motion of a point-mass is governed by Newton's and Kepler's laws which state, in effect, that the rate of change of momentum is equal to the total force acting on the point-mass. The total force F_T on a satellite, treated as a point-mass, is given by the following equation:

$$F_T = F_E + F_D + F_M + F_P \tag{1}$$

where

F_E = Most complete force field of the earth, such as the Jeffreys model of the oblate earth

F_D = Atmospheric drag

F_M = Force due to maneuvering rocket thrusts, such as those due to retro-rockets, ion-engines, etc., used for rapid *position* changes of the satellite

F_P = Perturbative forces due to solar photon pressure, lunar gravitational effects, rotation of earth, etc.

It is a fact that if F_T is completely known, and the initial values of position and velocity of the point-mass are known, the solution of the Newtonian force equation (1) is theoretically possible. However, uncertainties in several factors make it impossible to provide an ideal and true theoretical solution of the force equation and thus predict the position of the satellite point-mass. It is therefore necessary to make compromises in practice and be satisfied with the most reasonable, if not the best solution. The position and velocity accuracies of a satellite, based on such solutions, will be better for short flight times, as in Class I and Class II type systems, than for long flight times, as in Class III type of self-contained navigation systems. The mechanization of the computer for solving these equations depends also on the degree of sophistication employed. Two such methods are described below, the *open-loop* method for Class I and Class II types of satellite navigation and the *closed-loop* method for Class III type of navigation.

Open-Loop Method: For Class I and II types of self-contained navigation schemes for which the SCI is between 1 and 3, the explicit solution of the force equation (1) may be simplified by the help of some assumptions, such as the following:

1) That the earth is perfectly spherical,
2) That the earth's force field is perfectly central,
3) That the orbital elements of the satellite orbit may be accurately computed on the ground and transmitted to the satellite,
4) That there are no external perturbative, thrust maneuvering or atmospheric drag forces,
5) That the satellite is truly a point-mass.

The above simplifying assumptions yield a circular or elliptic orbit for the satellite. With known initial conditions, *i.e.*, with known position and velocity at any time of such an orbit, the subsequent orbital parameters after a time interval t may be easily computed algebraically rather than through a differential equation solution. The time difference t can be measured in the satellite by a simple clock or oscillator. The effects of earth oblateness, atmospheric drag, solar winds, etc., could be added, as refinements, in the form of empirical corrections. Specifically, the vehicle's position in spherical coordinates may be programmed into a small onboard function generator as trigonometric or power series. Thus, the latitude and/or longitude might be represented in the form

$$\theta = \theta_0 + \theta_1 t + \theta_2 t^2 + \theta_3 t^3 \cdots \quad (2)$$

or

$$\theta = \theta_0 + \theta_m \operatorname{Sin}(\omega t + \alpha). \quad (3)$$

The form of the function will be determined by the eccentricity, inclination of the orbit, altitude, etc. The block diagram for such a simplified navigation system is shown in Fig. 1. The output from the clock to the function generator serves to define the position and velocity. Position and velocity are channeled into the guidance computer along with vehicle attitude and thrust acceleration to be included in the guidance equations. Attitude control signals and thrust commands are generated by the guidance computer and are fed to the attitude control device and thrust actuating device, respectively. *A complex, high-data rate, general purpose computer is not necessary.*

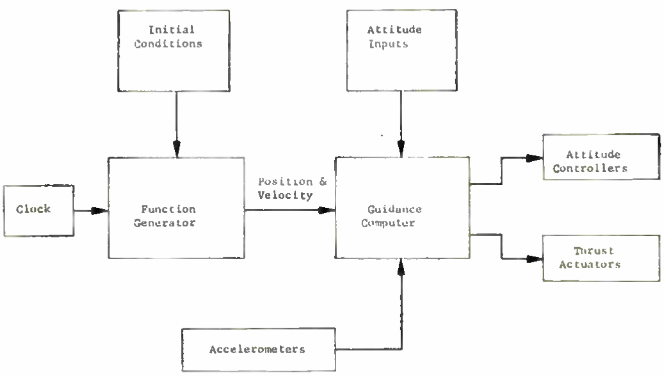


Fig. 1—Open-loop method—functional block diagram of self-contained navigation (Class I and Class II).

Closed-Loop Solution—Class III: In the previous method, position and velocity predictions are made by open-loop methods, with periodic updating from the ground. To obtain the accuracy and independence required for Class III type navigation systems, a closed-loop system will be necessary to improve system accuracy by up-dating refinements made by means of onboard measurements.

The basic refinement necessary for Class III type navigation schemes is the *explicit* solution of the force equation (1), taking into account as many perturbative factors as possible.

One such method has been presented by Satyendra and Bradford [3]. This method is built around the concept of sensing the satellite attitude by means of a prime reference established by a pair of star trackers; direct measurements, over a period of time, of the local geocentric vertical by means of a horizon tracker; and determining the orbital elements by a precision time reference and a data-processing computer.

The basic conceptual feature of this approach is that suitable stars are preselected and tracked. The initial star lock-on can be achieved by orienting the satellite from the injection guidance subsystem's gyroscopes. Once this lock-on is achieved, it is continually maintained to obviate a complex search and identification process implied by loss of orientation resulting from extended shutdowns of operation.

A typical design of a lightweight star tracker is described by Stevens [4]. Two star trackers are locked on two known stars, S_1 and S_2, as shown in Fig. 2. The selection of the stars S_1 and S_2 may be based on the techniques described by Stevens [5]. A coordinate reference frame consisting of three mutually perpendicular vectors, $\hat{C}_1$, $\hat{C}_2$, $\hat{C}_3$, is rigidly connected to the two star-tracker telescopes, which are represented by the vectors $\hat{P}_1$ and $\hat{P}_2$ in Fig. 2. With the star trackers locked on to the stars, S_1 and S_2, the orientation of the frame ($\hat{C}_1$, $\hat{C}_2$, $\hat{C}_3$) in inertial space is known. For convenience in analysis and mechanization, it is assumed that this orientation is such that the vector $\hat{C}_3$ is approximately normal to the plane of the orbit. The azimuth and elevation angles θ and ϕ between the inertial frame ($\hat{C}_1$, $\hat{C}_2$, $\hat{C}_3$) and the local vertical are measured by means of a horizon tracker.

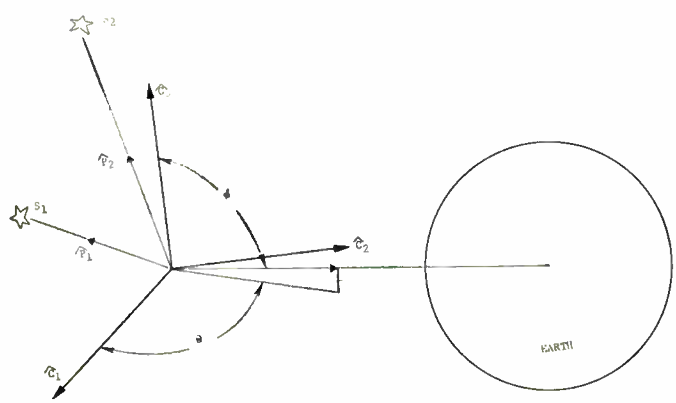


Fig. 2—Basic geometry of the Satyendra and Bradford method.

In the ideal case of perfectly accurate measurements and an exact knowledge of the force field about the earth, one would be able to determine an earth satellite's trajectory by measuring the direction of the local vertical at four distinct times. In such an ideal situation no new information could be obtained by collecting a greater amount of data. In reality, however, all measurements will be corrupted by noise; hence the above reasoning is not applicable to any practical situation. It can be clearly demonstrated that, if one has a large set of measurements of the local vertical, statistical decision techniques can be used to obtain a much more accurate estimate of the trajectory than could be obtained by using only four measurements. For this reason the use of approximately 1000 measurements of the local vertical is envisioned to obtain a good estimate of a satellite's orbit.

To estimate a trajectory it is necessary first to choose a model for the force field about the earth. Under such a choice, every possible trajectory will be completely characterized by six initial conditions—three of position and three of velocity. These six initial conditions will completely characterize a trajectory, for the period of time for which the model is applicable.

For this reason the problem of trajectory determination reduces to the problem of determining the six initial conditions of position and velocity.

The force model chosen for this method was the Jeffreys model of the oblate earth. This model neglects atmospheric drag and perturbations due to the influence of the moon; but these effects appear to be negligible for typical earth satellite trajectories for periods of time corresponding to at least several cycles of the satellite. If the only requirement is that the model give a satisfactory representation of a satellite's trajectory for a couple of periods, the assumed model is probably adequate for the present purpose.

The second step is to measure the various angles θ and ϕ, using the basic geometry shown in Fig. 2. Suppose that n measurements are taken of these angles. The data will then consist of the triples

$$(\theta_1, \phi_1, t_1),\ (\theta_2, \phi_2, t_2),\ (\theta_3, \phi_3, t_3),\ \cdots,\ \cdots,\ (\theta_n, \phi_n, t_n),$$

where t_i is the time at which the ith measurement was taken.

In the absence of any errors in the horizon tracker and inertial reference, the angles θ and ϕ would truly represent the direction of the geocentric vector. The graphs of θ or ϕ vs time would be straight lines or periodic curves for circular or elliptic orbits. The presence of instrument errors discussed by Jaramillo [6] results in data which are shown as points distributed along these curves for θ and ϕ as shown in Fig. 3.

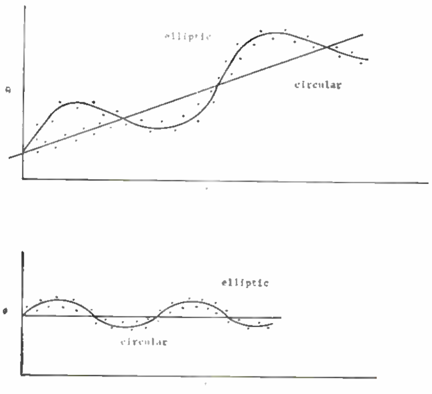


Fig. 3—Variation of θ and ϕ with time.

The basis of the method involves computing the difference between the calculated values of θ and ϕ from the force field model and the horizon tracker gimbal-pickoff measured values. These difference values are smoothed over as many full or fractional cycles to yield averaging values of θ and ϕ, from which the orbital elements can be calculated.

Error analysis based on a 400 statute mile circular orbit and assuming a 1σ error in angle determination of 1 arc minute, indicate the following 1σ errors in position and velocity over a 1 revolution smoothing period:

TABLE I
TYPICAL ACCURACIES OF 2-STAR HORIZON TRACKER METHOD

	1σ Position Error	1σ Velocity Error
X Component	335 ft	0.67 ft/sec
Y Component	674 ft	0.30 ft/sec
Z Component	291 ft	0.31 ft/sec

Of course, the accuracy figures depend somewhat upon initial conditions defining the orbital elements. These initial conditions can be established by ground-based measurements or by simple on-board techniques. Even with crude initial conditions, the smoothing techniques can be used to effect major improvements in updating the values.

The functional block diagram for a closed-loop Class III navigation system is seen in Fig. 4. The basic equipment consists of an astronertial platform, a horizon tracker, a clock and a computer.

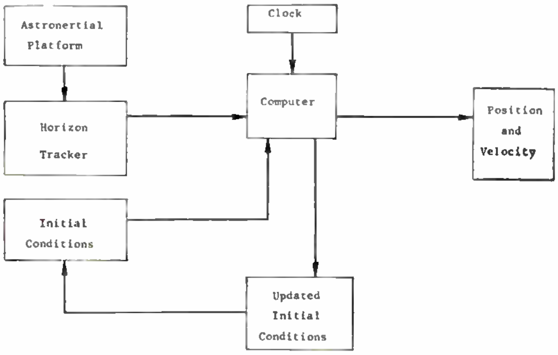


Fig. 4—Functional block diagram of a self-contained earth satellite navigation system (closed-loop method)—Class III.

The astronertial platform is used to maintain an inertial reference frame in terms of which the horizon tracker measures the angle describing the satellite-geocenter vector. The computer combines the measured angles with angles calculated from preliminary position predictions to obtain updated initial conditions. These updated initial conditions are used by the computer to make more accurate position and velocity predictions than can be made by an open-loop system. This method of periodic updating of initial conditions reduces inaccuracies due to errors in initial conditions and lack of completely accurate force field information.

The computer shown in the block diagram, Fig. 4, replaces the function generator shown in the block diagram of the open-loop system. In addition to this change from open-loop to closed-loop position and velocity determination, the system for the Class III satellites will require an attitude control system operable for a much longer time period than that required for Classes I and II.

The system described above is presented as an example of the methods of closed-loop position and velocity determination to be studied for Class III satellites. No attempt has been made here to describe a specific system, but rather the general method has been indicated. It is recognized that various instruments might be used, alternatively, in each of the instrument blocks.

Lunar Navigation

The navigational problems of lunar probes and manned lunar vehicles are receiving such detailed attention today that one cannot do justice to any given method or technique without a very detailed discussion. The purpose of this paper is not to go into such discussion but to point out the *trend* in lunar navigation which like earth-satellite navigation, once again shows a great emphasis on the concept of self-containment. Several types of lunar missions are in various stages of planning and progress. The guidance systems which have been used for space probes and those which are presently planned for other early programs have taken advantage of the current state of the art in systems and use off-the-shelf hardware. This generally results in a heterogeneous system consisting of independent launch and injection systems, in some cases separate for each boost vehicle stage; radio-controlled midcourse guidance with ground-based computation; and a variety of combination radio-optical-inertial systems for operation in the vicinity of the moon. The future *trend* will definitely be toward a nonradiating, fully integrated self-contained system. An excellent example is given by Hakes [7] in which the integrated guidance subsystem consists of an astronertial platform, a two-star tracker, wide-angle, and narrow-angle optical earth and moon trackers, an integrated computer-clock, and a radio altimeter. The utilization of these instruments by mission phase is shown in Table II (next page). Perhaps the most demanding of problems is the enhancement of navigational accuracies by self-contained methods. It is becoming increasingly clear that such accuracy enhancements can come more from employing statistical techniques than from over-design of sensors to achieve extreme accuracies, possibly jeopardizing cost and lifetime figures.

The statistical methods in the determination of the trajectory of a spacecraft are concerned with the problem of finding the "best" (in the sense of the statistically most significant) values of the position and velocity of a spacecraft, when the available data are redundant and, on account of the errors of measurement, contradictory.

As an example (Fig. 5), consider the radar R, at some point on the earth, which measures the range and direction (the azimuth and elevation angles) of a vehicle at V. From this data, a determination of the position of the vehicle (its x, y, and z coordinates, for example, in some reference frame) can be found. Here there is no problem, since one has a minimum of data. Three measurements, range, azimuth and elevation, are used to compute 3 position coordinates, and there is no redundancy.

Now, however, assume that another radar R', at another spot on the earth measures the corresponding quantities at the same time. From these measurements one can compute the x, y, and z coordinates of the vehicle in the same reference frame as before. Call these x', y', and z'.

Since each of these six measurements will inevitably be subject to error, no matter how carefully the measurements are made or how accurate the system, the position as calculated from R will not coincide exactly with the position calculated from R', and the question arises as to the choice of the position of the vehicle.

One simple answer would be to average the two positions:

$$\bar{x} = \tfrac{1}{2}(x + x')$$
$$\bar{y} = \tfrac{1}{2}(y + y')$$
$$\bar{z} = \tfrac{1}{2}(z + z').$$

Although this position would be better than either position separately, it does not generally represent the statistically best determination, *i.e.*, the most probable position in view of all available information.

Statisticians have worked with this problem, and have developed a simple criterion, referred to as the least squares criterion, by which the position of the spacecraft is so chosen that the sum of the squares of the errors is a minimum.

One complication immediately arises. The data include both distance and angle measurements. The question arises as to how to combine errors measured in feet with errors measured in degrees. To account for this difference, as well as possible differences in precision of various measurements of the same type, one minimizes the sum of the weighted residuals, where the weights are given by

$$w_i = \frac{1}{\sigma_i^2}$$

where w_i and σ_i are the weight and the standard derivation of the ith observation.

Suppose now that one makes a number of such observations at specified times. Since the vehicle during this interval must move in accordance with Newton's laws. the sum total of the known information consists of all the measured data and the laws of motion governing the trajectory, including all the forces acting on the vehicle.

TABLE II
EQUIPMENT REQUIRED BY MISSION PHASE

Subsystem	Mission Phase: Earth Launch	Earth Parking Orbit	Injection	Midcourse	Lunar Approach	Circumlunar Passage	Lunar Orbit	Lunar Landing	Lunar Launch	Return Midcourse	Earth Approach	Atmospheric Reentry
Astronertial Platform	X	X	X	X	X	X	X	X	X	X	X	X
Wide-Angle Earth-Moon Tracker		X		X	X	X	X	X	X	X	X	
Narrow-Angle Earth-Moon Tracker				X					X	X	X	
Computer-Clock	X	X	X	X	X	X	X	X	X	X	X	X
Radio Altimeter					X	X	X	X			X	X

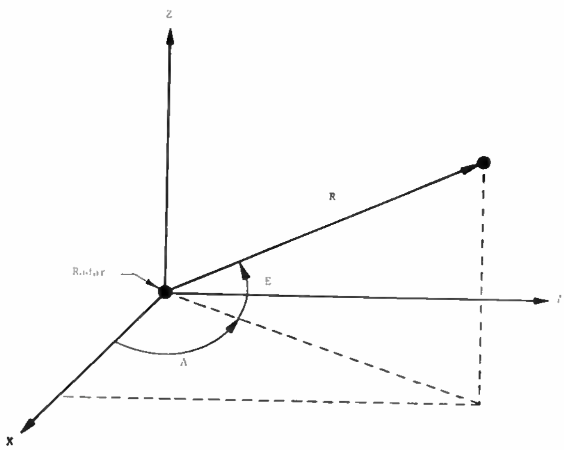


Fig. 5—Coordinates of spacecraft relative to ground-based radar.

A criterion for selection can thus be stated as follows: Select the trajectory which satisfies the laws of motion and for which the sum of the squares of the weighted residuals is a minimum.

In the case of a linear function, a least squares solution is a simple and straightforward computation to make. Unfortunately, in the case of a space vehicle, the differential equations of motion cannot be solved explicitly. This means that the trajectory must be calculated numerically and approximately, and the direct least squares solution cannot be made. Instead one is forced to expand the various functions in Taylor Series, of which one uses only the first term. For a given first approximation, the equations give an improved value for the trajectory. A second correction can give a still better estimate, and so on.

In many cases, the knowledge of the accuracy of the position and velocity determination is as important as the quantities themselves. Two methods are commonly used for finding these errors. The first, a numerical method, uses the Monte Carlo principle, in which a considerable number of trajectory cases are run with random errors added to the measurements. In the second, the usual formula for the propagation of errors is used. If the x-coordinate of position at a given time, for example, is thought of as a function of the measurements,

$$x = x(\psi_1, \psi_2, \cdots, \psi_n)$$

then the standard deviation of x, σ_x, is given by

$$\sigma_x^2 = \sum \left(\frac{\partial x}{\partial \psi_i}\right)^2 \sigma_{\psi_i}^2.$$

These calculations have been performed for a typical case of the re-entry vehicle from the moon. The Propagation of Errors Method gave the standard deviations in the six variables as

$$\sigma_x = \pm 155 \text{ ft} \qquad \sigma_{v_x} = \pm 0.50 \text{ ft/sec}$$
$$\sigma_y = 575 \text{ ft} \qquad \sigma_{v_y} = 0.20 \text{ ft/sec}$$
$$\sigma_z = 385 \text{ ft} \qquad \sigma_{v_z} = 0.52 \text{ ft/sec.}$$

The Monte Carlo method yielded

$$\sigma_x = \pm 172 \text{ ft} \qquad \sigma_{v_x} = \pm 0.41 \text{ ft/sec}$$
$$\sigma_y = \pm 438 \text{ ft} \qquad \sigma_{v_y} = \pm 0.22 \text{ ft/sec}$$
$$\sigma_z = \pm 411 \text{ ft} \qquad \sigma_{v_z} = \pm 0.48 \text{ ft/sec.}$$

This agreement is considered to be very good in view of the small number of samples.

Some Advanced Navigational Concepts

Advanced Navigation Requirements

The navigational requirements for aerospace vehicles differ from mission to mission and from phase to phase within a particular mission. In each mission, the navigation instruments, including the computer, must be capable of providing current position and velocity information and, in many cases, they must be capable of predicting the future position and velocity of the spacecraft. In addition, the navigation system in many cases is called upon to calculate maneuver ballistics and specify time, magnitude and direction of corrective thrusts.

The major variations in navigation requirements arise when one considers the accuracy and speed with which navigational data must be acquired and processed to provide the necessary guidance functions. In certain satellite missions, the navigational system must be capable of positional accuracies measured in feet, velocity accuracies of less than one ft/sec with a data acquiring and processing time measured in seconds. On the other hand, the navigational system of a deep space mission could tolerate midcourse positional errors measured in miles and take minutes, if not hours, to perform the necessary computations. However, in the interest of minimizing fuel requirements, maximizing payload, etc., it is desirable that the maximum accuracy and minimum time be obtained with due consideration for size, weight, power, cost and reliability. The consideration of these parameters and their trade-offs often lead to interesting and new navigational concepts. The trend is definitely pointed towards novel optical techniques which offer increased accuracy and reliability with reduced size, weight and power allocations.

Some Optical Navigational Techniques

An important distinction has to be made between an optical system in a vehicle within the atmosphere and a space vehicle. For example, a star tracker in a typical vehicle which has most of its flight profile within the earth's atmosphere, usually has a narrow field of view intended to minimize sky brightness and background noise and consists of a telescope system rigidly pointed towards a given star and held in track mode by a servo motor. The sky brightness and background noise are not so serious when the vehicle is above the atmosphere and hence the optical field of view can be increased, giving the opportunity to observe many celestial bodies. Mixed with celestial bodies are man-made objects which, in many instances, reflect solar energy, appearing like stars. Wide angle optical navigation systems must, therefore, have the ability to discriminate between natural celestial bodies and man-made moving targets. A typical servo control system which would enforce the pointing to a specified star of a tracker in an atmospheric vehicle, would create angular momentum unbalance in an exo-atmospheric spacecraft, unless compensated for. The need for preserving angular momentum equilibrium or balanced torques is therefore essential in spacecraft.

Some of these conflicting requirements and properties of space navigation systems suggest two possible configurations, a wide-angle nonrotating system and a rotating track-while-scan system.

Nonrotating Sensor Concept: In principle, this method considers the acquisition and monitoring of an adequate image of the entire spherical field of view available to a space vehicle. A conceptual sketch describing this concept is shown in Fig. 6. Such a device, which may be called a "Fly's Eye Camera" because of its ability to look through a wide angle and track different objects independently, may conceivably be developed by a mosaic type or sensor array in which multiple fields of view may be overlapped to give a broad picture of the celestial sphere. From a knowledge of the solar system and the celestial sphere, which may be carried in its memory, the Fly's Eye Camera could provide many valuable navigational data without serious control system requirements. There is no doubt that the on-board computer will have to be sophisticated because of the many functions and transformations it has to perform.

Track-While-Scan (TWS) Concept: This concept is an extension of the method of self-contained navigation presented by Lillestrand and Carroll [2]. In principle, the TWS instrument observes a portion of the celestial sphere by a rotating telescope. The instantaneous portion of the celestial sphere under observation is defined by a slit or slits in the image plane of the telescope. Fig. 7 illustrates a belt-like projection of the celestial sphere observed by a single scanning slit. The corresponding object space image of the slit is shown superimposed on the celestial sphere.

The azimuth angle between any two stars in the field of view may be measured in two ways:

1) If the platform supporting the rotating telescope remains fixed in space, an angle encoder may be employed and the angle of rotation of the telescope measured at the instant of transit of each star.
2) Alternatively, if the rotation rate ω is a constant, the periodic nature of the observed star pattern will permit an accurate calibration of this rate. The difference in azimuth angle $(\theta_2-\theta_1)$ between any two stars may then be determined by measuring the time interval (t_2-t_1) between each star crossing the slit, since $(\theta_2-\theta_1)=\omega(t_2-t_1)$.

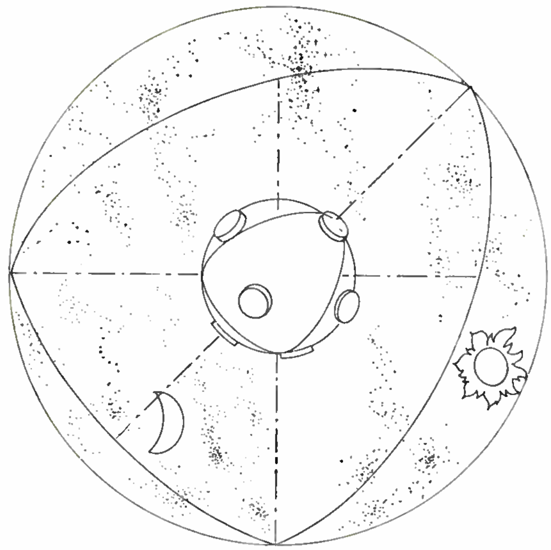


Fig. 6—Nonrotating tracking instrument.

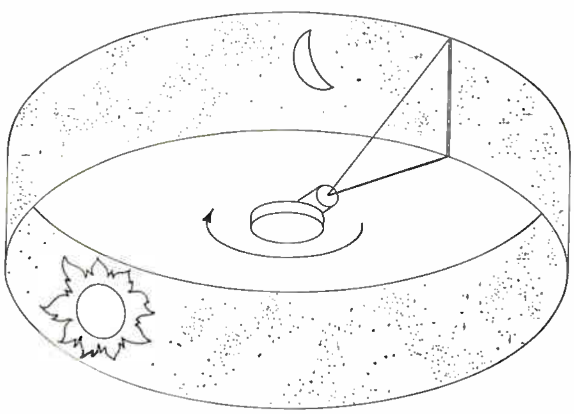


Fig. 7—Track-while-scan instrument.

The technique for precise measurement of azimuth angular separation of celestial bodies by means of a telescope rotating with a constant angular rate is similar to the well-known transit telescope. The aligned transit telescope is constrained so that its line of sight is always in the plane of the meridian. The constant angular rate, in the azimuth direction, is supplied by the angular rotation of the earth. During the manual operation of a transit telescope, the observer records the time at which each star crosses the instrument crosshair. In the TWS automatic instrument, the star transit across a narrow slit may be recorded as a photoelectric pulse from a photodetector placed behind the slit.

Advanced Concepts in Data-Processing Techniques

The successful deployment of self-contained navigation systems depends on successful development of reliable components of small size, weight and power allocations. The increased emphasis of accuracy and redundant logic, as well as explicit trajectory solution in completely self-contained systems, requires the use of a sophisticated, large-scale general-purpose data-processing machine. Some trends in data-processing computers may be pointed out.

Some factors which may be considered in optimizing the data processing are:

1) Equipment time sharing: The possibility of time sharing of an arithmetic module between two computational processes, where each process has a different time when peak load occurs.
2) Computer control: The problem may lend itself to unique addressing and instruction coding.
3) Input-Output: The realm of data conversion where perhaps the input-output unit can be greatly reduced. This appears particularly feasible if the system is self-contained. The problem then becomes one of designing the subsystems to operate using machine language inputs.

In addition to the factors listed above the design and development necessary to build a digital computer that will perform within high reliability standards, perform all desired functions with a speed and accuracy dictated by system specifications, and designed within critical weight and size limits, seems to center around the application of state-of-the-art considerations of packaging philosophy, electronic component microminiaturization construction, advanced circuit and logical design techniques, and a thorough testing and quality control program.

To obtain maximum packaging density a required amount of development must be expended in the area of efficient packaging of miniature and microminiature components. The state of the art in electronic packaging is changing at a rapid rate as shown in Fig. 8, which is a graphical representation of the variation of component packing density as a function of time. As indicated in Fig. 8, there are three methods of achieving microminiaturization of electronic equipment. All may be considered for application to space flight, and may be grouped as follows:

1) Single Dimension: This method uses thin films for resistors and capacitors and separate microcomponents for inductances and semiconductors. A single type of element (*i.e.*, resistor or capacitor) is deposited on each wafer. This method is currently in use and promises a density of up to 600,000 components per cubic foot.

2) Two Dimension: This method uses thin films for resistors and capacitors, as above, but combines them on a single wafer. The combining reduces external interconnections leading to an over-all increase in packing density. The disadvantage is that the yield goes down due to the multiple operations involved. When the problems are solved, packing densities of up to 2.5 million components per cubic foot can be realized.

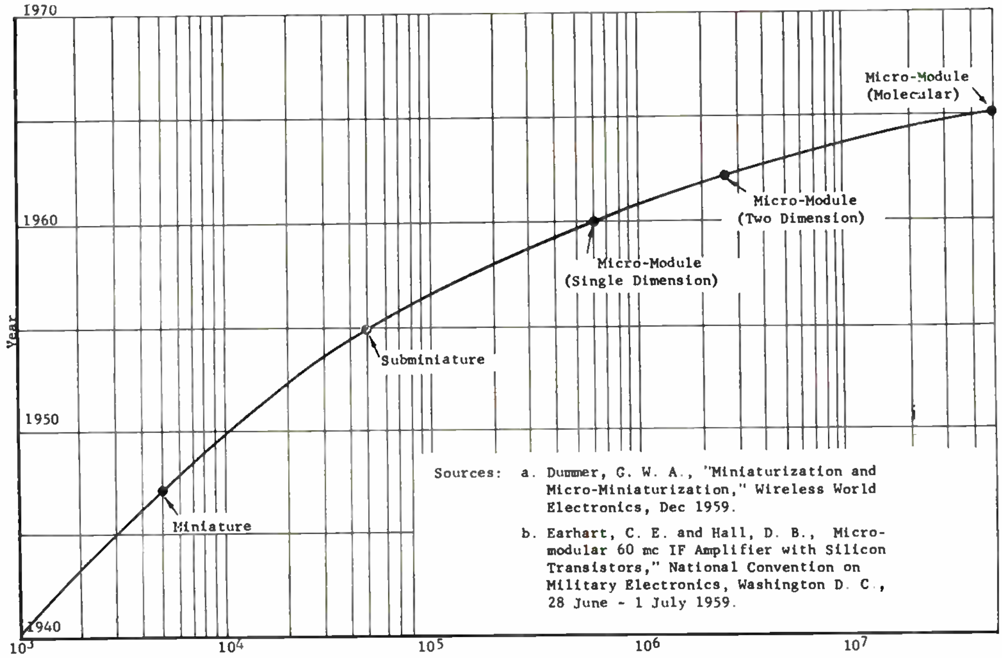


Fig. 8—Trend of packing density.

3) Molecular: In this method, resistors, capacitors, inductors, diodes, transistors, and other electronic devices are mounted together on one surface of a semiconductor wafer. Packing densities of from 20 to 100 million components per cubic foot are expected; however, major improvements in connectors, heat transfer methods, and intercommunication between circuits are required to fully utilize the high densities possible.

Advanced Concepts Using Optical Doppler Techniques and Lasers for Space Navigation

A comparison between radar and optical techniques for making velocity and distance measurements will soon convince the reader that the optical techniques offer higher accuracies, narrower beamwidths and all-round better performance than radar, especially in terms of size, weight and power requirements. A brief review is now presented of the use of optical Doppler techniques and lasers for space navigation.

Two papers by Franklin and Birx [8], [9] describe studies carried out for the USAF investigating the use of natural electromagnetic radiation for navigational purposes. The emphasis is on the measurement of velocity utilizing the Doppler phenomenon.

Franklin and Birx use the well-known fact that when there is relative motion between a source of radiation and an observer, the radiation observed is shifted an amount $\Delta\lambda$ which is proportional to the relative velocity according to the equation

$$\Delta\lambda = \lambda_1 - \lambda_2 = \frac{v\lambda_1}{c}$$

where

v = relative velocity of source with respect to observer
c = velocity of light
λ_1 = wavelength of radiation observed with no relative motion
λ_2 = wavelength observed with relative motion

from which one obtains the velocity equation

$$v = \frac{c(\lambda_1 - \lambda_2)}{\lambda_1}.$$

Franklin and Birx base their error analysis upon the velocity error equation

$$\delta v = \frac{2c\delta\lambda}{\lambda_1}$$

where $\delta\lambda$ = error in determination of wavelength. In their first paper [8] a method is described using a slit system constructed to allow the radiation from a spectral line to be split so that each half of the flux falls on a separate detector. The outputs of the two detectors would be balanced at zero velocity. Utilizing such a system, a limitation on accuracy of ± 60 ft/sec was the theoretical figure obtained. In their second paper [9], Franklin and Birx describe a method of computing the shift by comparing the spectral lines of the received radiation with a template consisting of a photographic negative of the spectral lines and measuring the displacement of the template necessary to return to a matched condition. The authors find that with such a system it should be possible to obtain sensitivities of from one to 10 m/sec but at the time of the writing of the paper, the work had not progressed far enough to provide evidence of this order of sensitivity.

It should be noted that all such methods of utilizing the Doppler phenomenon can give only the component of velocity in a radial direction from the source. Also, all measurements made thus far have been made with respect to the sun and great advancements in the state-of-the-art are required if radiation from other stellar bodies is to be used.

A paper by Norton and Wildey [10] is concerned with fundamental limitations to optical Doppler measurements for space navigation. The authors consider the result of various physical phenomena which produce asymmetric broadening of the spectral lines in the radiation from stellar bodies. The phenomena considered are: second-order Zeeman effect, quadratic Stark effect, pressure shift, sunspot effect, and variability of radius. Calculations by the authors using the sun as an example result in their concluding that the physical characteristics of stellar electromagnetic sources such as the sun vary in such a manner that a variability of approximately ± 200 ft/sec may be expected in the observer's measured optical Doppler velocity. Norton and Wildey express the opinion that it is this which imposes a fundamental limitation on the accuracy with which Doppler measurements may be made optically rather than the instrumentation.

One of the most interesting discussions on the determination of distance and velocity using optical maser principles is found in a paper by Ellis and Greenwood [11]. The particular features contributing to the advantageous application of optical masers to space navigation are:

1) The systems could be very small in size, and self contained,
2) Power requirements could be low even for long distances,
3) Velocity measurements will be very accurate.

The availability of the extraordinary degree of parallelism and the narrowness of frequency bandwidth achievable in the laser makes this instrument extremely noteworthy for space navigation applications. Another attractive feature of the laser is the extraordinary capacity of the laser receiver for discriminating against extraneous background light and to its extremely low internal noise level.

The determination of distance from a laser transmitter to some scattering object can be accomplished by standard methods similar to radar and sonar. Ellis and Greenwood present an example of a two-foot diameter transmitter and receiver optical maser consuming 66 watts of average power yielding an accuracy of one part in 100,000 for the measurement of distance between two spacecraft 160,000 km apart. The most serious disadvantage for many space applications is the requirement for transmitter and receiver beams to be trained in the same direction and its poor performance in the search mode.

Another interesting application of the laser is the high accuracy with which one can measure the altitude of a spacecraft above a very small spot on the surface of a planet. This is in contrast to the microwave situation where, because of unavoidable beamwidth, the measurement always constitutes an average of distances to a broad range of points below.

Ellis and Greenwood [10] extol the virtues of the laser as a Doppler velocity measuring device. They present an interesting example of a vehicle for soft lunar landing and point out that an accuracy of a tenth of one per cent in velocity is reasonable for lunar approach velocity of about 2300 knots at an altitude of about 1600 km above the moon's surface. Using a similar example for a planetary approach, they predict a self-contained system of about one cubic foot volume and about ten pounds weight.

The laser transmitter is now available; a space-worthy laser receiver is still not available. The future, however, holds promise for more advanced laser developments and hence their use in space navigation.

Advanced Concepts of Guidance Management

The majority of space boosters which are deployed for launching orbital payloads were developed originally for military missile applications. It is therefore quite common to see each stage of a multistage vehicle equipped with its own guidance and control system for thrust vectoring, booster control and engine cutoff. The payload seldom has a major part in the navigational functions of the individual stages of the booster which places the payload along its path. It seems possible that savings in weight, instrumentation costs, and ground-based equipment could be obtained by bringing the "payload into the picture," especially in advanced and future space missions for which both the payload and the boosters can be specially designed to fit each other's requirements. If one accepts the premise that future payloads will include the requirement of self-contained navigation, the final upper stage will have all the usual equipment, such as gyros, accelerometers, star and planet trackers, clocks, radio frequency equipment and a fairly large general purpose computer, necessary to guide the spacecraft during all phases of flight (Fig. 9). Such an integrated, self-contained system would seem very desirable for providing all the necessary navigation and guidance facilities for the boost phase, the orbital phase, midcourse, landing or return phase, etc. Thoughts such as these have been expressed by Cashmore [12].

A basic problem is created when the inertial sensors of the multistage system are located at a station or stations which are an appreciable distance from the center of gravity of the vehicle. During launch operations, the payload at the top of the booster does not experience the

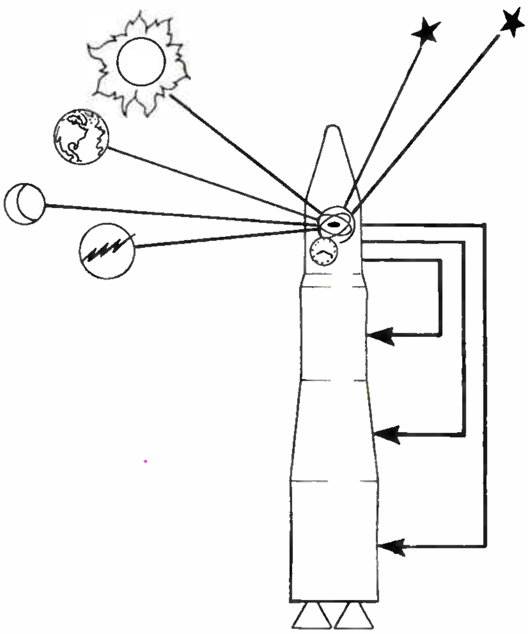


Fig. 9—Payload guidance management.

bending and c.g. separations and accelerations. This compensation or correction must be allowed for in the thrust-vector control and engine cutoff of the multi-stages. Secondly, the optical windows of the optical instrumentation contribute errors during launch and the tracking of celestial bodies is subject to refraction and shimmering until the vehicle approaches very high altitudes. Compensations for these errors will be necessary. Third, aborts on the launch pad are still better handled by ground crew than by the payload, and provisions for individual engine controls are necessary.

The future trend will, however, be an evolution of the guidance management concept by which the self-contained payload guidance and the ground-based command and control center jointly "manage" the guidance functions to achieve thrust-vector control and engine cutoff of the individual stages, as well as the navigation of the orbital spacecraft, all functions to be performed mostly by the use of and within the available instrumentation of the final upper stage itself.

Acknowledgment

A review paper such as the preceding necessarily reflects the concepts and creative efforts of many individuals, frequently informally transmitted and therefore difficult to acknowledge. Specifically cited, however, are the contributions of W. J. Reilly, Dr. J. Titus, G. D. Barnes, R. B. Freund, Dr. J. K. Robe, and R. C. Hakes.

References

[1] R. F. Hoelker, "The Evolution of Guidance Theory and Trajectory Analysis into a Single Scientific Discipline," presented at Inst. of Navigation Meeting, Williamsburg, Va.; June, 1961.

[2] R. L. Lillestrand and J. E. Carroll, "Self-Contained System for Interplanetary Navigation," presented at American Astronautical Soc. Meeting, August, 1961.

[3] K. N. Satyendra and R. E. Bradford, "A Self-Contained Navigational System for Determination of Orbital Elements of a Satellite," ARS Paper, No. 1495-60; December 5–8, 1960.

[4] F. Stevens, "Space navigation by star and planet tracking techniques," *Proc. 5th Annual National Space Electronics and Telemetry Symp.*, Washington, D. C., sect. 10.3; September, 1960.

[5] ——, "Application of Optical Techniques to Interplanetary Navigation," presented at 3rd AFOSR Symp., Los Angeles, Calif.; October, 1960.

[6] T. J. Jaramillo, "The Combined Non-Linear Effects of Earth Oblateness and Atmospheric Drag on a Near Earth Satellite," U. S. Dept. of Commerce, Washington, D. C., O.T.S. 161494, 250.

[7] R. C. Hakes, "Guidance and Navigation Systems for Lunar Missions," ARS Space Flight Rept. to the Nation, October, 1961.

[8] R. G. Franklin and D. L. Birx, "A study of natural electromagnetic phenomena for space navigation," Proc. IRE, vol. 48, pp. 532–541; April, 1960.

[9] ——, "Optical Doppler for Space Navigation," presented at ARS Guidance, Control and Navigation Conf., Stanford University, Calif.; August, 1961. ARS Preprint No. 1940-61.

[10] R. H. Norton and R. L. Wildey, "Fundamental Limitations to Optical Doppler Measurements for Space Navigation," presented at ARS Guidance, Control and Navigation Conf., Stanford University, Calif.; August, 1961. ARS Preprint No. 1942-61.

[11] C. B. Ellis and I. A. Greenwood, "Optical masers in space navigation," *Inst. Navigation J.* Also presented at 17th Annual Meeting; June 30, 1961.

[12] D. J. Cashmore, "The role of inertial equipment," *J. British Interplanetary Soc.*, vol. 18, p. 146; July—August, 1961.

Section 27
ULTRASONICS

Organized with the assistance of the IRE Professional Group on Ultrasonics Engineering

Uses of Ultrasonics in Radio, Radar and Sonar Systems*

WARREN P. MASON†, FELLOW, IRE

Summary—A short historical review is given of the development of ultrasonics. Uses in the low-amplitude range are sonar systems, dispersive and nondispersive delay lines, ultrasonic inspectoscopes, and crystal and mechanical filters. In all of these applications, the very low internal friction of metals and fused silica provides very low losses, while the guided wave properties of various shaped transmission paths provide either dispersive or nondispersive transmission. The dispersive transmission is useful for FM radar systems which are not limited in power by breakdown of the system, while nondispersive delay lines have uses in MTI radar systems, in pulse decoding systems and in computers. The low dissipation and frequency stabilities inherent in quartz crystal vibrations have been used extensively in the production of very selective filters and in the control of oscillators. Ultrasonics and hypersonics (frequencies about 10^9 cycles) are useful tools for investigating liquid and solid-state reactions.

I. Introduction and Review of Uses of Ultrasonics

The introduction and development of ultrasonics spans a time of about 80 years. Ultrasonics (originally called supersonics, a name now used to describe speeds faster than sound) can be said to have started in 1883 with the Galton whistle. Capable of producing sound waves of a frequency higher than can be heard by humans, such whistles were used to demonstrate the reflection, refraction, interference and propagation of sound. A minor use was found in summoning dogs without alerting humans.

No further uses were found for ultrasonics until the advent of World War I, when an intensive program on submarine detection was carried out. Since sound waves are the only type of wave propagation that can detect submerged submarines, considerable effort was spent generating and detecting such waves. The most sucessful effort [1] was carried out for the French government under the direction of P. Langevin. Using a sandwich made of quartz crystals cemented to steel plates, a directive sound beam was obtained which could be reflected off sunken objects and results were obtained which compared with those achieved up to the beginning of World War II. While this effort came too late to be of use in World War I, it laid the foundation for the extensive applications in World War II for both detecting and destroying submarines (by ultrasonic homing torpedoes) which contributed materially to Allied victory.

Between the two wars, this principle was applied to ultrasonic fathometers which quickly and accurately measure the depth of the ocean. This technique has been used to map the ocean bottom and has located deep trenches, incipient underwater plateaus, islands and mountain ranges. The ultrasonic inspectoscope was an application of the same principle to the inspection of solid pieces for flaws, and for determining the thickness of materials such as boiler plates for which only one side is accessible. In this country the first publications and patents were those of Firestone [2]. Such methods are among the best for detecting small imperfections and cracks in metal castings, glasses and ceramics and in inspecting the quality of adhesive bonds in automobile tires and other products.

* Received by the IRE, June 23, 1961; revised manuscript received, October 13, 1961.
† Bell Telephone Laboratories., Murray Hill, N. J.

The advent of World War II stimulated not only ultrasonic research in the underwater sound or sonar field but also ultrasonic adjuncts to radar transmission. It was during this time that the liquid and solid ultrasonic delay lines were developed for use as timing devices, anti-jamming devices and moving target [3] indicator radar systems which accentuate the appearance of moving targets by minimizing the appearance of reflections due to stationary targets. In this use one frame containing all the reflections during one pulse period is stored in a delay line having a delay time equal to the period between successive pulses. These pulses are made to balance out the radar reflections coming back from the next pulse with the result that fixed targets are minimized but moving targets are not balanced out. Such delay lines have also been used in pulse decoding systems [4], and for information storage for computers [5]. Recently dispersive delay lines have become of interest [6] for high-power radar systems. In this use they convert a short high-power density amplitude modulated pulse into a long low-power density frequency modulated wave which does not overload the final amplifier stage. The reflected frequency modulated wave is converted back into a short amplitude modulated wave by sending it through a network having the reverse dispersion characteristic of the sending network. This restores the accuracy of the distance and direction determination.

The high Q (low dissipation) inherent in mechanical vibrations in solids has been used in obtaining very selective filters and in the control of oscillators. The first and by far the largest use in filter networks is the use of quartz crystals as impedance elements in electrical filters [7]. These are used on all of the channel filters of the long distance carrier systems, microwave systems and submarine cables of the Bell Telephone System to separate a given conversation from a group of twelve similar conversations in a channel bank. Recently, such crystals have been grown synthetically [8] from small scraps of quartz or even from sand. This process eliminates the variability of the source of supply and produces crystals of an economic size for cutting which are free from the usual defects of natural crystals. An even larger use for the high Q present in quartz crystals is in controlling the frequency of oscillators and time standards. During World War II, over 30,000,000 crystals were produced for this application. The present effort [9] is directed at improving the frequency stability and long time aging effect present in such crystals. While the final stability has been eclipsed by that obtainable in an atomic standard, quartz crystals are still the most used secondary standard and serve as the fly wheel even for atomic standards. Considerable effort [10] has also been expended in using the high Q present in metals and other solids in all-mechanical filters driven by electromechanical transducers for converting from electrical energy into mechanical energy and vice versa. This is discussed further in Section IV.

All these devices can be considered as low-power devices since they utilize the linear ranges of the transmission media. A number of high-power devices are in general use. These employ nonlinear effects in the transmission media to produce processing effects. These are discussed in a companion paper by F. Massa.

Ultrasonic waves have also proved to be a convenient method for investigating the effects of imperfections and solid-state motions on the properties [11] of materials used by the radio engineer. These ultrasonic and hypersonic waves have proved to be useful tools in investigating the fundamental properties of liquids and solids. In this use the ultrasonic waves produce liquid or solid-state motions which react back on the attenuation and velocity of waves traversing the medium. For example, in a ferromagnetic crystal the stress waves produce motions of the Bloch walls that divide adjacent ferromagnetic domains. This motion produces an extra strain component, on account of the change in the magnetostrictive expansion due to the motion of the domain walls, which lowers the effective elastic constant of a demagnetized sample. When this sample is completely magnetized, the domain walls disappear and the elastic modulus increases. This effect is known as the ΔE effect and its variation with frequency [12] is correlated with the distribution of domain sizes. The motion of the domain walls produce eddy currents which act as energy loss mechanisms and produce the micro-eddy current effect.

For associated, nonassociated and ploymer liquids and rubbers there are alternate positions of molecules or chains of polymer molecules which are occupied successively under the action of thermal oscillations. An ultrasonic wave can produce stresses that alternately make one or the other of the positions more probable. When the frequency of the ultrasonic wave coincides with the frequency for which thermal agitation causes the molecules to alternate their position, an attenuation peak known as a relaxation occurs. Ultrasonic waves have located isolated relaxations and groups of relaxations in liquids which give information on the types of motion taking place. In solids relaxations have been measured connected with thermal effects, domain and grain boundary effects, and the motions of dislocations.

Recently ultrasonic waves have contributed materially to a knowledge of the motion of electrons at cryogenic temperatures. At room temperatures the motion of the lattice caused by ultrasonic waves can communicate energy and momentum to the conduction electrons but since the mean free path of electrons in the lattice at room temperature is so small, most of the energy is given back to the lattice in the same phase and is not lost to the vibration. As the temperature is lowered, the mean free path gets very large and in pure materials may become considerably larger than the wavelength. Fig. 1 shows [13] the attenuation of longitudinal sound waves for two directions and two frequencies in a pure tin single crystal. When the mean free path is smaller than the acoustic wavelength, the damping is equivalent to a viscous effect (*i.e.*, transfer of momentum) and the attenuation increases proportional to the square of the frequency. On the other hand when the mean free path is greater than the acoustic wavelength, theory shows that the attenuation is proportional to the frequency. The dividing line appears to be about 8°K from Fig. 1. As the temperature drops below the superconducting temperature of 3.73°K for tin, the attenuation rapidly approaches zero. The form of the curve agrees with that predicted by the Bardeen-Cooper-Schrieffer theory of superconductivity. By putting on a magnetic field, the normal conductivity can be restored and the top curve of Fig. 1 shows the extension of the normal attenuation down to lower temperatures.

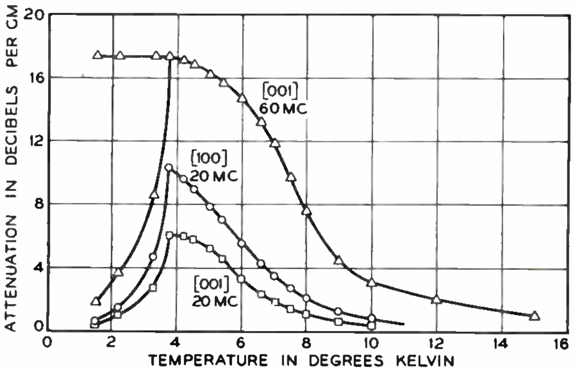


Fig. 1—Longitudinal wave attenuation for single crystal tin along the [001] and [100] axes.

Magnetic fields above the superconducting temperature have been used to delineate the shape of the Fermi surface, which is a constant energy surface which bounds the occupied states of electrons in momentum space. The electrical effects in a metal are primarily determined by the electrons of energy near the Fermi surface since these are the only ones free to move. For free electrons such surfaces are spherical in shape but in the band theory approximation, the Fermi surface for monovalent materials such as copper, gold and silver take the form [14] shown by the inside surface of Fig. 2. The outside surface is that for the Brillouin zone which delineates the surface in momentum space for which electrons suffer Bragg-type reflections. When the Fermi surface approaches the Brillouin zone a distortion of the spherical surface occurs as shown by Fig. 2. The effect of a magnetic field is to localize the electrons that can move onto a plane perpendicular to the magnetic field. It can be shown that the periodicity of the attenuation-H^{-1} curves can be related to the linear dimension of the Fermi surface perpendicular to the magnetic field and perpendicular in momentum space to the direction of wave propagation in real space. Measurements such as those shown [15] by Fig. 3 give details of the Fermi surface for different directions in momentum space.

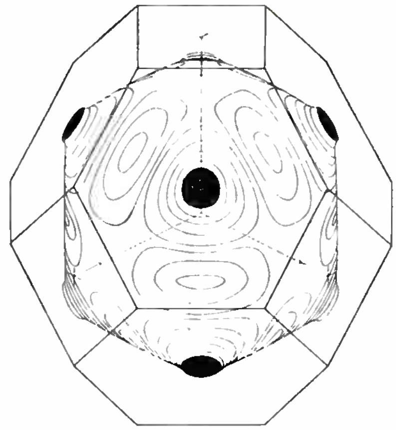


Fig. 2—Fermi surface and Brillouin zone for copper (after Pippard).

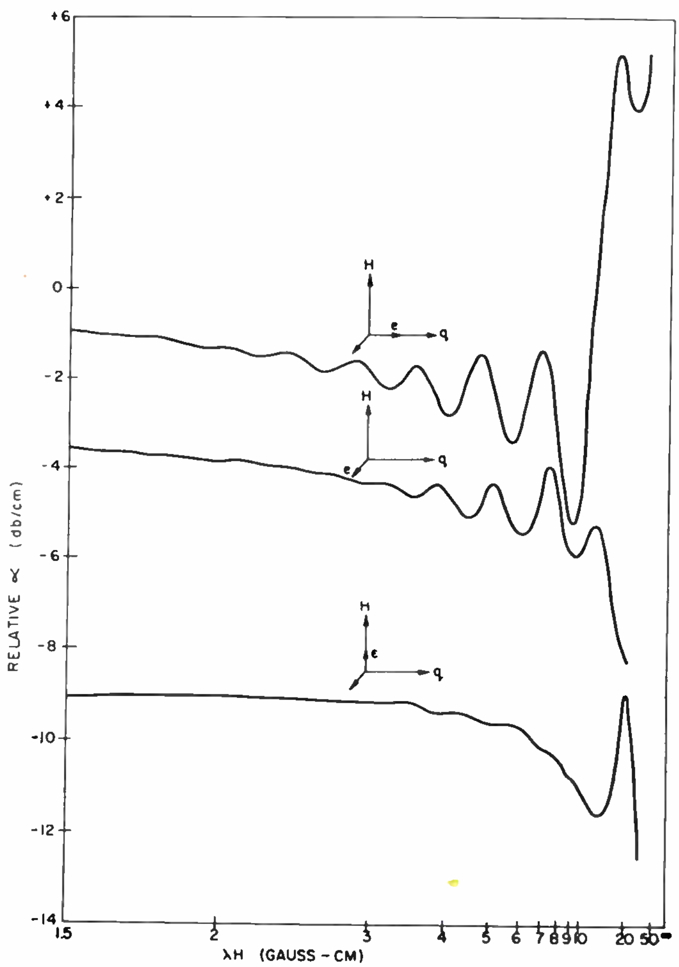


Fig. 3—Relative attenuation in pure single crystal copper as a function of the product of the wavelength times the magnetic field, for several orientations of the magnetic field and wave direction (after Morse and Gavenda).

Hypersonics has provided information on the conversion of sound waves to thermal energy which occurs in the form of quantized sound waves known as "Phonons." At low temperatures there are many fewer phonons and their mean free path and relaxation time τ between collisions becomes much larger. When the product of the angular frequency ω of acoustic waves propagating in the medium, and the relaxation time τ becomes greater than unity, the direct interchange of energy between acoustic waves and phonons gets much smaller and very little attenuation is experienced by the acoustic waves. Fig. 4 shows measurements due to Bömmel and Dransfeld [16] on the attenuation of shear waves in quartz for frequencies of 10^9 cycles and higher. Below about 15°K the attenuation becomes very small. This finding may have some practical uses in delay lines for storing large amounts of information since the bandwidth and hence the number of bits that can be stored at these frequencies can be very large for a given length. With the low attenuation possible at temperatures below 15°K, an enormous amount of information can be stored and retrieved at a subsequent time depending on the length of the rod.

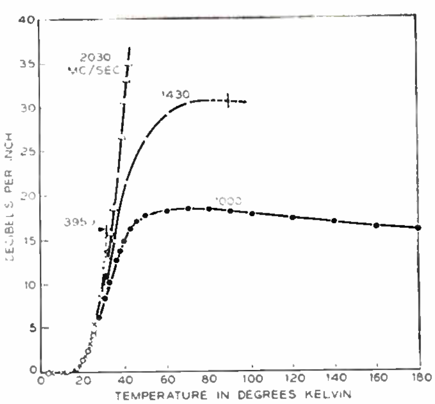


Fig. 4—Attenuation of shear waves in quartz at 1000 Mc (after Bömmel and Dransfeld).

II. Delay Lines

Section I gives a short summary of the various applications of ultrasonics. All of these probably affect the science of radio communication to some extent, but there are two applications very directly employed in radio, radar and sonar systems and these will be discussed in more detail in the last three sections. These applications are dispersive and nondispersive delay lines and ultrasonic mechanical filters. Neither of these types of transmission media can be connected with electrical circuits without the intermediary of an electromechanical transducer so that a short discussion of transducer properties is given. The properties of transducers are the determining factor in sonar systems.

There are of course all-electrical delay lines of the dispersive and nondispersive types but as the delays get larger and the frequency range covered becomes greater, the number of electrical elements becomes prohibitively large and it becomes cheaper and more satisfactory to resort to mechanical wave transmission. This follows because the velocity of transmission is only about 1 part in 10^5 as large as that for electrical waves in distributed conductors and because the Q (inverse of internal friction) is better for the best mechanical materials than for the best electrical lines in a comparable frequency range. The region for which it is better to shift from electical to mechanical probably is in the order of 10 μsec for bandwidths in excess of 1 Mc.

Delay lines initially used were usually of the water or mercury type but these types have mostly been replaced by solid delay lines. Such types are more compact, less subject to temperature changes and can stand considerably harder treatment. In the high-megacycle range solid lines are mostly constructed from strain-free fused silica, because this material has a very low attenuation and because sound scattering associated with a grain structure does not occur. According to the best measurements [17], the attenuation for longitudinal waves follows an equation of the form

$$\text{atten. (db per cm)} = 3 \times 10^{-4}F + 1.9 \times 10^{-5}F^n$$

where

$$n \doteq 2. \tag{1}$$

In this equation the attenuation is in db per cm and the frequency F in megacycles. For a delay of 1000 μsec which corresponds to a length of about 600 cm, the loss is about 11.5 db at 25 Mc. Shear wave losses have not been as carefully measured, but McSkimin finds that the Q in the low megacycle range is about the same as for longitudinal waves. The lower velocity (3.764×10^5 cm/sec compared to 5.968×10^5 cm/sec for longitudinal waves) gives a path length 63 per cent of that for longitudinal waves, but the loss is about the same.

Most of the fused silica delay lines are constructed in the form of many-sided polygons [18a] such as shown by Fig. 5, or wedges [18b]. In all of these space delay lines, shear waves are used because with the particle displacement parallel to the reflecting surface, only shear waves are reflected and at an angle of reflection equal to the angle of incidence. The dashed lines of Fig. 5 show the paths through the line from the transmitting crystal A to the receiving crystal J.

In these types of space lines, the beam spreads out as a function of distance and is much wider at the receiving crystal than at the transmitting crystal. Beyond the Fresnel region, which ends when the distance from the transducer d is

$$d > A^2/2\lambda, \tag{2}$$

where A is half the dimension of the crystal in a direction parallel to the principal face and λ the wavelength, the beam spreads out at a constant angle and

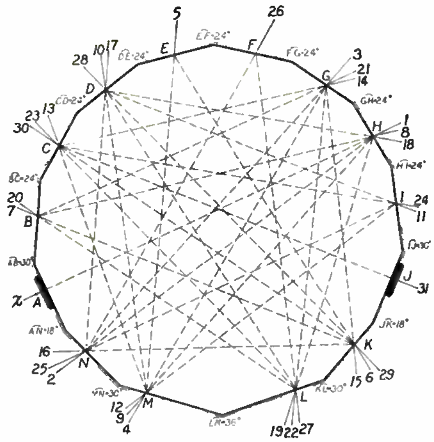


Fig. 5—Polygon shaped fused silica delay lines. Dashed lines show direction of wave propagation. Outside lines indicate damping regions to reduce spurious pulses (after Arenberg).

gives a loss in pressure approximately determined by the ratio

$$\frac{p}{p_0} = \frac{A^2}{d\lambda}. \tag{3}$$

The effect of this spreading is to produce unwanted reflections at times different from that for the main pulse. This can be minimized by putting absorbing material such as solder on the surfaces that are not required for the beam reflection.

The 15-sided polygon, which is used in most fused silica delay lines, proves to be almost an optimum compromise between maximum delay per unit area and maximum suppression of unwanted responses. With these type lines, center frequencies of 10 to 70 Mc, bandwidths up to about 50 per cent, losses of 40 to 70 db, suppression of unwanted reflections of from 30 to 40 db and delays up to 3500 μsec have been achieved with the largest fused silica pieces available.

The application of ceramic transducers [19a], [19b] to delay lines has resulted in the elimination of transducer loss since, as shown by Section IV, such transducers can be designed as band-pass filters with impedances which can be made to match the delay line. This has made possible a whole new field of delay line activity in producing short, compact low loss delay lines for pulse coding purposes [19b]. When thickness-shear ceramic transducers were developed [20], it became possible to operate a 700 μsec polygon line under matched termination conditions with a reduction in loss from 40 db to 20 db. The remaining loss is accounted for by absorption in the fused silica and by spreading losses.

The other type of line, that has been widely used, is the guided wave type such as the wire or strip lines. These types are useful not only for nondispersive lines but for dispersive lines as well [6]. Material such as 5052-H32 aluminum alloy has very low losses ($Q = 100,000$) while materials such as Ni-Span-C have a relatively low loss and a small change in velocity with temperature. Mathematical solutions for the propagation of longitudinal waves in rods were first given by L. Pochhammer and C. Chree [21a], but they were not studied extensively until the work of Bancroft [21b] and Hudson [21c]. The first systematic evaluation of all the modes in a wire or in a strip, including group velocity calculations, were made in References [6b] and [6c]. The first longitudinal mode in a wire has an inflection point, which leads to an approximately linear delay vs frequency characteristic. This is evident from the results of such a calculation [6b] as shown by Fig. 6. This figure shows the ratio of the phase velocity to the bar velocity $V_0(V_0 = \sqrt{Y_0/\rho})$, where Y_0 is the Young's Modulus and ρ the density, as a function of the diameter times the frequency divided by V_0. The result depends on the Poisson's ratio δ which is 0.33 for aluminum and 0.26 for Ni-Span-C. As can be seen, the ordinate should not exceed about 0.3 if a nondispersive transmission is desired. For frequencies of transmission up to a megacycle, the diameter should not exceed 0.15 cm. Wires of this type have been used in pulse decoding systems [4] and for information storage for computers [5]. On account of the small size, it is usually necessary to drive them by magnetostrictive means and loses of from 40 db to 50 db are common. Since such transducers cannot take all the energy out of the wave, it is generally necessary to terminate them in an extra length of highly damped material.

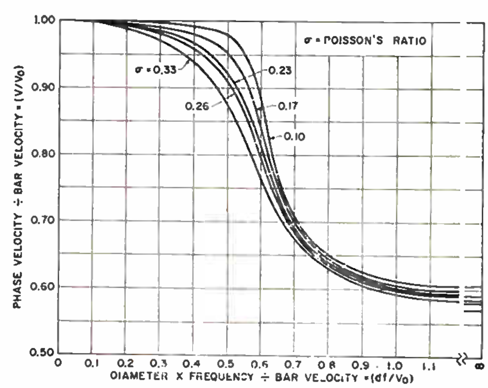


Fig. 6—Phase velocity as a function of the diameter frequency product for several values of Poisson's ratio (after May).

Magnetostrictive delay lines have the advantage that the delay can be made variable by changing the position of the transducers and it is fairly simple to obtain multiple outputs with different delays by incorporating more than two transducers. They are limited in frequency range by scattering by grains and surface imperfections.

Torsional waves can also be sent down wires and they have a mode which is nondispersive as long as the relation given by (4) is satisfied:

$$f < \frac{3}{4} \times \frac{5.136V}{2\pi a}, \tag{4}$$

where V is the torsional (shear) velocity of the wire and a the radius. Such torsional waves have been driven by two opposing longitudinal magnetostrictive wires [22a], by torsional barium titanate crystals [22b] and by the Wiedemann effect using [22c] a circumferential field in a magnetostrictive material to set up torsional oscillations. By sending a pulse of current down the magnetostrictive wire, combined with a preferred spacing of pick-up coils, it is possible [22d] to generate or decode a series of pulses.

Quite recently dispersive delay lines have come into considerable use [6] on FM radar systems. These are able to obtain power outputs over a good share of the repetition cycle, rather than just for the initial pulse. In principle, the action [6a] can be thought of as equivalent to dividing the width of the pulse frequency spectrum (equal approximately to $2/\tau$, where τ is the time of the sending pulse) up into a number of equal frequency ranges. If the dispersion is linear these ranges are then delayed by amounts which increase in equal increments for successive ranges with a total delay between the lowest and the highest frequency equal to the period of the radar repetition rate. In this way energy is being transmitted during the complete time interval. For the reflected radar pulse, the lowest frequencies return first and these are fed through a network which delays them most and the highest frequencies least. In this way the distance discrimination is restored. It can be shown [6a] that the discrimination factor D which in an amplitude modulation pulse system is defined as

$$D = T/\tau, \tag{5}$$

where T is the time between pulses and τ the time length of the pulse, becomes equal to, for a FM system,

$$D = (\delta_2 - \delta_1)(f_2 - f_1), \tag{6}$$

where $\delta_2 - \delta_1$ is the difference in delay between the lowest and highest frequency transmitted and $f_2 - f_1$ is the frequency range transmitted.

The dispersive properties of the wire or strip delay lines [6b–d] can be made use of in such systems. The phase velocity for a wire is shown by Fig. 6. The group delay is defined as

$$\delta = \frac{\partial B(\omega)}{\partial \omega}, \tag{7}$$

where $B(\omega)$ is the total phase shift and ω is 2π times the frequency f. Fig. 7 shows the delay as a function of frequency for the materials shown on Fig. 6. Similar delay curves [6c] are obtained for a flat strip driven by a longitudinal transducer. By mounting absorbing material along the thin edges, unwanted modes can be made as small as 55 db down with respect to the desired mode. When such lines are operated in the lower megacycle range a product of the time variation $(\delta_2 - \delta_1)$ times the bandwidth $f_2 - f_1$ of 100 has been achieved. It has been demonstrated that dispersive delay lines can replace the electrical version of the linear delay network containing hundreds of coils and capacitors and occupying 50 to 100 times the volume.

The nondispersive strip delay line of Meitzler [6d] using the shear mode has great potential for producing longer delays than possible with present fused silica delay line techniques while at the same time being very free of unwanted responses.

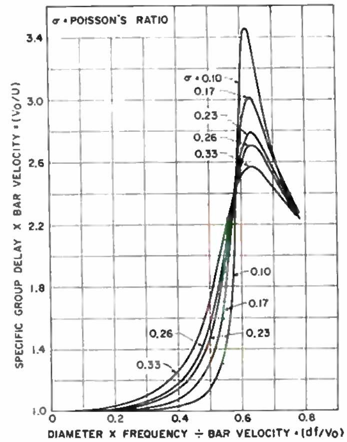


Fig. 7—Specific group delay as a function of the diameter-frequency product (after May).

III. Transducers and Resonators

All types of delay lines and indeed all electrical to mechanical ultrasonic devices require transducers to change electrical energy into ultrasonic energy and vice versa. There are two types of transducers in general use. These are the piezoelectric type (which includes a polarized electrostrictive material such as barium titanate as a special case) and magnetostrictive material of which the ferrites are the latest examples. The piezoelectric types obtain their conversion by producing a strain proportional to the applied field which reverses sign when the sign of the field reverses. Conversely, they produce an electric charge when they are strained. A magnetostrictive material has a strain which is proportional to the square or even power of the magnetic flux and it requires either a biasing field or a remanent polarization to produce a linear transducer.

The action of both types are well-known and this section is limited to discussing the condition for good efficiency of conversion and to giving the constants of the most used types of transducers. Equivalent circuits for transducers are useful for determining the efficiency and bandwidth possible for transforming electrical into mechanical energy and vice versa. While generalized circuits are available [23] for any terminating conditions, most of the transducers are of the half wavelength type which can be represented by the equivalent circuit of Fig. 8(a) for a piezoelectric type. For a polarized magnetostrictive transducer for which the eddy current loss is small (for example, for a ferrite) the circuit of Fig. 8(b)—which is the inverse of Fig. 8(a)—can represent

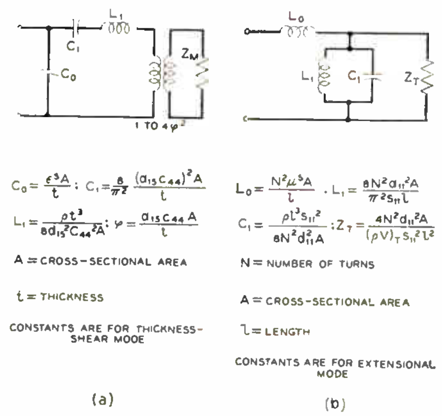


Fig. 8—Equivalent circuits for (a) piezoelectric driving crystal, (b) magnetostrictive transducer.

TABLE I

Crystal and Cut	Mode of Vibration	Elastic Compliance $\times 10^{11}$ meters²/Newton	Piezoelectric constant, $d \times 10^{12}$ coulombs/Newton	Dielectric Permittivity ϵ, $\times 10^{11}$ farads/meter	Electro-mechanical coupling factor, k	Density $\rho \times 10^{-3}$ kg/meter³
Quartz, $-18°$ X Cut	Ext.	$s_{22}^{E'} = 1.44$	$d_{21}' = 2.28$	4.06	0.094	2.65
Quartz, X Cut	T.L.	$1/c_{11}^E = 1.16$	$e_{11}/c_{11} = -2.0$	4.06	0.092	2.65
Quartz, Y Cut	T.S.	$1/c_{66}^E = 2.57$	$e_{26}/c_{66} = +4.4$	4.06	0.136	2.65
80 per cent $BaTiO_3$ 12 per cent $PbTiO_3$ 8 per cent $CaTiO_3$	T.L.	$1/c_{33} = 0.57$	$e_{33}/c_{33} = 51$	400	0.34	5.4
	T.S.	$1/c_{44} = 2.04$	$d_{15} = d_{33} - d_{31} = 160$	608	0.45	5.4
PZT 4	T.L. Ext.	$1/c_{33} = 0.664$ $s_{11} = 1.23$	$e_{33}/c_{33} = 160$ $d_{31} = -111$	1060 1060	0.61 0.31	7.5 7.5
PZT 5	T.L. Ext.	$1/c_{33} = 0.741$ $s_{11} = 1.48$	$e_{33}/c_{33} = 200$ $d_{31} = -140$	1200 1200	0.67 0.33	7.5 7.5
NbO_3 [K(0.5) Na (0.5)]	T.L.	$1/c_{33} = 0.71$	$e_{33}/c_{33} = 52.5$	235	0.41	4.4

Ext. Extensional; T.L. Thickness Longitudinal; T.S. Thickness shear

the relations. On account of the gyrator action of a magnetostricitive device [23], the mechanical elements, including the termination, have to be inverted to be used in this circuit.

The corresponding element values are given in terms of the fundamental data for longitudinal and shear vibrating quartz and four ceramics, the first three of which are widely used in transducers. These values are for low-amplitude motion. For larger amplitudes, changes [24] occur in some of the constants. (Table I)

The extensional modes have mostly been used for filters and underwater sound transducers, while the thickness shear and thickness longitudinal modes have been used for delay lines and oscillators.

For magnetostrictive material, the Piezoelectric Crystals Committee of the IRE has sponsored a "piezomagnetic" nomenclature which parallels the nomenclature of piezoelectric crystals. This nomenclature is followed in Table II, which lists several piezomagnetic materials. The MKS system is used throughout.

As an example of the use of such equivalent circuits, a fused silica delay line is considered. This is driven by a shear type piezoelectric ceramic. The widest band and the smallest loss result when the static capacitance C_0 is tuned by a series or shunt coil. The shunt coil results in the highest output impedance, which is usually preferred. All the elements are then shown by Fig. 9. This network has the configuration of a band-pass filter, which when properly terminated can transform electrical to mechanical energy without loss except for the usually small dissipation inherent in the elements. The design equations of such a network are

$$L_0 = \frac{(f_B - f_A)Z_0}{2\pi f_A f_B}, \qquad C_0 = \frac{1}{2\pi(f_B - f_A)Z_0},$$

$$L_1 = \frac{Z_0}{2\pi(f_B - f_A)}, \qquad C_1 = \frac{(f_B - f_A)}{2\pi f_A f_B Z_0}, \tag{8}$$

TABLE II

Material	Mode of Vibration	Elastic Compliance $\times 10^{11}$ meters^2/Newton (constant H)	Piezomagnetic constant, $d\times 10^9$, webers per Newton	Reversible Permeability $u\times 10^4$ Henries/meter	Electro-mechanical coupling factor, k	Density, $\rho\times 10^{-3}$ kg/meter^3
Nickel	Ext.	0.475	$d_{33}=-5.0$	0.98	0.232	8.9
45 per cent Permalloy	Ext.	0.625	$d_{33}=12.2$	7.36	0.101	8.17
2V Permindur 2 per cent V; 50 per cent Co; 48 per cent Fe	Ext.	0.435	$d_{33}=9.35$	3.54	0.123	8.9
4 per cent Co; 96 per cent Ni	Ext.	0.5	$d_{33}=13.0$	1.3	0.51	8.9
Ni$(1-x)$ Co(x) Fe_2O_3 $x=0.018$	Ext.				0.218	5.29

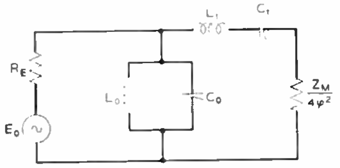


Fig. 9—Equivalent circuit for shear transducer driving a delay line.

where Z_0 is the image impedance at midband and f_A and f_B are the lower and upper cutoff frequencies when the transducer is considered as a band-pass filter. The bandwidth defined as $(f_B-f_A)/\sqrt{f_A f_B}$ is determined by the ratio of the shunt capacitance C_0 to the seies capacitance C_1. From (8) and Fig. 8(a)

$$\frac{f_B - f_A}{\sqrt{f_A f_B}} = \sqrt{\frac{C_1}{C_0}} = \frac{\sqrt{8}}{\pi} d_{15}\sqrt{\frac{c_{44}}{\epsilon^S}} = \frac{\sqrt{8}}{\pi}\frac{k}{\sqrt{1-k^2}}, \qquad (9)$$

where k, the electromechanical coupling factor for a shear crystal, is defined as

$$k = d_{15}\sqrt{\frac{c_{44}}{\epsilon^T}} = d_{15}\sqrt{\frac{c_{44}}{\epsilon^S}}(\sqrt{1-k^2}). \qquad (10)$$

For example for a shear wave transducer made from 80 per cent $BaTiO_3$, 12 per cent $PbTiO_3$, 8 per cent $CaTiO_3$, constructed as discussed by May [20], the fractional bandwidth ratio is 0.456. To get a 10-Mc bandwidth requires a mean frequency of 23 Mc. The other question of interest is how well the terminating load matches the transducer characteristic. Fused silica has a density ρ of 2.2×10^3 kg/M^3 and a shear velocity V_s of 3764 meters per second. The equivalent electrical termination is

$$R_E = \frac{R_M}{4\phi^2} = \frac{A\rho V_s}{4(d_{15}c_{44}A/t)^2} = \frac{\rho V_s t^2}{4(d_{15}c_{44})^2 A}, \qquad (11)$$

where t is the thickness of the transducer (1.3×10^{-4} meters) and A the area taken as 1 square cm (10^{-4} sq m). Inserting the numerical values from Table I, $R_E=5.66$ ohms. This should approximately match the design impedance of the transducer, which from (8) and Fig. 8(a) is

$$Z_0 = 2\pi(f_B - f_A)\rho_T t^3/8d_{15}^2 c_{44}^2 A = 15.2 \text{ ohms}. \qquad (12)$$

This provides a fair match over the band giving a slight rise in loss at the center.

Similar considerations apply to the design of other type transducers including those used for sonar. A recent paper [25a], describes a 200-inch diameter cylinder made by stacking up walls of barium titanate ceramic 12 inches long. This transducer is capable of radiating 2 Mw pulse power at 160 cps, at operating depths up to 10,000 feet. The lower frequency trend occurring since World War II makes use of the lower attenuation at low frequencies, and the greater freedom from the distorting effects of thermal gradients in the sea water media. The lower frequencies require much larger transducer sizes than were required for the kilocycle transmissions of World War II. The ferroelectric ceramics of Table I are one of the advances making such transducers possible. The low frequencies also improve the performance of magnetostrictive devices, and good efficiencies are found [25b] for flexural transducers employing nickel cobalt laminated structures.

IV. Filters and Oscillators Employing Mechanically Vibrating Elements

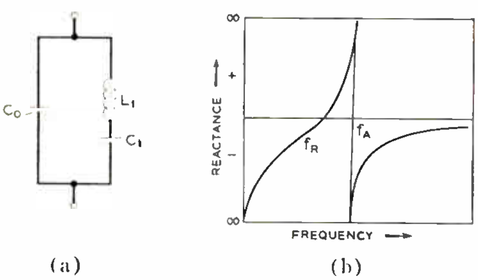
From the equivalent circuit of a half wave crystal on Fig. 8(a), it is seen that a crystal with a zero loading can be represented by a capacitance C_0 shunted by a tuned circuit having the element values shown by Fig. 10(a). The ratio between the shunt capacitance C_0 and the series capacitance C_1 is a function of the electromechanical coupling factor k. For the $-18.5°$ X-cut quartz which is largely used in crystal channel filters for separating one voice channel from another, this ratio is

$$\frac{C_0}{C_1} = \frac{\pi^2}{8}\,\frac{(1-k^2)}{k^2} = 139. \tag{13}$$

The impedance of a network of this type is shown by Fig. 10(b) and the separation between the resonance f_R and the antiresonance f_A is

$$\frac{f_A^2 - f_R^2}{f_R^2} = \frac{1}{r} \quad \text{or} \quad \frac{f_A - f_R}{f_R} \doteq \frac{1}{2r} = 0.0036. \tag{14}$$

The advantage of this type element over an equivalent coil and condenser is in the very high Q's possible in the motional arm (up to 20×10^6 at 100 kc if all the mounting loss is removed) and the low temperature coefficient of frequency inherent in most crystal cuts.



$$C_0 = \frac{\epsilon^S A}{l};\quad C_1 = \frac{8}{\pi^2}\,\frac{d_{21}^2 A}{s_{11}^E l};\quad L_1 = \frac{\rho l^3 s_{11}^{E^2}}{8 d_{21}^2 A}$$

$$\frac{C_0}{C_1} = \frac{\pi^2}{8}\,\frac{\epsilon^S s_{11}^E}{d_{21}} = \frac{\pi^2}{8}\,\frac{(1-k^2)}{k^2};\quad L_1C_1 = \frac{\rho s_{11}^E l^2}{\pi^2} = \frac{1}{\omega_R^2}$$

Fig. 10—Crystal resonator circuit and impedance characteristic.

Crystals alone are useful in narrow-band filters, but it can be shown that the bandwidth defined as in (9) cannot be larger than twice the value of (14). To get the wide bands necessary for voice communication, it was shown by the author [7], [26] that coils can be combined with crystals in lattice type filters. Fig. 11 shows a typical characteristic of a channel filter in a bank of 12 used in all the carrier, broad-band radio and submarine cable systems of the Bell System. With a source of supply guaranteed by synthetic quartz [8] and with the economics inherent in imperfection-free crystals of an economic size, there does not appear to be any present competitor likely to displace them.

Considerable work has been carried out in applying filters in the megacycle range. The principal difficulty is that at these frequencies it is necessary to use thickness modes of vibration of the crystal unit, and in these many spurious modes appear too near the desired mode to allow wide-band applications. Considerable theoretical effort [27a] has been applied to this problem and fairly close calculations have been made of the resonant frequency spectrum. On the experimental side, the spectrum [27b] of thickness AT shear vibrating crystals has been carefully measured. Contoured plates [28] and geometrical dispositions [29] of the driving plates have produced improvements. Fig. 12 shows a photograph of a contoured [28] fifth harmonic crystal mounted by rigid tabs to the side of the crystal. This crystal, which vibrates at 5.0 Mc, has a simple frequency spectrum, a Q consistent with quartz alone, and a very small aging rate.

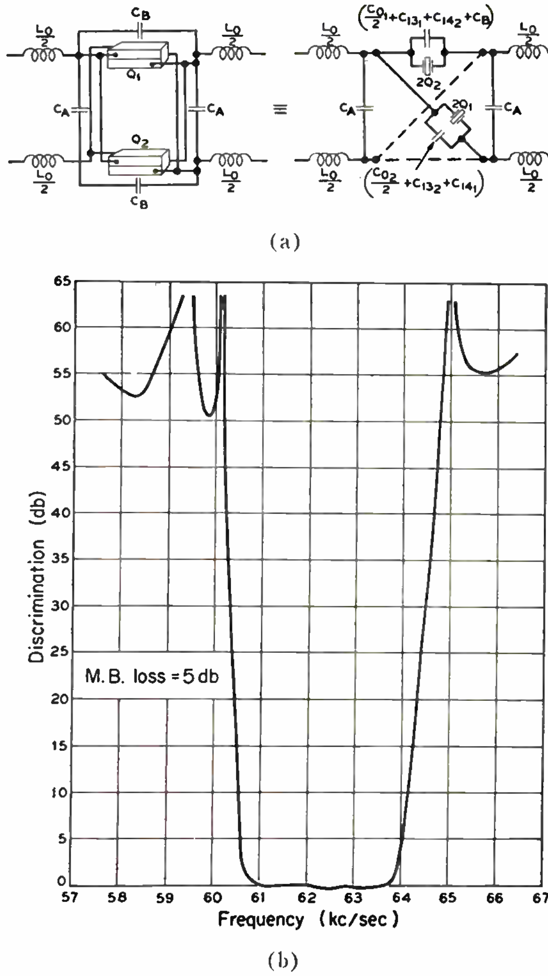


Fig. 11—Crystal channel filter. (a) Typical section showing use of divided plate crystals. (b) Attenuation characteristic for two sections of (a).

While such characteristics are consistent with good oscillator performance, it cannot be said that the resonator problem is completely solved. Research on oscillator crystals is directed toward mode studies, the effects of imperfections in the quartz on stability and aging, and on increasing the frequency range of crystal oscillators. The precision required has increased from 200

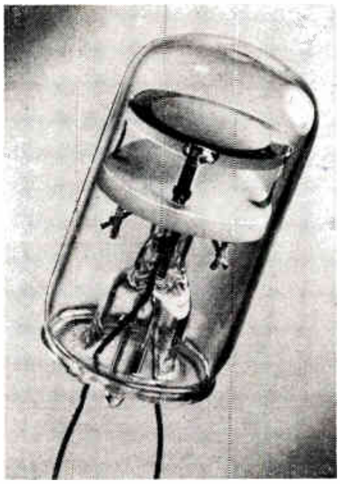


Fig. 12—Contoured fifth harmonic AT cut quartz crystal for use in frequency standards (after Warner).

parts in 10^6 during World War II to 20 parts in 10^6 at the present time. With the impending universal adoption of single-sideband communication, this tolerance [9] must be reduced by 2 orders more.

Another attack on the filter problem has been to use mechanical elements to produce the filtering, and transducers to couple the filter to the electrical circuit. Many types are possible [10], but the most widely used one appears to be the Collins mechanical band-pass filter whose constriction and performance are shown by Fig. 13. This usually works in the intermediate frequency range around 450 kc. Another type has recently been developed [30] which is capable of covering the frequency range from 10 kc to 60 kc with voice bands. Fig. 14 shows a typical characteristic. This filter, made of a low-temperature coefficient nickel iron composition, vibrates in a torsional mode. The side arms, which produce the filtering action, vibrate in flexural modes. The filter is driven by a transducer having 4 PZT ceramics on the two ends of the driving unit. These produce a second flexure having a node in the center, which is strongly coupled to the torsional filter. Such filters may be of interest for short haul carrier systems where the cost of modulating to higher frequencies would be prohibitive.

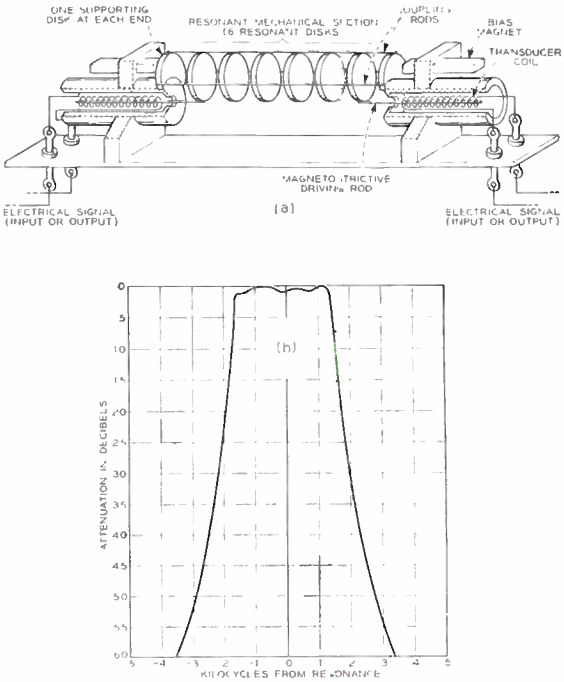


Fig. 13—Structure (a) and characteristic (b) of Collins mechanical filter.

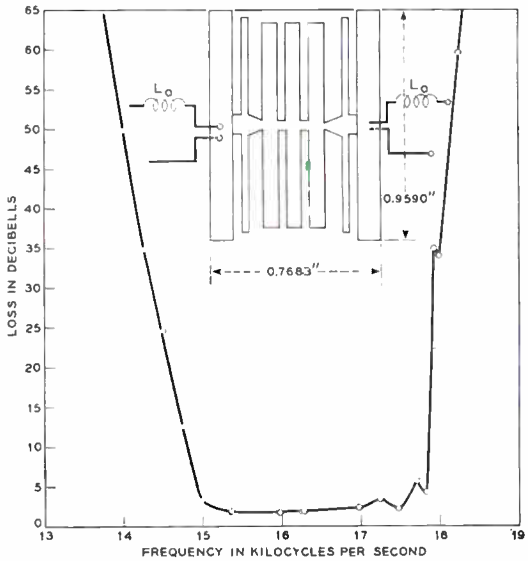


Fig. 14—Structure and characteristics of low-frequency mechanical filter (after Mason and Thurston).

References

[1] For a history of early uses of transducers and underwater sound systems see F. V. Hunt, "Electroacoustics," Harvard University Press, Cambridge, Mass., ch. 1; 1954.

[2] See F. Firestone, "The supersonic inspectoscope," *J. Acous. Soc. Am.*, vol. 17, p. 200; January, 1946, and U. S. Patent 2,280,226; April 21, 1942.

[3] A. G. Emslie and R. L. McConnel, "Radar Systems Engineering," McGraw-Hill Book Co., Inc., New York, N. Y., vol. I, ch. 16; 1947.

[4] H. Busignies, P. R. Adams and R. I. Colins, "Navar," *Elec. Commun.*, vol. 23, pp. 113–143; June, 1946.

[5] (a) E. Bradburd, "Magnetostrictive delay line," *Elec. Commun.*, vol. 28, p. 46; March, 1951.

(b) H. Epstein and O. Stram, "A high performance magnetostriction sonic delay line," *Rev. Sci. Instr.*, vol. 24, pp. 231–232; March, 1953.

[6] (a) J. R. Klauder, *et al.*, "Theory and design of chirp radars," *Bell Sys. Tech. J.*, vol. 39, pp. 745 July, 1960.

(b) J. E. May, "Wire type dispersive ultrasonic delay lines," IRE Trans. on Ultrasonics Engineering, vol. UE-7, pp. 44–53; June, 1960.

(c) T. R. Meeker, "Dispersive ultrasonic delay lines using the first longitudinal mode in a strip," IRE Trans. on Ultrasonics Engineering, vol. UE-7, pp. 53–58; June, 1960.

(d) A. H. Meitzler, "Ultrasonic delay lines using shear modes in strips," IRE Trans. on Ultrasonics Engineering, vol. UE-7, pp. 35–44; June, 1960.

[7] W. P. Mason, "Electromechanical Transducers and Wave Filters," D. Van Nostrand Co., Inc., Princeton, N. J., ch. 8, 1942; 2nd ed., 1948.

[8] R. A. Sullivan, "Growing quartz crystals," *Western Electric. Engrg.*, vol. 3, pp. 2–12; April, 1959.

J. M. Stanley, "The signal corps synthetic quartz program," IRE Trans. on Military Electronics, vol. MIL-4, pp. 438–447; October, 1960.

[9] E. A. Gerber and J. M. Haval, "Precision frequency control for military applications," IRE Trans. on Military Electronics, vol. MIL-4, pp. 424–437; October, 1960.

[10] (a) W. P. Mason, "Physical Acoustics and the Properties of Solids," D. Van Nostrand Co., Princeton, N. J., ch. 5; 1958.
(b) J. C. Hathaway and D. F. Babcock, "Survey of mechanical filters and their applications," Proc. IRE, vol. 45, pp. 5–16; January, 1957.
M. Borner, E. Kettel, and H. Ohnesorge, "Mechanische Filter für die Nachrichtentechnik," *Telefunken Zeitung*, vol. 31, p. 105; July, 1958.

[11] For properties of solids see, (a) W. P. Mason, "Physical Acoustics and the Properties of Solids," D. Van Nostrand Co., Princeton, N. J.; 1958.
For liquids see (b) K. F. Herzfeld and T. A. Litovitz, "Absorption and Dispersion of Ultrasonic Waves," Academic Press, New York, N. Y.; 1959.
For high polymers see (c) W. P. Mason, "Dispersion and absorption of sound in high polymers," *German Encyclopedia of Phys.*, vol. 11, pp. 361–417; 1961.

[12] R. M. Bozorth, W. P. Mason, and H. J. McSkimin, "Frequency dependence of elastic constants and losses in nickel," *Bell Sys. Tech. J.* vol. 30, pt. 1, pp. 970–987; October, 1951.

[13] W. P. Mason and H. Bommel, "Ultrasonic attenuation at low temperatures for metals in the normal and superconducting states," *J. Acous. Soc. Am.*, vol. 28, pp. 930–943; September, 1956.

[14] W. A. Harrison and M. B. Webb, Ed., "The Fermi Surface," John Wiley and Sons, Inc., New York, N. Y.; 1960.

[15] R. W. Morse and J. D. Gavenda, "Magnetic oscillations of ultrasonic attenuation in a copper crystal at low temperatures," *Phys. Rev. Lett.*, vol. 2, pp. 250–252; 1959.

[16] H. Bömmel and K. Dransfeld, "Excitation and attenuation of hypersonic waves in quartz," *Phys. Rev.*, vol. 117, pp. 1245–1252; March, 1960.

[17] (a) H. J. McSkimin, "Propagation of longitudinal waves and shear waves in cylindrical rods at high frequencies," *J. Acous. Soc. Am.*, vol. 28, pp. 484–494; May, 1956.
(b) M. Redwood and J. Lamb, "Attenuation in fused silica," *Proc. IEE*, pt. B, vol. 103, pp. 773–780; January, 1956.
(c) J. Lamb, M. Redwood, and Z. Shteinshlifer, "Absorption of compressional waves in solids from 100 to 1000 mc/sec.," *Phys. Rev. Lett.*, vol. 3, pp. 28–30; July, 1959.

[18] (a) D. L. Arenberg, "Ultrasonic solid delay lines," *J. Acous. Soc. Am.*, vol. 20, pp. 1–28; January, 1948.
——, "Ultrasonic delay lines," 1954 IRE National Convention Record, pt. 6, pp. 63–72.
(b) H. J. McSkimin, "Wedge geometry for ultrasonic delay lines," *J. Acous. Soc. Am.*, vol. 26, p. 146; January, 1954.

[19] (a) J. E. May, "Characteristics of ultrasonic delay lines using quartz and barium titanate ceramic transducers," *J. Acous. Soc. Am.*, vol. 26, pp. 347–355; May, 1954.
(b) C. A. Bieing, "Temperature controlled low loss ultrasonic delay lines with barium titanate ceramic transducers," *Proc. Electronic Components Symp.*, vol. 13, pp. 62–65; May, 1957.

[20] J. E. May, "Thickness-shear mode $BaTiO_3$ ceramic transducers for ultrasonic delay lines," IRE Trans. on Ultrasonic Engineering, vol. UE-7, pp. 7–12; February, 1960.

[21] (a) A. E. Love, "Theory of Elasticity," Cambridge University Press, Cambridge, Eng., p. 287; 1934. See for an account of the solution.
(b) D. Bancroft, "The velocity of longitudinal waves in cylindrical bars," *Phys. Rev.*, vol. 59, pp. 588–593; April, 1941.
(c) G. E. Hudson, "Dispersion of elastic waves in solid cylinders," *Phys. Rev.*, vol. 63, pp. 46–51; January, 1943.

[22] (a) G. G. Scarrot and R. Naylor, "Wire-type delay lines for digital storage," *Proc. IEE*, vol. 103, pt. 6; March, 1956.
(b) R. N. Thurston and L. M. Tornillo, "Coiled wire torsional wave delay lines," 1958 IRE National Convention Record, pt. 2, pp. 109–117.
(c) Mason, *op. cit.*, p. 139.
(d) W. P. Mason, patent application pending.

[23] Mason, *op. cit.*, ch. 6.
Mason, *op. cit.*, ch. 3.

[24] D. Berlincourt, "Transducer properties of lead titanate zirconate ceramics," IRE Trans. on Ultrasonics Engineering, vol. UE-7, pp. 1–6; February, 1960.

[25] (a) C. E. Green, "Ceramic cylinders for lower frequencies, lower impedances and greater depths," *J. Acous. Soc. Am.*, vol. 32, p. 1519; November, 1960.
(b) F. S. Woolett, "Magnetostrictive flexural bar transducers," *J. Acous. Soc. Am.*, vol. 33, p. 835; June, 1961.

[26] W. P. Mason, Patent No. 2,045,991.

[27] (a) R. D. Mindlin, "Mathematical theory of vibration of elastic plates and bars," *Proc. 12th Annual Symp. on Frequency Control*, Signal Corps, pp. 2–8; May, 1958.
(b) R. A. Sykes, "Quartz Crystals for Electrical Circuits," D. Van Nostrand Co., Princeton, N. J.; 1946.

[28] A. W. Warner, "Design and performance of ultra precise 2.5 Mc quartz crystal units," *Bell Sys. Tech. J.*, vol. 39, pp. 1193–1219; September, 1960.

[29] J. J. Vormer, "Crystal-plates without overtones," Proc. IRE, vol. 39, pp. 1086–1087; September, 1951.
J. A. Lewis, "The effect of driving electrode shape on the electrical properties of piezoelectric crystals," *Bell Sys. Tech. J.*, vol. 40, pp. 1259–1281; September, 1961.

[30] W. P. Mason and R. N. Thurston, "A compact electromechanical band-pass filter for frequencies below 20 kilocycles," IRE Trans. on Ultrasonics Engineering, vol. UE-7, pp. 59–70; June, 1960.

Ultrasonics in Industry*

FRANK MASSA†, FELLOW, IRE

Summary—**This paper presents a brief description of various successful industrial applications that have been made of ultrasonics. The applications are divided into three broad groups: low, medium, and high power. In the low-power group are included reduction of friction, inspection, remote control, flaw detection, and other types of gauging and nondestructive testing. The medium-power group includes ultrasonic machining, soldering, medical therapy, liquid level indicators, and similar types of depth gauges and instruments. The high-power group includes precision cleaning, emulsification, dispersion of metals, metallurgical, bacteriological, and other chemical and physical processing applications. It is only natural that the low- and medium-power applications received first commercial acceptance, while the high-power applications advanced more slowly awaiting the development of low-cost high-power ultrasonic transducers and low-cost ultrasonic power supplies. These latter requirements are at present reasonably well met, which means that great increases in industrial high-power ultrasonic applications can be expected in the immediate future.**

Introduction

Sound is a compressional wave in an elastic medium,[1] and because of its ease of propagation through liquids and solids, it is ideally suited for both examining and modifying the physical or chemical state of matter. Another property of sound is that it travels at a relatively slow speed (as compared with electromagnetic radiation), which makes it a useful tool for looking through solids and liquids in order to examine the medium for the presence of discontinuities, whether it be a submerged submarine or a hidden flaw inside a piece of steel.

One of the earliest developments in ultrasonics was the quartz sandwich transducer designed about 1920 for use as a navigational aid. In order to overcome the high-frequency limitations inherent in quartz, Langevin made up a quartz-steel sandwich consisting of several quartz plates cemented between two steel disks. This arrangement reduced the high resonant frequency of the quartz to the lower ultrasonic range and also provided a relatively large piston surface to achieve the desired radiation pattern.

During the latter part of the 1920's Wood and Loomis employed quartz crystal transducers to examine the effects of high-frequency ultransonic sound on the behavior of various liquids. Work in this field continued through the decade 1930 to 1940, during which time many investigators examined both chemical and physical changes in liquids caused by the introduction of ultrasonic energy. Because of the lack of suitable transducers and accurate instruments for the measurement of ultrasonic sound fields, the experiments were mostly qualitative in nature, and the reported results were primarily of academic interest.

During World War II, the tremendous success in enemy submarine warfare made it imperative that fundamental progress be made in underwater acoustics if this country were to survive. An intensive engineering development program initiated in 1940 produced in a few years a spectacular increase in fundamental knowledge in the design of underwater transducers capable of generating large amounts of ultrasonic power at any desired frequency. One of the most important factors contributing to this progress was the creation of standard hydrophones capable of measuring accurately underwater sound pressure over wide frequency ranges. Fig. 1(A) shows one of the earliest of these instruments that was developed during 1941 and employed Rochelle salt as the active element.[2] The inherent limitations of Rochelle salt, which restricted the range of application of this instrument, were overcome with the introduction of ammonium dihydrogen phosphate crystals which led to the improved measurement standard[3] shown in Fig. 1(B).

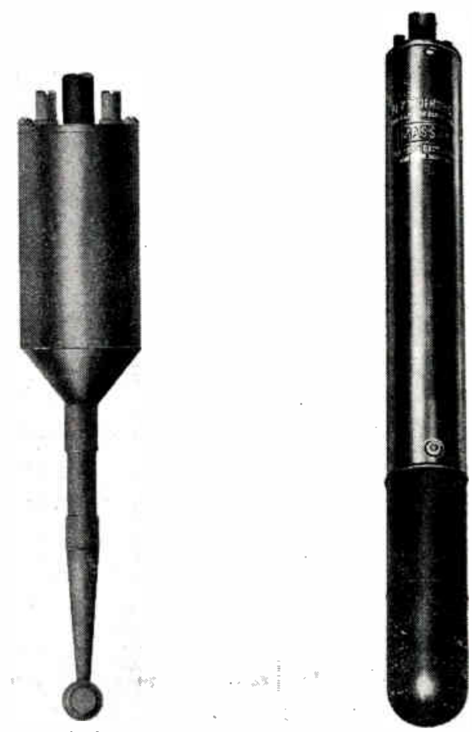

(A) (B)

Fig. 1—Sound pressure measurement hydrophones. (A) Early model using Rochelle salt. (B) Later model using ADP crystal.

* Received by the IRE, July 10, 1961; revised manuscript received, October 27, 1961.

† Massa Division, Cohu Electronics, Inc., Hingham, Mass.

[1] This is not meant to exclude shear waves in solid media.

[2] F. Massa, U. S. Patent No. 2,413,462; December 31, 1946.

[3] F. Massa, U. S. Patent No. 2,613,261; October 7, 1952.

By 1950 a wealth of basic engineering information on ultrasonic transducer design had been accumulated, and accurate instruments for measuring ultrasonic sound were generally available. With these new tools, the early qualitative researches on the effects of ultrasonic sound on the properties of matter took on renewed practical engineering interest.

The purpose of this paper is to acquaint the reader with the large variety of important uses to which ultrasonics has been applied and to give a broad understanding of the basic principles that are employed in the various fields of application. For the ultrasonic engineer who is interested in the technical details of the properties of ultrasonic waves and in the design of apparatus for the generation and measurement of ultrasonic energy, an excellent classified bibliography of all the contemporary literature published during the past 30 years may be found in the Cumulative Indexes of the *Journal of the Acoustical Society of America*. Each of the three separate 10 year volumes list thousands of references conveniently classified under various headings including: Ultrasonics, Waves and Vibrations, and Underwater Sound. In the latest two Indexes are also included digests of several thousand acoustical patents issued during the past 20 years. This condensed material presents an excellent picture to the research engineer of all the contemporary work that has been carried on in ultrasonics and related fields.

Low-Power Applications

It is quite natural that the first practical applications of ultrasonics in industry were those requiring low-power levels. A very important industrial application is ultrasonic inspection, which makes use of the simple fact that the propagation of sound through a solid is distriburbed when there is a discontinuity in the material. The two most common principles employed in the design of ultrasonic inspection equipment are based either on the indication of a reduction in the transmitted sound due to the presence of a flaw, or on the reflected echo from the flaw which may be used to indicate its location.

A. *Continuous Automatic Inspection for Internal Flaws*

During 1943 one of the earliest continuous mass-production ultrasonic inspection systems was developed by the writer to solve a serious bottleneck that had developed in the inspection of sticks of solid fuel propellant. The rocket fuel consisted of rubber-like extrusions several feet long and a few inches in cross section, and the problem was to detect the presence of small air bubbles that would sometimes be trapped into the material during the manufacturing process. The presence of the included air could cause an explosion during the burning of the fuel, which would usually be fatal to the plane using the jet assist rockets. X-ray inspection was unsatisfactory not only because the process was too costly and time consuming, but because of a lack of sufficient contrast between an air bubble and the rocket fuel to give reliable indication on the X-ray film.

A satisfactory solution to the difficult testing problem was achieved by the simple ultrasonic test equipment illustrated in Fig. 2. A number of transducers, T1, T2, etc., are arranged to transmit sound of the appropriate frequency through various portions of the material, and another group, R1, R2, etc., are arranged to receive the transmitted signals. The output from each receiving transducer is recorded as a separate graph on a multichannel strip chart recorder, while the position of the paper chart is synchronized with that of the test specimen as it passes by the transducer array. If a flaw intercepts any region being scanned by the transducer pairs, the sound intensity through the material will be reduced and show up as a dip in the graphs being recorded on the moving chart.

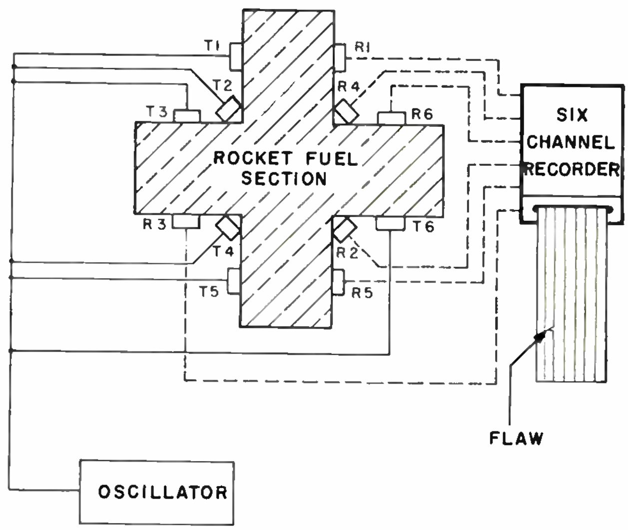


Fig. 2—Schematic sketch showing automatic ultrasonic inspection for hidden flaws in solid rocket fuel.

The system of ultrasonic scanning described above permitted the automatic recording of the homogeneity of large quantities of fuel sticks as they were continuously conveyed through the transducer array at the rate of 60 ft/minute. Air bubbles as small as $\frac{1}{16}$ inch diameter would produce a visible dip in the recorded charts. Hundreds of thousands of pieces were successfully inspected at negligible cost compared with the former less accurate and very expensive X-ray technique. Similar types of flaw-detection equipment have been developed and used for the continuous inspection of cemented or bonded joints and for inspecting for invisible discontinuities inside sheets such as are produced by foreign matter during the rolling of metals.

B. *Flaw Detection in Thick Materials*

The same principle employed in radar for locating airplanes by timing the arrival of a reflected pulse of energy has been successfully employed in the location of internal flaws in large metal parts by measuring the time of arrival of a reflected ultrasonic wave from the discontinuity. Firestone reduced this principle to practice about 1946;[4] it is now being widely used as an industrial inspection tool for examining the internal character of heavy metal parts better than can be accomplished by X-ray techniques. A schematic diagram of the ultrasonic flaw-detection system, which operates by timing the returned echo from a discontinuity in the material, is shown in Fig. 3. For a resolution of the order of $\frac{1}{8}$ inch in the location of the distance of the flaw below the surface, the echo time measurement has to be made in the order of 1 μsec which means that an ultrasonic transducer operating in the 5 to 10 megacycle range is required.

Referring to Fig. 3, a transducer is acoustically coupled by means of an oil film to a smooth prepared surface of the material under test while a pulse generator repetitively pulses the transducer at a rate of 60 cps. At the start of each pulse, a vertical deflection T_0 appears on the scope and, simultaneously, the horizontal time sweep is triggered in the oscillograph. The time base is adjusted such that the reflected echo from the opposite end of the specimen, T_1, fills the width of the tube. The pulses T_0 and T_1 appearing on the scope then represent the full height dimension of the specimen. As the transducer is moved over a region containing a flaw, an echo will be reflected from the flaw which will appear as T_2 on the scope. The relative position of T_2 along the base line will then represent the relative depth of the flaw inside the piece.

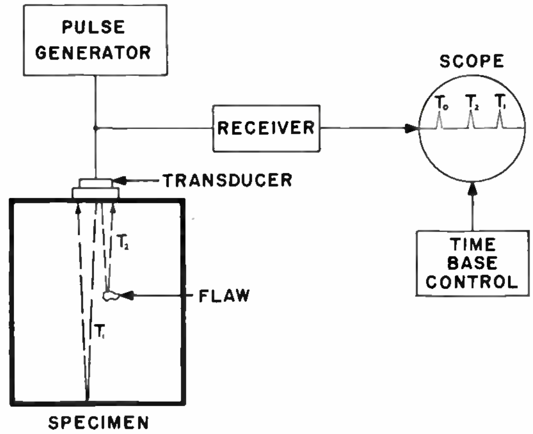


Fig. 3—Ultrasonic flaw detection by measuring reflected echo time.

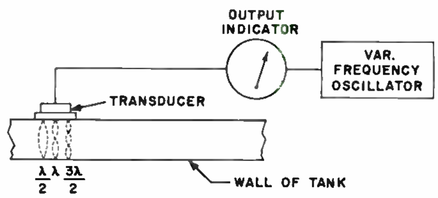


Fig. 4—Thickness measurement by ultrasonic resonance.

C. Thickness Gauging

The echo time principle for flaw detection may also be applied to the measurement of wall thickness where access is available from only one side such as in pipe lines and storage tanks. Ultrasonic gauging for these applications permits checking on internal corrosion without need for getting inside the pipe or tank. As might be inferred from the preceding section, it would be extremely difficult to measure echo time directly for gauging thin-walled materials because accurate time measurements of small fractions of a microsecond would be required.

For gauging thin-walled materials, a measurement of the thickness resonant frequency of the structure may be made, as illustrated in Fig. 4. A transducer is coupled to the outside wall of the material and is driven by a variable frequency oscillator. When the frequency supplied to the transducer corresponds with one-half wavelength of sound through the wall, the transducer will become effectively unloaded and the output meter will show an abrupt change in reading. A similar change will occur at multiples of the fundamental resonance, as illustrated by the standing waves $\lambda/2$, λ, and $3\lambda/2$, in Fig. 4. A necessary requirement in this method of thickness gauging is that the resonance frequency of the transducer be very much higher than the frequency corresponding to one-half wavelength in the thickness of the sheet.

[4] F. A. Firestone, "Supersonic reflectoscope, an instrument for inspecting the interior of solid parts by means of sound waves," *J. Acoust. Soc. Am.*, vol. 17, pp. 287–299; January, 1946.

D. Reduction of Friction

By injecting low-level ultrasonic energy into a material whose surface is in sliding contact with another, the coefficient of friction between the two bodies is very considerably reduced. Experimental work carried on by the author has shown reductions in the force required to move a heavy tool carriage over its polished guides by a factor of more than 10 to 1 in the presence of ultrasonics. In similar experiments in which paint was required to flow through a very thin slit, the injection of low-level ultrasonic energy into the nozzle maintained uniform flow of paint with slit openings as small as 0.001 inch, whereas without ultrasonics the slit had to be at least 0.007 inch to allow the paint to flow and then the flow would be intermittent due to clogging. For more detailed information concerning the equipment and its method of use, the reader may refer to U. S. Patents 2,738,173 and 2,746,813.

E. Remote Control Applications

One of the more recent large-scale applications for low-power ultrasonics is in the field of remote control. Hundreds of thousands of TV sets are being operated from across the room by means of a portable ultrasonic generator whose signal is picked up by a receiving transducer located at the set, which, in turn, actuates a relay to perform the required function. A similar application is for an automatic garage door opener which, in the simplest case, employs a transducer mounted on the front end of a car which sends out an ultrasonic signal on command to be picked up by a similar transducer located at the garage entrance, which actuates a relay to cause the door to open. Fig. 5 shows a photograph of inexpensive ultrasonic transducers as used in these applications.

Medium-Power Applications

A. Ultrasonic Machining

The fact that it is easily possible to ultrasonically vibrate the tip of a rod at amplitudes in the order of 100 g forms the basis for ultrasonic drilling. If a tool bit of any shape is vibrated at high frequencies and held with slight pressure against the surface of a hard brittle material, the tip will go through the material much like a vibrating air drill goes through a cement paving, except that at the ultrasonic frequencies, a polished finish is produced along the sides of the hole being formed. A fine abrasive mixed with a liquid to form a slurry is usually employed to achieve the best results.

A diagrammatic sketch of an ultrasonic drill is shown in Fig. 6 (next page). The transducer is fitted with a tapered tool post whose length is one-half wavelength of sound in the material at the frequency of operation. A tip of any desired shape is attached to the small end of the tool post and brought into contact with the work piece which is securely fastened to a table, as shown.

The arrangement illustrated in Fig. 6 has been used to "drill" any shaped hole through hard brittle materials that could not be otherwise machined. A similar type of ultrasonic tool has been used with the tool bit replaced by a die having an engraved face. By holding the vibrating die against the work piece, the imprint of the design is impressed into the hard material much as if a die were being pressed into the surface of wax.

B. Liquid Level Gauges

An accurate method of liquid level indication, which does not depend on the use of floats with their associated linkages and inaccuracies, makes use of the physical properties of sound-wave propagation through liquids. Since the speed of sound through a liquid is a fixed property of the liquid, it is possible to measure the transit time of a pulse of sound initiated from a transducer and passing through the height of the liquid in a tank, as illustrated in Fig. 7. A convenient and accurate manner for accomplishing the height reading is to choose a frequency such that the wavelength of the sound through the liquid corresponds to, say, 0.01 ft. An electronic counter in the timing circuit then merely counts the number of cycles of current passing through the transducer during the period from the start of the pulse until the reflected sound returns to the transducer after having travelled through the height of the liquid. The integrated number of cycles is a direct measure of the liquid height which can be directly displayed on an electronic counter in hundredths of a foot.

A basic requirement to the successful operation of the system is that the transducer have a very sharp beam angle and that it be free of significant secondary radiation at angles removed from its normal axis. A special transducer that achieves a 3° beam angle at 400 kc with radiation in secondary regions reduced by more than 30 db is shown in Fig. 8.

C. Medical Therapy and Diagnosis

It has long been realized that ultrasonic sound is easily conducted through body tissue and that strong localized heating may be produced by focusing the sound on discrete inner parts of the body. During the period immediately following World War II ultrasonic therapy was widely accepted among Western European countries. In West Germany alone several thousand ultrasonic therapeutic machines were in use, and reports on treatments in thousands of cases of numerous diseases indicated that a large majority were improved or completely healed. England and America remained skeptical over these early reports pending further medical research in this field and, therefore, ultrasonic therapy was more slowly accepted. At the present time there is renewed interest in this field, and it is noteworthy that President Kennedy was given ultrasonic treatment during his recent back illness.

In the diagnostic field, ultrasonics has been employed in locating brain tumors, kidney stones, and other "flaws" in the human body using techniques similar to those employed for flaw detection in metals.

D. Delay Lines

The physical principles and design technology of delay lines are covered in some detail in a companion article.[5]

E. Depth Indicators

Today tens of thousands of small boats are using ultrasonic depth indicators to give continuous information concerning the depth of water below the vessel. A typical low-cost depth sounder provides pulses of about 50 watts into a transducer which is usually mounted beneath the hull. The pulses are initiated by a contact switch synchronized with a rotating arm carrying a small neon bulb as the position of the bulb passes through zero. When the sound echo returns from the reflecting bottom the amplified received signal causes the neon bulb to flash. This process is repeated automatically several times per second so that the flashing bulb appears to remain at a fixed lighted position. A circular scale printed around the path of the rotating bulb is calibrated to show the depth of water in feet. This instrument will also show the presence of schools of fish, and it is so used as an aid in commercial fishing.

[5] W. P. Mason, this issue, p. 1374.

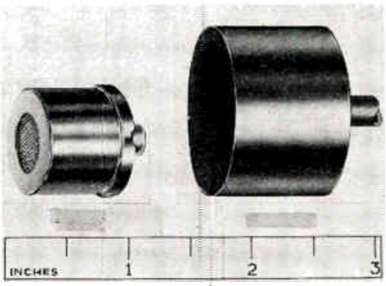


Fig. 5—Photograph of low-cost ultrasonic transducer used in remote control applications.

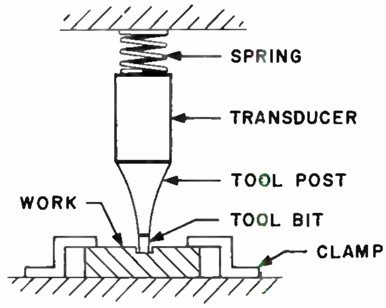


Fig. 6—Ultrasonic drill.

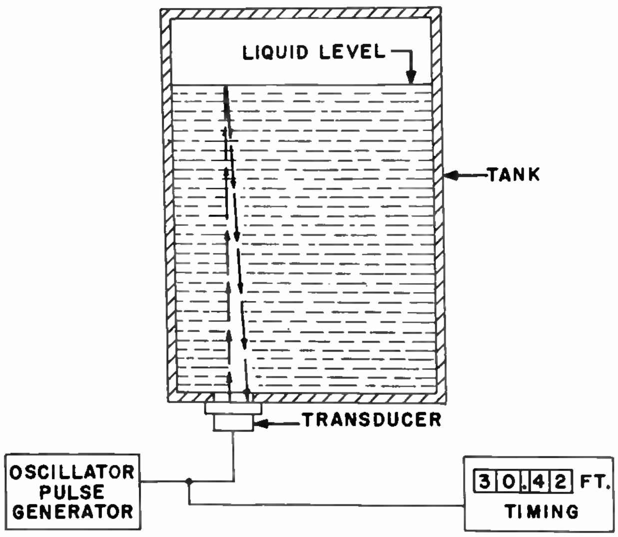


Fig. 7—System for ultrasonic liquid level indication.

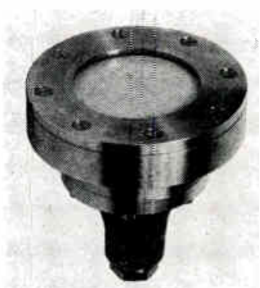


Fig. 8—Photograph of transducer for liquid level indicator (3° beam angle—negligible secondary lobes.)

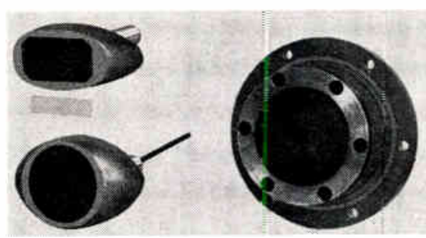


Fig. 9—Ultrasonic transducers for use in low-cost depth indicators.

Typical ultrasonic transducers designed for low-cost depth indicators are shown in Fig. 9. For deep-water navigational depth indicators, high-power equipment becomes necessary and larger transducers are required similar to those used in sonar systems.

High-Power Applications

A. Sonar

Sonar is to submarine detection and warfare what radar is to aircraft detection. In the field of high-power ultrasonic applications the greatest expenditures, reaching into the hundreds of millions of dollars, have been made in the development and production of sonar equipment. The successful development of the acoustically-guided torpedo during the latter stages of World War II probably saved us from the brink of disaster in which we were placed by the successful sinking of great numbers of our ships by enemy submarines. The vital needs of sonar in undersea warfare gave a tremendous impetus to the development of high-power ultrasonic transducers during the period of World War II.

We entered the war with inadequate sonar equipment. Two types of transducers were available: one used Rochelle salt with its inherent nonlinear characteristics and bad temperature stability, and the other used crude magnetostriction designs whose efficiencies were rarely higher than 10 per cent. These early transducers were generally piston-type radiators, one or two feet in diameter, operating at a kilowatt or two of power and were used as "searchlight" echo-ranging projectors covering about 10° in bearing at each position of search.

About 1942 ammonium dihydrogen phosphate crystals were made commercially available and removed all the limitations of Rochelle salt, resulting in the possibility of manufacturing high-efficiency, high-power transducers and extremely stable hydrophones. The searchlight-type sonar soon gave way to scanning sonar in which a high-intensity omnidirectional wafer-like beam of sound was sent out from a large cylindrical transducer, and the echo from a submarine would be received by an electronically-formed, rapidly-rotating beam. The scanning sonar transducer consists of a large number of vertical staves of elements arranged so that the outer radiating faces of the staves form a cylinder. When operated as a transmitter, all staves are connected in parallel and a high-intensity sound pulse of several wavelengths' duration is sent out in a 360° horizontal beam. Immediately after transmitting the pulse the transducer is connected as a receiver during which operation a group of staves are connected through time-delay networks such that the curvature of the stave fronts are effectively transformed in a plane surface, thereby creating a sharp horizontal beam for the receiving operation. The formed beam is electronically rotated at high speed by commutating the position of the lag line connections around the entire periphery of the transducer. The returning echo, whose pulse length is made longer than the scanning rate, is picked up by the rotating beam and the direction of arrival is indicated on a cathode-ray tube.

Scanning sonar transducers employing ADP crystals operate with powers of the order of 50 kw and have efficiencies of the order of 60 per cent. Improved magnetostriction transducers using staves made up of thin nickel laminations result in efficiencies of the order of 40 per cent. In applications requiring very accurate impedance and phase uniformity among individual staves (which are prerequisites for accurate beam information), ADP crystal has inherent advantages over the magnetostriction design. Recently, piezoelectric ceramics have become commercially available, and they are generally lower priced than either ADP or nickel. Barium titanate is the lowest cost and has found application in transducers where some degree of variation in impedance and sensitivity of the elements is permissible. Lead zirconate is the most recently available ceramic which has a much higher temperature stability than barium titanate and, therefore, permits higher power handling capacity. However, the new material still imposes a necessary variation in sensitivity and impedance among production elements.

Sonar transducers take dozens of forms depending on the particular application requirements. For a self-steering homing torpedo, for example, a multi-split-beam design is required to indicate whether the target is LEFT or RIGHT and UP or DOWN, so that automatic steering may be accomplished with the reception of each sonar signal. For underwater carrier telephony, an omnidirectional horizontal beam pattern is usually needed with a relatively small vertical beamwidth. Other transducers for special applications require special directional beam patterns. Today, transducers can be designed to deliver practically any amount of power at any desired frequency, and a wide choice of commercially available transducer materials permit the selection of the most economical design to meet the performance specifications.

B. Ultrasonic Cleaning

The basic advantages in the use of intense ultrasonic sound for thoroughly cleaning mechanical parts are well known.[6] However, premature attempts to market ultrasonic cleaning equipment with inadequate transducers and power supplies delayed general acceptance among the many potential industrial users of such equipment. Today, reliable high-power transducers may be produced, and economic power supplies can be built to permit large-scale installations for ultrasonic cleaning.

The chief merits of ultrasonic cleaning over the conventional methods are the great reduction of the cleaning time by the intense ultrasonic energy supplied to the cleaning liquid and the thoroughness of the cleaning. Under ultrasonic activation, the individual molecules of the cleaning liquid, vibrating at amplitudes corresponding to accelerations many times the value of gravity, literally bombard the microscopic foreign particles from the surface of the metal and allow a liquid film to get between the particles and the solid piece, thereby permitting all the microscopic deposit to be rinsed away by the cleaning liquid.

A striking example of the benefit of ultrasonic cleaning may be illustrated in an actual production application in which large aluminum plates had to be bonded with a high tensile strength joint. Using the best conventional degreasing and cleaning methods on the mechanical parts, the highest tensile strength realized at the bonded surface was 2500 psi under ideal controlled conditions. By introducing ultrasonic cleaning to the process before cementing the plates, the tensile strength of the bond was increased to over 6000 psi. A view of the cleaning tank and the plates being cleaned is shown in Fig. 10. This relatively small ultrasonic tank using 2 kw of driving power easily handled several hundred pieces per hour. To achieve the best results with ultrasonic cleaning, it is necessary that the acoustic power density at the surface of the transducer be in the range of about 10–50 watts per square inch, and that standing-wave regions in the cleaning solution be avoided by moving

[6] F. Massa, "Sound: a new basic industrial tool," *Elec. Production*, Cleveland Elec. Illuminating Co., Cleveland, Ohio, October, 1949.

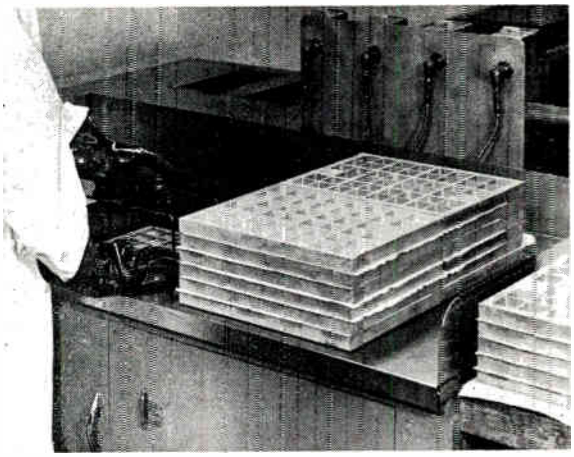


Fig. 10—Photograph of ultrasonic cleaning in production line.

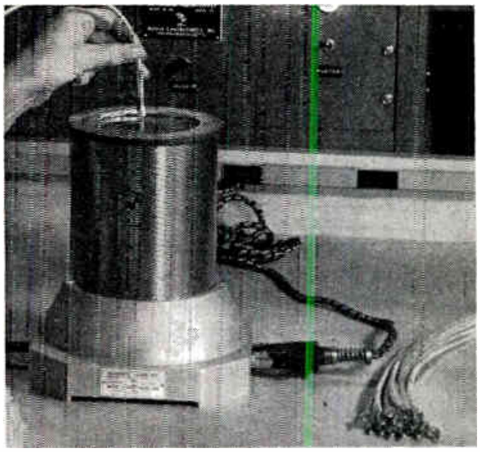


Fig. 11—Ultrasonic solder pot.

the parts relative to the position of the transducers.[7]

The spectacular advantages of ultrasonic cleaning extend into many related fields such as in the high-speed removal of heat treat scale from metals which can be accomplished in low-concentration pickling solutions operating at lower temperatures than is required by conventional methods. Another related application is micropolishing, which is accomplished by immersing the parts in an ultrasonic bath to which a fine abrasive is added. This method is particularly useful for polishing jewelry and precision parts of intricate shape where buffing methods are not feasible.

C. Metallurgical Applications

Ultrasonic energy has been employed to produce alloy-like mixtures of normally incompatible metals. For example, during the cooling of a molten mixture of iron and lead, the wide difference in melting points of the elements would result in a crude agglomeration of the lead-iron mixture. If the cooling takes place while the melt is subjected to intense ultrasonic activation, the two materials will solidify while they are held in a state of fine molecular dispersion, thereby producing a fine alloy-like mixture.

A similar application of ultrasonics has been used for the reduction of grain size in castings. By injecting intense ultrasonic energy into molten metals, it has been found possible to reduce the grain size of the solidified metal. Ultrasonics has also been effective in degassing molten metals, thereby producing denser castings.

The tinning or soldering of metals without the use of fluxes is made possible by simply dipping the metal parts into an ultrasonically-activated solder bath. The effect of the ultrasonic activation is to remove the oxide coating from the metal while it is submerged in the molten solder and immediately the metal becomes coated with the tinning solution. An additional beneficial effect of the ultrasonic activation is that the wetting action between the materials is increased and solder easily runs into tight crevices. By this procedure, dip soldering of printed circuits can be carried out without the fear of having "cold" solder joints and, also, because of the lack of necessity for using corrosive fluxes, fine wires may be ultrasonically soldered without fear of disintegrating the wire.

A commercial type of ultrasonic solder pot is shown in Fig. 11. In order to separate the ultrasonic transducer from the hot molten solder, a solid transmission line which is a multiple of one half the wavelength of sound in the material is used to transmit the sound energy from the transducer to the solder. The effectiveness of ultrasonic soldering is evidenced by the fact that a piece of untreated aluminum strip becomes brightly tinned by simply immersing the aluminum into the molten activated solder without the use of flux.

Ultrasonic equipment similar to that used for soldering applications can be used to improve the heat treatment of steel. For example, in the nitrogen hardening of stainless steels, a greater depth of penetration of hardening is achieved by ultrasonically activating the quenching bath.

D. Chemical Applications

The entire chemical industry will find the greatest variety of uses for ultrasonics. The terrific forces which are generated at the boundary region of two dissimilar materials when they are subjected to intense ultrasonic sound fields will cause such violent molecular activity that incompatible liquids such as mercury and water become emulsified. Solid metals immersed in ultrasonically-activated liquids may be broken down to become molecularly dispersed in the liquid. The tremendous increase in wetting action between materials in the presence of ultrasonic energy can be utilized for accelerating chemical changes or, in some cases, new chemical changes may take place that normally would not occur.

Fine-grained photographic emulsions have been produced ultrasonically much as fine-grained castings have been similarly produced. Artificially-produced colloidal-

[7] U. S. Patent No. 2,702,269; February 15, 1955.

like mixtures of metallic particles in liquids have been created by ultrasonic activation. High-frequency ultrasonic energy has successfully been used to control the molecular weight or viscosity of high polymers.

Electrodeposition of metals has been greatly influenced by ultrasonics. For example, it has been possible to chrome-plate titanium by ultrasonically activating the electrolyte. During electrolysis, the particles that pass from the anode to the cathode have been successfully broken down to finer size by subjecting the collecting cathode to ultrasonic vibration.

Future Industrial Applications

During the past decade a great many useful industrial applications for ultrasonics have been developed. In this same period, unfortunately, some equipment has been placed on the market that was not yet developed to operate reliably for the uses intended, which led to disappointing experiences in certain instances. On the other hand, there have been many successful pioneering industrial applications of ultrasonics that have become indispensible to the manufacturing processes in which they were used.

An important advancement of the past decade is the development of a variety of low-cost transducer materials that can be mass-produced in unlimited quantities to permit the economic production of reliable high-power transducers to fit practically any required application. Five primary materials are now available which will satisfy practically any industrial ultrasonic transducer requirement: quartz, ammonium dihydrogen phosphate crystals, magnetostrictive materials, polarized barium titanates, and polarized lead zirconates. Among these materials, quartz offers the greatest stability and uniformity, and its field of use is for such applications as frequency control, delay lines, and other high-frequency ultrasonic applications such as thickness gauges and flaw detection.

ADP crystal is stable and very uniform in its characteristics, and is primarily employed in the ultrasonic frequency range up to a few hundred kilocycles. Its greatest field of use is in applications requiring high precision such as in measurement standards and sonar applications requiring accurate beam-forming arrays. Magnetostrictive nickel and other alloys are useful for the lowest ultrasonic frequencies generally below 50 kc, and where extreme ruggedness is a major requirement. Of all the materials listed, only the magnetostrictive metals are free from tensile fractures under high imposed stress.

Polarized barium titanates are the lowest in cost among all the transducer materials, so it is natural to find it used where low cost is the predominant requirement in the design. Barium titanate loses some of its polarization and piezoelectric activity at temperatures above 175°F and it is also subject to deterioration with age. The most recently-developed material is the polarized lead titanate. These materials are fired ceramics similar to the barium titanates so that relatively low cost can be expected as more sources of supply become established. The lead zirconates are more stable than barium titanates and will operate at temperatures well beyond the boiling point of water. It seems very likely that the present increasing availability of lead zirconate as a source of low-cost high-power ultrasonic energy will enormously expand the uses of ultrasonics in industry. Sources of low-cost electrical power for operating high-power transducers as are now available with the new solid-state rectifiers and also with inexpensive rotating high-frequency alternators will also contribute to the widespread use of high-power ultrasonic equipment.

The greatest growth in industrial ultrasonics will undoubtedly be in the high-power field with large-scale industrial cleaning being perhaps the first big market. Up to the present time, high-power equipment operating from vacuum tube amplifiers has been too expensive to find widespread use. The availability of low-cost high-power ultrasonic equipment will also permit large-scale chemical and metallurigcal applications that heretofore were only of academic interest because of the prohibitive cost of the equipment.

In the low- and medium-power range, we shall see a greater acceptance of ultrasonic equipment in the medical field for both diagnostic and therapeutic applications. Continuing medical research with ultrasonics will find many ways to apply this equipment for the benefit of mankind. An indication of its increased acceptance by the medical profession is evidenced by the use of ultrasonic therapy by the President during his recent back difficulty. Many additional applications will be developed for new instrumentation and inspection techniques as well as in automatic and remote control systems in which controlled ultrasonic beams will efficiently and economically take over many menial chores now being performed by more laborious methods.

Present and Future Applications of Ultrasonics in Biomedicine*

WILLIAM J. FRY†, MEMBER, IRE

Summary—The current importance of ultrasonic energy for the investigation and modification of biological systems is reviewed and the immediately foreseeable potential apparent to the author is predicted. The uses of ultrasound in basic research studies of biological systems and medical applications are conveniently grouped into two major categories: passive uses, or those in which the acoustic field does not significantly modify either permanently or temporarily structure and/or function of the system, and active uses, or those in which modification of the system is the objective. Included within the first category are: absorption spectroscopy of solutions of macromolecular species, microscopy of cells and tissues, absorption characteristics of gross tissue, and visualization of soft tissue structure and its dynamics. The second category includes: the use of ultrasound in neuroanatomical and neurophysiological studies and the treatment of neurological disorders by the production of selective permanent or temporary changes in arrays of sites in the central nervous system, destruction of carcinogenic tissue, modification of endocrine glands, investigation of contractile and other properties of muscle, and the potentiation of ionizing radiation by simultaneous application of ultrasound. Applications of ultrasound in the biomedical field of a primarily technological nature are either mentioned only briefly or omitted entirely from the review.

The "socioeconomic" factors which determine the level of financial support, and thus the rate of scientific progress, for a field are also briefly mentioned.

INTRODUCTION

THE STUDY of the interactions of various forms of energy with a system furnishes information on its structure and dynamics. Appropriately controlled ultrasonic energy (arbitrarily defined as sound above 20-kc frequency) constitutes a tool of considerable power for the elucidation and modification of biological systems. This potential of ultrasound, for the investigation of biological material, has been glimpsed partially as a result of work already accomplished with the methodology and instrumentation that various investigators have developed under scantily supported research programs. This glimpse of the potential future is also indicated by ideas for immediate advances, a number of which are mentioned in this paper, but which are not active experimentally at the time of writing.

A review of this length cannot hope to be exhaustive and therefore some of the important work included under the title is either mentioned only briefly or in some cases entirely omitted. However, the specific directions of investigation discussed here include most of the major ones which appear fruitful to the author at the present time. These directions, both basic research (including animal and human) and technological applications (including medical therapy), are conveniently considered in two categories: 1) passive uses; that is, those employing acoustic field conditions which do not significantly modify either permanently or temporarily the biological system, and 2) active uses; that is, those employing acoustic field conditions which modify either permanently or temporarily the structure and/or dynamics of the biological system.

A paper of this type in which both past accomplishments in a field are scrutinized and future potential is predicted might well consider, in addition to the scientific and technological features of such endeavor, the associated "socioeconomic" factors. These latter are equal in importance to the former in determining the rate of progress that can be achieved in any discipline and their prominence in determining the present status of the field of ultrasound in biomedicine is strikingly apparent. Brief reference to these nontechnological factors is included here as a specific example of the way scientific progress in general is determined by the philosophy and/or modus operandi of an organization or social structure which provides the financial support for research.

I. PASSIVE USES OF ULTRASONIC FIELDS

Ultrasonic fields conditions, which do not disrupt or significantly modify biological systems, can be used to detect structure from microscopic to macroscopic levels and to elucidate function. Measurements of ultrasonic absorption and/or reflection characteristics usually constitute the basic information from which such knowledge is deduced. The study of biologically significant structure at the microscopic and submicroscopic levels is subdivided here into two categories: 1) ultrasonic absorption spectroscopy of solutions of biologically important macromolecular species, and 2) ultrasonic microscopy of cells and tissues. Similarly, the investigation of biological systems with passively employed ultrasound on the macro level is divided into two categories: 1) ultrasonic absorption characteristics of gross tissue, and 2) ultrasonic visualization of soft tissue structure and its movement. The present state of development in each of these subdivisions is considered briefly and future potential is indicated.

A. *Ultrasonic Absorption Spectroscopy*

From measurements of the ultrasonic absorption characteristics of macromolecular configurations and molecules of biological interest such as proteins, information on the states of aggregation and organization of these species in solutions and suspending media can

* Received by the IRE, September 5, 1961; revised manuscript received, November 8, 1961.

† Biophysical Research Lab., University of Illinois, Urbana, Ill., and Interscience Research Corp., Urbana, Ill.

be deduced. Work in this field is in the early stages and no *specific* mechanism to account for the major portion of the absorption has yet been formulated and verified. However, a phenomenological approach in which relaxation theory is applied to explain the frequency dependence of absorption of some of the materials on which data is available has received some attention. That the ultrasonic absorption coefficient might well constitute a sensitive indicator of configuration states of macromolecular species in solutions or embedding media is already suggested by the work of Schwan, Carstensen and collaborators. These investigators have reported the most precisely determined and comprehensive velocity and absorption results on erythrocytes in suspension, hemoglobin in solution, plasma, extracts from liver, etc., as a function of the concentration, temperature, and ionic content over the frequency range from 0.3 to 10 Mc [1], [2]. However, measurements over a considerably wider frequency range will probably be required before much information on structural features of biological macromolecular species in solutions and in cells can be deduced from such data.

The type of ultrasonic instrumentation employed in this work is illustrated in principle [3] in Fig. 1. The method makes possible the attainment of a high degree of accuracy in the measurement of absorption coefficients and velocity differences. The basic principle is the comparison of the acoustic properties of the medium of interest with those of water or other standard liquid. To achieve this, the sound tank is divided into two compartments by a diaphragm as illustrated. One compartment contains the standard solution, the other contains the medium. The ultrasonic transducers, which generate and detect the acoustic pulses, are moved by a precision sliding mechanism. Movement of this mechanism makes possible the substitution of any desired pathlength of the solution or suspension for an equal pathlength of the standard. Absorption coefficient values with an uncertainty as small as 0.002 cm^{-1} are computed from measurements of attenuation. By determining the phase of the signal at the receiver transducer with respect to the signal at the generator, it is possible to accurately measure acoustic velocity differences, and determinations to one part in 100,000 are reported.

The method just discussed for measuring absorption coefficients of biologically significant media is not useful if both the volume of material available and the magnitude of its absorption coefficient are small since the apparatus of Fig. 1 then requires a quantity of material of the order of 0.1 to 1 liter. When the volume of material available is small, for example, a few cubic millimeters, the thermocouple method illustrated in Fig. 2 can be used. This is an adaptation of the thermoelectric probe developed at this laboratory for measuring absorption coefficients of tissue and for determining acoustic field configurations [4]–[6]. From the recording of the transient rise in temperature of the medium surrounding the junction, following exposure to a pulse of ultrasound of rectangular envelope, it is possible to calculate the absorption coefficient of the embedding medium. This method, which requires that the absolute sound intensity and the heat capacity of the medium per unit volume be known to the same degree of accuracy as that desired in the absorption coefficient,

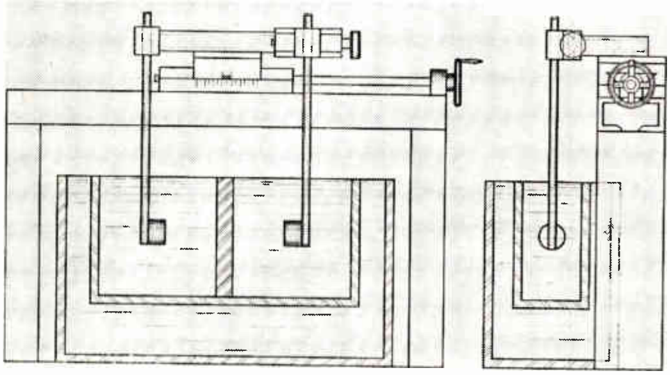


Fig. 1—Apparatus for precision measurement of ultrasonic absorption coefficients and velocity differences. Comparison of the acoustic properties of the medium of interest, which fills the left compartment of the sound tank, with a reference or standard liquid filling the right compartment of the tank, is achieved by this type of instrument. After choosing an appropriate spacing for the transmitter-receiver separation, the relative amplitude and phase of the received signal are determined as a function of the path length of the medium substituted for an equal path length of the standard by movement of the transmitter-receiver assembly relative to the two-compartment sound tank. The duration of the plane wave pulses and the dimensions of the tank are chosen so that reflection of ultrasonic energy from the tank walls does not interfere with the measurements.

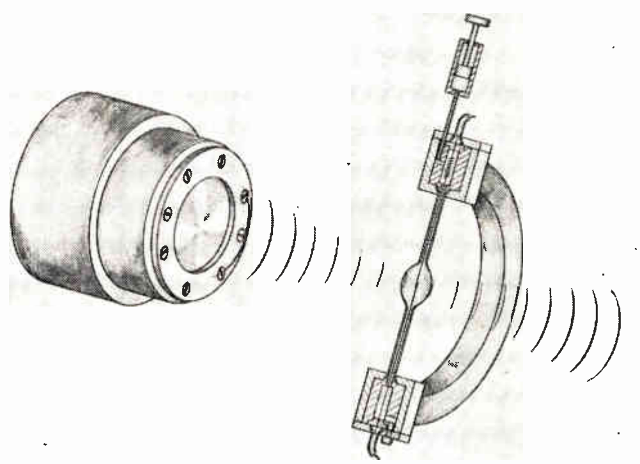


Fig. 2—Method of measuring the ultrasonic absorption coefficients of materials on samples as small as a few cubic millimeters. The material is supported in a tubular shaped container, formed of thin plastic sheet, in which a thermocouple is mounted in approximate coincidence with the long axis. The ultrasound pulses of rectangular envelope are focused in the volume of the medium surrounding the thermocouple junction and the voltage generated across the couple leads indicates the transient change in the temperature of the medium in the immediate neighborhood of the junction. From an analysis of the form of the temperature response, that part caused by absorption of the ultrasound in the body of the medium can be deduced, and from this and a knowledge of the heat capacity of the medium per unit volume and the absolute sound intensity, the acoustic absorption coefficient can be computed.

promises to be extremely useful in the further development of the field of ultrasonic absorption spectroscopy of solutions and suspensions of biological significance.

Absorption of ultrasound occurs in a *homogeneous* liquid when the changes in density are not in time phase with the changes in pressure. This type of behavior is produced by a variety of mechanisms classified into the two categories: viscosity or frictional lag and relaxation processes. The mechanism of the first category results from the fact that liquids exert resistance to shearing forces and for a Newtonian liquid this force is proportional to the velocity gradient and to the *constant* coefficient of shear viscosity. In non-Newtonian liquids the value of the viscosity coefficient depends upon the velocity gradient. The term "relaxation" process is used to include heat conduction, which is relatively unimportant for biological systems, and thermal and structural relaxation mechanisms [7]. Thermal relaxation results when the temperature of a fluid is changed due to the propagation of a sound wave, and the transfer of energy between external and internal degrees of freedom of components of the medium requires a time interval comparable to the period of the wave. The out-of-phase relations which result cause a conversion of acoustic into thermal energy. Structural relaxation results when a time interval comparable to the period of the sound wave is required for redistribution of the mutual orientation or degree of association of the components of the medium in response to the pressure changes produced by the sound field. Since configuration changes entail energy transfer, some acoustic energy is converted into heat when the rate of the redistribution process is too slow to follow the variations in pressure produced by the sound field without appreciable lag. If the equilibrium state of a chemical reaction is dependent to a sufficient degree upon the pressure of the reactants then acoustic absorption of the relaxation type takes place.

Absorption of ultrasound can occur as the result of relative motion between suspended structures and imbedding medium [8]–[11], *e.g.*, in cell suspensions or tissue. Such relative motion results since the densities of the suspended particles or their constituent parts are not, in general, equal to that of the suspending matrix. This relative motion gives rise to friction forces which cause absorption of acoustic energy.

Liquids can be subdivided into three classes on the basis of their ultrasonic absorption characteristics [12]–[14]. Members of the first group have ultrasonic absorption coefficient values close to those calculated on the basis of shear viscosity losses. Monatomic liquids such as mercury are in this group as well as some diatomic liquids such as oxygen and nitrogen. As far as is known at the present time, no biologically interesting materials are in this class. The second group is characterized by a positive temperature coefficient of absorption; that is, the magnitude of the absorption coefficient increases as the temperature increases and reaches values from three to four hundred times those calculated on the basis of a shear viscosity mechanism. Unassociated liquids such as benzine and carbon tetrachloride are in this group. The "excess" absorption of these liquids may be the result of a slow rate of exchange of energy between the external and internal degrees of freedom. The third group, the associated polyatomic liquids, exhibit a negative temperature coefficient of absorption. Liquids of this type include water and alcohols. The blood proteins exhibit such a negative temperature coefficient of absorption [15] but nerve tissue exhibits a positive temperature coefficient [7].

It is not possible in a review of this length to consider in any detail the methods of analysis leading to the identification and separation of shear viscosity mechanisms from relaxation processes and the identification of specific relaxation mechanisms and their characteristic frequency distributions. (See, for example, reference [7] for a discussion.)

The primitive state of knowledge of this field is apparent when it is compared with the present status of electromagnetic spectroscopy of atoms and molecules. In the ultrasonic case only a single molecular species—hemoglobin—has been studied in any detail and the underlying mechanism of absorption of not a single macromolecular structure has been worked out. When it is realized that classical electromagnetic spectroscopy (infrared, visible and X ray), which has yielded such a vast amount of information on the structure of molecules, atoms and crystals, can furnish little essential information on the macromolecular features of structures of interest here, and that electron microscopy is not readily adapted to the examination of solutions or suspensions, it is apparent that ultrasonic spectroscopy methods, in which the values of the absorption coefficient are critically dependent upon the macromolecular configurations in the embedding media, should receive major attention. It is emphasized here and it will be increasingly apparent from subsequent sections of this article that ultrasonic methods of the type described are particularly useful for the study and examination of structure on the biologically significant level.

B. Ultrasonic Microscopy of Cells and Tissues

Since different protein solutions and presumably other macromolecular species of biological significance are characterized by different values of the ultrasonic absorption coefficient [2], it is possible to develop an ultrasonic absorption microscope [7] to determine distributions of molecular species on an intracellular level. Since the mechanisms of absorption of ultrasonic and electromagnetic energy are completely different in general, components of cells and tissues would not exhibit the same differential absorption of these two forms of energy. Consequently, it is reasonable to expect to detect and visualize structures ultrasonically which are not seen by microscopes using visible light or electromagnetic energy in other portions of the spectrum and vice versa.

The principle of operation of one form of an ultrasonic absorption microscope is illustrated in Fig. 3. Acoustic waves are generated in a coupling medium by a piezoelectric element such as an X-cut quartz plate vibrating in thickness mode and this liquid conducts the sound to the specimen which is interposed between the source and a small thermoelectric probe or array of such probes. This array detects the acoustic energy transmitted through the portions of the specimen adjacent to the individual probes and thus an acoustic image of the structure can be reproduced by moving the specimen relative to the probe array [7], [16].

It should be noted here that this same device provides a method for measuring acoustic absorption coefficients of solutions or suspensions in the kilomegacycle frequency range. For this use the fluid medium, whose absorption coefficient is to be measured, fills the chamber and a single thermocouple probe is moved along the direction of propagation of the sound in the medium. The transient change in the temperature at the thermocouple junction, resulting from conversion of a fraction of the energy of the acoustic pulse into heat in the immediate neighborhood of the junction, measured as a function of the spacing distance between the source and probe, provides the necessary information to permit a calculation of the absorption coefficient of the medium. Work now in progress at this laboratory has shown that such a device can be used to measure absorption coefficients from the Mc frequency range into the kMc frequency range. At the present time it appears that an operating frequency spectrum of four orders of magnitude will be feasible.

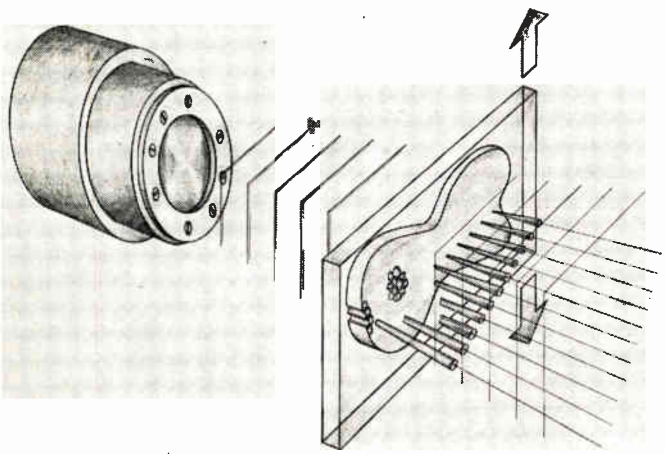


Fig. 3—Principle of operation of an ultrasonic microscope capable of resolving structure on an intracellular level. Plane pulsed waves of sound are generated by a piezoelectric plate element vibrating in thickness mode at any of a series of odd harmonics of the fundamental resonant frequency. After passing through the structure under examination, which is coupled in close proximity to the transducer element by a fluid medium, the acoustic energy passes into a medium imbedding an array of thermocouple junctions immediately adjacent to the specimen. Since differential absorption of the sound occurs in the specimen it is apparent that the temperature distribution defined by the different thermocouple outputs is a manifestation of structure in the specimen. By moving the specimen in its plane and in a direction at right angles to the direction determined by the linear array of junctions it is possible to obtain a two dimensional ultrasonic picture of microstructure of the specimen. The resolution of such a device is dependent on the size of the thermocouple junction, the thermal properties of the couple leads and of the imbedding medium, the ultrasonic pulse duration and intensity and the magnitude of differential absorption characteristic of the microstructural features.

Instruments designed on the basis of the principle illustrated in Fig. 3 have not yet been employed to examine or study the micro detail of cells or tissue structures since such devices are still in an early stage of development at this laboratory and much additional sophistication will be required before the microstructure of biologically interesting materials can be studied. However, a form of this device employing a single thermocouple junction is in use on the measurement of absorption coefficients of liquids from the low Mc into the kMc frequency range.

Calculations of the resolving power of an ultrasonic absorption microscope of the type discussed have been made [7], and it appears feasible on the basis of the analysis to expect a resolution of a few tenths of a micron at the higher operating frequencies. High resolving power requires operation in the kMc frequency range, not because of wavelength limitations, but rather because its attainment requires the use of short pulse lengths to minimize the effects of thermal conduction. The use of increasingly shorter pulse lengths requires increasingly higher ultrasonic absorption coefficient values in order to achieve a sufficient temperature change to be detectable, and since the absorption coefficient per unit pathlength increases with the frequency it is necessary to operate in the UHF range.

It should be noted here that an ultrasonic microscope based on lenses or reflector focusing principles would not be practical at the frequencies necessary to obtain high resolution since the acoustic absorption coefficient values per unit pathlength at these frequencies in liquid media are too high for the pathlengths which would be required in such designs.

Ultrasound employed in the manner indicated in this section would appear to have considerable potential for the examination and elucidation of biological structure on a microscopic level. It not only constitutes a new way of detecting structural features but it is also well suited for the examination of tissue or cells *in vivo*.

C. Ultrasonic Absorption Characteristics of Tissue

The absorption and the propagation characteristics of tissue for ultrasound must be determined to provide information basic to the utilization of this form of energy for: examination of the gross anatomy and normal dynamics of organs *in vivo*, diagnosis of malfunction and pathological states, modification or destruction of internal structure without incising of the intervening tissue in animals and humans. Quantitative information is essential for a basic understanding of the manner in which proteins and other macromolecular species contribute to the absorption in tissue and for an understanding of absorption mechanisms related to tissue structure as contrasted with that resulting from processes occurring on the "molecular" level.

As will be indicated in Section I-D of this paper, the ultrasonic visualization of soft tissue structure offers enormous potential for the examination of the anatomy of intact organisms and the associated dynamics but the acoustic propagation characteristics of such tissue must be understood in order to achieve the potential. In addition, the entire field concerned with the modification of tissue structure, reviewed in Section II, is dependent upon an accurate knowledge of the propagation characteristics of ultrasound in tissue. Both velocity [17] and absorption characteristics [18], [19] must be known.

In view of the incompleteness of the acoustic propagation data [20] in this field it is apparent that much experimental work and theoretical analysis must be accomplished before the absorption characteristics of a variety of tissue structures in the ultrasonic frequency range are both known and understood [8], [9], [11], [21]. A recent example of the type of data and analysis required for a basic understanding of absorption mechanisms in tissue in the megacycle frequency range is that reported from this laboratory on the absorption in lung [22].

Not only is it essential that the acoustic propagation characteristics of normal tissue be measured and understood but it is also necessary that the characteristics for pathological tissue be determined. This latter is particularly important if ultrasound is to be useful passively for diagnostic purposes.

D. *Ultrasonic Visualization of Soft Tissue Structure*

Nearly all soft tissue components possess essentially the same X-ray density, and therefore structural features of soft tissue cannot be visualized directly by using this radiation. By contrast, since soft tissue structure is not acoustically homogeneous (slight differences in acoustical impedance are present), small fractions of the ultrasonic energy incident on interfaces within or between such tissues are reflected. Since under the ultrasonic irradiation conditions required, no deleterious effects are produced in tissue, this form of energy provides a method of visualizing tissue structures (both normal and pathological) in three dimensions [23], [24]. In addition, their dynamic characteristics can be studied; for example, it is possible to observe the movements of major blood vessels and characteristics of heart motion [25]–[29]. Ultrasonic visualization methods have been applied in ophthalmology [30], and tumor localization and diagnosis have received the attention of a number of investigators [31]–[33].

The wavelength of the radiation determines the resolution which is obtainable and since the velocity of sound in soft tissue is close to that of water it is necessary to operate at frequencies of 1 Mc and above if resolution of the order of a millimeter is desired. At these frequencies the absorption per centimeter of tissue path is relatively high so that the received acoustic energy reflected from interfaces at different tissue depths is considerably affected in magnitude by this factor. By the use of focused beams and short pulse lengths it is possible to obtain resolution in both azimuth and range.

The instrumentation required for visual presentation of the reflected ultrasonic signals as a reconstruction of the tissue structure and for detection and processing the signals for the presentation system is quite sophisticated. It is necessary because of the tremendous range of amplitudes of the returning reflected signals to design the system with gain compensation. Since the orientation of the reflecting interface with respect to the direction of the incident beam is important in determining the direction of the reflected acoustic energy, and consequently the amount of energy which is returned to the receiving transducer, it is necessary to view each site in the structure from a number of different directions. The importance of this is apparent when it is observed that, for example, an angular shift from normal incidence of 10° can cause a reduction of the amplitude of the received signal by a factor of 100. Compound scanning techniques thus improve the picture detail and eliminate artifacts (see, for example, reference [7] for further details).

The ultrasonic method of soft tissue structure visualization just described—that is, one depending upon existing small differences in the acoustic impedance at tissue interfaces to provide reflected energy—although furnishing a tremendous amount of information not obtainable previously, does not permit the visualization of all tissue structures of interest. In addition, increased contrast between structures now detectable is extremely desirable in many cases. A considerable advance in these two directions can be expected by a new ultrasonic visualization method recently proposed and evaluated analytically by the author. The advantages offered by the new method can be illustrated by the following specific comparison. Present ultrasonic visualization methods do not detect boundaries between any internal brain structures since the acoustic impedance differences are not of sufficient magnitude to provide detectable amounts of reflected energy. However, it is presently possible to detect the ventricular system since the impedance difference between brain tissue and cerebrospinal fluid is sufficiently great to provide the necessary reflection. Since the study and modification of brain structure and function requires the placement of beams of radiation, electrodes, cannulae, etc., at specified sites in particular structures it is extremely important to have a system which permits visualization of all major brain structures (that is the boundaries between the major gray and white matter masses). This can be achieved and presently undetectable structural details in other tissues and organs of the body can be visualized by designing an instrument to induce temporary impedance differences at tissue interfaces. Such impedance changes can be induced at interfaces where the ultrasonic absorption coefficient changes in value; for example, in the brain the white matter exhibits an ultrasonic absorption coefficient approximately 1½ to 2 times that characteristic of gray matter.

The basic principle underlying the method proposed by the author is illustrated in Fig. 4. Two pulsed properly synchronized scanning beams of ultrasound are employed, one (the modifier) of relatively long pulse duration and the other (the analyzer) consisting of a series of pulses short compared to the first. These sound beams are, in general, of different frequencies. The long-pulse focused beam or modifier is used to produce in a relatively small volume of tissue (the site within the specimen receiving immediate attention) transient changes in temperature by conversion of acoustic energy into heat. In the boundary regions between tissue structures characterized by different values of the ultrasonic absorption coefficient transient steep gradients of temperature would thus be produced. Since the velocity of sound is temperature dependent corresponding gradients in the acoustic velocity would thus exist temporarily in these boundary regions. The second focused beam, the analyzer, would then be partially reflected from regions exhibiting such gradients. Calculations demonstrate that sufficient acoustic energy can be reflected when temperature differences which would not produce damage to the tissue are employed. This new method will achieve not only visualization of structures which are undetectable by instruments employing the present methodology, but will also permit increased contrast to be achieved between structures detectable with the present equipment, but for which improved contrast would be extremely advantageous. The new method is well suited to the centering of attention on chosen small volumes of tissue to provide enhancement of echoes from this region compared to those from surrounding or embedding structure. This is achieved since the temperature gradients which either induce or increase acoustic impedance differences at boundaries are of appropriate magnitude only in the region of the focus of the modifying beam. Thus, three major advantages over present methods of ultrasonic visualization of soft tissue structure are provided: visualization of structures not now detectable, increased contrast between presently detectable structures, and relative suppression of reflected energy from embedding and intervening tissue.

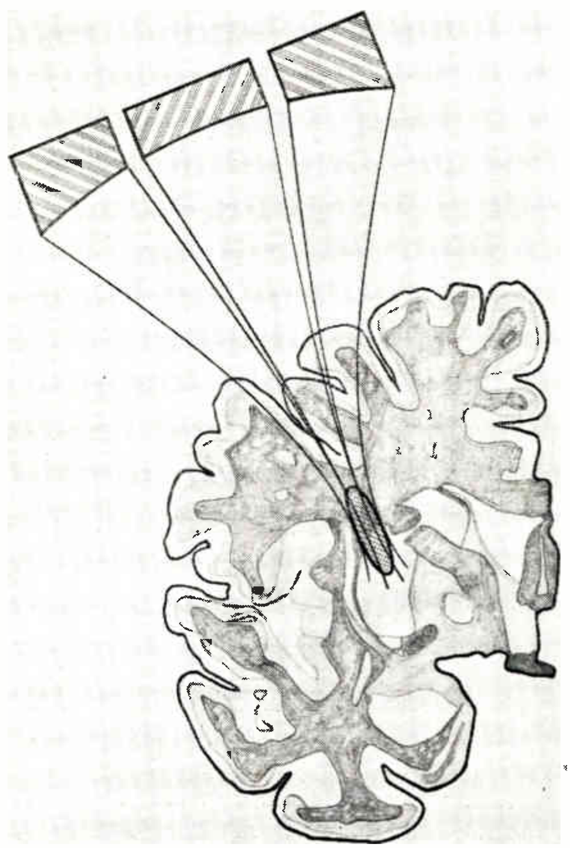


Fig. 4—Schematic diagram of ultrasonic method of visualizing soft tissue structure by employing one focused and pulsed beam of sound (included within the two concentric cones which surround the inner beam illustrated in the diagram) to induce temperature gradients of small nondamaging amplitude at tissue interfaces (brain, gray-white matter interface in the figure) and other structural features at which changes in acoustic absorption occurs. Since the velocity of sound is temperature dependent acoustic velocity gradients are thus induced in the regions of structural change. These gradients are then detected and their geometric positions determined by a second pulsed and focused ultrasonic beam (included within the innermost cone of the diagram) appropriately synchronized with the first. Both the temperature modifying and the examining beams are moved to scan the specimen. The duration of the pulse length of the former is relatively long (0.1 to 1 sec) in order to induce velocity gradients of appropriate amplitude without the use of extremely high ultrasonic intensities and the pulse length of the latter is very short (1 μsec) in order to provide appropriate resolution (0.1 to 1 mm).

The potential of ultrasound for examination of tissue is apparent when it is observed that the present widespread use of X rays in diagnosis is based on the visualization of a much more restricted range of structure than that which will be possible with instruments based on ultrasonic methods. The implications can again be emphasized by reference to the neurological field. No present method can achieve visualization of internal brain structures (nuclei and fiber tracts) and a variety of landmarking methods have been devised to provide systems of reference which at best furnish meager information. The direct visualization of all major brain structures, which the new ultrasonic method outlined briefly here is expected to provide, now appears to be within our grasp.

II. Active Uses of Ultrasound

Modification of structure and/or function at deep sites in tissue both for basic research on experimental animals [34]–[36] and humans [37], [38] and for routine treatment of disorders in medical therapy can be achieved with ultrasound. By focusing it is possible to produce changes at any desired depth without damage to intervening tissue and by appropriate control of the irradiation parameters selective action on components of tissue structure can be accomplished [34], [39]–[46]. Either permanent or temporary [34], [47] changes can be produced as desired. The central nervous system has been the object of the most intensive study using precisely controlled intense ultrasound and the present status of this area of activity will be briefly reviewed and future potential indicated.

A. Central Nervous System

In the central nervous system the selectivity that can be realized at the present time by employing ultrasonic field parameters corresponding to pressure amplitudes of 5 to 50 atmospheres and appropriate durations of exposure constitutes an extremely powerful tool for investigation of normal structure and function. By moving the focused beam about with an appropriate precision positioning system changes in regions of arbitrary sizes, shapes and orientations can be produced in deep sites without damaging intervening tissue. Microscopic studies of irradiated tissue show that blood vessels are the part of the brain most resistant to the action of the sound and also demonstrate that it is possible to interrupt selectively the neural components of white matter (the regions which contain only nerve fibers) without damaging gray matter (the regions containing the nerve cell bodies) subjected to the same irradiation conditions. No evidence of long term cumulative action comparable to that produced by ionizing radiation and ionizing particles exists.

The minimum size region that can be affected is determined by the volume of the focal region and this is, in turn, dependent upon the wavelength of the sound. At 1 Mc the wavelength in brain tissue is $1\frac{1}{2}$ mm and therefore the minimum size tissue volume that can be affected in easily reproducible fashion in homogeneous tissue is a few cubic millimeters. At higher frequencies changes can readily be restricted to smaller volumes of tissue if so desired. For example, at 4 Mc the tissue volume affected for a single exposure position can be as small as a few hundredths of a cubic millimeter. However, it is not possible to decrease the wavelength indefinitely, by increasing the operating frequency, to affect smaller and smaller tissue volumes in depth because as the frequency increases the absorption per unit path length also increases—linearly in the range of interest (at 1 Mc the pressure absorption coefficient is 0.1 cm^{-1}).

The procedure employed at the present time for the irradiation of brain structures requires that a portion of the skull be removed since bone has a much higher absorption coefficient [19] than soft tissue and, therefore, if left in place, thermal damage to underlying tissue would occur as the result of heat conduction from the irradiated bone. In addition, the irregularities of skull thickness and orientation would modify the beam shape and reduce the accuracy of positioning the focus. When the new ultrasonic visualization method discussed in the previous section is in operation, it may be feasible to develop instrumentation to eliminate the present requirement of bone removal since it will then be possible to view directly the focus of the ultrasonic beam, used for modification or disruption of neural structure, in the specific sites to be affected. The attainment of such a goal must however await the development of very elaborate instrumentation. Obviously all possible effort should be bent toward this objective since its realization would make the precision modification of brain structure for the treatment of a variety of neurological and other disorders in the human extremely simple to achieve in practice since no preliminary or auxiliary procedures such as the present surgical and ventriculographic preparations would be required.

Much work has been carried out on experimental animals to investigate the types of tissue changes which are produced (see references listed in the introduction to this section) and as a result ultrasonic methods are currently being employed at this laboratory in fundamental research studies of a neuroanatomical and behavioral nature. Anatomical investigations [35] heretofore not feasible are now possible since changes can be confined to selected volumes of tissue without *any* damage to other structures. When the desired "lesion" is achieved there is no ambiguity regarding interpretation of results such as ordinarily arises because of damage to other tissue. Fig. 5 illustrates in cross-section lesion arrays of various configurations in cat brains produced by focused ultrasound.

The advantages for neurophysiological and behavioral studies of the ultrasonic method of producing *permanent* changes in brain structures are similar to those for neuroanatomical investigations; that is, the production of selective lesions of desired sizes and shapes in combinations of structures without damage to other regions.

The extensive work carried out on experimental animals provides the basis for employing ultrasonic methods to study and modify the symptoms of human neurological disorders and a variety of such disorders has received the attention of the author and collaborators up to the present time [34], [37], [38]. These include: the tremor and rigidity of Parkinson's disease; involuntary movements of cerebral palsied individuals; intention tremors and nonpatterned movements; intractable pain, hyperesthesias and dysesthesias following cerebral vascular accidents; and phantom images and pain in amputees. The symptoms indicated can be affected by modifying various brain structures and the ultrasonic work already accomplished has provided much new information on the identification of specific structure subserving the underlying mechanisms. The flexibility and absence of stress during irradiation of the awake human patient makes possible the subtle and consistent step-wise modification of the neurological status as determined by continuous interview and examination during the procedure.

The instrumentation necessary for attaining the flexibility, reproducibility and accuracy for producing *permanent* changes at desired arrays of sites in the hu-

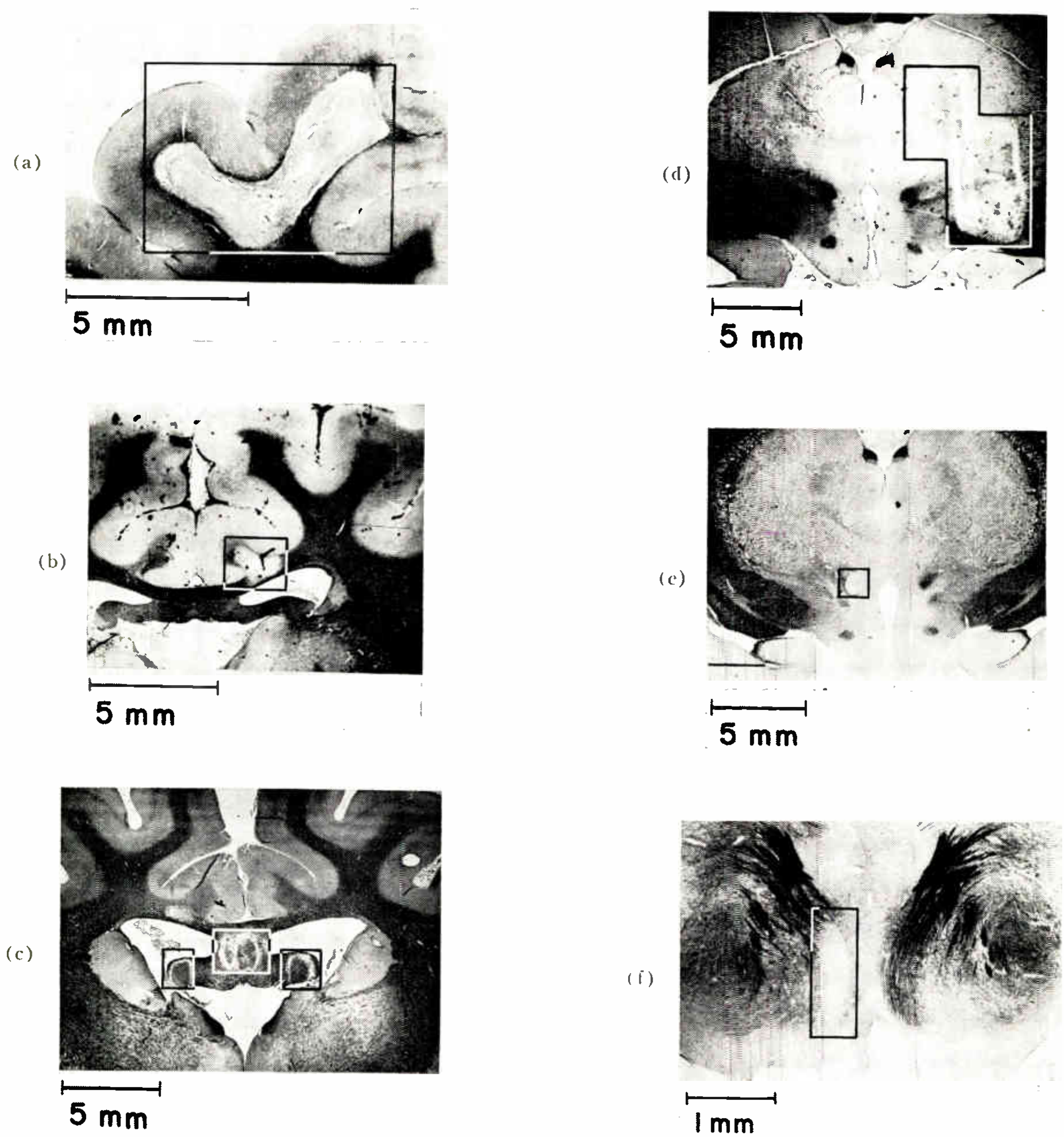


Fig. 5—Stained tissue sections of cat brain showing various shaped ultrasonic lesions in a variety of structures. The white matter or nerve fiber tract regions are stained dark and the gray matter or nerve cell body regions are relatively unstained. (a) Lesion in the subcortical white matter with no invasion of the immediately adjacent cortical gray matter. The brain cross section shown exhibits the maximum dimensions of the lesion. Its extent perpendicular to the cross-section is small, It was produced by moving the focus (1 Mc sound) in the plane of the section with a 1 mm spacing between adjacent sites. (b) Lesion placed to interrupt longitudinally running fibers—the cingulum running perpendicular to the plane of the section—in the white matter of the cingulate gyrus. The lesion was produced by placing the focus at an array of sites spaced 0.2 mm apart, 4 Mc sound. (c) Three lesions in the subcallosal fornix. The longitudinally running fibers are interrupted in the lateral half of the structure bilaterally and the fibers in the medial portion of the structure in the plane of the section are destroyed. No damage was produced in the corpus callosum and underlying thalamic structures: 0.2 mm spacing between adjacent sites of the array, 4 Mc sound. (d) Two thin rectangular sheet lesions in the thalamus and subthalamus. These lesions are thin in the direction perpendicular to the section shown, 4 Mc sound used. (e) Interruption of the mammillothalamic tract: 0.2 mm spacing, 4 Mc sound. (f) Lesion in the medial part of the medial mammillary body. The lesion extends throughout the length of this structure, *i.e.*, for 1½ mm in the direction perpendicular to the section shown. The lesion is shown in cross section: 0.2 mm spacing, 4 Mc sound.

man brain is quite elaborate and not inexpensive. The apparatus presently employed, which is illustrated in Fig. 6, includes: a stereotaxic head holder incorporating X-ray equipment; ultrasonic focusing transducers; a positioning system for placing the focus of the beam at desired predetermined arrays of sites; electronic equipment for supplying the necessary electric excitation to the transducers and for accurately controlling the sound level and duration of exposure; calibration instrumentation for precisely determining acoustic field configurations and for measuring absolute sound levels; and auxiliary equipment for control of factors such as coupling liquid temperatures [34], [36], [48].

Intense ultrasound can also be used to produce temporary, that is, reversible changes in the central nervous system. This is demonstrated by results obtained on the visual system of the cat [34], [47]. For example, when an eye of the animal is subjected to flashes of light or if the optic nerve is stimulated electrically, electrical events occur in regions of the brain concerned with the receipt and processing of incoming visual information. The pattern and complexity of these electrical events in the visual cortex are dependent upon a number of factors and various components of the evoked response can be identified depending upon the level and type of anesthesia, the temporal sequence of exciting stimuli, the temperature of the brain, etc. When the focus of the sound beam is placed in structures (*e.g.*, lateral geniculate nucleus) concerned with the processing and transfer of information and an appropriate set of irradiation parameters are employed, it is possible to produce changes in the magnitudes and latencies of the components of the potentials evoked in the visual cortex. A focused sound beam employed in this fashion constitutes an analyzer for the complex circuitry of the brain. That is, nerve pathways and centers over which information is transmitted and processed can be located and studied by observing the temporary changes induced by focused ultrasound in, for example, electrical responses manifested in brain structures following stimuli of various types and configurations.

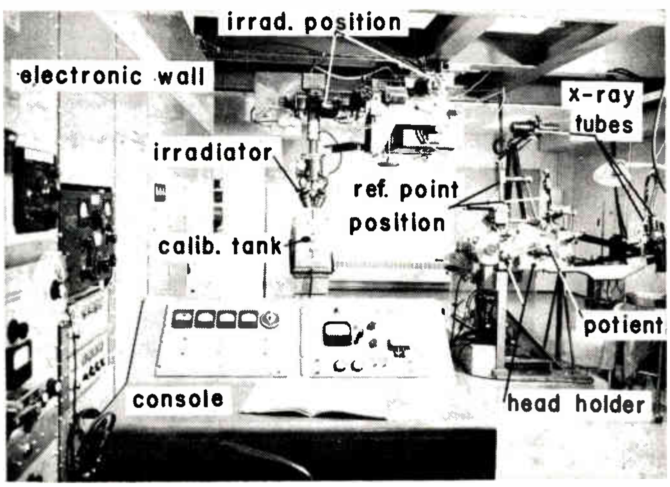


Fig. 6—View of the irradiation room for modifying brain structures in the human. An awake comfortably positioned patient is shown with head supported in the holder by four stainless steel rods whose rounded tips engage hemispherical indentations in the outer table of the skull. The tips of these rods can be accurately repositioned in space by the micrometer adjustments provided on the supporting posts in order to permit a number of irradiation procedures if desired without the necessity of determining the coordinates of the internal landmarks more than once. The three X-ray tubes and cross hair positioning system employed in the coordinate determination of landmarks are shown mounted on the head holder. The four beam irradiator, which has provided the focused intense ultrasound for the human work accomplished up to the present, is in position over the calibration tank. It swings into place over the head of the patient after a hopper which supports the coupling liquid is positioned over and engaged in water tight connection with the patient's scalp. The placement of the focus of the ultrasound in the sites of the array of positions to be modified in the brain is accomplished by the overhead structure. Since the scalp is intact during irradiation (an appropriate portion of the skull bone is removed at an earlier procedure) there is no necessity for sterile procedure once the small scalp incisions at the sites where the supporting rods engage the skull are sealed. The electronic instrumentation is mounted on the panels on the left wall and in the control console in the foreground.

From the work that has been accomplished it is apparent that three-dimensional mapping of brain function of a type and scope heretofore completely unattainable will be possible by moving appropriately controlled focused beams of ultrasound through brain structures and observing resultant changes. Some of these possibilities are apparent when one considers that the resolution obtainable with ultrasonic methods can make a moving sound focus the equivalent of 10,000 to 100,000 electrodes capable of disturbing in controlled temporal sequences threshold relations of neural events in essentially any combination of structures in a single brain. Such methods will be extremely useful both in fundamental research studies on experimental animals and on humans and in human therapy. In the latter application it will not be necessary, for example, to destroy any brain structure to locate regions involved in specific malfunction or pathological behavior.

If will be possible to investigate relations between complex behavior and brain mechanisms with some hope of obtaining a basic understanding of the operation of the maze of inter-connections and inter-relations which exist. However, to bring such a program to fruition, major developments in instrumentation must be forthcoming. The instrumentation which is required (at an estimated cost of several million dollars for a single installation) will include: data storage equipment; data analyzers, programming apparatus, presentation systems for results of data analysis; anatomical and functional presentation systems; a multiplicity of focusing transducers; positioning systems; ultrasonic visualization instrumentation; and a variety of types of control equipment.

The study and correction of malfunctions of the human brain—the most complex mechanism yet identified by man—would appear to warrant the development and use of instrumentation which is at least comparable in sophistication to that currently in use in investigations of the "physical" world. We have a long way to go but it is to be hoped that support for such activity can be obtained so that the major contributions promised by new methodology, for example the ultrasonic procedures indicated herein, are forthcoming.

B. Other Structures

High-level ultrasound can be used to modify tissue other than that of the central nervous system but other structures have not received comparable attention. The application of intense ultrasound to the treatment of soft tissue tumors has been reported [49], [50]. An unfocused beam of large cross-sectional area (50 cm^2) and intensity levels up to 500 w/cm^2 were used to modify and destroy malignant tissue. The treatment of deep tumors by high intensity ultrasound will require the development of focusing transducers capable of irradiating relatively large volumes of tissue simultaneously. This is necessary in order to confine the duration of the procedure to practical time limits. The "effective" sizes of the foci employed for the current work on the central nervous system are much too small (0.1 to 10 mm^3) for treating large tissue masses (1 to 10 cm^3). The development of focusing transducers with focal regions large in cross sectional area uniform, within a few per cent, in acoustic field parameters presents difficulties. Some headway on the development of appropriate focusing systems using modified cylindrical lenses has been achieved at this laboratory [51], [52].

It should be noted here that work accomplished during the 1940's on the use of relatively low intensity ultrasound to modify tumors with enhancement reported in some cases and suppression in others does not constitute an appropriate basis on which to judge the possible applicability of high-level ultrasound for the favorable modification of carcinogenic tissue. This view is reasonable since desirable effects may only be achieved above certain threshold levels, as has been demonstrated in the case of the central nervous system. Investigations into the possible production of selective changes in carcinogenic cells, that is, for example, the possible action of high-level ultrasound to *selectively* destroy such material when embedded in normal cellular masses is extremely desirable and should receive major attention in the immediate future.

The endocrine status of both experimental animals and humans has been modified by irradiation of the pituitary in work accomplished by the author and collaborators [53], [54]. This work has application not only to the study of normal endocrine balance but also to the modification of endocrine function in, for example, patients with endocrine tumors—metastasizing breast tumors [54]. From the work accomplished it is apparent that the endocrine state of an organism can be drastically modified by hypophyseal irradiation with ultrasound and made to undergo a series of changes with various recovery rates either to pre-irradiation or altered endocrine status. The histological work of this laboratory [53] indicates that it may be possible to alter the ratios of the gland cellular populations of the pituitary and thus the interesting possibility of redesigning the hypophysis by the application of controlled dosages of ultrasound is suggested. If this can be achieved it will constitute a powerful procedure for the study and modification of endocrine physiology.

Another type of tissue, which has been the object of basic investigative work employing ultrasound in attempts to obtain fundamental information, is muscle [55]–[58]. Such work continues and it is hoped that increased knowledge of muscle structure and its relationship to the contractile and electrical characteristics of the tissue will be forthcoming.

The action of ultrasound on cells and tissues other than those already mentioned has received the attention of various investigators. For examples of work reported within the last ten years see [59] through [61] for liver, skin and some embryonic tissues and reference [62] for a specific plant tissue.

Before concluding this section it is of interest to call attention to another intriguing use of ultrasound in the study of biological systems, that is, in combination with ionizing radiation. An extremely small amount of work has been accomplished in this direction and it has been shown that X-ray dosages necessary to irreversibly destroy some types of tumor tissue can be reduced by a factor of two or more by simultaneous irradiation with relatively low-intensity ultrasound [63], [64]. The mechanism of such action should receive attention and studies must be extended to the use of high-level ultrasound in combination with ionizing radiation.

III. Technological Applications

It has not been the intention of this review of the present status and future potential of ultrasound in biomedicine to consider various developments of a strictly technological nature. A number of these, both apparatus and nonresearch applications, have already been developed and more are expected in the future including equipment for use in basic research and for medical technology. Present examples of equipment development include: ultrasonic apparatus which produces cavitation to disrupt cells and tissues in order to free intracellular contents in a form unaffected by chemical agents, equipment for separation of layers of tissue by the use of ultrasonically induced cavitation [65] and ultrasonic blood flowmeters [66], [67]. In medical therapy the use of ultrasound at low level as a diathermy procedure might well be classified as technological in character [68], [69]. These few examples serve to indicate the range of developments that have been thus far conceived and applied. It is, of course, practically impossible to predict at any one time the specific apparatuses which will be invented for future use in research and technology. However, it does appear possible to predict investigative directions which will produce fruitful research and to indicate new fields of applications which can be expected to arise therefrom. Such has been the theme of this paper.

Concluding Statement

It is apparent from the basic research work in progress and accomplished and current applications that precisely controlled ultrasound constitutes an extremely powerful tool for the investigation and modification of structure and function in biological systems. This form of energy is particularly useful for perturbing or disrupting structure on the micro-level characteristic of biological organization. Although some of the results already achieved are rather exciting when compared with past capabilities, it is apparent that the field is in a primitive state compared to the future potential. Although it is possible at the present time to outline at least some of this extensive potential for basic biological and biomedical investigations and medical technology, its rate of attainment is determined to a major extent by the economic status of the field as decided by agencies which provide the financial support. It is to be hoped that the "socioeconomic" status of this field will soon change to permit the necessary increased support to be provided so that a major fraction of the potential now visualized is realized within this decade. A support level at least two orders of magnitude greater than that currently available to the field will be necessary in order to achieve the indicated research objectives and initiate the consequent applications during this period of time. Application of the results of the basic research in routine medical practice will require additional funds both for the acquisition of appropriate sophisticated instrumentation and the training of individuals to utilize it.

References

[1] E. L. Carstensen and H. P. Schwan, "Absorption of sound arising from the presence of intact cells in blood," *J. Acoust. Soc. Am.*, vol. 31, pp. 185–189; February, 1959.

[2] E. L. Carstensen and H. P. Schwan, "Acoustic properties of hemoglobin solutions," *J. Acoust. Soc. Am.*, vol. 31, pp. 305–311; March, 1959.

[3] G. M. Piersol, H. P. Schwan, R. B. Pennell and E. L. Carstensen, "Mechanism of absorption of ultrasonic energy in blood," *Arch. Phys. Med.*, vol. 33, pp. 327–332; June, 1952.

[4] W. J. Fry and R. B. Fry, "Determination of absolute sound levels and acoustic absorption coefficients by thermocouple probes—Theory," *J. Acoust. Soc. Am.*, vol. 26, pp. 294–310; May, 1954.

[5] W. J. Fry and R. B. Fry, "Determination of absolute sound levels and acoustic absorption coefficients by thermocouple probes—Experiment," *J. Acoust. Soc. Am.*, vol. 26, pp. 311–317; May, 1954.

[6] W. J. Fry and R. B. Fry, "Temperature changes produced in tissue during ultrasonic irradiation," *J. Acoust. Soc. Am.*, vol. 25, pp. 6–11; January, 1953.

[7] W. J. Fry and F. Dunn, "Ultrasound: Analysis and experimental methods in biological research" in Physical Techniques in Biological Research," W. L. Nastuck, Ed., Academic Press, New York, N. Y., vol. 6, 1962.

[8] O. A. Angerer, G. Barth and W. Güttner, "Über den Wirkungsmechanismus biologischer Ultraschallreaktionen," *Strahlentherapie*, vol. 84, pp. 601–610; May, 1951.

[9] W. J. Fry, "Mechanism of acoustic absorption in tissue," *J. Acoust. Soc. Am.*, vol. 24, pp. 412–415; July, 1952.

[10] O. Brandt, H. Freund and E. Hiedemann, "Schwebstoffe im Schallfeld," *Z. Phys.*, vol. 104, pp. 511–533; March, 1937.

[11] P. S. Epstein, "On the absorption of sound waves in emulsions and suspensions," in "Theodore von Karman Anniversary Volume," California Inst. Tech., Pasadena, pp. 162–188; 1941.

[12] J. M. M. Pinkerton, "Absorption of ultrasonic waves in liquids and its relation to molecular constitution," *Proc. Phys. Soc. (London) B*, vol. 62, pp. 129–141; February, 1949.

[13] J. J. Markham, R. T. Beyer and R. B. Lindsay, "Absorption of sound in fluids," *Rev. Mod. Phys.*, vol. 23, pp. 353–411; October, 1951.

[14] K. F. Herzfeld and T. A. Litovitz, "Absorption and dispersion of ultrasonic waves," Academic Press, New York, N. Y.; 1959.

[15] E. L. Carstensen, K. Li and H. P. Schwan, "Determination of the acoustic properties of blood and its components," *J. Acoust. Soc. Am.*, vol. 25, pp. 286–289; March, 1953.

[16] F. Dunn and W. J. Fry, "Ultrasonic absorption microscope," *J. Acoust. Soc. Am.*, vol. 31, pp. 632–633; May, 1959.

[17] G. D. Ludwig, "The velocity of sound through tissues and the acoustic impedance of tissues," *J. Acoust. Soc. Am.*, vol. 22, pp. 862–866; November, 1950.

[18] T. F. Hueter, "Messung der Ultraschallabsorption in tierischen Geweben und ihre Abhängigkeit von der Frequenz," *Naturw.*, vol. 35, pp. 285–288; September, 1948.

[19] T. F. Hueter, "Messung der Ultraschallabsorption im menschlichen Schädelknochen und inre Abhängigkeit von der Frequenz," *Naturw.*, vol. 39, pp. 21–23; January, 1952.

[20] D. E. Goldman and T. F. Hueter, "Tabular data of the velocity and absorption of high frequency sound in mammalian tissues," *J. Acoust. Soc. Am.*, vol. 28, pp. 35–37; January, 1956.

[21] H. E. von Gierke, H. L. Oestreicher, E. K. Franke, H. O. Parrack and W. W. von Wittern, "Physics of vibrations in living tissues," *J. Appl. Physiol.*, vol. 4, pp. 886–900; June, 1952.

[22] F. Dunn and W. J. Fry, "Ultrasonic absorption and reflection by lung tissue," *Phys. in Med. and Biol.*, vol. 5, pp. 401–410; April, 1961.

[23] J. J. Wild and J. M. Reid, "The application of echo ranging techniques to the determination of structure of biological tissues," *Science*, vol. 115; pp. 226–230; February, 1952.

[24] D. H. Howry, "Techniques used in ultrasonic visualization of soft tissue structures of the body," 1955 IRE National Convention Record, pt. 9, pp. 75–88.

[25] I. Edler and C. H. Hertz, "The use of ultrasonic reflectoscope for the continuous recording of the movements of heart walls," *Kungl. Fysiogr. Sallsk. I. Lund. Forh.*, vol. 24, NR 5; March 10, 1954.

[26] D. H. Howry, "Techniques used in ultrasonic visualization of soft tissues," in "Ultrasound in Biology and Medicine," E. Kelly, Ed., Amer. Inst. Biol. Sci., Wash., D. C., pp. 49–65; 1957.

[27] S. Satomura, "Ultrasonic Doppler method for the inspection of cardiac functions," *J. Acoust. Soc. Am.*, vol. 29, pp. 1181–1185; November, 1957.

[28] I. Edler and A. Gustafson, "Ultrasonic cardiogram in mitral stenosis," *Acta. Med. Scand.*, vol. 159, pp. 85–90; 1957.

[29] S. Effert and E. Domanig, "The diagnosis of intra-atrial tumors and thrombi by the ultrasonic echo method," *Ger. Med. Monthly*, vol. 4, pp. 1–3; January, 1959.

[30] G. Baum and I. Greenwood, "The application of ultrasonic locating techniques to ophthalmology," *A.M.A. Arch. Ophthalmol.*, vol. 60, pp. 263–279; August, 1958.

[31] J. J. Wild and J. M. Reid, "Progress in the techniques of soft tissue examination by 15 Mc pulsed ultrasound," in "Ultrasound in Biology and Medicine," E. Kelly, Ed., *Amer. Inst. Biol. Sci.*, Wash., D. C., pp. 30–48; 1957.

[32] D. H. Howry, D. A. Stott and W. R. Bliss, "The ultrasonic visualization of carcinoma of the breast and other soft tissue structures," *Cancer*, vol. 7, pp. 354–358; March, 1954.

[33] Y. Kikuchi, R. Uchida, K. Tanaka, and T. Wagai, "Early cancer diagnosis through ultrasonics," *J. Acoust. Soc. Am.*, vol. 29, pp. 824–833; July, 1957.

[34] W. J. Fry, "Intense ultrasound in investigations of the central nervous system," in "Advances in Biological and Medical Physics," J. H. Lawrence and C. A. Tobias, Eds., Academic Press, New York, N. Y., vol. 6, pp. 281–348; 1958.

[35] W. J. Fry, R. Krumins, F. J. Fry, G. Thomas, S. Borbely and H. Ades, "Origins and distributions of several efferent pathways from the mammillary nuclei of the cat," *J. Comp. Neurol.*, to be published.

[36] W. J. Fry and F. J. Fry, "Fundamental neurological research and human neurosurgery using intense ultrasound," IRE Trans. on Medical Electronics, vol. ME-7, pp. 166–181; July, 1960.

[37] W. J. Fry, R. Meyers, F. J. Fry, D. F. Schultz, L. L. Dreyer and R. F. Noyes, "Topical differentia of pathogenetic mechanisms underlying parkinsonian tremor and rigidity as indicated by ultrasonic irradiation of the human brain," *Trans. Am. Neurol. Assoc.*, pp. 16–24; June, 1958.

[38] R. Meyers, F. J. Fry, W. J. Fry, R. C. Eggleton and D. F. Schultz, "Determination of topological human brain representations and modifications of signs and symptoms of some neurologic disorders by the use of high level ultrasound," *Neurology*, vol. 10, pp. 271–277; January, 1960.

[39] W. J. Fry, W. J. Wulff, D. Tucker and F. J. Fry, "Physical factors involved in ultrasonically induced changes in living systems: I. Identification of non-temperature effects," *J. Acoust. Soc. Am.*, vol. 22, pp. 867–876; November, 1950.

[40] W. J. Fry, D. Tucker, F. J. Fry and V. J. Wulff, "Physical factors involved in ultrasonically induced changes in living systems: II. Amplitude duration relations and the effect of hydrostatic pressure for nerve tissue," *J. Acoust. Soc. Am.*, vol. 23, pp. 364–368; May, 1951.
[41] W. J. Fry, "Action of ultrasound on nerve tissue—A review," *J. Acoust. Soc. Am.*, vol. 25, pp. 1–5; January, 1953.
[42] W. J. Fry, W. H. Mosberg, Jr., J. W. Barnard and F. J. Fry, "Production of focal destructive lesions in the central nervous system with ultrasound," *J. Neurosurg.*, vol. 11, pp. 471–478; September, 1954.
[43] J. W. Barnard, W. J. Fry, F. J. Fry, and R. F. Krumins, "Effects of high intensity ultrasound on the central nervous system of the cat," *J. Comp. Neurol.*, vol. 103, pp. 459–484; December, 1955.
[44] W. J. Barnard, W. J. Fry, F. J. Fry and J. F. Brennan, "Small localized ultrasonic lesions in the white and gray matter of the cat brain," *Arch. Neurol. Psychiat.*, vol. 75, pp. 15–35; January, 1956.
[45] W. J. Fry, "Ultrasound in neurology," *Neurology*, vol. 6, pp. 693–704; January, 1956.
[46] H. T. Ballantine, T. F. Hueter, W. J. H. Nauta and D. M. Sosa, "Focal destruction of nervous tissue by focused ultrasound: Biophysical factors influencing its application," *J. Exptl. Med.*, vol. 104, pp. 337–360; September, 1956.
[47] F. J. Fry, H. W. Ades and W. J. Fry, "Production of reversible changes in the central nervous system by ultrasound," *Science*, vol. 127, pp. 83–84; January, 1958.
[48] W. J. Fry, F. J. Fry, R. Meyers and R. C. Eggleton, "The use of ultrasound in neurosurgery," *Proc. Third Internatl. Conf. on Medical Electronics*, to be published.
[49] A. K. Burov, "High intensity ultrasonic oscillations for treatment of malignant tumors in animals and man," *Doklady Akad. Nauk SSSR*, vol. 106, pp. 239–241; January, 1956.
[50] A. K. Burov and G. D. Andrewskaya, "The effect of ultra-acoustic oscillations of high intensity on malignant tumors in animals and man," *Doklady Akad. Nauk SSSR*, vol. 106, pp. 445–448; January, 1956.
[51] W. J. Fry and F. S. Brunschwig, "Study of methods for producing uniform high intensity ultrasonic fields," Annual Progress Rept. on Grant G7421 with Natl. Science Foundation, pp. 1–17; July, 1958.
[52] W. J. Fry and F. W. Brunschwig, "Study of methods for producing uniform high intensity ultrasonic fields," Annual Progress Rept. on Grant G7421 with Natl. Science Foundation, pp. 1–17; August, 1960.
[53] R. Krumins, F. J. Fry, W. J. Fry and E. Kelly, "Use of high intensity ultrasound for study of the hypophysis," *Abstracts of Fifth Annual Meeting of the Biophys. Soc.*, St. Louis, Mo.; February, 1961.
[54] R. C. Hickey, W. J. Fry, R. Meyers, F. J. Fry and J. T. Bradbury, "Human pituitary irradiation with focussed ultrasound in advanced breast cancer—An initial report," *AMA Arch. Surgery*, vol. 83, pp. 620–633; October, 1961.
[55] W. Welkowitz and W. J. Fry, "Effects of high intensity sound on electrical conduction in muscle," *J. Cellular Comp. Physiol.*, vol. 48, pp. 435–458; December, 1956.
[56] E. Kelly and W. J. Fry, "Isometric twitch tension of frog skeletal muscle as a function of temperature," *Science*, vol. 128, pp. 200–202; July, 1958.
[57] J. W. Gersten, "Nonthermal neuromuscular effects of ultrasound," *Am. J. Phys. Med.*, vol. 37, pp. 235–236; August, 1958.
[58] R. G. Busnel, J. Gligorijevic, P. Chauchard and H. Mazoue, "Contribution a l'etude des effets et des mechanismes d'action des ultrasons sur le systeme neuro-musculoire," *Der Ultraschall in der Med.*, vol. 6, pp. 1–25; March, 1953.
[59] E. Bell, "Some changes in liver tissue which survives irradiation with ultrasound," in "Ultrasound in Biology and Medicine," E. Kelly, Ed., Amer. Inst. Biol. Sci., Wash., D. C., pp. 203–225; 1957.
[60] E. Bell, "The action of ultrasound on the mouse liver," *J. Cellular Comp. Physiol.*, vol. 50, pp. 83–104; August, 1957.
[61] E. Bell, "Action of ultrasound on adult and embryonic organ systems," *Am. J. Phys. Med.*, vol. 37, pp. 184–191; August, 1958.
[62] J. F. Lehmann, E. J. Baldes and F. H. Krusen, "Destructive effects of high intensity ultrasound on plant tissues" in "Ultrasound in Biology and Medicine," E. Kelly, Ed., Am. Inst. Biol. Sci., Wash., D. C., pp. 177–190; 1957.
[63] K. Woeber, "Studium der wirkung von hyperthermie durch überwarmungsbad, ultrakurzwellen und ultraschall auf krebsgewebe sowie biologische grundlagen einer kombinierten röntgen und ultraschallbehandlung oberflachlicher tumoren," *Strahlentherapie*, vol. 95, pp. 333–366; November, 1954.
[64] K. Woeber, "Die bedeutung des ultraschalls für die dermatologie und seine anwendung bei hauttumoren in kombination mit röntgenstrahlen," *Strahlentherapie*, vol. 98, pp. 169–185; September, 1955.
[65] E. Bell, M. E. Kaighn and L. M. Fessenden, "The role of mesodermal and ectodermal components in the development of the chick limb," *Developmental Biol.*, vol. 1, pp. 101–124; April, 1959.
[66] E. J. Baldes, W. R. Farrall, M. C. Haugen and J. F. Herrick, "A forum on an ultrasonic method for measuring the velocity of blood," in "Ultrasound in Biology and Medicine," E. Kelly, Ed., Amer. Inst. Biol. Sci., Wash. D. C., pp. 165–176; 1957.
[67] D. L. Franklin, R. M. Ellis and R. F. Rushmer, "Aortic blood flow in dogs during treadmill exercise," *J. Appl. Physiol.*, vol. 14, pp. 809–812; September, 1959.
[68] Proceedings of the 4th annual conference on ultrasonic therapy," *Am. Inst. Ultrasonics in Med.*, pp. 41–46; 1955.
[69] "Proceedings of the international conference of ultrasonics in medicine," *Am. J. Phys. Med.*, vol. 37, p. 152–156; June, 1958.

Section 28

VEHICULAR COMMUNICATIONS

Organized with the assistance of the IRE Professional Group on Vehicular Communications

The History of Land-Mobile Radio Communications*

DANIEL E. NOBLE†, FELLOW, IRE

Summary—**This paper on land-mobile radio communications was prepared for the noncommunications and nonmobile radio specialist. The early pioneering work of the police is covered, and after a brief statement about the early use of mobile transmitters, the Connecticut State Police two-way FM system is described. From the introduction of FM into the Connecticut system to the present, engineering efforts have been directed toward the improvement of selectivity, the elimination of spurious responses and spurious radiation of the equipment, and in general, toward technical improvements which would make possible both an increase in spectrum utilization and an increase in channel loading. In the more recent embodiments of equipments, transistors have been used for the power supply, for audio amplifiers, and for complete receiver circuit design. While the use of single-side-band modulation and the general characteristics of random-access systems are under investigation, advantages have not been disclosed which would justify a switch from FM to some other modulation system. The efficiency of frequency utilization must be increased, and any system which offers a substantial improvement over FM will be given careful attention.**

Land Vehicular Communications

THE EARLY history of land-mobile radio communications is a history of police pioneering. The police of the U. S. have always been identified with the innovators in the communications field. As early as 1877, the Albany, N. Y., Police Department installed five telephones in the Mayor's office, connected to the precinct police stations; this was only two years after Alexander Graham Bell developed the telephone. In 1880 the Chicago Police Department installed the first police call box on the street, and three years later the Detroit, Mich. Police Department installed one police telephone, at a time when there were only seven telephones in the entire city.

In the radio field, spark transmitters were used by the New York Harbor Police in 1916 to communicate with their boats and other boats in the harbor. The Pennsylvania State Police installed point-to-point radio telegraph between headquarters and posts on 250 kc, back in 1923.

Just for a bit of perspective, remember that in 1904 Indianapolis was the automobile manufacturing center of the world, with Stutz, Marmon, Cole, National, and Dusenberg in the area, and it was during this year that the Indianapolis Police replaced the horse-drawn paddy-wagons and automobiles. The police began to use their first motorcycles in 1909, and in 1917 the Detroit Police began using automobiles with two men per car parked at police telephone booths along the streets. They did not patrol, but stayed in the booth between calls. A telephone in each booth supplied communications; a potbellied stove in one room, the coal bin in another, took care of the winter temperatures.

To a man of vision, the need for direct communication with moving police automobiles was clear. That man of vision, who refused to accept repeated failures in his attempts to establish practical radio communications with moving police cars, was Commissioner William P. Rutledge, of the Detroit Police Department. In 1921, only four years after the city pioneered the use of automobiles for police work, Commissioner Rutledge purchased a Western Electric 1-A, 500-w broadcast transmitter and installed it in police headquarters. This was before the days of crystal control, and the

* Received by the IRE, July 17, 1961.

† Executive Vice President, Communications, Semiconductor, and Military Electronics Divisions, Motorola Inc., Chicago, Ill., and Phoenix, Ariz.

unit was a self-excited oscillator-modulator combination with two WE-212D's as oscillators, and another pair for modulators, and with a 1500-v motor generator for the power supply. For a period of six years, Commissioner Rutledge and his organization tried to develop a practical system to provide satisfactory voice communications to the moving cars. Their point of failure was the receivers. They could not build receivers which would work reliably in the police cars. Both voice and radio telegraph were tried, but the basic problem of receiver instability and lack of sensitivity limited the coverage. With each new year they found new approaches, but all were failures. The accumulation of frustration became so strong that, in 1927, the station was shut down and the radio room was locked up. During the six-year period, they had also battled a merry-go-round of license changes. When the station was first opened on May 23, 1921, it operated under the amateur license W8BNE, working on 200 m. But on August 16 of the same year, the license was changed to experimental W8XAB and the wavelength changed to 375 m. In 1922 another change put it on a provisional limited commercial license KOP on 360 m, which was changed in 1923 to 286 m, again in 1924 to a Class-A broadcast station on 277.8 m, and back again in 1926 to a limited commercial license WCK on 144.8 m. Even then the peregrinations were not complete. Later in 1926 the station was moved back to 200 m with a Class-A broadcast license; it was changed again in 1927 to a limited commercial grant for Station WCK on 144.8 m. According to one of the rulings in 1922, the Detroit Station KOP was required to provide broadcast entertainment during regular hours, with the police calls interspersed as required. Finding suitable performers for the broadcast programs was difficult, and the police band was given a notable workout.

Although Commissioner Rutledge closed down the radio station in 1927, he did not give up. He was convinced that the modern automobile had given the criminal an advantage in speed that could not be overcome by the use of police cars controlled by telephone. The prohibition era of crime flourished in the nation, and big city bootleg gangs were in ruthless control; slayings and corruption paced a nationwide major increase in crimes of all types. With the automobile, the criminal could strike hard and speed away with little chance of interference from the police cars awaiting calls at the telephone police booths. Patrolling without communications was a futile exercise. Commissioner Rutledge was desperate when he closed the radio room in 1927, but he did not give up.

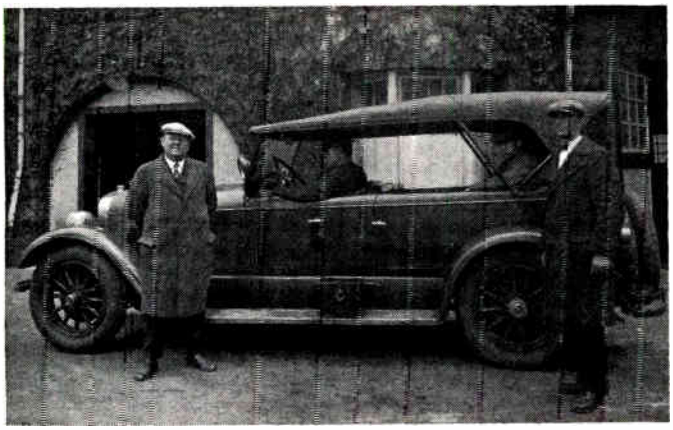
Remember that 1927 was still the early days of radio. Most home radios used B and C batteries for power and A batteries for storage, and even the car broadcast radio did not make its appearance until 1930. Short-wave radio was still 200 m, so Commissioner Rutledge was pressing the state of the art with his proposed police system on 144.8 m.

During the summer of 1927, Robert L. Batts, a young student from Purdue University, was working at a radio parts store in downtown Detroit. He developed quite a following of do-it-yourself customers who were purchasing kits to build radio sets. One of his customers was a young Detroit motorcycle policeman, named Kenneth Cox, and he and Bob Batts frequently discussed the possibility of making a radio receiver work in a police car. Batts had been using a superheterodyne receiver with a loop in a Dodge truck for tracking radio interference. To him the problem of car reception seemed simple. In the fall, when Batts went back to Purdue, he and Cox continued to communicate by mail, and Batts sent along suggestions and sketches for the construction of a police radio receiver. Late that fall, Cox went to Commissioner Rutledge and told him he could make a radio work in a police car. He had built a breadboard model thoroughly cushioned with foam rubber, and he deliberately dropped the receiver on the floor in Rutledge's office to show how rugged it was. It still worked. Cox received the assignment from Commissioner Rutledge to develop the receiver and he began a campaign immediately to entice Batts back to Detroit.

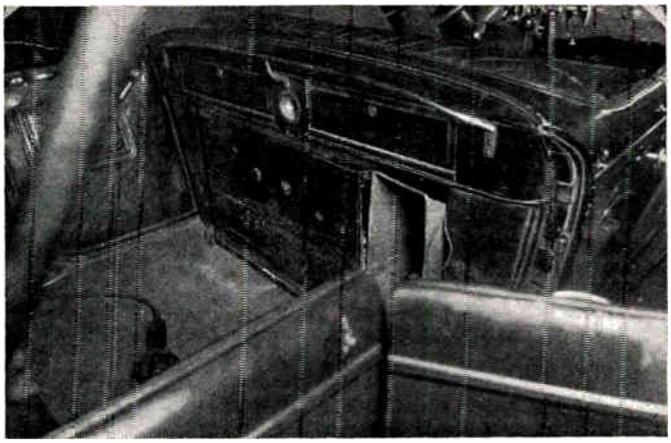
After some jousting with the Federal Radio Commission, a new construction permit was granted on February 4, 1928, to move the equipment to Belle Isle for operation on 144.8 m, and later on 94 m for the starting of new tests. Perhaps we should describe this 94-m assignment as the UHM of the early radio days.

Batts was finally prevailed upon to return to Detroit as a patrolman (that was the only way Commissioner Rutledge could pay him), to start work on the development of the new receivers. Bernard Fitzgerald and Walter Vogeler were given the job of rebuilding the Western Electric 1-A transmitter from a self-excited unit to a crystal-controlled MOPA, and Batts started on the new receiver design. The new receiver consisted of three stages of tuned RF, using the newly-available screen grid type-322, a 200-A detector, a 201-A first audio amplifier, and a 112-A output tube. Copper shielding compartments in the RF and detector sections, with locked-tuning capacitor adjustments, contributed to the electrical and mechanical stability. Heavy duty 135-volt B-batteries were used, and a 6-volt storage battery mounted on a running board provided the A-power. The A-battery drain was 1.1 amperes, and it was necessary to switch the storage batteries every four days.

Station W8FS first went on the air April 7, 1928, transmitting to a new receiver in cruiser No. 5 (Fig. 1) —and it worked! The receiver stayed tuned and reception was satisfactory all over the city. After nearly seven years of persistent effort, Commissioner Rutledge's dream had come true, and the improved communications system soon proved its value for the apprehension of criminals. The radio-equipped cruisers caught hold-up men, car thieves, and burglars, sometimes seconds after the call was reported to the dispatchers at police



(a)



(b)

Fig. 1—Detroit police Lincoln cruiser #5. The first successful mobile radio installation, April 7, 1928. (a) Car showing running board battery installation. (b) Receiver box and speaker back of the front seat.

headquarters. At times, the car was just around the corner or almost in front of the place from which the call originated. With this initial success to spur him on, Batts carried on an extensive program of field testing and built new and improved receivers to equip the entire fleet of police cars.

The pioneering of this early police radio system was truly the beginning of the land-mobile radio communications industry. The Detroit Police Radio System drew world-wide publicity, and visitors arrived from all over the world to inspect the system. Other city police departments planned radio systems, and like Detroit, they were forced to build their own receivers. In September, 1929, the Cleveland Police Department was the second system to go on the air with a few cars. Batts, who had moved on to Indianapolis in October, 1929, put that city police department on the air on December 24 of the same year.

There were other firsts. The pioneering record would not be complete without a brief mention of the first freight train installation. Arthur Batcheler, Supervisor of Radio for the Department of Commerce Radio Division, took part in a test demonstration April 23, 1928 on a New York Central freight train. A train consisting of a locomotive, 125 freight cars and 2 cabooses, was equipped with an experimental General Electric 50-watt radiotelephone installation to furnish communication between locomotive and caboose, or between either locomotive or caboose and the signal towers which the train passed en route. Batcheler reported, in a letter to the Department of Commerce, "The demonstration was a most successful one and communication was carried on continuously over the entire route." He also stated, "The Federal Radio Commission no doubt will have to provide channels for operation of this service, which, in the future, is destined to come into general use." For the freight train applications the problems of cost, installation, and maintenance outweighed urgency, and the work was dropped. For the police use of radiotelephone, urgency was the paramount consideration, and the number of installations increased rapidly.

The pioneering work of the Detroit City Police had broken through the barrier to successful radio communications to the police cars. Equivalent systems were established in many cities. The State Police became interested also, and on May 17, 1930, the Federal Radio Commission granted the State of Michigan the license for 5000 w on 1642-kc daytime and 1000 w at night. The Michigan State Police Station was on the air by October, 1930.

As soon as the use of broadcast transmission to police cars became routine, it became apparent that there was an urgent need for acknowledgment and talk-back. Lt. Vincent Doyle, the radio man for the Bayonne, N. J., Police Department, decided to do something about the two-way communications matter; in March, 1933, he was on the air with the first two-way police radio system, using REL AM equipment on 33.1 Mc. Four police cars were equipped. The units used superregenerative receivers and noncrystal-controlled MOPA transmitters with a pair of 210's in the final. The stations were operated on the temporary experimental license W2XCJ.

Bob Batts had not lost his pioneering spirit. In 1933 he had the Indianapolis Police Radio Division on the air with a 200-w base station and two 20-w mobile two-way units. The mobile transmitter units were high powered compared to the ones used in the initial installations in other parts of the country. Batts teamed his 20-w transmitters with 6-tube VHF AM superhets of good sensitivity, with 6-v tubes. The base station was equipped with a 100-foot antenna, and very reliable two-way communications were established despite the fact that quartz crystals were not available for either the transmitter or the receiver control. The lack of crystals and the high battery drain discouraged a further expansion of the system, even though the test results were excellent. The Indianapolis Police Department also pioneered the installation of superheterodyne receivers on police motorcycles. In 1931 there were thirty-two solo motorcycles in the Indianapolis Police Department equipped with AM superheterodyne receivers operating on 2442 kc, using P. R. Mallory vibrator power supplies. Although a few commercial receivers were available, the equipment used until 1935 in nearly all of the police radio systems was of the do-it-yourself variety—and excellent equipment it was, too, for the state of the art. Late in 1931 the Bosch Corporation offered the first commercial compact superheterodyne receiver with a separate dynamotor power supply, which eliminated the B batteries. The Sparton Company in Michigan offered receivers for a brief period. The introduction of the Mallory vibrator power supply in 1931 was an important step forward, with a resulting decrease in operating costs and an increase in reliability. RCA and GE entered the mobile radio field in 1936, with F. M. Link following, and Motorola began selling police radio receivers in 1937. As a result of hearings held in 1936, the Federal Communications Commission issued Order No. 19, dated October 13, 1937, which allocated twenty-nine VHF channels to police departments in the band 30.58 to 39.9 Mc. This order was a milestone in the development of two-way VHF police communications. The use of crystal control for both transmitters and receivers became universal, and in 1938 the FCC established the maximum allowable frequency tolerance of 0.05 per cent on frequencies above 30,000 kc.

Interest in two-way communications spread to the power companies, and in 1938 George Underhill established a system for the Central Hudson Gas and Electric Company of Poughkeepsie, N. Y. Earl Glatzel established the second system a few months later for Detroit Edison. In 1939 the FCC defined police communications as emergency service and recognized the power utilities with provisions for special emergency licenses. By then, a number of commercial brands of equipment were available, with RCA, F. M. Link, GE, and Motorola as the leaders.

Two-way VHF communications were standard for city installations, but state police were still operating on one-way broadcast systems. It remained for Colonel Edward J. Hickey, Commissioner of the Connecticut State Police, to establish the first two-way state police system in the country, and it was the privilege of the author of this paper to pioneer the system as the first police FM installation. The success of the Connecticut system started the nationwide switch from AM to FM.

The details of the Connecticut State Police FM radio system design may be found in the November and December, 1940, issues of *Electronics*. The shift from AM to FM was a radical step, but the use of FM, or more properly PM (phase modulation), in the Connecticut system only partially accounted for the success of the system. In nearly all city two-way AM installations, the base station transmitters and receivers were located at the most convenient spot, which was usually the City Hall or some other police-controlled facility. An ambient noise level at the base station receiver sometimes reached several microvolts, and only strong signals from the mobile transmitters could get through. For the Connecticut system, each tentative base station location was checked for noise level, and a test run was made with a mobile unit maintaining communications with a temporary base station installation. The optimum location was picked for both low noise level and high elevation, and telephone wire connections were provided to control the station from the barracks radio dispatch office.

It would be obvious to any engineer that the best place for the mobile antenna would be in the middle of the rooftop of the automobile, and so, to gain maximum communications range, this was the location selected. Tapered steel-tubing fishing rods were purchased and mounted on a spring held in place by a molded collar. Coaxial cable connected the antenna to the transmitter and receiver in the car trunk. This combination of the selection of low-noise and high-elevation base station locations, and the use of efficient mobile antenna installations assured the success of the system. While good coverage would probably have been achieved with AM, the unique capture effect of FM made it possible for ten individual police-troop base stations to operate simultaneously on the same frequency assignments without significant interference. Two frequencies were used, 39,400 kc for the base station transmitter and 39,180 kc for the mobile transmitters. The use of two frequencies was a unique approach, designed to make it unnecessary for the low-powered mobile transmitters to compete with the adjacent area 250-w base stations transmitting from high antennas. To provide car-to-car communications, provisions were made for switching extra crystals into the mobile transmitters for operation on the base station frequency. Since there was only a difference of 220 kc, the switch could be made without retuning the transmitter and with only a nominal loss of power. As a practical matter, it was found that the adjacent base stations could go on the air simultaneously, and each base station would capture its own mobile stations. Where there were exceptions to this rule in the case of a preferred transmission path existing between the mobile unit and the opposing station, it was only necessary for the mobile unit to move along slowly and stop on the peak of the standing wave to capture the desired station. In general, it may be said that each base station operated as an independent system; but that every base station was capable of communicating with every other base station, and of taking over communications with the patrol cars in an adjacent area when a base station failure occurred. This two-frequency, three-way design provided flexibility for state-wide communications with generous factors of safety and with a minimum of equipment complexity. An alternate system using individual frequencies for each area either would have eliminated interarea communications, making it impossible for a car to switch from one base station to another as it traveled across the state, or would have made it necessary to equip each car and base station with complex multichannel equipment.

The major virtue of the use of phase modulation in the Connecticut system was not the wideband FM noise-reducing characteristic which is usually identified with the FM broadcast service. The virtues of FM noise reduction are realized only when the signal exceeds the noise level, but in the mobile communications systems, the chief noise and the most serious noise limitation are produced by the radiating automobile ignition systems. The ignition-pulse interference nearly always exceeded the signal level and, at times, pulses would reach peak values several hundred times greater than the level of the desired signal. Each pulse would take over the limiter temporarily and, in effect, would gate the receiver for the duration of the pulse. This was an inherent brute-force pulse noise suppressor, and it outperformed the pulse noise gates sometimes used in AM receivers. In fairness, it should be stated that there existed one AM pulse noise suppressor system which followed the envelope of the received modulated wave and cut the pulse at the instantaneous amplitude of the associated wave. With this envelope-following pulse clipper, an AM receiver offered competition for FM in the one category of ignition noise suppression. There were, however, other important attributes of the FM system which combined to outclass the general performance of AM for mobile communications. It was often necessary to maintain communications with very weak signals. With the AM system, the moving car passing through the standing waves would produce a flutter of modulation as the signal dipped from a usable level to an unusable level. It seemed impracticable to design the AM receiver with time constants necessary to control the distortion caused by the rapid changes in the radio and the subsequent audio signal levels. With FM the audio volume was not a function of the signal level, and as a result the standing-wave modulation produced some noise flutter as the signal dipped below the reception level, but the standing-wave modulation would yield readable signals where communication was lost in an equivalent AM system. The sensitivity of an FM receiver could also be maintained at a very high level with generous factors of safety to compensate for tube aging and for circuit detuning. With a limiter functioning normally at $\frac{1}{4}$ v, it was possible to design a receiver which would amplify the front-end ambient noise to 15 or 20 v at the limiter. Tube aging and circuit detuning could reduce this limiter voltage substantially with no apparent effect upon the reception performance. Perhaps the final point of advantage of FM over AM was the superior squelch system. The AM squelch consisted of a simple gate which would be opened when the received signal was strong enough to trip the blocking voltage on the grid of the audio amplifier. Unfortunately, the receiver did not recognize the difference between noise energy and carrier energy, so that a strong noise would also open the squelch. The police officer in the car would understandably adjust the squelch so that the noise would not open it, and as a result the sensitivity of the receiver would become the rather low sensitivity of the squelch control. The FM equipment employed in the Connecticut State Police system was manufactured by the F. M. Link Company and was, in general, an adaptation of high-quality AM standard units but with the necessary change-over to FM to provide increased gain with an effective limiter and a frequency detector, but the squelch system used was the same as that employed in the AM equipment.

When the first Motorola FM mobile communications equipment was designed, a new squelch system was introduced, which became the standard of the FM industry. This circuit, which became known as the "differential squelch," balanced the voltage change produced by the characteristic FM noise reduction in the presence of a carrier, against the voltage resulting from the rectification of received noise. With this arrangement, the squelch would open with a very-low-level signal (0.15 μv) when there was little noise present, but as the noise increased, the signal level required to open the squelch also increased. The system was so balanced that the loudspeaker would always be activated at a signal-to-noise ratio so low that the voice was unreadable. With this squelch system, noise bursts would close, rather than open, the squelch gate, and the sensitivity of the receiver was no longer limited by the sensitivity of the squelch. A further receiver design refinement substantially increased the effective sensitivity of the receiver by decreasing the inherent first mixer noise. The final receiver design could provide excellent communications in low-noise areas with a signal input voltage of less than $\frac{1}{2}$ μv. A signal generator input to the receiver of 0.4 μv would produce a 20-db receiver noise reduction. The receiver squelch system, the high sensitivity, and the reserve gain were the keys to the extended range and the reliability of performance made possible when the equipment was used in systems with the base stations installed at high-elevation and low-noise-level locations.

The mobile transmitter designs were straightforward, with a crystal oscillator feeding a modulator, which, in turn, activated two quadruplers, a doubler, and a final amplifier. Perhaps it should be noted that, with the PM transmitter using the same power amplifier tube as that used for AM, the average power could be increased to a value approximating the peak power of the AM output, or to maximum tube rating. The 20- to 25-w AM transmitter became a 35–40 w FM transmitter, with an 807 tube in the final amplifier. Two tubes in the final amplifier and an increased power supply provided a 50- to 60-w output with no modulator change.

The first FM two-way mobile communications installations were characterized by exceptional reliability of communications and exceptional range of communications as compared to the usual AM installations.

In the early days of FM equipment design, the receiver selectivity was wide open, with an IF selectivity 50 kc wide, 6 db down, and 200 kc wide, 60 db down. The receiver's spurious and image responses were only 40 or 50 db down. The deviation of the transmitter was uncontrolled, with a modulation sensitivity varying all the way from 18 kc to 45 kc for the same signal input. The final test for modulation was a voice test with a microphone, and the deviation was not measured. The spurious radiation of the transmitters was practically uncontrolled, with substantial spurs radiating interference. These faults, however, were important only in a crowded spectrum and they did not limit or interfere with the performance of isolated systems. In 1940 there were only a few thousand transmitters on the air, and interference was not a major factor, but by 1948 the number had grown to 86,000 transmitters, with a leap to 695,000 in 1958, and the number projected for 1963 is 1,390,000 transmitters. The technical history of mobile communications equipment design since 1940 is a history of development to improve the efficiency of spectrum utilization.

While the police pioneered the development of the first practical mobile communications system, the engineers of the radio communications industry carried out a continuous program of refining the equipment design to increase the spectrum loading. As spectrum crowding developed, bringing with it the inevitable interference problem, the needed reduction of transmitter spurious radiation, the reduction of receiver spurious responses, and a substantial improvement in receiver selectivity followed rapidly. Additional refinements reduced the transmitter noise, introduced instantaneous deviation control, and provided receiver design to minimize desensitizing and intermodulation interference. Without going into laborious detail, a look at the specifications for modern "transistorized" units will partially define the progress which has been made.

General

Two-Way Mobile Radio Station: single case ($3\frac{1}{4} \times 13\frac{3}{8} \times 17\frac{1}{2}$ inches; 29 pounds).

Transmitter

Power Supply: Transistor oscillator type "dc transformer."

Power Output: RF 30 or 50 w, 40 Mc; 25 w, 160 Mc.

Transmitter Spurious: More than 85 db below carrier.

Frequency Stability: ± 0.0005 per cent of assigned frequency. $-30°C$ to $+60°C$ with $+25°C$ reference.

Modulation: As specified, ± 5 or ± 15 kc for 100 per cent at 1000 cps.

Circuits and Multipliers: Phase modulator and deviation control; doubler; tripler; doubler-driver; power amplifier.

FM Noise: -50 db below ± 3.3-kc deviation at 1000 cps.

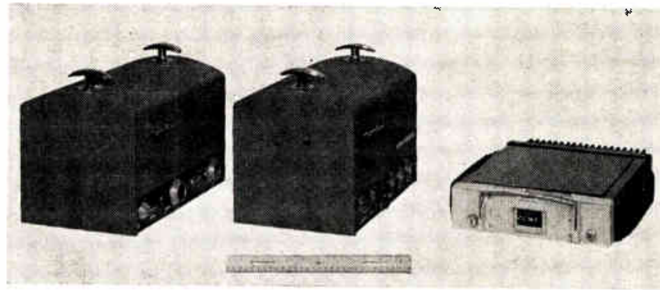


Fig. 2—1940 separate FM transmitter and receiver "breadloaf" packages compared to the 1961 single-packaged transmitter-receiver with transistor-type receiver.

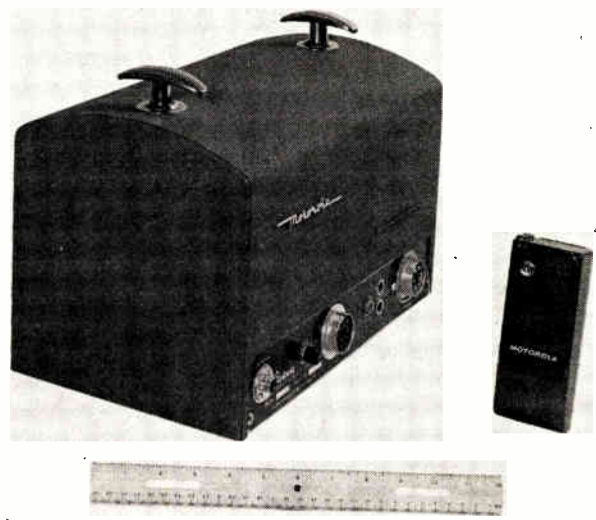


Fig. 3—1940 receiver compared with the equivalent 1961 transistor-type receiver, which is complete with speaker and battery supply.

Fig. 4—Quarter-wave cavity selector used for base station control of intermodulation and desensitizing in the 160-Mc band.

Receiver

Selectivity: -100 db at ± 15 or ± 32 kc, as specified.
Channel Spacing: Designed for original 40- or 60-kc spacing and, for split channels: 20, 30, 40, or 60 kc.
Sensitivity: Depends upon frequency, etc. 0.3 μv to 0.6 μv for 20-db quieting with 50 ohms of RF input impedance.
Stability: ± 0.0005 per cent of reference over temperature range -30 to $+60°$C.
Squelch: Noise compensated. Threshold sensitivity, 0.15 to 0.25 μv. Also tone-coded squelch available.
Audio Output: 5-w to 3-ohm load.

Of all the problems which have limited spectrum loading, desensitizing and intermodulation have been the two most difficult to control. The design of 455-kc IF selectivity providing 100-db protection at the edge of the channel was readily accomplished, but desensitizing and intermodulation are related to receiver front-end RF rather than IF selectivity. A mobile unit parked in the shadow of an opposing or undesired station may attempt to receive a 1-μv signal from its parent base station. Since the opposing station, which may be on a frequency several channels removed from the desired channel, can pass signals through the acceptance band of the RF section of the mobile receiver, this high signal level can blank out the desired 1-μv signal from the distant station even though the nearby station cannot be heard because of the tight IF selectivity. In a somewhat similar pattern, two nearby transmitters, radiating strong signals on, for example, the second and third channels removed from the referenced desired signal, can mix in the front end of the receiver to generate a mixing product directly on the frequency of the desired signal, and strong enough to wipe out the desired distant station signal. The only immediately apparent solution to such interference problems is RF selectivity comparable to that obtained in the IF section of the receiver, but of course, at 50, 100, and 450 Mc, this is completely beyond the art. By stressing the linearity of the RF amplifier circuit design and by resorting to multicircuit high-Q front-end filter stages, the engineers have provided a modest 65- to 70-db protection.

Particular cases of interbase-station interference problems were solved by the use of very-high-Q quarter-wave cavities. Obviously, such cavities were too large for mobile station use and even this radical approach failed to provide the 100- to 150-db adjacent-channel protection necessary for adequate spectrum loading. While the use of cavities in the receiver base station front-end design did not offer complete solutions, it did increase the efficiency of spectrum utilization and solved many specific problems which developed as new systems were added in crowded areas. Fortunately, the recent introduction of the varactor and its promised application to the up-converter and down-converter circuits offers a potential solution to both desensitizing and intermodulation, with the expectation that the needed 100- to 150-db protection may be achieved in the future.

Spectrum loading for maximum utilization is a three-dimensional problem involving frequency, space, and time. Frequency division has been the traditional approach to radio spectrum utilization. The available channels may be increased by narrowing each channel so that radio systems are operating with less frequency separation. The original FCC rule making established 40-kc bandwidths for the 25- to 50-Mc band, 60-kc bandwidths for the 148- to 162-Mc band, and 100-kc bandwidths for the 450- to 470-Mc region. Subsequent channel splitting changed these bandwidths to 20 kc, 30 kc, and 50 kc, respectively. With the development of intermodulation and desensitizing protection of 100 db, or greater, channel splitting to 15 kc for the 160-Mc band, and 25 kc for the 450-Mc band, may be expected to provide additional spectrum loading.

As the bandwidths are narrowed, the automobile ignition-pulse-noise interference in the lower bands becomes a limiting factor for systems range. The inherent pulse-noise suppression characteristics of FM receivers employing 40-kc and 60-kc bandwidths made the use of additional pulse-noise receiver gates unnecessary, but with the split channels, especially in the 25- to 50-Mc band, the range of communications is decreased by the ignition noise and the effectiveness of a system can be restored only by the use of a self-gating pulse-noise suppressor. These pulse suppression systems are beginning to be used in industry, and a typical application may be found in the Motorola low-band transistor receivers which utilize the following straightforward design: While the pulse is delayed by a line section in the front end of the receiver, a parallel circuit amplifies and shapes the pulse for use as a gating signal which temporarily blanks the receiver input at the instant of the arrival of the opposite-number delayed pulse. Since we are dealing with pulse durations of a few microseconds, the receiver gate may be closed for the duration of the pulse without destroying the intelligibility of the voice transmission. In actual tests, with weak signals and with strong high-repetition rate ignition-pulse noises completely destroying the intelligility of reception, the activation of the pulse-gating system reduced the noise to such a degree that voice intelligibility was restored. It may be said in general that the increase of spectrum loading by splitting channels to narrower bandwidths results in a degradation of the signal-to-noise ratios of the operating systems. The use of effective pulse-noise-gating circuitry becomes necessary to substantially restore both the range and the quality of reception.

While the development of the land-mobile communications systems has been almost wholly concerned with the use of FM or PM, tests have been carried out to explore the potential of both single-sideband and random-access systems for this service. There are two major disadvantages to the use of single-sideband in mobile applications:

1) The complete suppression of the carrier is not at present practicable because of the problem of generating restored carrier in the receiver within the tolerance requirements for low-speech distortion. The necessary maintenance of the locally generated carrier at ±30 c at 150 Mc would demand design tolerances which, at the very least, are beyond the economic tolerance of the industry.
2) Single-sideband reception cannot tolerate ignition pulse-noise interference. The range and performance of ordinary single-sideband systems in comparison to ordinary FM systems is inferior to the point of unacceptability without the use of some form of single-sideband pulse-noise suppression.

Obviously, both objections may be limited by the extension of the art. The vestigial carrier can be transmitted with the sideband to synchronize the local receiver carrier generator. It is also possible to establish community carrier-frequency reference transmitters which would be used to provide the synchronizing frequency for the restoration of the carrier in all receiver systems operating in the area.

The ignition-noise problem is a temporary problem, since a normal development of the automobile may be expected to eliminate such noise-making transmitters; but in the meantime, it is feasible to equip single-sideband receivers with pulse-noise suppressors which will bring them into performance competition with FM receivers.

There seems to be no information about the equivalent of the flutter noise (standing-wave modulation effect) as applied to single-sideband systems.

One significant advantage of the use of single sideband transmission should be stressed. It should be obvious that, because of the reduced average power transmitted, potentially interfering intermodulation products will be greatly reduced.

So far, there has been no strong trend in the mobile industry to move toward single sideband, because the equipment is, in general, more costly, more critical in terms of adjustment and maintenance, and without substantial advantages within the present art for either improved performance or for increased channel loading. This is not the final word, however, and the pressure for radio spectrum space may eventually force the development of systems which may be operated without interference and with no more than 5-kc channel spacing. It should be noted in passing that the generation of many single-sideband channels properly coordinated can be accomplished at a given base station; but by the very nature of the mobile station, each car installation must be regarded as a separate and isolated system, thus limiting the use of such multiplexing to the fixed stations.

While the single-sideband approach to increased channel loading will be scrutinized constantly by the mobile radio industry, a second approach of interest is the wide-band time-sharing system sometimes referred to as the "random-access" system. At present, the application of the random-access discreet address system seems to have more potential interest for military communications than for use as a substitute for the FM police and industry communications systems. In the random-access approach, each transmitter sends short pulses which carry both the voice and the essential coding. Simultaneous conversations in the same spectrum band are possible because each transmitter has a relatively low-duty cycle transmitting the short pulses, and different transmitters share time on the same channel. Since the transmitters are not synchronized, there will be occasional coincidence of pulses. To minimize the resulting interference, the pulses from each transmitter are coded to particular receivers, so a receiver recognizes only the pulses addressed to it. The coding, or addressing, is accomplished by using redundant transmission, with each single transmitting pulse replaced by several pulses which carry the same modulation, but which are transmitted on several carrier frequencies and/or in different time slots. A receiver may have a fixed address, and a transmitter may have a variable code so that it can change the address to reach different receivers. The required bandwidth for a single RF channel is somewhere between 1 and 10 Mc, depending upon how many simultaneous talkers are to be accommodated in a given communication system. The maximum permissible number of simultaneous messages per Mc depends upon the selected modulating and addressing scheme, and on the geographical distribution of the stations. Several simultaneous messages per Mc may be possible. In a wide band-system, three or four hundred stations might find it possible to operate simultaneously, providing instant communication to associated addressee mobile units. The random-access approach substitutes time division for frequency slicing. Since we are now reaching frequency division limits which make it difficult to achieve improved efficiency through the use of further frequency slicing, the potential for increasing efficiency of spectrum utilization offered by other systems approaches must be thoroughly explored. It is too early to evaluate the competitive significance of random access and other similar time division systems for industrial and special mobile service applications, but the use of such systems would seem to be limited to special operational requirements of the type encountered in military operation.

In the frequency, time, and space dimensions of radio spectrum communications loading, there is the time factor which involves the time on the air for any transmitter as a factor independent of any modulation system used. The taxi radio dispatch service has succeeded in maintaining high-channel loading by operating systems which share time on the same channel. It is probable that the effectiveness of channel time-sharing could be greatly increased by the use of automatic selection and lock-out techniques to insure the protection of the station using the air. Either carrier control or wire line interconnections could be used, with proper monitoring by a computer memory and logic system to provide for an equitable distribution of time on the channel for the sharing systems.

The third dimensional factor, space, is concerned with a balance of many elements associated with the problem of geographical separation as a means for allowing channel assignment duplication without interference. It is obvious that elevation, antenna height, radiated power, tropospheric and ionospheric transmission, and FM capture-effect are all pertinent to the influence of space upon channel loading. Of the factors mentioned above, the FM capture effect has contributed the greatest positive force for permitting channel duplicate assignments with a reasonable geographical separation between systems. Unfortunately, the FM capture effect is reduced as the channel width is decreased; and while there is a gain in the total number of channels available for assignment, there is a loss in the number of duplicate installations for a given area because of the necessity for increasing the geographical spacing between systems operating on the same channel. Until adequate protection is provided against intermodulation and desensitizing, the splitting of channels may not add as many interference-free systems as might be possible with the wider bandwidth assignment in areas where modest geographical separations permit same-channel operation.

The single-sideband system does not provide a capture effect, but where the carrier is eliminated, the radiated energy follows the random surges of energy radiation which is characteristic of voice communications, and the average interfering power radiated is low. It would be expected that compared to the equivalent performance of a pair of FM systems, single-sideband systems operating on the same channels would cause less intelligibility-destroying interference. Every proposed new system of modulation displays both advantages and disadvantages. Until there is some *substantial* advantage in the use of the frequency, time, and space dimensions for one mobile radio modulation system over all others, there is very little incentive for changing the present FM pattern of systems design.

The change from AM to FM did not alter the physical characteristics of mobile equipment significantly. The transmitter was still a combination of frequency multipliers and a power amplifier, with the phase modulator substituted for the amplitude modulator. The receiver became a double-conversion superheterodyne with RF and IF amplification, but with the addition of limiters and a frequency detector which differentiated it from the AM unit. The first major construction change came with the adoption of the single-chassis construction for both transmitter and receiver, as opposed to the earlier approach where the transmitters and receivers were separate units. The single chassis reduced the size of the equipment, but perhaps the real advantage was the reduction in manufacturing costs.

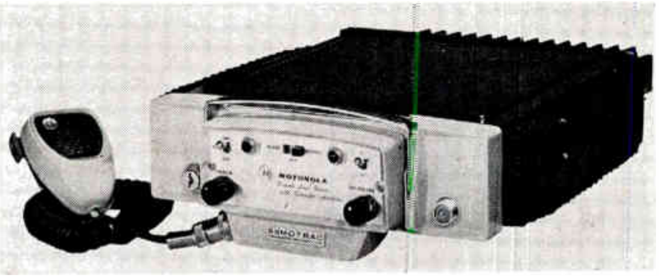


Fig. 5—The 1961 mobile FM station with transistorized receiver and power supply.

The most radical change in equipment construction came with the application of the transistors and solid-state diodes. The first application of transistors provided a power amplifier for the receiver output to increase the audio level. High audio level in police cars is particularly desirable because of the necessity for overriding the wind noise produced at the time of a high speed chase.

The second solid-state application was the use of transistors in a power supply oscillator circuit to step up the 6- and 12-v battery supplies to the high voltages required for receiver and transmitter operation. Both the transmitter power amplifiers and the transistor power supplies proved to be more reliable than the tube, vibrator, and motor generator equivalents. Size, weight, heat, and cost reduction were additional bonuses realized as a result of this application of solid-state devices. The present trend is toward the greater use of transistors in all mobile and portable equipment. There has been a complete change to transistors for both receivers and transmitters used in some portable applications. For the mobile units, the change-over has provided complete receivers with transistors and with no tubes, but the lack of high-power transistors for VHF amplification has limited the application to transmitter design. The use of transistors in the receiver has made additional reduction in the equipment, heat, size, and weight possible. Reports from the field, concerning the maintenance of new equipment built with transistors, show a substantial improvement in reliability over the record established with tube apparatus. The trend toward the utilization of transistors in equipment design will continue, and transistors will be applied to transmitter construction as soon as transistors with the needed high-power output characteristics become available.

The common carrier land-mobile radio service, pioneered by AT&T in 1946 with the St. Louis installation, has always used selective calling, which has also been available for use in the private systems. While the squelch provides protection against noise, it does not provide protection against unwanted carriers on the same frequency. Although selective calling would provide such protection, it also tends to slow down the operational response characteristics of the system, and for this reason, many systems owners have rejected selective calling. In operations where speed is less important, the use of selective calling has been discouraged by the substantial additional cost of the required accessory equipment. A compromise approach has been the use of selective squelch or tone-operated squelch. When the selective squelch system is used, the base station is equipped with a coded tone, usually operating at a subaudible level for the system, and the squelch systems in all of the associated mobile units will open only when activated by the base-station coded tone. The system requires that squelch code tones be allocated to avoid duplication in any given area. In the taxi service particularly, the use of coded squelch has been found to eliminate a high percentage of the unwanted and unrelated voice communications from other systems operating on the same channel. The coded squelch is a means for increasing the practicability of channel time-sharing.

While this paper has been primarily concerned with the pioneering work of the police and with the technological history of the land-mobile radio telephone systems development, the story would not be complete without some mention of the rule-making established by the Federal Radio Commission and, later, by the FCC, and of the qualification of new users and the growth of the land-mobile radio communication industry. The first rules in the format as we know them today (Part 10—Emergency Radio Service, and Part 11—Miscellaneous Radio Services) were adopted October 3, 1933, by the Federal Radio Commission. The FCC was established in 1934. The definitive standards for FM in the land-mobile services were not adopted by the FCC until it implemented the Atlantic City International Treaty of 1947, which was ratified by the U. S. on June 18, 1948. These technical standards were the same as those incorporated in our present-day rules up to the time of the split-channel rule-making proceedings of 1958. The 1948 rule-making established the land-mobile services on a fully licensed basis, extended the qualifications for licenses to many new users, and in general, established the foundation upon which the modern mobile communications industry has grown. From the original tentative allocation of frequencies for special emergency service, formal licensing has been extended to include, in addition to the public safety radio services, such services as land transportation, industrial radio, common carrier, and citizens band.

The continuing substantial expansion of land-mobile radio communications is inevitable. The EIA report on, "Private Land Mobile Systems in the United States," June 5, 1961, predicts that there will be 1,390,000 transmitters in service in 1963; 2,650,000 in 1968, and 5,000,000 in 1978. Both increased spectrum allocations and the use of new spectrum-stretching techniques must be adopted if growth is to be accommodated without an associated rise of interference to intolerable levels. Fortunately, the mobile radio engineers will have problems to keep them busy until 2012 A.D.

Acknowledgment

The assistance of R. L. Batts in the preparation of the early history of police mobile communications is gratefully acknowledged.

Future Developments in Vehicular Communications*

AUSTIN BAILEY†, FELLOW, IRE

Summary—**Starting from current trends in the field of land mobile radio-telephone services, primarily in the United States of America, forecasts of development over the next ten to twenty years have been made. It is estimated that by 1970 there will be two and a half million mobile telephones in the United States and over thirteen million by 1980. There is a discussion of the interrelated political, economic and engineering factors which will affect the future of vehicular communications. To meet user needs, coordinated broad-band systems are predicted for use cooperatively and on a common carrier basis with great efficiency of spectrum utilization. Currently, a broad-band system of 1000 channels could be accommodated in about 75 Mc of contiguous frequency spectrum in the lower part of the UHF band. It is conceivable that within the next two decades many more channels could be accommodated in a given bandwidth. Reducing the cost of mobile telephones is both necessary and capable of attainment. Pocket-size for a mobile telephone is indicated as probable during this period. New features for user convenience and better systems performance, such as full-duplex operation, push buttons for signaling, improved frequency stability, automatic channel selection and mobile station identification are predicted in coordinated systems within the next twenty years.**

Introduction

VEHICULAR communications include signaling and talking to or from vehicles which move on land, swim in water, or fly through the air and even in outer space. The only available means for connecting with vehicles is radio. If vehicular communications are to develop and expand, adequate provision in the limited spectrum available for all communication purposes will have to be made. It seems self evident that those regions in the radio spectrum best suited to vehicular communications and presently employed for other kinds of services soon will have to be made available to meet the demand of a people who are more and more a mobile population.

A discussion of the future of vehicular communications might properly include all the many ways of transmitting information and the adaptation of these ways to all kinds of vehicles. By far the largest number of mobile communications, however, are by telephone with and between vehicles traveling on land. Therefore, the scope of this paper will be limited to a discussion of telephone communication with land vehicles.

Looking toward what is yet to come always requires a glance backward to obtain perspective. We are living quite literally in the presence of an explosion of knowledge. It seems prudent to consider in this paper only those future developments which may be expected in the next ten to twenty years.

* Received by the IRE, June 26, 1961.
† Consulting Engineer, Millburn, N. J.

" . . . what's past is prologue;"

Public safety agencies were the first to find a real need to communicate with land vehicles. At first the communications were sent only one way. As the technology advanced and the use of somewhat higher frequencies became feasible, radio transmitters began to be installed in automobiles. The first two-way mobile telephone system for a police department was placed in service about the middle of 1933. By the end of that decade there were about 10,000 radio transmitters in police cars—all of which employed amplitude modulation.

About the middle of the 1930's, Edwin H. Armstrong made a discovery which was destined to change completely the techniques employed for land mobile radiotelephony. Although first applied to entertainment broadcasting, the impact of frequency modulation has been even greater in land mobile services. At the beginning of 1940, the Federal Communications Commission announced that it would accept applications for use of FM in services other than broadcasting. This was the beginning of a new era in land mobile communications. So valuable was FM in this type of service that within the period of only six years, substantially all transmitters were FM. FM not only reduces noise but helps reduce co-channel interference, signal flutter and variable intelligibility within the coverage area.

Beginning in 1945, with about 18,000 land mobile transmitters in service in the United States, the number has almost tripled every five years. Statistics from France indicate that the number in service there is doubling about every five years. Along with this phenomenal growth there have been advances in techniques resulting in lower costs, increased reliability, decreased weight and size, and generally improved attractiveness. But there have been few changes in the type of communications in the last decade from the standpoint of the user. Whether he is telephoning on a private system or employing the service furnished by a common carrier, this service is still largely one of dispatchers or land telephone subscribers talking to people in cars and trucks whenever these people happen to be in the right vehicles and within range of an associated base station. While it is true that technical advances have resulted in more satisfaction to the user because of reduced maintenance, lower battery drain and less channel sharing as a result of split-channel operation, something more than this is needed to make the service as convenient and valuable as is desirable. This will be discussed later.

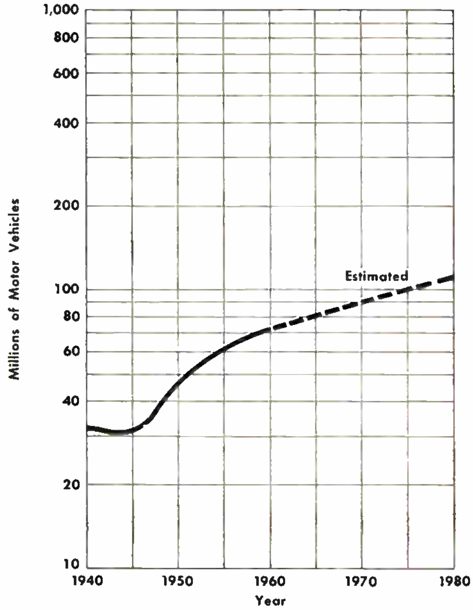


Fig. 1—Millions of motor vehicles registered in the United States based on Automotive Manufacturers Association data and the projection to 1970 made by General Motors Corporation.

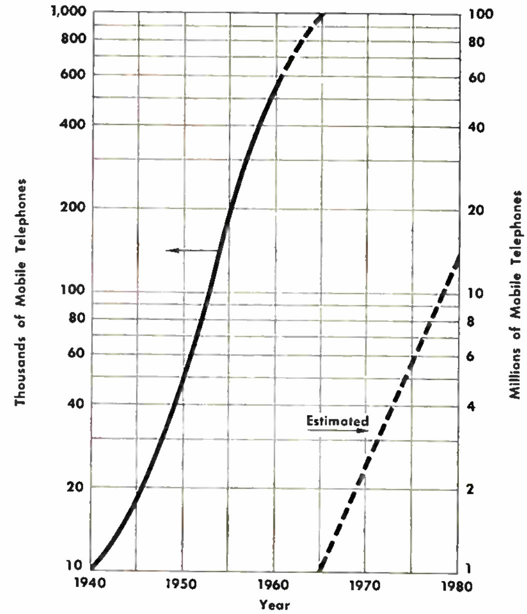


Fig. 2—Thousands of mobile telephones in the United States based on FCC annual reports and EIA projections to 1970.

" . . . what to come, is your and my discharge."

What kind of a world will it be in the next ten to twenty years? The world population at present is multiplying about four times in every 100 years. In the United States the rate, including immigration, is even greater—about fourfold every 70 years. The 1960 census in the United States counted some 179,323,175 people. The U. S. Census Bureau estimates that this figure will be about 203,000,000 in ten years and by 1980 about 231,000,000. Along with such a growth rate has come a substantial decentralization of industry and a redistribution of urban population into suburban and country areas. This is possible today because of the extensive communications network now available. The trend will certainly continue and means of communication will have to keep pace with the times.

Another change in the pattern of living in the United States is that individuals are increasingly on the move. This is apparent in the growth of annual passenger miles by public transportation and in the increase in the number of vehicles registered in the United States each year. Both of these are increasing at a rate greater than the rate of population increase. Fig. 1 shows the present growth rate in the number of vehicles registered in the United States to be about fourfold in forty years. However, it is estimated that this rate of growth will not continue but will gradually decrease in the next two decades. In Fig. 1 it is estimated that the total number of vehicles registered in 1970 will be about 90 million and in 1980 about 102 million. Logarithmic presentation is used because the slope of the curve indicates a growth rate independent of the numerical value of the ordinate. (Fig. 2 is on the same scale.)

With population, extra-urban development, mobility of individuals, the convenience of communications and the number of motor vehicles all increasing at astounding rates, what will be the growth in the number of land mobile telephones in use in the United States at the end of the next ten and twenty years? The answer to this question depends also on many political, economic and technical factors, and all of these factors are interrelated. Currently, in this country there are about 70,000,000 land telephones and only about 700,000 mobile telephones. It is estimated that by 1970 there will be about 2,500,000 mobile telephones in use and by 1980 over 13,000,000. Fig. 2 shows an estimate of the growth in the number of mobile radio transmitters used in the land mobile services.

Political Considerations

Nature has provided a limited frequency spectrum suitable for mobile services. Because this limited space is one of the natural resources, it is and should be controlled by government. There has to be a proper balance between the values of various services in the assignment of this frequency space. Furthermore, each service must be efficient in its use of frequency space.

There are a number of ways in which better use could be made of the frequency space available. Sharing, both geographically and on a time basis, is employed somewhat in services with similar characteristics such as those in the various categories of land mobile services. In the frequency range above 50 Mc a great deal more could be done.

Already there is substantial geographical sharing in the 150-Mc band; probably the greatest application of this has been in the common carrier services. To the extent that it proves useful, it is to be expected that computers will be used to determine the geographical distribution of frequency assignments. Data on the interference range of existing stations will be stored in the memory of these computers and new systems will be engineered in on the basis of noninterference with existing systems.

Time-sharing is already in use to a limited extent among those stations in the same classes of service where the time-value of information is not too critical. There is no doubt but that more of this is possible. However, a great deal will be accomplished when the load factors of different types of services are taken into account. As an example of this, many of the land mobile services which are employed primarily for commercial purposes have load factors with high peaks during certain hours of the business day and very small demand during the evening and night periods. On the other hand, the audiences for broadcasting and television are smaller during the business day and much greater during the evening hours. It seems self evident that in the not too distant future there will be time-sharing of frequencies by different classes of services.

Engineers have accomplished much to alleviate the regulatory problems. In the mobile services, improvement in techniques have already enabled channel assignments to be placed more closely together. In the 30- to 50-Mc region, channel spacings in the United States have been reduced from 40 to 20 kc. In the 150- to 160-Mc range, channel spacings are currently 30 kc in the same geographical area and 15 kc in adjacent areas. The region between 450 and 470 Mc has also been reduced to 50-kc spacing, and a further split to 25 kc is proposed.

The consensus of engineering opinion at this time is that little more can be accomplished in the way of channel splitting in this general frequency region with the present individual channeling plan. The most promising solution is the general use of broad-band systems. With such systems, a plurality of voice communication channels from a common base station is possible with freedom from the coordination problems present with large numbers of individual channels. Accordingly, they would probably be operated by common carriers for service to many categories of users. A broad-band system which has been proposed and awaits only governmental approval contemplates up to 1000 channels, in up to about 75 Mc of frequency spectrum in the lower part of the UHF band. This system gives each call access to a group of these channels, and this greatly improves the efficiency with which channels are utilized. This *group operation factor*, or the gain in message handling capacity obtained by having access to a group of channels rather than to only a single channel, depends on a number of variables such as calling rate, holding time and the allowable average delay or permissible probability of delay. It can amount to a fourfold increase where access is provided to somewhere in the order of 5 to 10 channels. Probably, the broad-band system will embody access to a larger number of channels for greater efficiency.

To obtain even greater channel utilization for mobile telephone frequencies in densely built-up metropolitan areas, a cellular plan for broad-band systems has been proposed. This plan contemplates the use of a considerable number of low-power base stations with quite limited coverage from each. By this arrangement it would be possible to utilize the same frequency assignment simultaneously for several conversations in the metropolitan area. Of course, to make this plan effective electronic switching associated with memory circuits will be essential to maintain continuity of the telephone connection while a conversation is in progress.

Early construction of and experience with the proposed broad-band systems, together with advancements in several phases of the art over the next two decades, could lead to further improved spectrum utilization. It is possible to envision a 4-Mc band of spectrum providing about 250 two-way telephone circuits which, with utilization of the group operation factor, might give service equivalent to that of more than 1000 individual channels today.

As a further step in making greater message-carrying capacity available, digitized voice transmission, or some other new and better method of transmitting the intelligence in speech on a reduced bandwidth, may come into general use. Such a new development might result in a further increase in channel availability of over tenfold.

Economic Considerations

The mobile telephone user needs an equipment which is not only low in first cost but also one that will require low maintenance. Transistorization may lead to units with these characteristics in a few years. However, low cost and low maintenance are not the only features needed. Attention will be directed later to other changes from present practice to improve system operation or to increase the user's convenience. All of these improvements will require extensive development work in this mobile radio field.

Development costs are bound, in one way or another, to become a part of the price of the equipment. Only if good mobile telephone equipment can be manufactured in sufficient quantity will the price be low. Automated manufacturing methods tend to reduce the cost of fabrication, but the tooling expense is considerable. Tooling costs for automatic production prove in only when the quantity to be produced is relatively large. Both the development and manufacturing costs, on which so much of the price to the user depends, will be reduced only if there is assurance that a market exists for the product.

At present, installation costs for mobile telephones in cars are material. They add to what the user must pay to get a telephone in his vehicle. There are two approaches to the reduction of this item of expense. The first approach will most certainly be factory installation of the mobile telephone as an accessory in the motor vehicles. This alone would reduce the installation cost to about a tenth of the present figure. There seems to be no valid reason why this should not be done now in fleets of cars for private systems and in automobiles for users having a need for telephone service through the common carriers.

The other approach one may estimate will probably materialize in the period between 1970 and 1980. With subminiaturization, the mobile telephone may become more an item of apparel than an automotive accessory. However, motor vehicles may well be factory equipped with standardized arrangements to carry the mobile telephone when it is to be used in an automobile. A radiotelephone *holster* will provide connection to the car battery and connection to an antenna installed on the automobile. Moreover, it may provide additional features.

From the point of view of economics, mobile telephones will have to be removed from the class of luxury items and become commonplace tools of industry. Certainly, the user will be willing to pay more for access to mobile communication facilities than to fixed ones, but how much more? That all depends on the value the potential user places on the mobility factor in communications. Proportionately more people will prove interested in mobile telephones as the costs trend toward those of telephones in fixed locations. However, it is likewise a proven fact that more people will want mobile telephones as convenience in the use of such facilities approaches the convenience of telephones at fixed locations.

Engineering Considerations

Recent advances in solid-state physics and all the variety of diodes, transistors, thin metallic films, new magnetic materials, new dielectric compounds, unheard of semiconductors and *molecular circuitry* are but the harbingers of the developments to be expected over the next two decades. This must be well understood in considering future developments in the engineering of mobile radiotelephones because the changes from present practices and techniques will be almost revolutionary.

Coordinated Broad-Band Systems

As previously mentioned, coordinated broad-band systems are the best hope for the spectrum development essential to satisfying the growing demand for mobile telephones for both cooperative and common carrier uses. Engineering studies have shown this in a most convincing manner. Not only is a coordinated system the logical way to share equitably the suitable RF space, but it seems to be the only way to obtain more mobile telephone channels with less interference and lower cost to the user. Moreover, most of the complexity required, including multiple receiver locations, will be in the shared base station facilities, and thus the costs would be divided among the users.

It must be appreciated that the development of the coordinated broad-band systems with the efficiency of spectrum utilization envisioned in the section on "Political Considerations" will not come into being all at once but will evolve from extensive development. To be sure, when first used, the efficiency of spectrum utilization may be only a small number of times greater than split channels can give but they have far greater ultimate capacity.

Without further consideration of the engineering aspects of coordinated broad-band systems or the base stations used therein, some of the features that may be developed for mobile equipments operating as a part of such systems can be outlined.

Size Reduction

Over the past two decades there has been a substantial reduction in the size of mobile telephone equipments. The past history and the trend of the future is shown in Fig. 3. In the photograph of Fig. 4 this trend is further illustrated by comparison of a two-unit equipment of only a few years ago with a current single-unit equipment which is transistorized in part. Unfortunately, the price of transistors has also increased the cost of mobile telephones, but this situation is only temporary. Transistors are being produced in ever-increasing quantities with improved quality at lower cost. In a few years, as the methods for manufacturing are improved, the cost of transistorized equipment may be expected to be less than that of equipment employing vacuum tubes.

Some transistors are now being produced in quantity occupying a volume of only 4×10^{-11} cubic inches, and this is probably not as small as they can be. However, reduction in size of the equipment has not been accomplished solely as a result of the smaller size of transistors. The efficiency of transistors has also made it possible to reduce the size of the equipment and still obtain adequate heat dissipation. This problem of heat dissipation will probably set the lower limit on the size of mobile telephones but current practices in engineering design are a long way from that limit.

It is estimated that the micromodular approach to the problem of size will result in a 100 to 1 reduction over present-day components. Reduction by a factor of five might result from the use of thin films deposited on substrates to form the necessary electric circuits. It has been predicted that this will probably be in use by 1970. This seems not to be the limit in size reduction, since it is known that integrated circuits on common semiconductor substrates will yield a further 4 to 1 decrease in size. Two-stage amplifiers have already been demonstrated which have a volume no more than three ten-thousandths of a cubic inch. Moreover, it has been claimed that the unconventional electrical principles of

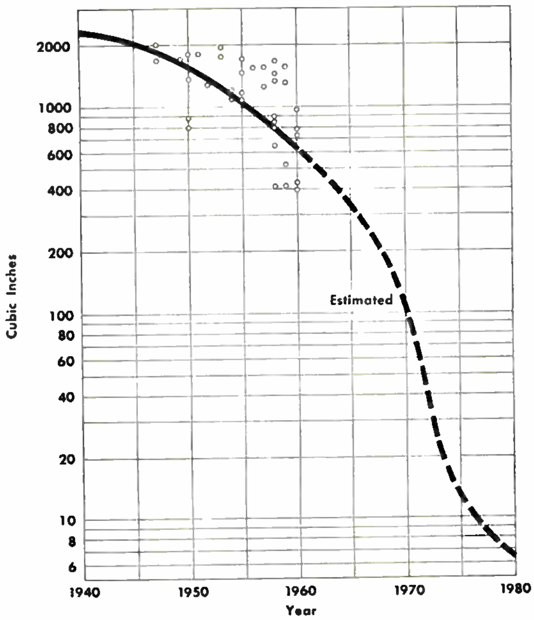


Fig. 3—Volume in cubic inches of two-way mobile radio telephone equipment—10–30 watts RF output—from data furnished by General Electric Company.

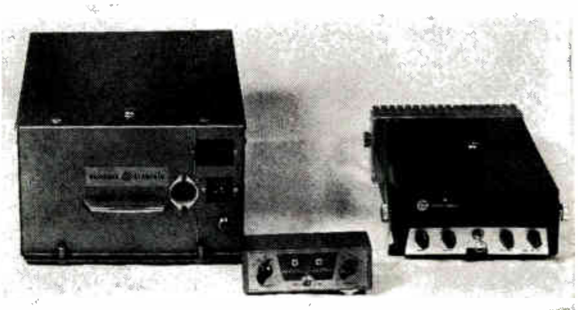


Fig. 4—General Electric Company mobile telephone equipments showing size reduction between 1950 and 1960.

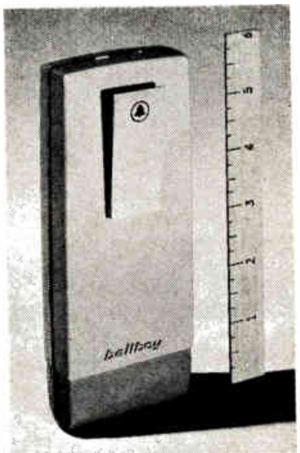


Fig. 5—Bell System's "Bellboy"—a 152- to 162-Mc radio receiver for adjacent channel operation with selective signaling and self-contained rechargeable battery.

molecular circuitry will make possible a reduction in equipment volume to one ten-thousandth of the size currently used for mobile telephone equipment. Obviously not all of these reductions will be needed to obtain in a package of about 6.5 cubic inches the equipment now contained in about 650 cubic inches.

Pocket-carried radio receivers with about 7 cubic inches volume, including batteries, are available now from several suppliers. These are employed in certain paging services. Radio transmitters with lower power and limited range of about this same size are likewise on the market. Illustrated in Fig. 5 is a selective signaling receiver which operates in the 152- to 162-Mc frequency range. This may be thought of as a mobile telephone bell. All of these devices perform but one of the several functions needed in the mobile telephone of the future. It seems quite reasonable to expect that before 1980 all of the functions then required for a two-way mobile telephone can be provided in a unit no larger than one of the paging receivers.

One thing not covered by the foregoing discussion of size nor illustrated by Figs. 3 or 4, is the power supply necessary to run the equipment. Some way must be found to reduce the size of the power supply. While the use of transistors has already done much to reduce the power requirement of mobile telephones, there is a limit when consideration is given to the RF power that must be generated and radiated. It is expected that significant advances will be made in the power supply field by 1970. A simple alternative would be a large reduction in mobile transmitter power with a correspondingly large increase in the number of radio receivers associated with each base station transmitter. Such proliferation of radio receiving locations may be found to be economically feasible in any event.

Duplex Operation

Most mobile telephone systems today are operated on a push-to-talk basis. The common carriers, however, are using push-to-talk in the vehicle only. Where the general public, such as passengers on railroad trains, has access to the telephone in the vehicle, and separate transmit and receive frequencies are employed, no push-button is involved.

Push-to-talk operation requires some experience in proper coordination between hand and voice. There are available today arrangements which permit operation of the mobile units without using manual switching. It seems likely that soon push-buttons will be outmoded, except for car-to-car communications or single frequency operation.

Mobile Frequency Control

Single-sideband operation, with its accompanying efficiency of spectrum utilization in coordinated broadband systems, requires a higher order of frequency stability than currently in use. In such systems there are numerous ways this might be accomplished. An innovation which may be introduced is that all radio frequencies needed in the mobile set will be derived from control frequencies sent out from the base station transmitter. By this means all of the frequencies in the coordinated system might be held within a tolerance two orders of magnitude better than present practice. This can be done without costly equipment in the vehicle.

The techniques, to provide much better frequency stability at the base station, are available at the present time. Application of these techniques will increase the base station cost, but again this incremental increase on a system basis is divided among the large number of mobile units served by the system. It seems likely that control of mobile station frequencies from the associated base station will come into extensive use even before the change is made to single-sideband operation. Simple cost considerations indicate that this will be the case.

Signaling and Switching

A definite trend in mobile equipment today is toward the inclusion of some kind of selective calling. The very nature of radio makes a mobile telephone a party-line telephone, and it is desirable in nearly all cases to communicate with but one party at a time. Common carrier systems have used selective signaling since the very beginning.

The selective signaling apparatus in the mobile unit will be used also to switch the mobile equipment automatically to some designated channel in the group which, at that time, is not being employed for communication. Moreover, this same apparatus also will function automatically to send its identifying signal back on the designated channel. Arrangements to do this have been demonstrated but are not yet in commercial use even on the few channels now available. It seems likely that the value of automatic channel selection and mobile station identification will make them practically essential on coordinated systems.

One of the features essential to economy in common carrier systems will be selective signaling in both directions on substantially all calls. Several installations of such *two-way dialing* systems have already been made and are in successful operation.

In the next several years there will be further development in the signaling arrangements used for selective calling, station identification, automatic channel selection and two-way dialing. The use of push buttons in the vehicle instead of the conventional telephone dial is indicated. User demand for convenience in operation will be the controlling factor.

A new signaling system could retain the desirable characteristics of frequency shift keying so essential to reliability but eliminate the undesirably slow speed of a dial-pulse system. It may be essentially a start-stop system, possibly employing a four-unit code. The code signal for a single digit should be sent in about a tenth of a second; just a little faster than the time required for a person to push successive buttons in a group to send a sequence of digits. Either electronic or mechanical means may be used, but the apparatus would differ materially from that now in mobile telephone service.

Conclusions

There exists, most certainly, a need for great expansion in vehicular communications over the next two decades. Large-scale production of equipment for this expansion would tend to reduce costs. Technological advances are providing new tools with which to go further in the directions indicated by current trends. Adequate spectrum space to support this expansion must be assured, however, if the need is to be met.

Acknowledgment

The author wishes to express his appreciation for the information and assistance he has received from J. A. Anderson, General Motors AC Spark Plug Division, J. E. Dingman, AT&T Company, Admiral F. R. Furth, IT&T Corporation, H. Van Aken, General Electric, and from their associates in each of these companies. Particular thanks are due R. P. Gifford, GE Company and J. W. Stubner, AT&T Company, for their valuable help during the preparation of this paper.

Contributors

M. Donald Adcock (SM'56) was born in Tolladega, Ala., on August 16, 1921. He received the A.B. degree in physics from Howard College, Birmingham, Ala. He also studied radar at Massachusetts Institute of Technology and Harvard University, Cambridge. in 1944.

In 1942, he joined M.I.T.'s Radiation Laboratory, under Dr. L. C. Van Atta, where he designed, supervised, and operated the first primary-pattern apparatus which was the forerunner of many such equipments; he assisted in the development of the original ground-based microwave fire-control antenna and worked on the first cut paraboloid antenna design. In 1944 he joined the Navy as a Radar Officer; upon his release he worked for the Antenna Research Branch, Naval Research Laboratory, Washington, D. C. There he developed broadband antennas, and did research with Dr. A. E. Marston on helical radiators. In 1951 he was employed by Hughes Aircraft Company, Culver City, Calif., where he was Head of the Microwave Laboratory. He is responsible for many antenna developments, such as two-dimensional slot arrays, low-frequency very-large antenna systems, and reconnaissance antennas. He was co-founder of American Systems, Inc., Inglewood, Calif., where he heads the Electromagnetic Systems Division.

❖

Samuel N. Alexander (A'44–M'55–SM'55–F'56) was born in Wharton, Tex., on February 22, 1910. He received both the A.B. degree in physics and the B.S. degree in electrical engineering, in 1931, from the University of Oklahoma Norman, and was awarded a Tau Beta Pi Fellowship for graduate studies. In 1933 he received the M.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, where he continued his doctoral studies under a Coffin Foundation Research Fellowship, in 1934.

From 1935 to 1940 he worked as a Laboratory Engineer for the Simplex Wire and Cable Company, Cambridge, and during 1939 and 1940 was a Research Assistant at M.I.T. From 1940 until 1943, he was a Physicist with the Navy Department working on electronic instrumentation and from 1943 to 1946 was Senior Project Engineer with Bendix Aviation Corporation, Towson, Md., working on development of military telemetering equipment. In 1946 he joined the staff of the National Bureau of Standards, Washington, D. C. As Chief of the Electronic Computers Laboratory he was engaged in the development of digital computer technology and had over-all supervision of the design and construction of SEAC and DYSEAC. Since 1954 he has been Chief of the Data-Processing Systems Division of NBS. This activity grew out of the earlier computer development program and now includes the application of digital techniques to data processing and information storage and retrieval and the application of combined analog and digital techniques to automatic instrumentation and dynamic control systems.

Mr. Alexander is a member of the AIEE, ACM, American Physical Society, American Documentation Institute and Washington Academy of Sciences. He is also a member of Phi Beta Kappa, Sigma Xi, Tau Beta Pi and Sigma Tau. In 1956 he received a silver medal from the Royal Academy of Engineering Science, Stockholm, Sweden.

❖

Edward W. Allen, Jr. (M'44–F'53) was born in Portsmouth, Va., on February 14, 1903. He received the E.E. degree from the University of Virginia, Charlottesville, in 1925, and the L.L.B. degree from George Washington, University, Washington, D. C., in 1933.

He joined the Federal Communications Commission, Washington, D. C., in 1935. He became Chief of the Technical Research Division in 1946, and since 1951 has been Chief Engineer of the FCC. He has been active in the development of technical standards and radio allocations throughout his association with the FCC.

Mr. Allen, a member of the Electrical Standards Board of the American Standards Association, has participated in electrical coordination work for many years. He is also a member of the U. S. National Committees of URSI and the International Electrotechnical Commission. He is a member of the Executive Committee of the International Radio Consultative Committee (CCIR) of the International Telecommunications Union (ITU) and has served as a delegate to seven meetings of the Union, at two of them as Delegation Chairman. He has been a member of the IRE Wave Propagation Committee since 1949, and received the IRE Harry Diamond Award, in 1958. He is a member of Tau Beta Pi and of Phi Alpha Delta.

Gene M. Amdahl (A'53–M'57) was born in Flandreau, S. Dak., on November 16, 1922. He received the B.S. degree in physics from South Dakota State College, Brookings, in 1946, and the M.S. and Ph.D. degrees in physics from the University of Wisconsin, Madison, in 1949 and 1952, respectively.

From 1952 to 1956 he was associated with the IBM Development Laboratory, Poughkeepsie, N. Y., where he was Project Engineer and Systems Designer of the IBM 704, 709 and Stretch computers. From 1956 to 1960 he worked at the Thompson Ramo Wooldridge Corporation and at Aeronutronic Systems, Inc., where he held the position of Manager of Data Processing Engineering. In 1960 he rejoined IBM in Research, Yorktown Heights, N. Y., where he held the position of Director of Experimental Machines, and is now the Advanced Systems Design Manager in the Data Systems Division, Poughkeepsie, N. Y.

Mr. Amdahl is a member of the American Physical Society.

❖

Morton M. Astrahan (S'45–A'50–M'51–SM'57) was born in Chicago, Ill., on December 5, 1924. He received the B.S.E.E. degree from Northwestern University, Evanston, Ill., in 1945, the M.S.E.E. degree from the California Institute of Technology, Pasadena, in 1946, and the Ph.D. degree from Northwestern. in 1949.

He joined the IBM Corporation, Endicott, N. Y., in 1949. In 1951 and 1952 he participated in the planning and logical design of the 701 Computer at Poughkeepsie, N. Y. From 1953 to 1955 he was in charge of the IBM Systems Planning on the AN/FSQ-7 Computer for the SAGE Air Defense System. He also directed input-output development and prototype testing for the AN/FSQ-7. In 1956, he moved to the San Jose Research Laboratory of IBM, Calif. He was Manager of systems research and responsible for applications and specifications of experimental research devices for future data-processing systems, 1956 to 1958. In 1959 he joined the San Jose Laboratory of the newly formed IBM Advanced Systems Development Division. He is currently Assistant Manager of a systems development project.

Dr. Astrahan was the first Chairman of the IRE Professional Group on Electronic Computers, in 1952 and 1953. He was Chairman of the National Joint Computer Committee, from 1956 to 1958, and Chairman of the NJCC delegation visiting computers in the USSR, in 1959. He is currently Chairman of the Finance Committee for the American Federation of Information Processing Societies. He is a member of the AIEE, Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.

❖

Stephen S. Attwood (SM'44–F'61) was born in Cleveland, Ohio, on May 29, 1897. He received the B.S.E. degree in mechanical engineering, in 1918, and the M.S. degree in electrical engineering, in 1923, both from The University of Michigan, Ann Arbor.

During the years 1918 to 1919 he served as Ensign, U.S.N.R., aboard naval vessels with the responsibility of Assistant Engineer Officer. The year 1919 was devoted to construction work for the Packard Motor Car Company. He returned to The University of Michigan in 1920 as Instructor in Electrical Engineering, and was promoted through the grades to Professor, in 1937. He was a member of the U. S. Government Mission to England on Radio Wave Propagation in late 1943. From 1944 to 1945 he was Director of the Wave Propagation Group, at Columbia University, New York, N. Y., for the Office of Scientific Research and Development. In 1953, at The University of Michigan, he was named Chairman of the Electrical Engineering Department, and in 1958 he was appointed Dean of the College of Engineering.

Mr. Attwood is a member of Tau Beta Pi, Sigma Xi, Eta Kappa Nu, and is a Fellow of the AIEE.

❖

William S. Bachman (S'32–A'36–M'45–SM'54–F'56) was born in Williamsport, Pa., on October 29, 1908. He received an E.E. degree from Cornell University, Ithaca, N. Y., in 1932.

From 1934 to 1946 he was associated with the General Electric Company, as Engineer in the Radio Receiver Engineering Department, having made contributions to the development and design of loudspeakers, radio and phonograph combinations, and phonograph pickups. Among his developments, the most familiar is the G.E. variable reluctance pickup. In 1946 he joined Columbia Records, Inc., as Director of Engineering and Research, and in this capacity, contributed to the development and production engineering of Lp records, the heated stylus disk recording technique, viscous damped reproducer arms and program controlled automatic groove-spacing equipment.

Mr. Bachman has been active on numerous industrial committees and is a Past Chairman of the Technical Committee of the Record Industry Association of America. He has served for many years on the Board of Editors of the IRE. He was granted the Audio Engineering Society Berliner Award for outstanding developments in disk recording and reproduction. He is a Fellow of the Audio Engineering Society.

❖

Austin Bailey (A'22–M'25–F'36–L) was born in Lawrence, Kan., on June 9, 1893. He received the A.B. degree from the University of Kansas, Lawrence, in 1915, and the Ph.D. degree from Cornell University, Ithaca, N. Y., in 1920.

In 1918 he was made a Second Lieutenant of the U. S. Army Signal Corps. After his university work and following military service, he was first employed as Superintendent of the Apparatus Division, Corning Glass Works, Corning, N. Y., and then became Assistant Professor of Physics at the University of Kansas. In June, 1922, he accepted a position with the American Telephone and Telegraph Company, New York, N. Y., where he worked on radio engineering problems, retiring from the Bell System as Mobile Radio Engineer on June 30, 1958. Since then he has been an independent Engineering Consultant.

Dr. Bailey is a member of Sigma Xi, Alpha Chi Sigma, and a Fellow of AIEE.

❖

Ralph R. Batcher (J'16–A'18–M'22–SM'43–F'50) was born in Dysart, Iowa, on June 7, 1897. He received the B.S. degree in electrical engineering from Iowa State College, Ames, in 1920.

While a radio amateur (starting in 1909), he organized in 1915 an early broadcasting service in the middle west which sent daily weather and crop reports from station 9YI. During World War I, he became a Radio Inspector at the Port of New York, and also taught at the Marconi Institute and Signal Corps classes at the College of the City of New York. In 1920 he became a Circuit Designer Engineer on machine switching systems at the Western Electric Company, New York, N. Y. In 1924 he joined A. H. Grebe and Company, L. I., N. Y., as receiver Designer and a Consultant to their broadcasting stations WAHG, WBOQ and WABC.

While there he invented the SLF variable condenser.

In 1928 he helped organize and was an officer of the Decatur Manufacturing Company, Brooklyn, N. Y., manufacturers of radio speakers. In 1935 he joined the Allen D. Cardwell Company, Brooklyn, as a Consultant, and later became Director of Engineering until 1944, when he became Consultant on the publications of the Caldwell-Clement Corporation, New York, N. Y., on the magazines *Electronic Industries, Electronic Instrumentation* and *Tele-tech.* While at Cardwell, he developed a number of printed circuit applications and the use of plated circuits for electronic assemblies. He also designed a completely automatic production calibrator for military frequency meters that "read" the dials at more than 100 readings per minute and printed the calibration books automatically. This calibrator saved many hundred man-years of calibrator's time during the production interval.

In 1950 he became Chief Engineer of the RETMA (now Electronic Industries Association), where he handled their Data Bureau in charge of tube registrations and more than one hundred technical committees dealing with system and product standardization work. He was a Consultant to the Joint Electron Tube Engineering Council (JETEC), and a Consultant to the Department of Defense Advisory Group on Reliable Electronic Equipment (AGREE). At present he provides a consulting service to manufacturers on product design and production problems, and automatic control.

Mr. Batcher is a Fellow of the Radio Club of America and a member of Tau Beta Pi, and a registered professional engineer in the state of New York. He was Chairman of the Board of Directors of the 1954 National Electronics Conference. From 1954 to 1957 he served as Chairman of the IRE Professional Group on Production Techniques. He has been active on many IRE committees, including the Annual Review Committee (Chairman, 1949–1954), Public Relations (Chairman, 1950–1951), Board of Editors, Meetings and Papers, Standards, Symbols, and Television. From 1951–1952 he served as IRE representative to the ASA Sectional Committee (C16) on Radio (Secretary, 1952).

❖

Max C. Batsel (A'21–F'27) was born in Fulton, Ky., on June 16, 1894. He received the B.M.E. degree from the University of Kentucky, Lexington, in 1915.

He then immediately joined the Western Electric Company, Hawthorne Works, Chicago, Ill., as a Student Engineer. Upon completing this course of training, he joined the staff of the Bureau of Standards, Washington, D. C. He was commissioned as Second Lieutenant in the Signal Corps during

World War I and assigned to the Radio Development Section. In 1920 Mr. Batsel resigned from the Signal Corps and joined the Westinghouse Electric and Manufacturing Company, where he participated in the design of the first receivers specifically designed for receiving broadcasted programs in 1921. The Radio Department was established in 1922, and he was appointed Section Engineer responsible for broadcast receiver development and design. In 1928 the responsibilities of the section were increased to include the development and design of sound motion-picture equipment, both 16 and 35 mm.

In 1929 he was appointed Chief Engineer of the RCA Photophone Division, a separate RCA Subsidiary Company with headquarters in New York City and a branch office in London, England. In 1932 the Photophone Company was merged with the RCA Manufacturing Company in Camden, N. J., and Mr. Batsel was appointed Manager of the Sound Section of the Engineering Division. The engineering and manufacturing operations of the Sound Division were moved to Indianapolis, Ind., in 1941, and he became Chief Engineer of the Indianapolis Plant. He served in this capacity until recalled to Camden to assume the duties of Chief Engineer of the Engineering Products Department of which the Sound Division became a Section. In 1955 he was appointed Chief Engineer of the Defense Electronic Products. At present he is retired and acting as Consultant.

Mr. Batsel is a Fellow of the Society of Motion Picture and Television Engineers, and a member of the American Society of Naval Engineers and of Tau Beta Pi. He received the Modern Pioneers Award given by the National Association of Manufacturers in 1940.

❖

Benjamin B. Bauer (S'37–A'39–SM'44–F'53) was born in Odessa, Russia, on June 26, 1913. He received the B.S. degree in industrial electrical engineering from Pratt Institute, New York, N. Y., in 1932, and the E.E. degree from the University of Cincinnati, Ohio, in 1937.

In 1936 he joined Shure Brothers, Inc., Chicago, Ill., manufacturers of microphones and electronic components; he became Chief Engineer, in 1940, and Vice President, in 1950, and continued in this capacity until 1957 when he was appointed Head of Audio and Acoustics at CBS Laboratories, Stamford, Conn. In 1958 he became Vice President of CBS Laboratories taking charge of the Acoustics, Magnetics, and Instrumentation Research Departments. During this time he was responsible for the development of a great many electroacoustical instruments such as microphones, disk recording and reproducing transducers, and magnetic recording devices. He has over thirty patents in these fields and has frequently been concerned with the task of carrying various products through research, design and into the production engineering phases.

Mr. Bauer is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi, and a Fellow of the Acoustical Society of America and the Audio Engineering Society. He recently received the Eta Kappa Nu Award of Merit.

❖

Arnold H. Beck (S'36–A'39–M'55–F'59) was born in Norfolk, England, on August 7, 1916. He was educated at Gresham's School, Holt, and at University College, London, where he received a B.S. degree in engineering, in 1937.

After two years work on electronic navigation with Henry Hughes & Sons, he was seconded to the Admiralty Signal establishment until the end of World War II. In 1947 he joined Standard Telephones & Cables to establish the valve laboratory then located at Enfield. He worked on a broad range of valves and gas tubes for communication systems at S.T.L. In 1958 he joined the academic staff of Cambridge University where he is a Lecturer in Electrical Engineering.

Mr. Beck is an associate member of the IEE, and the author of three books on electronics: "Velocity Modulated Thermionic Tubes," "Thermionic Valves" and "Space-Charge Waves."

❖

Vitold Belevitch (SM'54–F'61) was born in Helsinki, Finland, on March 2, 1921. He received the E.E. and M.E. degrees, in 1942, and the D.Sc. degree, in 1945, all from the University of Louvain, Belgium.

From 1942 to 1957 he was in the Wire Transmission Department of Bell Telephone Mfg. Cy., Antwerp, Belgium.

At present he is Director of the Comite d'Etude et d'Exploitation des Calculateurs Electroniques, Brussels, Belgium, and a Professor at the University of Louvain.

Dr. Belevitch is the author of two books on network theory (in French) and of numerous articles.

Leo L. Beranek (S'36–A'41–SM–45–F'52) was born in Solon, Iowa, on September 15, 1914. He received the B.A. degree from Cornell College, Mt. Vernon, Iowa, in 1936, the M.S. and Sc.D. degrees in applied physics from Harvard University, Cambridge, Mass., in 1940, and the D.Sc. (Hon.) degree from Cornell College, in 1946.

At Harvard he was an Instructor from 1940–1941, Assistant Professor from 1941–1943, Director of Research on Sound Control from 1940–1943, and Director of the Electro-Acoustics and Systems Research Laboratories from 1943–1946. At the Massachusetts Institute of Technology, Cambridge, he was Associate Professor of Communication Engineering from 1947–1959, Technical Director of the Acoustics Laboratory from 1947–1953, and has been a Lecturer since 1959.

Dr. Beranek is Chairman of the Acoustical Standards Board of the American Standards Association and a Fellow of the American Physical Society, the Acoustical Society of America, the American Academy of Arts and Sciences, the AAAS, and the Audio Engineering Society.

❖

Harold H. Beverage (A'15–M'26–F'28–L) was born in North Haven, Me., on October 14, 1893. He received the B.S. degree in electrical engineering in 1915, and the honorary degree of Doctor of Engineering, in 1938, both from the University of Maine, Orono.

He was associated with Dr. E. F. W. Alexanderson of the General Electric Company in 1916, specializing in the development of receiving systems. He developed the wave antenna in collaboration with C. W. Rice and E. W. Kellogg. In 1920 he was transferred to the Radio Corporation of America, and continued to specialize in the development of receiving systems for VLF and HF transoceanic communication, at Riverhead, N. Y. He collaborated with H. O. Peterson and J. B. Moore in the development of the space diversity receiving system for high frequencies. He became Chief Research Engineer for RCA Communications, Inc., in 1929, and was appointed Vice President in Charge of Research and Development in 1940. When the Communications Research Laboratory became part of RCA Laboratories in 1942, he was appointed Director of Radio Research. He retired in 1958, and is presently a Consultant on Communications to the Radio Corporation of America.

Mr. Beverage is a Fellow of the AIEE

and the AAAS. He has received many honors, including the IRE Morris Liebmann Memorial Prize, 1923; the IRE Medal of Honor, 1945; the IRE PGCS Achievement Award, 1958; the AIEE Lamme Medal, 1957; the Signal Corps Certificate of Appreciation, 1944; and the Presidential Certificate of Merit, 1948.

❖

Frank J. Bingley (A'34–M'36–SM'43–F'50) was born in Bedford, England, on November 13, 1906. He received the B.Sc. degrees in mathematics and physics (both with honors), in 1926 and 1927, respectively, from the University of London, England.

Since then he has been continuously associated with television. He was employed by the Baird Television Company of London, in 1927 and was in charge of their New York Laboratories for two years. In 1931 he joined the Philco Corporation of Philadelphia, Pa., where he has been associated with the development of transmitting and receiving equipment. as well as with television-systems engineering. He has worked extensively in recent years in the field of color television, and is currently in charge of a program of industrial and military television equipment design.

Mr. Bingley is a member of the Franklin Institute, has served on many industry committees concerned with the development of television standards, and is the recipient of the 1956 Zworykin Television Award of the Institute for his work in colorimetry.

❖

Henry P. Birmingham (M'53–SM'56) was born in Newburgh, N. Y., on March 17, 1920. He received the Ed.B. degree from Rhode Island College, Providence, in 1942, and did further study at the University of Pennsylvania, Philadelphia, and the American University, Washington, D. C.

He served in the U. S. Air Force as an Instructor in Radar Maintenance and as a classification specialist during World War II. He later joined the staff of Brown University, Providence, as a Research Psychophysiologist under NDRC. He joined the U. S. Naval Research Laboratory in 1945, where he is presently Head of the Engineering Psychology Branch in the Applications Research Division.

He organized the IRE PGHFE and served as Chairman during the first year of its existence. He received the Naval Ordnance Development Award for work in fire control, and the Meritorious Civilian Service Award for work in the development of man-machine control system theory.

Mr. Birmingham is a member of the Aerospace Medical Association, the Human Factors Society, and the Eastern Psychological Association.

❖

Henry G. Booker (SM'45–F'53) was born in Barking, Essex, England, on December 14, 1910. He became a naturalized American Citizen, in 1952. He received the B.A. degree, in 1933, specializing in pure and applied mathematics, and the Ph.D. degree, in 1936, specializing in ionospheric physics, both from Cambridge University, England.

In 1935 he became a faculty member of Christ's College, Cambridge, England, in which capacity he conducted research in radio wave propagation and taught applied mathematics and theoretical physics. From 1937 to 1938 he was a Visiting Scientist at the Department of Terrestrial Magnetism of the Carnegie Institute, Washington, D. C. At this time he evolved the theory of propagation of radio waves through a stratified doubly-refracting ionosphere, demonstrating the twisted unsymmetrical paths followed by wave packets. During World War II he was in charge of theoretical research at the Radar Research Establishment of the Royal Air Force, Malvern, England; here, he was involved in the development of new ideas in antennas and propagation. He holds patents on the resonant slot and cosecant antennas, and has conducted radio meteorological investigations. After the war he returned to Cambridge, where he lectured in applied mathematics and conducted ionospheric research. In 1948 he became a Professor of Electrical Engineering at Cornell University, Ithaca, N. Y., where he is now the Director of the School of Electrical Engineering; he is also a member of the faculties of physics and engineering physics. In 1959 he became Associate Director of Cornell's Center for Radiophysics and Space Research.

Dr. Booker is the recipient of the Smith Prize, three Premiums of the Institute of Electrical Engineers (London), and a Guggenheim Fellowship, in 1954, to conduct research on the theory of radio reflections from aurorae and to visit various European scientific institutions. He has been the Chairman of two International Commissions of the International Scientific Radio Union. He is a Fellow of the Royal Meteorological Society and is a member of the American Geophysical Union, the American Astronomical Society, the American Meteorological Society, the AIEE, and the American Association of Physics Teachers. In 1960 he was elected to the National Academy of Sciences.

❖

Virgil E. Bottom (M'47–SM'60) was born in Douglass, Kan., on January 6, 1911. He received the A.B. degree from Friends University, Wichita, Kan., in 1931, the M.S. degree from The University of Michigan, Ann Arbor, in 1938 and the Ph.D. degree from Purdue University, Lafayette, Ind., in 1949.

He was Professor of Physics at three institutions, Friends University, Colorado State University, and McMurry College, Abilene, Tex. During the years 1943 to 1945 he was employed by the Signal Corps, first in OCSIGO, Crystal Section, Washington, D. C., and later at SCEL, Fort Monmouth, N. J. From 1953 to 1958 he was employed by the Semiconductor Products Division of Motorola, Inc., Phoenix, Ariz., where he was engaged in the development of semiconductor devices. From 1948 to 1953 he was a member of the subpanel on Frequency Control Devices of the Research and Development Board, Department of Defense, Washington, D. C. He is presently Professor of Physics at McMurry College.

Dr. Bottom is a member of the American Physical Society, the American Association of Physics Teachers, Sigma Xi, Sigma Pi Sigma and a Fellow of the Texas Academy of Science.

❖

James T. Brothers (A'29–SM'45–F'59) was born in New York, N. Y. on April 30, 1908. He attended the Washington Square College of New York University, and did graduate work at Columbia University, New York, N. Y., the City College of New York, and the University of Detroit, Mich.

From 1926 to 1927 he worked on component research and development at the Polymet Manufacturing Company, New York, N. Y. In 1927 he was employed by Pilot Radio Company, Brooklyn, N. Y. He was with the A. H. Grebe Company, Richmond Hill, N. Y., in 1928, working on ac tubes and receivers, and he also worked as a Consultant on sound on film recording, and on electronic equipment for medical purposes. In that same year he returned to the Pilot Radio and Tube Corporation, where he worked on vacuum tube and component design, sound systems, aircraft, automobile, and home entertainment radio and television transmitters and receivers. From 1930 to 1933 he was a Consultant to licensees on receiver circuitry, tube design, measureurement techniques and evaluation at the RCA Patent Department, New York, N. Y. Since 1933 he has been employed by the Philco Corporation, Philadelphia, Pa., as a Consultant on component and tube design, evaluation and application, circuitry and measurement techniques. He is now in charge of standards and technical liaison at Philco.

Mr. Brothers has been a member of numerous government and industrial committees, Chairman of the Advisory Group on Electronic Parts Working Group on Capacitors, of the Research and Development Board Subpanel on Electro-Mechanical Devices, and of the Electronic Industries Association Safety Committee. He is also a member of the Philadelphia District Council of the American Society for Testing Materials, Chief Technical Advisor and Delegate to the International Electrotechnical Commission Technical Committee No. 12.

❖

Christopher Buff (A'45–M'48–SM'55) was born in Camden, N. J., on January 6, 1917. He took engineering courses at Union Junior College, Cranford, N. J., and at Hofstra College, Hempstead, N. Y.

From 1941 to 1946 he was employed in various engineering capacities by Press Wireless Inc., New York, N. Y., and was instrumental in the development of some of the early FSK equipment. In 1946 he joined Mackay Radio and Telegraph Company, New York, N. Y., as an Engineer engaged in the development of HF communications equipment. Subsequently he was active in systems engineering for the company, making several trips abroad. He has also participated in the work of CCITT study groups at Geneva. He is now Chief Engineer of the parent company, American Cable and Radio Corporation, New York, N. Y. He has been an active licensed radio amateur since 1931.

❖

Charles R. Burrows (A'24–M'38–SM43'–F'43) was born in Detroit, Mich., on June 21, 1902. He received the B.S.E. in electrical engineering and the professional degree of E.E. from The University of Michigan, Ann Arbor, in 1924 and 1935, respectively. He received the A.M. and Ph.D. degrees in physics from Columbia University, New York, N. Y., in 1927 and 1938, respectively.

From 1924 to 1945 he was a member of the technical staff of the Bell Telephone Laboratories working in the research department, specializing in radio wave propagation. He participated in the transatlantic experiments both on long and short waves,

being responsible for the analysis of the latter. He supervised early ultra-short wave propagation experiments which uncovered many of the laws of ultra-short wave propagation. Just prior to the war he worked on the proximity fuse. Early in the war he contributed to the development of radar. Later, as Chairman of the Committee on Propagation of NDRC, he was responsible for all the propagation research affecting the war effort. At the end of the war he accepted the appointment as Professor of Electrical Engineering and Director of the School of Electrical Engineering at Cornell University, Ithaca, N. Y., where he developed research in radio astronomy, radio wave propagation, the ionosphere and vacuum tubes. He is currently Vice President of Engineering and Research at Datronics Engineers, Inc., Bethesda, Md.

Dr. Burrows has been active in URSI, having served many years on the U. S. National Committee as Vice Chairman and Chairman; he was head of the U. S. delegation to the 10th General Assembly of URSI in Sydney, Australia, in 1952. From 1948–1954 he served as International President of Commission II on Tropospheric Propagation. From 1946–1954 he was President of the Joint Commission on Radio-Meteorology of the International Council of Scientific Unions.

Dr. Burrows has received the Presidential Certificate of Merit; and a certificate as a Distinguished Alumnus of the College of Engineering at The University of Michigan. He is a Fellow of the AIEE, and the APS, and a member of the American Geophysical Union, American Astronomical Society, American Rocket Society, Sigma Xi, Tau Beta Pi and Eta Kappa Nu.

❖

Marvin Camras (S'41–A'42–SM'48–F'52) was born in Chicago, Ill., on January 1, 1916. He received the B.S.E.E. degree from the Armour Institute of Technology, Chicago, in 1940, and the M.S. degree from the Illinois Insitute of Technology, Chicago, in 1942.

Since 1940 he has been a staff member of the Armour Research Foundation, Chicago, where he has been concerned with research in the Electronics Division on such projects as remote control, high-speed photography, magnetostriction oscillators, and static electricity.

He has contributed developments which are used in modern magnetic tape and wire recorders, including high-frequency bias, improved recording heads, wire and tape materials, magnetic sound for motion pictures, multitrack tape machines, and binaural sound reproduction.

Mr. Camras is a member of the Acoustical Society of America, AIEE, SMPTE, AAAS, Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. He is presently Editor of the IRE TRANSACTIONS ON AUDIO, and received the John Scott Medal, in 1955, for his work in magnetic recording.

❖

Philip S. Carter (S'48–A'50–M'56), 1896–1961 was born in Glastonbury, Conn., on July 22. He received the A.B. degree in mechanical engineering from Stanford University, Calif., in 1918.

He was employed by the General Electric Company in Dr. E. F. W. Alexanderson's radio laboratory, and was associated with H. H. Beverage in the early experiments leading to the development of the wave antenna. In 1920 he transferred to the Radio Corporation of America and assisted in the design of the VLF antennas at Rocky Point, N. Y. In 1925 he joined the Transmission Laboratory at Rocky Point, where in association with C. W. Hansell and N. E. Lindenblad, he pioneered in the development of long-wire types of antennas for high frequencies. In 1942 the Rocky Point Laboratory became a branch of RCA Laboratories. For over 40 years, he was active in the development of antennas and well known as an antenna expert.

❖

Sheldon S. L. Chang (SM'53–F'61) was born in China, on January 20, 1920. He attended Yenching University, Peking, National Southwest Associated University, Kunming, and Tsinghua University, in Kunming, and received the B.S. and M.S. degrees in physics in 1942 and 1944, respectively. He won a scholarship by competitive examination for studying electrical communication at Purdue University, Lafayette, Ind., in 1945, and received the Ph.D. from Purdue, in 1947.

He was an Instructor at Purdue, from 1947 to 1948, and Research and Development Engineer with Robbins and Myers, Inc., Springfield, Ohio, from 1948 to 1952. He joined the faculty of New York University, in 1952 and is presently a Professor of Electrical Engineering.

Dr. Chang is a member of AIEE, ASEE, the American Physical Society, Eta Kappa Nu, Sigma Xi and Tau Beta Pi.

Colin Cherry was born in St. Albans, England, on June 23, 1914. He received the B.Sc. degree, in 1936, from the University of London, and the M.Sc. and the D.Sc. degrees from London University, in 1940 and 1942.

During World War II he was with the British Telecommunications Research Establishment doing research on radar and in this area was mainly concerned with flying trials. Since 1949 he has been on the faculty of the Electrical Engineering Department of the Imperial College of Science and Technology of London University, England. In 1958 he was appointed a Professor in that University.

Prof. Cherry is a member of the Institution of Electrical Engineers, London.

❖

Harold Chestnut (SM'57) was born in New York, N. Y., on November 25, 1917. He received the B.S.E.E. and the M.S.E.E. degrees from the Massachusetts Institute of Technology, Cambridge, in 1939 and 1940, respectively.

He has been with the General Electric Company since 1940, when he joined the Test Course. In 1942 he supervised the G.E. Electrical B Advanced Engineering Program. Upon completion of this program in 1943, he joined the Aeronautics and Ordnance Systems Division. From 1943 to 1956 he had assignments in AOS in Systems Development, Advanced Development, Guided Missiles, and Marine Systems sections. In 1954 he was appointed Ordnance Engineer, and in 1955, Project Engineer F-104. Since 1956 he has been Control Systems Engineer in the General Engineering Laboratory of General Electric, Schenectady, N. Y.

Mr. Chestnut has been Guest Instructor in the Extension Division of Union College, Schenectady, a member of the Research and Development Board Committees, Chairman of the AIEE Feedback Control Systems Subcommittee on Component Specifications, Vice Chairman of the North American Control Council, first President of the International Federation of Automatic Control, and Chairman of the AACC Control Advisory Committee. He is currently President of the AACC.

❖

Lan Jen Chu (A'39–VA'39–SM'48–F'52) was born in Hweiying, Kiangsu, China, on August 24, 1913. He received the B.S. degree

in electrical power from Chiao Tung University, Shanghai, China, and the S.M. and the Sc.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1935 and 1938, respectively.

He served as a Consultant to the Radiation Laboratory and the Radio Research Laboratory, Harvard University, Cambridge, Mass., from 1942 to 1945, on various electromagnetic problems. He joined the staff of the Radiation Laboratory, M.I.T., in 1942 and during the last years of the war he supervised research and design of many special antennas for radar and communication applications. In 1945 he served as an Expert Consultant to the Secretary of War and was sent to China to head the Advisory Specialist Group to the U. S. Armed Forces in China. He is now Professor of Electrical Engineering at M.I.T. and a staff member of the Research Laboratory of Electronics.

Dr. Chu is a Fellow of the American Physical Society and the Academia Sinica, and a member of Sigma Xi.

❖

James D. Cobine (A'34–SM'44–F'57) was born in Oklahoma City, Okla., on May 10, 1905. He received the B.S.E.E. degree from the University of Wisconsin, Madison, in 1931, and the M.S.E.E. degree from the California Institute of Technology, Pasadena, in 1932. From the latter school he also received the Ph.D. degree in electrical engineering and physics, in 1934.

From 1933 to 1934 he was an Assistant at the California Institute of Technology, and from 1934 to 1935 he was a Graduate Assistant at the Harvard University Graduate School of Engineering, Cambridge, Mass. He remained at Harvard as Instructor, 1935 to 1938, Faculty Instructor, 1938 to 1941, and Assistant Professor, 1941 to 1945. From 1943 to 1945 he was Group Leader of the Harvard Radio Research Laboratory (OSRD). In 1945 he became a Physicist for the General Electric Research Laboratory where he is presently working. He was a member of the Advisory Board of the Research Council of Rutgers University, 1950 to 1959. His field of specialization is electrical discharges in gases, high current arcs and noise in gas discharges.

Dr. Cobine is a Fellow of the American Physical Society and a member of AIEE, NSPE, ASEE, AAAS, Sigma Xi, Tau Beta Pi and Eta Kappa Nu. He has received the Army-Navy Certificate of Appreciation.

Ivan S. Coggeshall (A'26–M'29–F'42–L) was born in Newport, R. I., on September 30, 1896. He was President of the IRE in 1951, and Director, 1941–1944, and 1951–1953. An early wireless amateur, he joined Western Union, New York, N. Y., in 1917. He retired in 1959 as Assistant Vice-President, International Communications.

His activity in the IRE and the AIEE spans 30 years; he is a Fellow in and contributor of papers to both. He wrote the article "Radiotelegraphy" in *Encyclopaedia Britannica*, and the chapter "Wire Telegraphy" in the *Radio Engineering Handbook*. He holds the honorary degree of Doctor of Engineering from Worcester Polytechnic Institute, Mass., which he attended from 1914 to 1917. He retired as Commander, U. S. Naval Reserve, after service in both World Wars. He is now of the Headquarters Staff, AIEE, and a member of IRE's History Committee.

❖

Paul D. Coleman (A'45–M'55) was born in Stoystown, Pa., on June 4, 1918. He received the B.A. degree from Susquehanna University, Selinsgrove, Pa., in 1940, the M.S. degree in physics from Pennsylvania State University, University Park, in 1942, and the Ph.D. degree in physics from the Massachusetts Institute of Technology, Cambridge, in 1951.

He was employed as a Physicist with the U. S. Signal Corps, Dayton, Ohio, and subsequently with the U. S. Air Force at Wright Air Development Center, Dayton, from 1942 to 1946. During this period he worked on electromagnetic theory, and received the AAF Meritorious Civilian Award, in 1946, for his contribution to aircraft antenna theory. From 1946 to 1951 he was a Physicist with the U. S. Air Force Cambridge Air Research Center, Bedford, Mass., and later a Research Associate in Physics at the Research Laboratory of Electronics, M.I.T., where he was concerned with the generation of submillimeter waves.

In 1951 he became an Associate Professor of Electrical Engineering at the University of Illinois, Urbana, where he established the Ultramicrowave Group in the Electrical Engineering Research Laboratory. He is presently a Professor on the graduate electrical engineering staff, teaching and directing research on submillimeter wave generation, detection, and propagation.

Dr. Coleman is Chairman of the Board of Technological Counselors of FXR, Inc., Woodside, N. Y. He is a member of Sigma Xi, the American Physical Society, and Pi Mu Epsilon.

Edmund Harry Cooke-Yarborough (M'45) was born at the village of Campsall near Doncaster, Yorkshire, England, on December 25, 1918. He was educated at Canford School and at Christ Church, Oxford, where he received the M.A. degree in physics, in 1944.

In 1940 he joined the Telecommunications Research Establishment at Dundee, and moved with the Establishment, first to Swanage and then to Malvern, where he worked on airborne radar and radar countermeasures during the war and on weapon guidance and nuclear instrumentation after the war. In 1949 he joined the Electronics Division of the Atomic Energy Research Establishment at Harwell, and became Head of the Electronics Division in 1957.

Mr. Cooke-Yarborough is a member of the IEE and a Fellow of the Institute of Physics.

❖

Stanley F. Danko (M'46) was born in New York, N. Y., on January 6, 1916. He received the B.S. degree in electrical engineering from Cooper Union, New York, N. Y., in 1937, and pursued graduate studies in UHF techniques at the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, and at Rutgers University, New Brunswick, N. J.

He was employed in the U. S. Army Signal Corps Inspection Laboratory in Philadelphia for the period 1940–1946, joining the U. S. Army Signal Research and Development Laboratory at Fort Monmouth, N. J., in 1946. He has been associated with the electronic component development activities at the Signal Research and Development Laboratory, specializing in the development of coaxial cables, RF connectors, inductive components and miniaturization techniques, including printed circuits and packaging of electronic assemblies. He was a principal in the establishment of the dip-soldered printed wiring process (Auto-Sembly) in use today. In 1957 he received a Department of the Army Award for his miniaturization contributions. He is currently Chief of the Electronic Parts and Assemblies Branch of the Electronic Components Department of the U. S. Army Signal Research and Development Laboratory.

Mr. Danko is Secretary of the Professional Group on Military Electronics, an Associate Member of the Working Group on Planning and Requirements of the Advisory Group on Electron Devices, and a registered professional engineer in Pennsylvania.

Paul S. Darnell (SM'52–F'61) was born in Philadelphia, Pa., on November 14, 1901. He received the B.S. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1922, and the M.A. degree from Columbia University, New York, N. Y., in 1925.

In 1922 he joined the Engineering Department of The Western Electric Company, New York, N. Y., which became the Bell Telephone Laboratories in 1925. From that time until 1960 he worked principally in development and design of electronic component parts and devices for both the Bell System and military applications. He was appointed Director of Transmission Apparatus Development in 1953, and in 1960 he became Director of Military Reliability Engineering. He has served on many industrial and government committees. During 1958–1960 he chaired the Ad Hoc Study Group on Parts Specification Management for Reliability, sponsored by the Department of Defense.

❖

Edward E. David, Jr. (A'48–M'56–SM'58) was born in Wilmington, N. C., on January 25, 1925. He received the B.E.E. degree from the Georgia Institute of Technology, Atlanta, in 1945, and the Sc.D. degree from Massachusetts Institute of Technology, Cambridge, in 1950.

During the years 1945 to 1950 he was a Research Assistant at the M.I.T. Research Laboratory of Electronics, Cambridge, working with microwave vacuum tubes and noise theory. He joined the Bell Telephone Laboratories, Murray Hill, N. J., in the Fall of 1950 and was concerned with communication acoustics and underwater sound. He is presently Director of Visual and Acoustics Research and is concerned with the perception and coding of sensory information and with human information processing.

Mr. David was selected by Eta Kappa Nu as one of the country's Outstanding Young Engineers in 1954. In 1958 he received the George W. McCarty Award from the Georgia Institute of Technology as the outstanding young alumnus of the year. He is the author of two books on the physics and physiology of speech communication: "Man's World of Sound," and "Waves and the Ear." He contributed a chapter entitled "Phasing of R-F signals," in "Crossed-Field Microwave Devices." He is a Fellow of the Acoustical Society of America, and a member of Eta Kappa Nu and Sigma Xi.

Glenn L. Dimmick was born in Macon, Mo., on December 7, 1905. He received the B.S.E.E. degree from the University of Missouri, Columbia, in 1928.

Upon graduation he joined General Electric Company as an Engineer on talking motion-picture equipment. In 1930 he was transferred to the RCA Victor Plant in Camden, N. J., as a Development Engineer in the field of sound recording. During the 20 years from 1928 to 1948 he was active in the fields of sound recording, infrared, telephony, monochrome and color television, and high-vacuum evaporation. This work resulted in improved photographic recording methods, optics for sound recording, infrared detection, television, sound powered telephones, recording galvanometers, and magnetic recording devices. In 1948 he was made a Group Manager, and in 1953 he became Section Manager of the Optics, Sound and Special Engineering Section of RCA, working on television and other electronic equipment. Early in 1955 he was appointed Manager of the General Engineering Development Section of Defense Electronic Products; in 1956 he was made Chief Development Engineer, General Engineering Development Section; and in 1957 he became Chief Development Engineer of the Special Systems and Development Department. He is presently on the staff of the Chief Engineer of the Surface Communications Division, Defense Electronic Products.

Mr. Dimmick is a Fellow of the Society of Motion Picture and Television Engineers. His awards include the SMPTE Progress Medal Award in 1941, the RCA Victor Award of Merit in 1949, the Award of Merit for outstanding achievement presented by the Academy of Motion Pictures and Sciences in 1952, and the Missouri Honor Award for Distinguished Service in Engineering in 1954.

❖

Harold E. Dinger (A'27–M'43–SM'43–F'58) was born in Barberton, Ohio, on May 7, 1905. He attended the University of Akron, Ohio, and the University of Maryland, College Park.

From 1929 to 1939 he was Chief Instructor at the McKim Technical Institute, Akron, Ohio, and during 1939 and 1940 was Transmitter Engineer for the Ohio Broadcasting Company. Since 1940 he has been with the U. S. Naval Research Laboratory, Washington, D. C., specializing in radio propagation and radio interference studies.

Mr. Dinger received the Navy's Meritorious Civilian Service Award in 1947 and has received certificates of commendation from the Department of State, the ASA and the RDB. He has been a U. S. delegate to meetings of the IEC and CISPR in Lucerne (1947), Paris (1950), and Philadelphia (1961); the URSI in Sydney (1952), Boulder (1957), and London (1960); and the CCIR in London (1953), Warsaw (1956), Geneva (1958), and Los Angeles (1959). He is a member of the U. S. Executive Committee, CCIR, the U. S. National Committee, URSI, and is a Past Chairman of URSI Commission IV. He is a member of the RESA and AFCEA, and a Fellow of the AIEE, and AAAS. He is Chairman of the IRE-PGRFI for the term 1961–1962.

❖

Robert D. Elbourn (S'40–A'41–M'48) was born in Indianapolis, Ind., on October 5, 1919. He received the B.S.E.E. degree, in 1940, and the M.S.E.E., in 1949, from Purdue University, Lafayette, Ind. In 1940 and 1941 he worked on magnetic recording and electronic instrumentation for C. G. Conn, Ltd. in Elkhart, Ind. From 1941 to 1947 his work in the U. S. Naval Ordnance Laboratory, Washington, D. C., was chiefly concerned with the development of mines and torpedoes, with about one year devoted to the design of an electronic digital computer. In 1947 he became a member of the professional staff of the Electronic Computers Section of the National Bureau of Standards, Boulder, Colo., which became the Data Processing Systems Division, in 1954. Here he continued work on digital magnetic recording, participated in the design of SEAC (Standards Eastern Automatic Computer) and was engaged in research and development for improving circuits and components for high-speed electronic digital computers. He became Chief of the Components and Techniques Section, in October, 1954, and has been supervising research on information retrieval techniques and the design and development of advanced computer components and circuitry.

Mr. Elbourn is a member of the AIEE, American Physical Society, Association for Computing Machinery, Association for Symbolic Logic, Philosophical Society of Washington and the Washington Academy of Sciences. He is also a member of Tau Beta Pi and Eta Kappa Nu.

❖

L. Essen was born in Nottingham, England, on September 6, 1908. He received the B.Sc. degree from Nottingham University, in 1928. Later he earned the London Ph.D. and D.Sc. degrees for published work.

After some post-graduate work on the diffusion of electrons in gases, he joined the National Physical Laboratory, Teddington, Middlesex, England, in 1929 and is now a Deputy Chief Scientific Officer there. He was awarded the Vernon Boys Prize for his work on the velocity of light, the Tompion Gold Medal of the Worshipful Company of Clockmakers for his work on quartz clocks, and the Popov Gold Medal for his work on atomic clocks.

Dr. Essen is a Fellow of the Royal Society of London.

❖

Ronald M. Foster (SM'53–F'54) was born in New York, N. Y., on October 3, 1896. He received the S.B. degree (*summa cum laude*) from Harvard University, Cambridge, Mass., in 1917, and the Sc.D. degree (honorary) from Fairleigh Dickinson University, Rutherford, N. J., in 1960.

He was an Instructor in Mathematics at Harvard University from 1921 to 1922. He joined the research and development department of the American Telephone and Telegraph Company (1921 to 1934), where he developed the Foster Reactence Theorem, published in 1924. This work represented the first major break-through in the synthesis of electrical networks. In 1934 he joined the Bell Telephone Laboratories. He remained there until 1943 when he became Head of the Mathematics Department at the Polytechnic Institute of Brooklyn, N. Y.

He has done notable work on the mathematical basis of antenna arrays, ground return circuits and network theory. In addition to his work on the Foster Reactance Theorem, he is co-author (with Dr. G. A. Campbell) of the Campbell-Foster Tables, and the author of "Fourier Integrals for Practical Application." He resigned his position as Department Head in June, 1961 but continues to teach. He is currently on a year's leave of absence from his teaching duties.

Dr. Foster is a member of Phi Beta Kappa, Sigma Xi, the London Mathematical Society, the Edinburgh Mathematical Society, and a Fellow of the AAAS.

❖

William J. Fry (A'42–M'55) was born in Johnstown, Pa., on April 14, 1918. He received the B.S. and M.S. degrees from Pennsylvania State University, University Park, in 1940 and 1941, respectively.

From 1939 to 1941 he was a Graduate

Assistant in Physics at Penn. State, and from 1941 to 1946 he was a Research Physicist at the Naval Research Laboratory, Washington, D. C. Since 1946 he has been with the University of Illinois, Urbana, first as a Research Assistant Professor, and since 1954 as Research Professor and Head of the Biophysical Research Laboratory, which he established at the University. Included in his personal research are contributions to the following fields: quantitative neuroanatomy, analysis of cardiovascular systems, physical mechanism of the action of ultrasound on biological systems, ultrasonic human neurosurgery, mechanism of acoustical adsorption in tissue, vision mechanisms, reversible changes in the central nervous system produced by ultrasound, design of ultrasonic transducers, and ultrasonic interferometry.

Prof. Fry is a Fellow of the Acoustical Society of America, and a member of the American Physiological Society, the American Academy of Neurology, American Mathematical Society, Mathematical Association of America, Society for General Systems Theory, American Institute of Physics, American Institute of Ultrasonics in Medicine, the Biophysical Society, and Sigma Xi.

❖

Herman Garlan (M'46) was born in New York, N. Y., on December 5, 1907. He received the B.S. degree from the College of the City of New York, in 1929, and the E.E. degree from Columbia University, New York, in 1936.

He has been with the Federal Communications Commission since 1940, starting as a Field Engineer at the Chicago office, where he was concerned with radio station inspections and enforcement matters. In 1945 he was transferred to Washington, D. C., where he worked on regulatory matters dealing with the land mobile communications services, including station licensing, frequency allocations, promulgation of technical standards, and other rule-making activities. He transferred to the Office of the Chief Engineer in 1953, assigned to coordinate and unify the technical standards in the FCC rules. Since 1954 he has been Chief of the RF Devices Branch in the Office of the Chief Engineer where he is responsible for all aspects of the regulatory program in Parts 15 and 18 of the FCC Rules which regulate interference from nonlicensed devices such as diathermy, industrial RF heating equipment, receiver radiation, and other low-power devices.

Mr. Garlan is a member of Tau Beta Pi and Sigma Xi. He has been a member of the Administrative Committee of PGRFI since 1960. He was Vice Chairman of the 2nd and 3rd National Symposia on RFI. He was a member of the U. S. delegation to the 1958 (The Hague, Netherlands) and 1961 (Philadelphia, Pa.) Plenary Sessions of CISPR—the Special International Committee on Radio Interference.

❖

Peter C. Goldmark (A'36–M'38–F'42) was born in Budapest, Hungary, on December 2, 1906. He studied at the Universities of Berlin, Germany, and Vienna, Austria, receiving the Ph.D., in physics from the latter.

In 1936 he joined the Columbia Broadcasting System, Inc., as Chief Television Engineer, later becoming Director of the Research and Development Division. The first practical color television system was developed under his direction in the CBS Laboratories, and on August 27, 1940, the first color broadcast in history was made from the CBS television transmitter in New York. During the war CBS Laboratories were responsible for many military developments in the field of electronic countermeasures and reconnaissance, under his supervision. After the war the long-playing record was developed by Dr. Goldmark and his associates at CBS Laboratories. In 1954 he became President of CBS Laboratories and Vice President of CBS, Inc. He is also a Visiting Professor for Medical Electronics, University of Pennsylvania Medical School, Philadelphia.

Dr. Goldmark is a Fellow of the AIEE, the SMPTE, the Audio Engineering Society, and the British Television Society. In 1945 he was awarded the Television Broadcasters Association Medal for his color television pioneering work. He is the only member of the IRE to hold both the Morris Liebman Memorial Prize (1946), and the Vladimir K. Zworykin Television Prize (1961). In 1960 he received the Achievement Award from the IRE Professional Group on Audio.

❖

John N. Grace (S'51–A'52–M'54) was born in Pittsburgh, Pa., on May 27, 1924. He studied electrical engineering at Carnegie Institute of Technology, Pittsburgh, and received the D.Sc. degree in 1951. Since finishing graduate school he has been employed at Westinghouse Bettis Atomic Power Laboratory, Pittsburgh. He started in Instrumentation and Control and became Supervisor of Reactor Plant Kinetics. In 1955 he became Manager of the Advanced Development Reactor Section and in 1957 was assigned responsibility for S1W, S2W, S3W and S4W Physics. He later was responsible for Physics and Kinetics in the Natural Circulation Reactor Project and presently is Manager of Physics and Kinetics on the Destroyer Project.

Dr. Grace is a member of the American Nuclear Society and AIEE.

❖

Alfred R. Gray (A'27–VA'39–SM'56–F'61) was born in Farmington, N. H., on March 12, 1906. From 1924 to 1948 he studied courses at university extensions; from 1950 to 1960 he attended Massachusetts Institute of Technology, Capital Radio Engineering Institute, and the University of California, for advanced study.

He is President and Chief Engineer of Astronics Reliability, Inc., and Executive Vice President and General Manager of Astronics of Florida, both in Orlando, Fla. He has been an active member of the IRE Professional Group on Production Techniques, now Product Engineering and Production. He was Editor of this Group's TRANSACTIONS from 1954 to 1960. He is presently Chairman of the Administrative Committee. During 1960–1961 he also was Chairman of the Orlando Section of the IRE.

Mr. Gray is a Fellow of the Radio Club of America.

❖

Elliot L. Gruenberg (A'43–SM'50) was born in New York, N. Y., on March 16, 1918. He received the B.E.E. degree from the College of the City of New York, in 1938.

His first work was with the U. S. Army Signal Corps, as Junior Engineer inspecting radio and navigation systems such as instrument landing localizers. In 1944 he joined the Technical Devices Corporation, Roseland, N. J., as Quality Control Engineer. From 1947 to 1951 he was Senior Engineer with the J. H. Bunnell Co., Brooklyn, N. Y., working on the design of military radio receiving and transmitting equipment.

From 1951 to 1958 he was with the W. L. Maxson Corporation. As Manager of Staff Engineering, he directed activity in radar systems, ECM, ECCM, telemetry, and navigation applied to defense and control systems. In 1958 he joined the Federal Systems Division of the IBM Corporation, where he has been Technical Leader on advanced defense concepts studies concerned with ballistic missile defense. He has contributed to studies in radar, communications, and telemetry for space tracking and surveillance, missile defense and control, intelligence, and air traffic control. He is presently Manager of the Space Communications Department of the Communications Systems Center, Rockville, Md.

Mr. Gruenberg holds patents on classified electronic devices and radar systems improvements. He is co-inventor of the FASTAR system. He is presently serving as Editor-in-Chief of the "Handbook on Telemetry and Remote Control," McGraw-Hill Book Company, Inc., New York, N. Y.

❖

Ernst A. Guillemin (A-41–SM'48–F'49) was born in Milwaukee, Wis., on May 8, 1898. He received the B.S. degree in electrical engineering from the University of Wisconsin, Madison, in 1922, and the S.M. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1924. He was awarded the 1924 Saltonstall Traveling Fellowship for study at the University of Munich, Germany, where he received the Ph.D. degree in 1926.

He returned to M.I.T. as an Instructor in 1926, became an Assistant Professor in 1928, Associate Professor in 1936, and Professor of Electrical Communications in 1944.

He was appointed Consultant to the Microwave Committee of the National Defense Research Committee, in 1940. In this capacity, approximately half of his time was devoted to consultation with various groups in the Radiation Laboratory, M.I.T., on a large variety of problems dealing with the design of electrical networks for special applications. One of the results was development of a network for production of radar pulses. In 1941 he also took over the administrative responsibilities of the Communications Option, Department of Electrical Engineering, at M.I.T. In 1960 he was appointed Webster Professor of Electrical Engineering, the chair which he now holds.

In 1948 Dr. Guillemin was awarded the President's Certificate of Merit for his outstanding wartime contributions and in 1960 he was awarded the IRE medal of honor. He is a Fellow of the AIEE and the American Academy of Arts and Sciences, and a member of the American Society for Engineering Education.

Raymond F. Guy (A'25–M'31–F'39) was born in Hartford, Conn., on July 4, 1899. He is an engineering graduate of Pratt Institute, New York, N. Y.

After 42 years with NBC-RCA, New York, N. Y., he retired to enter the field of engineering consultation and representation in AM, FM, TV, and international broadcasting, with headquarters in Haworth, N. J. During the past year he has been Consulting Engineer to the United States Operations Missions in Saigon and Rabat.

Mr. Guy was President of the Institute of Radio Engineers in 1950, the Broadcast Pioneers in 1956, the De Forest Pioneers in 1958 and the Veteran Wireless Operators Association in 1960. He is Secretary of the Broadcast Pioneers, Chairman of the Engineering Subcommittee of the Voice of America, a member of its Industry Advisory Committee, and a Science Consultant for the Agency. He is a Fellow of the AIEE and the Radio Club of America. He was awarded the Marconi Gold Medal by the VWOA, a special plaque and citation by the Broadcast Pioneers, the 1961 Annual Engineering Achievement Award by the National Association of Broadcasters, and was cited by the Radio and Television Executives Society for having the longest continuous experience in broadcasting of any person in the world. For many years he was Chairman of the Engineering Committee of the TV Broadcasters Association and the Engineering Advisory Committee of the NAB. He is Chairman of the IRE Professional Group on Broadcasting.

❖

Raymond A. Heising (A'20–F'23–L) was born in Albert Lea, Minn., on August 10, 1888. He received the E.E. degree from the University of North Dakota, Grand Forks, in 1912, the M.S. degree from the University of Wisconsin, Madison, in 1914, and the honorary D.Sc. degree from the University of North Dakota, in 1947.

He entered the Research Department of the Western Electric Company, New York, N. Y., in 1914 and was one of four men assigned to initiate research in radio. He conducted the first Bell System laboratory work on multiplex carrier on wires. He designed and constructed the first Bell System radio transmitters at Montauk, N. Y., and Washington, D. C., from these, radio telephonic demonstrations were made in April and September, 1915, respectively; the latter involved transcontinental and transoceanic demonstrations. He was responsible for the development of the first practical radio telephone transmitters, which took the form of radio telephone sets for aircraft and naval ships. He independently discovered the *E* and *F* layers and their movements by "radar" type pulse technique. Since retiring from the Bell Telephone Laboratories in 1953, he has been active as a Consulting Engineer.

Dr. Heising received the Morris Liebmann Memorial Prize in 1921 and the Founders Medal in 1957. He has served on many IRE Committees and was President of the IRE in 1939. He promoted the division of work between the Executive Committee and the Board of Directors in 1942. His latest contribution to the IRE was the devising of the Professional Group System.

❖

Keith Henney (A'18–M'26–SM'43–F'43) was born in McComb, Ohio, on October 28, 1896. He received the B.A. degree from Western Reserve University, Cleveland, Ohio, in 1921, and the M.A. degree from Harvard University, Cambridge, Mass., in 1925.

In 1912 he had his first experience with radio via a crystal detector and a two-slide tuner. He had his first experience in publishing in Marion, Ohio, where he was a cub reporter on the anti-Harding daily paper. His radio experience continued by means of rotary spark gaps, Thordarson 1-kw transformers, and glass plate condensers in his amateur station 8ZD. He is still a radio amateur, operating W1QGU in New Hampshire and K2BH in New York.

In 1923 he joined the technical staff of the Western Electric Company. The next five years were spent developing a radio laboratory for Doubleday Doran & Company, publishers of *Radio Broadcast*. In 1929 his first book, "Principles of Radio" was published. It is now in its sixth edition. He became Associate Editor of *Electronics* upon its founding by the McGraw-Hill Book Company, Inc., New York, N. Y., in 1930. He became Managing Editor in 1934; Editor-in-Chief in 1935; Consulting Editor in 1946, and Editor of *Nucleonics* in 1947. He retired from McGraw-Hill, in 1961.

Mr. Henney edited the "Radio Engineering Handbook" (1933) now in its fifth edition; published "Electron Tubes in Industry" (1934) and "Color Photography for the Amateur" (1938), and was co-editor of the "Handbook of Photography" (1939). He is a member of the Harvard Club of New York City, a Fellow and Past President of the Radio Club of America, and an Associate of the Photographic Society of America. He was a Director of the IRE from 1945 to 1948. He was awarded a plaque at the 1944 Rochester Fall Meeting for "His many years of unselfish service to the radio and electronic industries through the technical press."

Karl G. Hernqvist (SM'60) was born in Boras, Sweden, on September 19, 1922. He received the E.E. and the Licentiate of Technology degrees from the Royal Institute of Technology, Stockholm, Sweden, in 1945 and 1951, respectively, and the Ph.D. degree from the same institution, in 1959.

He was an Assistant Engineer in the Royal Swedish Air Force during 1945 and 1946, where he worked in the fields of radar and microwave instrumentation. From 1946 to 1952 he was concerned with microwave measurements and electron tube research with the Research Institute of National Defense, Stockholm. During 1948 and 1949 he was a trainee under the sponsorship of the American-Scandinavian Foundation at RCA Laboratories, Princeton, N. J., where he performed research on high-density electron guns. He has been a member of the technical staff of the Electronic Research Laboratory at RCA Laboratories since 1952, where he has specialized in the gaseous electronics field. During the past several years he has headed a group working on thermionic converters.

Dr. Hernqvist is a member of the American Physical Society and Sigma Xi.

❖

Edward W. Herold (A'30–M'38–SM'43–F'48) was born in New York, N. Y., on October 15, 1907. He received the B.Sc. degree in physics from the University of Virginia, Charlottesville, in 1930, the M.Sc. degree from the Polytechnic Institute of Brooklyn, N. Y., in 1942, and an honorary D.Sc. degree, also from the Polytechnic Institute of Brooklyn, in 1961.

From 1924 to 1926 he worked at Bell Telephone Laboratories, and from 1927 to 1929 with E. T. Cunningham, Inc. He was with RCA at Harrison, N. J., from 1930 to 1942, and from 1942 to 1959 at RCA Laboratories, Princeton, N. J., where he became Director of the Electronic Research Laboratory. In 1959 he joined Varian Associates, Palo Alto, Calif., as Vice President of Research. He has made his career in research and development, specializing in electron tubes, semiconductor physics and devices, signal-to-noise problems, and plasma physics.

Dr. Herold is a member of Phi Beta Kappa and Sigma Xi, and has been active in IRE affairs, including membership on the Board of Directors, from 1956 to 1958.

❖

Julia F. Herrick (M'46) was born in North St. Paul, Minn., on September 14, 1893. She received the B.A. and M.A. degrees from the

University of Minnesota, Minneapolis, in 1915 and 1919, respectively. After several years of high school and college teaching, she returned to the University of Minnesota where she received the Ph.D. degree in biophysics, in 1931.

She was appointed a member of the staff of the Mayo Clinic and Mayo Foundation, Rochester, Minn., as a Consultant in Biophysics at the Institute of Experimental Medicine in 1936. She became an Assistant Professor of Physiology at the Mayo Foundation in 1938, an Associate Professor in 1945, and Professor in 1958. Her designation was later changed to Professor of Biophysics.

Her studies of blood flow in various mammalian blood vessels have resulted in some modifications in the Rein thermostromuhr which extended the use of this instrument to the measurement of the flow of blood under more normal conditions.

In 1942, at the request of the War Department, Dr. Herrick joined the Signal Corps Engineering Laboratories, Fort Monmouth, N. J., where her work was concerned largely with radio direction-finding. Returning to the Mayo Foundation in 1946, her research activities were devoted to the biologic effects of microwaves and ultrasound, physiologic thermometry, and the circulation of the blood. In 1959 she joined the Cardiovascular Laboratory, Department of Medicine, University of Wisconsin, Madison, as a Research Associate, and since 1960, she has been a Senior Scientist at the Vista Laboratory, California Institute of Technology, Pasadena.

Dr. Herrick has worked extensively on the development of an ultrasonic flowmeter, and with the assistance of electronic engineers and physicists, a first model of the flowmeter has been constructed. The instrument is presently being used for measuring the velocity of blood outside the natural circulation in the body; further development is necessary for measurements in the living body.

Dr. Herrick is a Fellow of the New York Academy of Sciences, and a member of the American Physiological Society, the American Physical Society, AAAS, Sigma Xi, Phi Beta Kappa, Alpha Gamma Delta, and Mortar Board. She has been chairman of the IRE Professional Groups on Medical Electronics and Ultrasonics Engineering, and is former Editor of IRE TRANSACTIONS ON MEDICAL ELECTRONICS.

❖

Ralph E. Hiatt (M'47–SM'58) was born in Portland, Ind., on April 12, 1910. He received the B.A. degree in physics from Indiana Central College, Indianapolis, in 1932, and the M.A. degree in physics from Indiana University, Bloomington, in 1939. He later took several graduate courses in mathematics

at Boston University, Mass.

He taught mathematics and science in Indiana public schools until 1942. From 1942 to 1945 he was employed by the Radiation Laboratory, Massachusetts Institute of Technology, Cambridge. As a Staff Member in the Radiation Laboratory Antenna Group, he worked on microwave radar antennas; he organized the Ipswich Antenna Test Station and was Chief of this installation in 1944 and 1945. He joined the Air Force Cambridge Research Center in 1945 and was Chief of the Ground Antenna Section until 1956. From 1956 until 1958 he was Chief of the Antenna Laboratory (later the Electromagnetic Radiation Laboratory). In 1958 he joined the Radiation Laboratory of The University of Michigan, Ann Arbor, as Associate Head of the laboratory and he organized an experimental group in the laboratory. He became Head of the Radiation Laboratory in 1961. He is currently engaged in work on radar scattering, in antenna research, and in experimental plasma studies.

Mr. Hiatt is a member of the American Physical Society, of Sigma Xi, and the American Institute of Physics. He was President of the AFCRC Branch of the Research Society of America in 1957–1958, and is listed in *American Men of Science*.

❖

John K. Hilliard (A'25–M'29–SM'43–F'52) was born in Wyndmere, N. Dak., on October 22, 1901. He received the B.S. degree in physics from Hamline University, St. Paul, Minn., in 1925, and did graduate work in electrical engineering at the University of Minnesota, Minneapolis. He received the D.Sc. degree in engineering from Hollywood University, Calif., in 1951.

He was with the Metro Goldwyn Mayer Studios, Culver City, Calif., for 14 years, where he was engaged in the development of recording and reproducing film and tape equipment, as well as designing microphones and loudspeakers for theaters. He worked then at the Radiation Laboratory of the Massachusetts Institute of Technology, Cambridge, as Project Engineer. For the past 17 years he has been with the Altec Lansing Corporation, Anaheim, Calif., where he is concerned with transducers and communication equipment. His present position is Vice President and Director of the Ling-Altec Research Division.

Dr. Hilliard is a Fellow of the Acoustical Society of America, the Audio Society, SMPTE, and is a member of AIEE, Eta Kappa Nu, the Institute of Environmental Engineers, and of the Armed Forces Committee on Hearing Bioacoustics. He is Acoustic Consultant to the Brain Institute, UCLA Medical School.

❖

Albert S. Hoagland (S'50–A'54–SM'57) was born in Berkeley, Calif., on September 13, 1926. He received the B.S. degree, in 1947, the M.S. degree, in 1948, and the Ph.D. degree, in 1954, all in electrical engineering, from the University of California at Berkeley.

He was Project Engineer on the California Digital Computer Program from 1948 to 1950. From 1954 to 1956, he was a Consultant for IBM, San Jose, and Assistant Professor at the University of California, teaching senior and graduate courses in switching theory and computers. He joined the staff of the IBM Research Laboratory, San Jose, Calif., in 1956, where he was a Senior Engineer in charge of advanced mass memory development. In 1959 he was appointed Manager of the Engineering Sciences Department. During 1960 he was on special assignment to the random access memory product development group in San Jose and is now a Systems Manager in the Advanced Systems Development Division.

Dr. Hoagland is a member of AIEE, Phi Beta Kappa, Sigma Xi, Eta Kappa Nu, and Tau Beta Pi. He is a registered professional engineer in the state of California.

❖

Conrad H. Hoeppner (SM'47–F'58) was born in Spooner, Wis., on March 12, 1918. He received the B.S.E.E. degree, in 1939, and M.S.E.E. degree, in 1940, both from the University of Wisconsin, Madison. He did graduate work at the Massachusetts Institute of Technology, Cambridge, and received the professional E.E. degree, in 1947, from the University of Wisconsin.

Thereafter, he was employed by the U. S. Naval Research Laboratories, Washington, D. C.; he was Director of the Electronics Laboratory, Glenn L. Martin Co., Baltimore, Md.; Director of the Engineering Products Department, Raytheon Manufacturing Company, Waltham, Mass.; Vice President and Director of General Electronic Laboratories, Manager of the Electronics Division of the W. L. Maxson Corporation, New York, N. Y.; Manager of the Development Division, Stavid Engineering

Plainfield, N. J., and Chief Scientist and Director of Research at Radiation, Inc., Melbourne, Fla. He is presently President of the Electronetics Corporation, Melbourne, Fla.

Mr. Hoeppner is a member of Sigma Xi, Pi Mu Epsilon, Phi Kappa Phi, Phi Eta Sigma, and Eta Kappa Nu, and is a Fellow of Tau Beta Pi and the Wisconsin Alumni Research Institute. He also served with the Research and Development Board of the Department of Defense and he holds over forty patents.

❖

Linwood S. Howeth, Captain, USN (Ret.), was born in Hurlock, Md., on February 19, 1902. He was appointed Ensign, U. S. Navy, in 1925, and received the B.S. degree from the United States Naval Academy, Annapolis, Md.

Prior to World War II he served in various assignments, afloat and ashore, most of which were in the field of communications. During World War II his major assignment was in communications in the Pacific Theater. Following the war he served in various communications billets in the Office of the Chief of Naval Operations, Washington, D. C. In 1952 he became Deputy Director, Naval Communications, Washington, D. C. He retired from that assignment in 1955. Since retiring he has devoted the major portion of his time to the preparation of "A History of Communication-Electronics in the United States Navy," which is to be published in the near future.

❖

Edward C. Jordan (S'36–A'39–SM'45–F'53) was born in Edmonton, Alberta, Canada, on December 31, 1910. He received the B.Sc. degree, in 1934, and the M.Sc. degree, in 1936, both in electrical engineering, and both from the University of Alberta, Edmonton, and, in 1940, he received the Ph.D. degree from Ohio State University, Columbus.

He was Control Operator at radio station CKUA for seven years, 1928 to 1935, Electrical Engineer for International Nickel Company for 2 years, 1935 to 1937 and since 1940 he has taught electrical engineering at Worcester Polytechnic Institute, Mass., Ohio State University and the University of Illinois, Urbana. At the latter two universities he supervised research on antennas and radio direction finding. Presently he is Professor and Head of the Department of Electrical Engineering at the University of Illinois.

Dr. Jordan is author or co-author of four books on antennas and wave propagation, and radio and electronics. He is a member of the USA National Committee of URSI and Past-Chairman of Commission VI on Radio Waves and Circuits. He is Past Chairman of the IRE Professional Group on Antennas and Propagation.

❖

John E. Karlin (M'56) was born in Johannesburg, Transvaal, South Africa, on February 28, 1918. He received the B.A. and M.A. degrees from Capetown University, in 1938 and 1939, respectively, and the Ph.D. degree from the University of Chicago, Ill., in 1942. His formal education emphasized both electrical engineering and experimental psychology.

From 1942 to 1945 he was engaged in military communications research at Harvard University's Psycho-Acoustic Laboratory, Cambridge, Mass. In 1945 he joined the technical staff of Bell Telephone Laboratories, Murray Hill, N. J. to study psychophysical problems concerned with transmission. Since 1952 he has been Head of the Human Factors Research Department studying communication properties of man from the viewpoint of man-machine coupling in communication systems.

Dr. Karlin is a member of the Acoustical Society of America and the American Psychological Association.

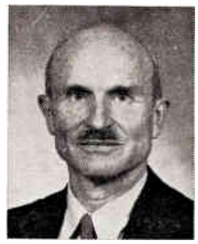
❖

Arthur C. Keller A'36–VA'39–SM'51–F'61) was born in New York, N. Y., on August 18, 1901. He received the B.S. degree, in 1923, from Cooper Union, New York, N. Y., the M.S. degree, in 1925, from Yale University, New Haven, Conn., and did graduate work at Columbia University, New York, N. Y., from 1925–1929.

He is at present Director of Switching Apparatus Development at Bell Telephone Laboratories Inc., New York, N. Y. Previously he was Director of Component Development and Director of Switching Systems Development. His interests have been in the fields of electromechanical devices, sound recording and reproducing equipment, sonar, switching apparatus, electronic heating, sputtering equipment and complete telephone systems. For contributions in World War II he received two U. S. Navy citations for his work in sonar systems and devices. At present he is a Consultant for the Department of Defense (Research and Engineering) and the Chairman of a Working Group.

Mr. Keller is a member of the AIEE, American Physical Society, Acoustical Society of America, SMPTE, Yale Engineering Association and Society for Experimental Stress Analysis, and is a licensed professional engineer in New York State. He is listed in "American Men of Science," "Who's Who in the East," "Who's Who in Engineering," etc.

❖



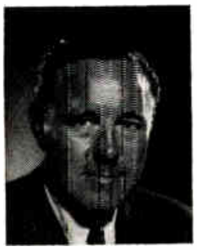
Ronold W. P. King (A'30–SM'43–F'53) was born in Williamstown, Mass., on September 19, 1905. He received the A.B. degree, in 1927, and the S.M. degree, in 1929, both from the University of Rochester, N. Y., where he majored in physics. He was awarded the Ph.D. degree by the University of Wisconsin, Madison, in 1932, after having done graduate work at the University of Munich, Germany, and Cornell University, Ithaca, N. Y.

He served as Teaching and Research Assistant at the University of Wisconsin in 1932–1934 and as Instructor and Assistant Professor of Physics at Lafayette College, Easton, Pa., in 1934–1937. The year 1937–1938 he spent in Germany as a Guggenheim Fellow. In 1938 he joined the faculty of Harvard University, Cambridge, Mass., where he advanced to the rank of Professor in 1946. He is now Gordon McKay Professor of Applied Physics at Harvard University. In 1958 he studied and traveled abroad as a Guggenheim Fellow. His research has been primarily in the field of antennas, transmission lines, and microwave circuits.

Dr. King is a Fellow of the American Physical Society and the American Academy of Arts and Sciences, a corresponding member of the Bavarian Academy of Sciences, and a member of the American Association of University Professors and the AAAS. He is also a member of Phi Beta Kappa and Sigma Xi.

❖

Winston E. Kock (S'45–F'52) was born in Cincinnati, Ohio, on December 5, 1909. He received the E.E. degree and the M.S. degree from the University of Cincinnati, in 1932 and 1933, respectively, and the



Ph.D. degree in physics from the University of Berlin, Germany, in 1934. He was a Teaching Fellow at the University of Cincinnati from 1934 to 1935, attended the Institute for Advanced Study, Princeton, N. J., from 1935 to 1936, and was a Fellow of the Indian Institute of Science at Bangalore, in 1936.

In 1936 he became Research Engineer and later Director of Electronic Research at the Baldwin Piano Company. In 1942 he joined the Radio Research Department of the Bell Telephone Laboratories, where he conducted research on microwave antennas. He was appointed Director of Acoustics Research in 1951, in which capacity he directed the research on the Navy Jezebel-Caesar project, and in 1956, he became Director of Audio and Video Systems Research. He joined the Systems Division of the Bendix Corporation, Ann Arbor, Mich., as Chief Scientist in late 1956, and became Director and General Manager of the Research Laboratories Division, Southfield, Mich., in January, 1958.

In 1958 Dr. Kock received the award of Outstanding Young Electrical Engineer from Eta Kappa Nu, and in 1952 he was awarded the honorary degree of Doctor of Science by the University of Cincinnati. He is a Fellow of the American Physical Society and the Acoustical Society of America, and a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi. He is also a member of the Governing Board of the American Institute of Physics and the Board of Directors of Eta Kappa Nu.

❖

Granino A. Korn was born in Berlin, Germany, in 1922. He received the B.S. degree in physics and mathematics, from Brown University, Providence, R. I., in 1942, and the M.A. and Ph.D. degrees in physics from Columbia University, New York, N. Y., and Brown, in 1943 and 1948, respectively.

After wartime service in the U. S. Navy, he was successively a Project Engineer with Sperry Gyroscope, Great Neck, N. Y., Head of the Analysis Group at Curtiss-Wright, Columbus, Ohio, a Staff Engineer with Lockheed Aircraft Company, Burbank, Calif., and a Consultant under his own name. Since 1957 he has been a Professor of Electrical Engineering at the University of Arizona, Tucson, where he teaches courses on electronic computers and communications theory.

Dr. Korn is a member of Sigma Xi and Simulation Councils, Inc.

Ezra S. Krendel (SM'58) was born in New York, N. Y., on March 5, 1925. He received the B.S. degree from Brooklyn College, N. Y., in 1945, the S.M. degree in physics from the Massachusetts Institute of Technology, Cambridge, in 1947, and the M.A. degree in social relations from Harvard University, Cambridge, Mass., in 1949.

He has been with the Franklin Institute Laboratories, Philadelphia, Pa., since 1949, and presently he is the Manager of the Engineering Psychology Laboratory. His professional experience has included diverse operation research and human factor studies including human dynamics, EEG analysis, visual research, human mechanical power, communications and information flow in large man-machine systems, and the detailed analysis of many weapon systems. With D. T. McRuer, he was awarded The Franklin Institute's Louis E. Levy Medal for a quantitative engineering description of human dynamic behavior (1960).

Mr. Krendel is a member of the American Psychological Association, Operations Research Society of America, the Ergonomics Research Society, the Committee E-18 of the ASTM, and the IRE-PGHFE Administrative Committee.

❖

Benjamin Lax was born in Hungary, on December 29, 1915, and came to the U. S. in August, 1926. He ceived the B.S. degree in mechanical engineering, in 1941, from Cooper Union, New York, N. Y., and the Ph.D. degree in physics, in 1949, from the Massachusetts Institute of Technology, Cambridge.

After entering the U. S. Army in 1942, he went through OCS, and then through radar school at Harvard University, Cambridge, Mass., and the Massachusetts Institute of Technology, Cambridge. He was a Radar Officer assigned to M.I.T. Radiation Laboratory from 1944 to 1946, when he became a Consultant for Sylvania Electric Company. From 1949 to 1951 he carried out research in microwave gas discharge for the Geophysical Directorate of Cambridge Research Center. In November, 1951 he joined the Solid State Group at Lincoln Laboratory, where he became head of the Ferrites Group, in 1953, and of the Solid State Group, in 1955. In 1957 he became Associate Head of the Communications Division at Lincoln, in charge of solid-state physics in several laboratory groups, and was appointed Head of the Solid State Division when it was established, in 1958. He has also been appointed Director of the new M.I.T. National Magnet Laboratory, as of July, 1960.

Dr. Lax received the 1960 Oliver E. Buckley Prize from the American Physical Society for his fundamental contributions to microwave and infrared spectroscopy of semiconductors. He is a Fellow of the American Physical Society, and has been a member of the Executive Committee of the Solid-State Division for the American Physical Society. He is a member of Sigma Xi, of the Advisory Panel for Solid-State Science of the AF Office of Scientific Research, and of the Advisory Committee of the National Bureau of Standards. He was Associate Editor of the *Journal of Applied Physics* from 1957 to 1959, and is at present Associate Editor of the *Physical Review* and of the *Microwave Journal*.

❖

Arthur W. Lo (S'43–A'50–SM'56) was born in Shanghai, China, on May 21, 1916. He received the B.S. degree in physics from Yenching University, China, in 1938, the M.A. degree in physics from Oberlin College, Ohio, in 1946, and the Ph.D. degree in electrical engineering from the University of Illinois, Urbana, in 1949.

He taught physics and electronics in several colleges before joining the Radio Corporation of America, in 1951. At RCA Laboratories he developed a number of basic techniques on using solid-state devices for logic and memory. He was in charge of a switching research group responsible for developing various basic millimicrosecond switching techniques using parametric phase-locked oscillators and tunnel diodes, and for developing a unified view on digital components and circuits. He joined the International Business Machines Corporation, Poughkeepsie, N. Y., in 1960, and served as Manager, Advanced Technical Development, in the Data Systems Division. He is presently in charge of Exploratory Digital Technology in the Components Division.

In his field of work Dr. Lo has contributed a number of technical papers and lectures, and U. S. and foreign patents. He has also co-authored a textbook, "Transistor Electronics." He is a member of Sigma Xi.

❖

Harold W. Lord was born in Eureka, Calif., on August 20, 1905. He received the B.S. degree in electrical engineering, in 1926, from the California Institute of Technology, Pasadena.

Since 1926 he has been employed in several Engineering Departments and Laboratories of the General Electric Company doing research and advanced development work in various fields of electronic and magnetic circuitry and combinations thereof, including radar pulse modulators, television sweep circuits, magnetic amplifiers, broad-band transformers, ballasts for arc discharge devices and high-speed transient studies of arc ignition. For the invention of thyratron synchronous timing circuits for resistance welders, he received the Coffin Award in 1933. He is presently in the Applied Physics Section of the General Electric Research Laboratories, Schenectady, N. Y., and is also a Consultant to the U. S. Department of Defense.

Mr. Lord is a Fellow of the AIEE.

❖

Lee B. Lusted (A'45–SM'53–F'59) was born in Mason City, Iowa, on May 22, 1922. He received the B.A. degree from Cornell College, Mount Vernon, Iowa, in 1943, with a major in mathematics and physics.

Following graduation he joined the Radio Research Laboratory, Harvard University, Cambridge, Mass., as a Special Research Fellow to work on radar countermeasures. In 1944, at the request of General Eisenhower, he went to the European Theatre of Operations as a member of the Petticoat Mission to supervise the installation of countermeasures equipment on ships of the British Home Fleet prior to D-Day.

Following World War II he entered the Harvard Medical School and received the M.D. degree, in 1950. He took specialty training in radiology at the University of California, San Francisco, where he was Assistant Professor of Radiology until 1956, when he joined the Diagnostic X-ray Department at the National Institutes of Health. At the present time he is Professor of Biomedical Engineering and Associate Professor of Radiology at the University of Rochester, N. Y.

He was Chairman of IRE-PGME, in 1956–1957, and he has been Editor of the IRE Transactions on Bio-Medical Electronics, since 1959. In 1960 and 1961 he was a United States representative on the Council of the International Federation for Medical Electronics and is now a member of the Executive Committee. He is Chairman of the National Institutes of Health Advisory Committee on Computers in Research and a member of the National Academy of Sciences-National Research Council Committee on the Use of Electronic Computers in the Life Sciences.

Dr. Lusted is a Fellow of the New York Academy of Sciences and the AAAS. He is a Diplomate of the American Board of Radiology and a member of the American Roentgen Ray Society, the Radiological Society of North America and Sigma Xi.

R. Stuart Mackay (SM'56) was born in San Francisco, Calif., on January 3, 1924. He received the B.A. and Ph.D. degrees in physics from the University of California at Berkeley, in 1944 and 1952, respectively.

He was a Teaching Assistant in both physics and electrical engineering at the University of California from 1944 to 1949. He was responsible for the University's first electron microscope, 1946–1951. From 1949 to 1957, he was Lecturer and then Assistant Professor of Electrical Engineering. During this period he organized some of the first courses in electronic pulse techniques and in electron microscopy. At the University of California Medical Center, San Francisco, he organized and was Director of the Research and Development Laboratory, from 1954 to 1958. He was also Lecturer in Biophysics and Associate Research Biophysicist, from 1954 to 1957. In the summer of 1956 he was a Deputy Project Officer at the AEC tests in the Marshall Islands.

During 1957–1958, he was a Guggenheim Fellow at the Karolinska Institute, Stockholm, Sweden, where later he was Visiting Associate Professor of Medical Physics in the summer of 1959. From 1958 to 1960 he was Associate Clinical Professor of Experimental Radiology and Associate Research Physicist in the Radiology Department of the UC Medical Center. From 1960 to 1961 he was Visiting Professor of Biophysics at the University of Cairo, Egypt, and also Fulbright Lecturer. He is presently Associate Clinical Professor of Optometry and Associate Research Biophysicist at University of California at Berkeley.

Dr. Mackay is a member of the Mine Advisory Committee of the National Research Council of the National Academy of Sciences, and of Phi Beta Kappa, Sigma Xi, and Pi Mu Epsilon. He is former Vice Chairman of the IRE Professional Group on Bio-Medical Electronics.

❖

Laurence A. Manning (S'43–A'45–SM'55) was born in Palo Alto, Calif., on April 28, 1923. He received the A.B. degree in electrical engineering from Stanford University, Calif., in 1944, and the M.Sc. and Ph.D. degrees, in 1948 and 1949, respectively.

He joined the war-time Radio Research Laboratory at Harvard University, Cambridge, Mass., as a Special Research Associate. He is now a Professor of Electrical Engineering at Stanford. His principal interests are in the fields of ionospheric radio propagation and engineering education.

Dr. Manning is a member of Sigma Xi, Phi Beta Kappa, the American Geophysical Union, the American Meteorological Society, and the U.S.A. National Committee of the International Scientific Radio Union.

❖

Jesse Marsten (M'25–SM'43–F'57) was born in Hungary, on January 1, 1897. He received the B.S. degree, in 1917, from the College of the City of New York, N. Y.

From 1917 to 1919 he was an Assistant Engineer at the Marconi Wireless Telegraph Company of America, Aldene, N. J. He then joined the Radio Corporation of America, New York, N. Y., as a Broadcast Transmission Engineer. He was Chief Engineer at the Freed-Eisemann Radio Corporation, Brooklyn, N. Y., during the years 1925 to 1930. In 1930 he joined the International Resistance Company, Philadelphia, Pa., where he was Vice President and then Senior Vice President. He retired from this firm in 1961, but continues to serve as a Consultant.

❖

Warren P. Mason (A'36–F'41) was born in Colorado Springs, Colo., on September 28, 1900. He received the B.S.E.E. degree from the University of Kansas, Lawrence, in 1921, and the M.S. and Ph.D. degrees from Columbia University, New York, N. Y., in 1924 and 1928, respectively.

He has been principally engaged in investigating the properties of solids. He has also studied wear, fatigue in metals, and the joining of materials in solderless wrapped connections and thermocompression bonds. He is in charge of the Applied Mechanics Research group of the Bell Telephone Laboratories, Murray Hill, N. J.

Dr. Mason is a Fellow of the American Physical Society and of the Acoustical Society of America. He is a member of the Rheological Society and of Sigma Xi and Tau Beta Pi.

❖

Frank Massa (A'30–VA'39–M'55–F'59) was born in Boston, Mass., on April 10, 1906. He received both the B.S. degree in electrical engineering and the M.Sc. degree from the Massachusetts Institute of Technology, Cambridge, in 1927 and 1928, respectively. From 1927 to 1928, he was a Swope Fellow at M.I.T.

From 1928 to 1940 he was employed by the Victor Talking Machine Company, Camden, N. J., where he did pioneering work in electroacoustics, including the creation of earliest acoustic test procedures and the initial basic development of many types of microphones, loudspeakers, phonograph pickups, etc., widely used in radio, phonograph and sound-motion-picture reproduction. He was associated with the Brush Development Company, Cleveland, Ohio, from 1940 to 1945, where he was Director of Acoustical Engineering in charge of basic research and development of sonar transducers. He was personally responsible for the development of more than fifty new underwater transducers used during World War II in connection with anti-submarine warfare and acoustic mine equipments. During the years 1945 to 1958 he was President and Director of Engineering at Massa Laboratories, Inc., Hingham, Mass. There he supervised the development of numerous electroacoustic products and instruments, including sonar transducers for use in acoustically guided torpedoes, acoustic mines, antisubmarine warfare applications, and other electroacoustic products for industrial applications. In 1958 he became President of the Massa Division, Cohu Electronics, Inc., Hingham, Mass., which is continuing the work of Massa Laboratories. He is also Vice President and Director of Cohu Electronics, Inc. During World War II, he served under U. S. Navy auspices on several scientific committees dealing with undersea warfare, hydrophonic standardization, calibration of hydrophones, underwater sound measurements, etc. He also served on numerous committees of the American Standards Association and the Acoustical Society of America where he dealt with problems in electroacoustics.

Mr. Massa is a Fellow of the Acoustical Society of America, and a member of AIP. He is the author of textbooks, papers, and patents mostly dealing with electroacoustics.

❖

John G. Mavroides was born in Ipswich, Mass., on December 29, 1922. He received the B.S. degree in electrical engineering from Tufts University, Medford, Mass., in 1944, and the M. S. and Ph.D. degrees in physics, in 1951 and 1953, respectively, from Brown University, Providence, R. I. While at Brown, he was Research Assistant in Ultrasonics, in 1950 to 1951, and R. G. D. Richardson Fellow, in 1951 to 1952.

In 1944 he was commissioned an Aviation Officer in the U. S. Naval Reserve. After attending the Navy Electronics and Radar Schools at Bowdoin College, Brunswick, Me., and the Massachusetts Institute of Technology, Cambridge, he served in the Pacific Theatre as a Technical and Tactical Radar Officer. In 1946 he joined the Transducer Group of the U. S. Navy Underwater Sound Laboratory, New London, Conn., where he was engaged in the research, design and development of solid-state sonic and ultrasonic electromechanical devices for underwater communications and sonar.

Since 1952 he has been a member of the staff of M.I.T. Lincoln Laboratory, Lexington, where he has been concerned with various phases of solid-state research, including semiconductor devices and the investigation of the electronic band structure of solids employing galvanomagnetic, magnetoacoustic resonance and infrared magnetospectroscopy techniques. At present he is leader of the Solid-State Spectroscopy Group.

Dr. Mavroides is a member of the American Physical Society, Tau Beta Pi, and Sigma Xi.

❖

Eleanor M. McElwee (M'51–SM'53) was born in New York, N. Y., on May 15, 1924. She received the B.A. degree in english and mathematics from Ladycliff College, Highland Falls, N. Y., in 1944. She has also taken courses at the Cooper Union Evening School of Engineering and at the Technical Publications Center of Fordham University.

She worked for the Western Electric Company, N. Y., from 1944 to 1947, as an Assistant Engineer of manufacture, and for the Sylvania Product Development Laboratories from 1947 to 1951, first as an Engineering Analyst and then as a Technical Editor. Since 1951 she has been an Engineer in the Commercial Engineering activity of the RCA Electron Tube Division, Harrison, N. J.

Miss McElwee has had several papers published in Sylvania and RCA company magazines and in various IRE publications. She has also taught courses in writing and editing at RCA and at Fordham University. She was the organizer of the IRE Professional Group on Engineering Writing and Speech and has served as Secretary since its inception. She is also a member of the Professional Groups on Education and Engineering Management.

❖

Brockway McMillan (SM'54–F'62) was born in Minneapolis, Minn., on March 30, 1915. He received the B.S. and Ph.D. degrees in mathematics from the Massachusetts Institute of Technology, Cambridge, in 1936 and 1939, respectively.

Prior to going on active duty with the Navy in 1943, he was an Instructor at M.I.T., and a Procter Fellow and H. B. Fine Instructor at Princeton University, N. J. During his tour with the Navy, he served at the U. S. Naval Proving Ground, Dahlgren, Va., and at the Los Alamos Laboratory, N. Mex. In 1946 he joined Bell Telephone Laboratories, Murray Hill, N. J., as a Research Mathematician; in 1955 he became Assistant Director of Systems Engineering, and in 1959 he was named Director of Military Research.

Dr. McMillan has served as occasional Consultant to a number of offices in the Department of Defense, to the Office of Defense Mobilization, to the Weapons Systems Evaluation Group, and with the Office of the President's Special Assistant for Science and Technology. In June, 1961, he was sworn in as Assistant Secretary of the Air Force for Research and Development.

Dr. McMillan is a member of the Society for Industrial and Applied Mathematics (of which he is a Past President), the American Mathematical Society, Mathematical Association of America, Institute of Mathematical Statistics, Operations Research Society, and the AAAS.

❖

Duane T. McRuer (S'47–A'50–SM'54) was born in Bakersfield, Calif., on October 25, 1925. He received the B.S. and M.S.E.E. degrees from the California Institute of Technology, Pasadena, in 1945 and 1948.

Since 1957 he has been President of Systems Technology, Inc., Inglewood, Calif. His previous associations have been with Northrop Aircraft, from 1948 to 1954, and with Control Specialists, Inc., from 1954 to 1957. Professionally, he has covered most aspects of control systems engineering with particular emphasis on the development and design of manual and automatic guidance and flight control systems. With E. S. Krendel, he was awarded, in 1960, the Franklin Institute's Levy Medal for work on the description of human dynamic behavior.

Mr. McRuer is a member of SAE Committee A-18 (Aerospace Vehicle Flight Control, and the IRE-PGHFE Administrative Committee. He is an Associate Fellow of the Institute of Aeronautical Sciences, a senior member of the American Astronomical Society, and a member of the American Rocket Society, the American Physical Society, and the AIEE.

Stewart E. Miller (M'46–SM'53–F'58) was born in Milwaukee, Wis., on September 1, 1918. He received the S.B. and S.M. degrees in electrical engineering from Massachusetts Institute of Technology, Cambridge, in 1941.

Since 1941 he has been employed by the Bell Telephone Laboratories, Red Bank, N. J., and has been concerned with coaxial transmission system development, radar system development, and microwave communication research. At present he is Director of Guided Wave Systems Research. He has served the U. S. Air Force in several advisory capacities.

Mr. Miller is a member of Eta Kappa Nu and Tau Beta Pi and is an associate of Sigma Xi.

❖

Captain Alton B. Moody was born in Thatcher, Ariz., on July 28, 1911. He was graduated from the U.S. Naval Academy, in 1935, and is an experienced air and sea navigator.

He has taught navigation at the Naval Academy, and has written or edited several widely-used books and a number of articles on navigation. He was chairman of a group that in 1961 conducted a long-range scientific study of navigation for NATO. He has received the Thurlow Award for the outstanding contribution to the science of navigation, and other recognitions. In February, 1962 he became Chief, Navigational Systems of the National Aeronautics and Space Administration.

Captain Moody is Past President of the Institute of Navigation and a member of the Board of Governors of the American Polar Society, the New York Academy of Sciences, the AAAS, and the Scientific Research Society of America.

❖

George A. Morton (A'39–SM'46–F'51) was born in New Hartford, N. Y., on March 24, 1903. He received the B.S. degree in electrical engineering, in 1926, the M.S. degree in electrical engineering, in 1928, and the Ph.D. degree in physics, in 1932, all from Massachusetts Institute of Technology, Cambridge.

He was a Research Associate and Instructor at M.I.T., from 1927 to 1933. He joined the Research Laboratory of RCA Manufacturing Company, Camden, N. J., in 1933, as a Research Engineer. He was transferred to RCA Laboratories, Princeton, N. J., in 1941, as a Research Section Head and in 1954 became Associate Director of the Physical and Chemical Research Laboratory. He is now Director of the Conversion Devices Laboratory, Electron Tube Division, RCA Laboratories, Princeton.

During World War II Dr. Morton was a Section Member, National Defense Research, and a member of the AAF Advisory Board, NDRC, RDB Panel IR. He received the IEE Overseas Premium Award and Air Force, Navy, and Army Certificates for War Research. He is the co-author of "Television" (1940, 1954), "Electron Optics and the Electron Microscope" (1945). He is a Fellow of the American Physical Society and a member of Sigma Xi, the AIEE, and the AAAS.

❖

Daniel E. Noble (A'25–SM'44–F'47) was born in Naugatuck, Conn., on October 4, 1901. He received the B.S. degree, in 1929, from Connecticut State College (now the University of Connecticut), Storrs. His Connecticut training was supplemented by graduate work at the Massachusetts Institute of Technology, Cambridge, and by summer work at Harvard University, Cambridge, Mass.

Over a period of seventeen years, he combined the teaching of mathematics and electrical engineering at the University of Connecticut with consulting work and radio and electronic experimentation. He joined Motorola, Inc., in 1940, as Director of Research. Before leaving the University of Connecticut, he pioneered FM radio broadcast and mobile radio development work and, as Consultant to the Connecticut State Police, he was responsible for the design and development of the first statewide two-way mobile radio communications state police system and the first FM system. With Motorola, he advanced over the years from Director of Research to General Manager of the Communications and Electronics Division, to Vice President and Director, and finally, to his present position as Director and Executive Vice President in charge of the three Motorola Divisions dealing with communications, semiconductor, and military electronics products. While Manager of the Communications Division of Motorola, he directed the development of equipment improvements which paced the universal change-over of land-mobile radio systems from AM to FM. He particularly emphasized engineering effort directed to the solution of the problems related to the need for substantial increases in the efficiency of spectrum utilization.

With the announcement of the invention of the junction transistor, his interest turned to the semiconductor and solid-state electronics fields as the most promising areas for corporate expansion. He initiated programs of transistor circuitry research and transistor applications in the Communications and Military Electronics Divisions, and established the Corporate Semiconductor Research Organization, which was later expanded into the Motorola Semiconductor Products Division. He also organized the Solid-State Electronics Department. This new department extended the research into the areas of specialized solid-state devices, integrated circuits, and informational processing and controls. He planned that this department should also serve as the point of dissemination for new solid-state techniques for use in the development of new products and product improvement in all other corporate divisions. He has been particularly active in the organization and coordination of corporate research and development programs concerned with both thin film and morphological, or molecular, integrated circuits. His experience and interests have included not only research and development, which he characterizes as major problems of selection and emphasis, but also the over-all responsibility for the design, mass production, sales, service, and customer relations associated with a complex industrial electronics manufacturing operation.

He has served either as chairman or as a member of many industrial and professional, technical and advisory committees, including the Radio Technical Planning Board (Chairman of Panel 13), and the National Color Television Systems Committee. He holds three patents on electronic devices, including the basic FM squelch patent.

Mr. Noble is a member of the Board of Directors of the Arizona State University Foundation, Chairman of the Foundation Advisory Committee on Engineering and Science Education, and a member of the Army Scientific Advisory Panel. He received the honorary D.Sc. degree from Arizona State College, Tempe, in 1957.

❖

Kenneth A. Norton (A'29–M'38–SM'43–F'43) was born in Rockwell City, Iowa, on February 27, 1907. He received the B.S. degree in physics from the University of Chicago, Ill., in 1928, and continued his studies at Columbia University, New York, N. Y., from 1930 to 1931.

During 1929 he was with the Western Electric Company, Chicago, Ill. From 1929 to 1934 he was in the radio section of the National Bureau of Standards, Washington, D. C. He then joined the technical information section of the Federal Communications Commission, Washington, D. C. from 1934

to 1942, and was responsible for a technical study of clear-channel broadcasting and the initial technical studies leading to the allocation of frequencies to television broadcasting. He was Assistant Director of the operational research group and a Consultant on radio propagation in the Office of the Chief Signal Officer, Washington, D. C., from 1942 to 1943 and from 1944 to 1946. He also served as a radio and tactical countermeasures analyst in the operational research section of the Eighth Air Force in England, 1943 to 1944. Since 1946 he has been in the Central Radio Propagation Laboratory of the National Bureau of Standards where he organized and was Chief of the Frequency Utilization Research Section. At present he is Chief of the Radio Propagation Engineering Division, Boulder, Colo., and is currently concerned with radio guidance systems for missiles and satellites, long-range radio communications systems involving transmission via satellites, tropospheric scatter and propagation at very low frequencies.

He has been a delegate to several international radio conferences, including the Provisional Frequency Board, Geneva, Switzerland, 1948, and the High-Frequency Broadcasting Conference, Mexico City, 1948. He was Vice Chairman of the United States delegation to the 1950 meetings of Television Study Group 11 of the International Radio Consultative Committee (CCIR) in the United States, France, The Netherlands, and the United Kingdom. He was a United States delegate to the Interim Study Group Meetings of the CCIR, Geneva, 1958, and to the IXth Plenary Assembly of the CCIR, Los Angeles, Calif. He was also a delegate to the XIth and XIIIth General Assemblies of the International Scientific Radio Union (URSI), The Hague, The Netherlands, 1954, and London, England, 1960, respectively, as well as Chairman of the Local Arrangements Committee for the XIIth General Assembly, Boulder, Colo., 1957.

Mr. Norton was awarded the Stuart Ballantine medal by the Franklin Institute, for his work on radio propagation and FM and television frequency allocations. In 1960 he received the IRE Harry Diamond Memorial Award, the highest award offered to a government employee in the field of radio and electronics. He was cited for "contributions to the understanding of radio wave propagation." In 1962 he received the Exceptional Service Award of the U. S. Department of Commerce for "outstanding contributions and leadership in the field of radio propagation research." He is a Fellow of the American Physical Society, AAAS, AIEE, and a member of the Scientific Research Society of America, the American Geophysical Union, the American Mathematical Society, the Institute of Mathematical Statistics, and the American Statistical Association.

James D. O'Connell (SM'53–F'57), Lieutenant General, USA (Ret.), was born in Chicago, Ill., on September 25, 1899. He received the B.S. degree, in 1922, from the United States Military Academy, West Point, N. Y. Prior to World War II he served in a variety of troop assignments and commands overseas and in the United States. During this period he attended the Signal School, Northwestern University, Evanston, Ill., Sheffield Scientific School, Yale University, New Haven, Conn. (where he received the M.S. degree in communications engineering); and the Command General Staff School.

During the war he served in Washington, in the North African theater and later in the European theater. Following the war, he served as Commanding Officer, Signal Corps Engineering Laboratories; Signal Officer, Eighth U. S. Army in Japan; Signal Officer, Second U. S. Army; Deputy Chief Signal Officer and Chief Signal Officer, U. S. Army. He holds both foreign and U. S. Army decorations, including the Distinguished Service Medal. He retired in May, 1959, and later that year became a Vice President of General Telephone, and Electronics Laboratories.

General O'Connell is a member of the Joint Technical Advisory Committee of the IRE-EIA, and Chairman, JTAC Ad Hoc Committee 60.2 Space Communications.

❖

B. M. Oliver (S'40–A'40–M'46–SM'53–F'54) was born in Santa Cruz, Calif., on May 27, 1916. He received the B.A. degree in electrical engineering from Stanford University, Calif., in 1935, and the M.S. degree from the California Institute of Technology, Pasadena, in 1936. Following a year of study in Germany under an exchange scholarship, he returned to the California Institute of Technology where he received the Ph.D. degree, in 1940.

From 1939 to 1952 he was employed at the Bell Telephone Laboratories, Murray Hill, N. J., in television research and radar development. He is now Vice President in Charge of Research and Development at the Hewlett-Packard Company, Palo Alto, Calif.

Dr. Oliver is a member of the American Astronautical Society.

❖

Harry F. Olson (A'37–VA'39–SM'48–F'49) was born in Mt. Pleasant, Iowa, on December 28, 1902. He received the B.E.

degree, in 1924, the M.S. degree, in 1925, the Ph.D. degree, in 1928, and the professional E.E. degree, in 1932, all from the University of Iowa, Iowa City. He received the honorary degree of D.Sc., in 1959, from Iowa Wesleyan, Mt. Pleasant.

In 1928 he joined the Radio Corporation of America, and worked in several of their divisions until 1941, when he joined the RCA Laboratories, Princeton, N. J. He is presently the Director of the Acoustical and Electromechanical Laboratory of the RCA Laboratories. From 1939 to 1943 he was a Lecturer in Acoustical Engineering at Columbia University, New York, N. Y.

For his contributions to the field of audio engineering, he received the Modern Pioneer Award of the National Association of Manufacturers, in 1940, the John H. Potts Medal of the Audio Engineering Society, in 1952, the Samuel Warner Medal of the Society of Motion Picture and Television Engineers, in 1955, the John Scott Medal of the City of Philadelphia, in 1956, and the Achievement Award of the IRE Professional Group on Audio, in 1956. In 1959 he was honored by election to the National Academy of Sciences.

Dr. Olson is a Fellow of the American Physical Society, the Acoustical Society of America, the Audio Engineering Society, and SMPTE, and a member of Tau Beta Pi and Sigma Xi. He was President of the Acoustical Society of America, in 1952, and is currently President of the Audio Engineering Society. He has served as Chairman of the Professional Group on Audio (1957–1958), Director of the IRE (1959–1960), and on numerous IRE committees.

❖

Clure H. Owen (A'32–VA'39–M'55) was born in Aldrich, Mo., on June 5, 1904. He received the B.S. and M.S. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1930 and 1936.

From 1927 to 1931 he worked as a Radio Operator at Broadcast Station WSB, Atlanta. He served as Radio Inspector and Radio Engineer with the Federal Communications Commission in Atlanta and Washington, 1931 to 1935. Since 1945 he has been with the American Broadcasting Company, New York, N. Y., as Allocations Engineer and Business Manager of the Engineering Department.

Mr. Owen was on the Administrative Committee of the IRE Professional Group on Broadcasting from its inception until 1960, and was its Chairman in 1958 and 1959.

Alvin L. Pachynski (SM'59), Major General, USAF (Ret.), was born in Chicago, Ill., on October 10, 1904. In 1927 he was appointed Second Lieutenant, Signal Corps, and received the B.S. degree from the United States Military Academy, West Point, N. Y. In 1928 he received the M.S. degree from Yale University, New Haven, Conn.

There followed 12 years of experience in Army communications, primarily in the engineering of fixed radio stations. In 1940 he was assigned to Air Force staff and command positions in which he was variously responsible for the organization, planning, operation, programming and research and development of communications-electronics systems; these culminated with his appointment in 1953 and 1956, respectively, as Air Force Deputy Director and Director of Communications-Electronics. He retired in 1958, and at present is a Consultant.

Major General Pachynski is an honorary Life Member and a National Director of AFCEA.

❖

Robert M. Page (SM'45–F'47) was born in St. Paul, Minn., on June 2, 1903. He received the B.S. degree in physics from Hamline University, St. Paul, in 1927, the M.S. degree in physics from George Washington University, Washington, D. C., in 1932, and was awarded an Honorary Doctor of Science degree by Hamline University, in 1943.

Since 1927 he has been at the Naval Research Laboratory, Washington, D. C., where his work has been precision instrumentation in the field of electronics. In 1934 he completed the first pulse radar in the world for detection of aircraft, for which he received the U. S. Navy Distinguished Civilian Service Award, the Presidential Certificate of Merit, an IRE Fellowship, the Harry Diamond Memorial Award of IRE, and the Stuart Ballantine Medal of the Franklin Institute. In recognition of technical and scientific achievement in research and development for the United States Navy, he has been given the Captain Robert Dexter Conrad Award. In March of 1960 he was presented with the President's Award for Distinguished Federal Civilian Service. He is now Director of Research at the U. S. Naval Research Laboratory.

Mr. Page is a Fellow of the AAAS and the American Scientific Affiliation. He is a member of the Washington Academy of Sciences, the Scientific Research Society of America, and the Armed Forces Electronics and Communications Association.

George W. Patterson (J'31–A'49–SM'54–F'59) was born in Rochester, N. Y., on May 6, 1912. He received the B.S.E.E. degree from the University of Vermont, Burlington, in 1934, the M.A. degree in physics from Columbia University, New York, N. Y., in 1936, and the Ph.D. degree in mathematics from the University of Pennsylvania, Philadelphia, in 1958.

After teaching in public schools (1935–1937), he joined the staff of the National Bureau of Standards, initially on the staff of WWV, 1938 to 1939, and later on the Bat-Pelican project, 1953 to 1956. He received the Naval Ordnance Development Award for distinguished civilian service from the U. S. Navy for his contribution to this pioneer guided missile development. In 1946 he joined the staff of the Moore School of Electrical Engineering of the University of Pennsylvania, where he designed the logic for the EDVAC arithmetic circuits and coordinated the logic design for the entire machine. From 1950 to 1955 he was with the Research Center of the Burroughs Corporation, Paoli, Pa., after which he rejoined the faculty of the Moore School. Currently he is Associate Professor and Section Head for systems logic research there.

In addition to teaching and directing research on switching theory, Prof. Patterson has been active in computer standardization, having been a member of the Electronic Computers Committee since 1949, and currently Vice Chairman. He was also a member of the ACM committee that wrote the first programming glossary, and a member of the ASA Y32-14 task group that developed standards for computer logic symbols. He is now a member of ASA sectional committee X3 on data-processing systems. He is also a member of the American Mathematical Society, the Franklin Institute, Phi Beta Kappa, Sigma Xi, and a Fellow of the AAAS.

❖

Richard L. Petritz (A'54–M'60–SM'62) was born in Rockford, Ill., on October 24, 1922. He received the B.S. and M.S. degrees in electrical engineering, in 1946 and 1947, respectively, and the Ph.D. degree in physics, in 1950, all from Northwestern University, Evanston, Ill.

While completing his graduate studies at Northwestern University where he was a Walter P. Murphy Fellow, from January, 1948 to June, 1949, he held various positions in the fields of physics, electronics and nuclear weapons. From 1950 to 1954 he was a member of the Physics Department of the Catholic University of America, Washington, D. C. Prior to joining Texas Instruments, Incorporated, Dallas, Tex., in 1958, he was Chief of the Semiconductor Research Branch at the U. S. Naval Ordnance Laboratory, White Oak, Md. At present he is Director of the Semiconductor Research Laboratory, Corporate Research and Engineering at Texas Instruments.

Dr. Petritz received the IRE Browder J. Thompson Memorial Prize, in March, 1954, and the U. S. Navy honored him with the Meritorious Civilian Service Award in December, 1954. He is a member of the American Association of Physics Teachers, the American Association of University Professors, the American Physical Society, Eta Kappa Nu, Sigma Xi, Tau Beta Pi, Triangle Fraternity, the Washington Philosophical Society, and a Fellow of the Texas Academy of Science.

❖

John R. Pierce (S'35–A'38–SM'46–F'48) was born in Des Moines, Iowa, on March 27, 1910. He received the B.S. degree, in 1933, and the Ph.D. degree, in 1936, from the California Institute of Technology, Pasadena, and went to the Bell Telephone Laboratories the same year.

From 1936 to 1955 he worked on high-frequency electron tubes, and particularly on traveling-wave tubes. He became Director of Electronics Research in 1952 and is now Executive Director, Research—Communications Principles Division, a field including research in radio, television, electronics, acoustics and vision, mathematics, and group behavior. The Bell Laboratories' work on Project Echo has been carried out in his department.

In 1948 he received the IRE Fellow Award for his "many contributions to the theory and design of vacuum tubes." He was the recipient of the Eta Kappa Nu Outstanding Young Electrical Engineer Award for 1942 and the IRE Morris Liebmann Memorial Prize for 1947. He was Editor of the PROCEEDINGS OF THE IRE 1954–1955. In 1960 he received the Stuart Ballantine Medal.

He is the author of "The Theory and Design of Electron Beams," "Traveling-Wave Tubes," "Electrons, Waves and Messages," "Man's World of Sound" (with E. E. David, Jr.), and "Symbols, Signals and Noise," and in addition has written many popular science articles for various magazines.

Dr. Pierce is a Fellow of the American Physical Society, the Acoustical Society of America, and the British Interplanetary Society, and a member of the National Academy of Sciences.

Leon Podolsky (A'30–M'46–SM'53–F'57) was born in Philadelphia, Pa., on November 3, 1910. He received the E.E. degree from the Drexel Institute of Technology, Philadelphia, in 1934, and he did graduate work in medical physics at Temple University Medical School, Philadelphia, from 1934 to 1937.

His entire career in the electronics industry has been associated with components. He has acted as a Consultant to the Armed Forces and has been chairman of numerous military-industry task forces in the electronic components field, including D.O.D. AGREE Task Force V. He is Chairman of the EIA Component Parts Panel, Chairman of ASA Committee C83, and Chief Technical Advisor on Electronic Components to the U. S. National Committee of the International Electrotechnical Commission. As the Chief U. S. Delegate on Electronic Components, he attended meetings of the International Electrotechnical Commission in Europe and Asia for the years 1952 to 1960. Presently, he is Technical Assistant to the President of the Sprague Electric Company, North Adams, Mass.

Mr. Podolsky is a Fellow of the AIEE and is a registered professional engineer in the states of Massachusetts and New York. He has received numerous government and industry awards for outstanding service and was the recipient of the IRE-EIA Fall Meeting Award in 1957.

❖

Insley B. Pyne (S'49–A'50–M'55) was born in New York, N. Y., on July 26, 1919. He received the A.B. degree in physics and the M.S.E. degree in electrical engineering, both from Princeton University, N. J., in 1948 and 1950, respectively.

From 1951 to 1961 he served on the faculty of the Electrical Engineering Department at Princeton, first as Instructor (1951 to 1952) and later as Assistant Professor (1952 to 1961). In 1961 he joined the staff of the Princeton Computer Center. His work has been in the field of analog and digital computers, and switching theory.

Mr. Pyne is a member of the AIEE and Sigma Xi.

❖

John R. Ragazzini (A'41–M'46–SM'52–F'55) was born in New York, N. Y., on January 3, 1912. He received the B.S. and E.E. degrees from the College of the City of New York, in 1932 and 1933, respectively, and the A.M. and Ph.D. degrees from Columbia University, New York, N. Y., in 1938 and 1941, respectively.

He began his teaching career in the School of Technology at C.C.N.Y., later going to Columbia University where until recently he was Professor of Electrical Engineering and Chairman of that department. In July, 1958, he left Columbia University to become Dean of the College of Engineering at New York University. He directed research at Columbia University during World War II in the area of fire and bombing control by means of electronic simulators. This work led to many of the early developments in electronic analog computers and simulators. His later research work was in the field of control systems and, more specifically, in sampled-data control systems. He has been a Consultant in the fields of computers, control and guidance systems for a number of industrial organizations.

Dr. Ragazzini is a member of Phi Beta Kappa, Tau Beta Pi, Sigma Xi, and Eta Kappa Nu.

❖

Edward G. Ramberg (M'53–SM'53–F'55) was born in Florence, Italy, on June 14, 1907. He attended Reed College, Portland, Ore., and received the B.A. degree from Cornell University, Ithaca, N. Y., in 1928, and the Ph.D. degree in theoretical physics from the University of Munich, Germany, in 1932.

He joined the Electronic Research Laboratory of the RCA Manufacturing Company, Camden, N. J., as a Junior Engineer, in 1935, and was transferred to the RCA Laboratories at the David Sarnoff Research Center, Princeton, N. J., in 1942. Except for three interruptions—in Civilian Public Service, 1943 to 1945, as Visiting Professor at the University of Munich, in 1949, and as Fulbright Lecturer at the Technische Hochschule, Darmstadt, Germany, 1960 to 1961, —he has worked since then as a Research Physicist at the RCA Laboratories in the fields of electron microscopy, electron optics and optics relating to television, theory of thermoelectric refrigeration, and other problems in electron physics.

Dr. Ramberg is a Fellow of the American Physical Society and a member of the Electron Microscope Society of America, the American Association of Physics Teachers, the AAAS, Sigma Xi, and Phi Kappa Phi. In 1959, he was named a Fellow of the Technical Staff of RCA Laboratories.

❖

Simon Ramo was born in Salt Lake City, Utah, in May, 1913. He received the B.S. degree in electrical engineering from the University of Utah, Salt Lake City, in 1933, and the Ph.D. degree in electrical engineering and physics from the California Institute of Technology, Pasadena, in 1936.

At this time he joined General Electric's research staff where his work was in the areas of microwaves, electron optics and other phases of physics and engineering. From 1946 to 1953 he was with the Hughes Aircraft Company, Culver City, Calif., where he first held the positions of Director of Research, Electronics Department, and Director of Guided Missile Research and Development and later the position of Vice President and Director of Operations. He was Executive Vice President of The Ramo-Wooldridge Corporation (which he and Dr. Wooldridge founded in September, 1953) prior to the merger of that Corporation with Thompson Products, Inc. At present, he is Vice Chairman of the Board and Director of Thompson Ramo Wooldridge, Inc., Los Angeles, Calif.

He is a Regents' Lecturer in Engineering at the University of California and a Research Associate at the California Institute of Technology. He was also the W. Rupert Turnbull Lecturer, in 1956, and the Steinmetz Memorial Lecturer, in 1959. He received the honorary degree of Doctor of Engineering, in 1960, from Case Institute of Technology, Cleveland, Ohio, and, in 1961, from the University of Utah.

Dr. Ramo is a Fellow of the AIEE, the American Physical Society, the American Rocket Society, and the Institute of Aeronautical Sciences. He was a recipient of the Eta Kappa Nu Honorable Mention Award, 1941; the IRE Electronic Achievement Award of the Pacific Region, 1953; the Raymond E. Hackett Award, 1955; and the Arnold Air Society Paul T. Johns Trophy, 1961.

❖

Oscar Reed, Jr., (S'39–A'41–SM'47) was born in Washington, D. C., on April 2, 1917. He received the B.E.E. degree from the Catholic University of America, Washington, D. C., in 1939.

He then joined

the firm of Jansky & Bailey, Washington, D. C., where he is presently a Consulting Radio Engineer. He has specialized in broadcast-engineering allocating developments, practice before the FCC, and broadcast-station design and evaluation work.

Mr. Reed is a past member and former Chairman of the Professional Group on Broadcasting, and currently serves as Chairman of the AIEE Technical Committee on Broadcasting. In 1949 and 1950 he represented the U. S. Government as an industry advisor to the North American Regional Broadcast Conference, Montreal, Canada, and Washington, D. C. He is a registered professional engineer in the state of Maryland and in the District of Columbia.

❖

Walther Richter (SM'44–F'53) was born in Heidelberg, Germany, on May 17, 1900. He obtained the B.S. degree in electrical engineering at the Institute of Technology in Karlsruhe, Germany.

He joined the Engineering Department of the A. O. Smith Corporation, in 1925, where he became Director of Electrical Engineering and Research, in 1930. From 1939 to 1943, he engaged in private consulting practice. From 1943 to 1951, he was associated with the Engineering Development Department of the Allis-Chalmers Manufacturing Company, Milwaukee, Wis., where he was in charge of the electronics laboratory of this department. In 1951 he returned to consulting. He is now a Consultant in Electronics to the staffs of the Research Center of the Outboard Marine Corporation, the Electrical Research Laboratory of the A. O. Smith Corporation, and the Development Department of Cutler-Hammer, Inc., Milwaukee, Wis. He taught courses in industrial electronics from 1935 to 1950 in evening sessions at the University of Wisconsin, and from 1954 to 1960 at Marquette University, Milwauee.

Mr. Richter is a Fellow of the AIEE.

❖

Preston Robinson (M'54) was born in Boston, Mass., on June 30, 1902. He received the S.B. degree, in 1922, and the S.M. degree in chemical engineering, in 1923, both from the Massachusetts Institute of Technology, Cambridge. He received the Ph.D. degree in physical chemistry from the University of California at Berkeley, in 1925.

From 1923 to 1925 he was a Teaching Fellow in Chemistry at the University of California, and the following four years he was a Research Metallurgist at Guggenheim Brothers Laboratories, New York, N. Y. Since 1929 he has been with the Sprague Electric Company, North Adams, Mass., first as Director of Research and Engineering, and since 1953 as Consultant. His recent assignments have been as Consultant to the Department of Defense in the capacity of Chairman of the Working Group on Electronic Materials, and Consultant to the Materials Advisory Board of the National Academy of Sciences.

Dr. Robinson is a member of the American Chemical Society, the American Physical Society, Sigma Xi, and several other technical societies.

❖

John D. Ryder (A'29–SM'45–F'52) was born in Columbus, Ohio, on May 8, 1907. He received the B.E.E. degree, in 1928, and the M.S. degree, in 1929, both from the Ohio State University, Columbus, and the Ph.D. degree in electrical engineering, in 1924, from Iowa State University, Ames.

From 1929 to 1931 he was with the General Electric Company, Schenectady, N. Y., in vacuum and gas tube development. In 1931 he joined the Bailey Meter Company, Cleveland, Ohio, as Supervisor of the Electrical and Electronic Section of the Research Laboratory. In 1941 he was appointed Assistant Professor of Electrical Engineering at Iowa State College, was made Professor, in 1944, and in 1947 assumed the position of Assistant Director of the Iowa Engineering Experiment Station. In September, 1949 he became Head of the Department of Electrical Engineering at the University of Illinois, Urbana, and in July, 1954 became Dean of the College of Engineering at Michigan State University, East Lansing.

He was President of the NEC, 1953, President of the IRE, 1955, and President of Eta Kappa Nu, 1956 to 1957, and also served on the Boards of Directors of these organizations. He was the Editor of the IRE, in 1958 and 1959, and a member of the Army Scientific Advisory Panel, 1957 to 1959. He has served since 1957 on the Signal Corps Research and Development Advisory Board, and since 1955 as Advisor to the National Research Council for the Office of Ordnance Research.

Dean Ryder is a Fellow of the AIEE and the AAAS, and a member of Tau Beta Pi, Eta Kappa Nu, Phi Kappa Phi, Sigma Xi, and Pi Mu Epsilon.

❖

Clifford M. Ryerson (SM'55) was born in Brooklyn, N. Y., on September 23, 1915. He received the B.S. degree in physics from Stetson University, De Land, Fla., in 1937,

and the M.S. degree in electronics from Duke University, Durham, N. C., in 1938.

Later, on the technical staff at Duke, he managed three cosmic-ray expeditions, and then won a Meritorious Civilian Service Award for a war-time instrumentation program for the Navy. In 1946 he organized and managed the Applied Physics Branch of the Naval Gun Factory, Washington D. C. In 1951 he became a system designer, and soon advanced to Project Manager of the RCA Moorestown Plant. Expanding his earlier work on product reliability, he then advanced to the main Camden Plant, where he developed the RCA Reliability Engineering Program. He then expanded this effort into a broad program of control for product assurance, which included the management and integration of all control specialties, such as reliability maintainability, value engineering, quality assurance and system effectiveness. He is currently concerned with making consulting and technical audit service in this field available to all industry and to the government. At present he is Vice President of the EL-TEK Divisions of EL-TRONICS, Hawthorne, Calif.

Mr. Ryerson is a registered professional engineer, a senior member of the American Society for Quality Control, an official of the National Reliability Symposia, the organizing chairman of the National Reliability Training Course, and a participant in many military-industry programs, such as AGREE, BIMRAB, PSMR, etc.

❖

Peter C. Sandretto (A'30–M'40–SM'43–F'54) was born in Pont Canavese, Italy, on April 14, 1907. He received the B.S. and E.E. degrees from Purdue University, Lafayette, Ind., in 1930 and 1938, respectively. He also graduated from the Army's Command and Staff School, and performed graduate work at Northwestern University, Evanston, Ill.

He began his more than 30-year association with aeronautical electronics, in 1930, as a member of the technical staff of Bell Telephone Laboratories, where he was engaged in the design of some of the first radio equipment for commercial transports. From 1932 to 1942, while Head of the Communications Laboratory of United Air Lines Transport Corp., he participated in the early air traffic control efforts.

During World War II, he served with the United States Air Force as Assistant Chief of the Radar Division, Headquarters, Army Air Force; as U. S. Signal Liaison Officer in the Air Ministry, London, England; as Chief of the Electronics Test Section, Air Proving Ground Command, Eglin, Fla.; and as Chief of the Electronics Division, Army Air Force's Pacific Ocean Areas, where his staff section was responsible for the establishment of the Air Force's traffic control system for the Twentieth Air Force. He holds the rank of Brigadier General in the U. S. Air Force Reserve.

In 1946 he joined the International Telephone and Telegraph Corporation's Aviation Department and attended many meetings of the International Civil Aviation Organizations, which were concerned with air navigation and traffic control facilities. He has served on committees of the Radio Technical Commission for Aeronautics since its inception and was a member of Special Committee 31 which established the common system for air navigation and traffic control. He became Technical Director and Vice President of IT&T Laboratories, in 1954 and 1956, respectively, and is presently Deputy Executive of IT&T's U. S. Defense Group, Nutley, N. J. He has written and lectured extensively in many parts of the world on the subject of air navigation and traffic control, and, as a consequence, has become internationally recognized for his work in this field. In 1958 he was Honorary (Keynote) speaker at the Ausschuss für Funkortung in Berlin. He is the author of "Principles of Aeronautical Radio Engineering," first published in 1942, and "Electronic Navigation Engineering."

Mr. Sandretto is a member of Eta Kappa Nu and the Institute of Navigation, and an associate member of the IEE.

❖

K. N. Satyendra (SM'61) was born in Mysore, India, on May 19, 1924. He received the B.Sc. (Hons.) degree in physics from the University of Mysore, India, in 1943 and the E.E. degree from the Indian Institute of Science, in 1946. He was awarded gold medals for outstanding work at these two institutions. He received the M.S. and Ph.D. degrees from Illinois Institute of Technology, Chicago, Ill., in 1948 and 1954, respectively.

From 1948 to 1954 he served as an Assistant Professor of Electrical Engineering at Illinois Institute of Technology. During the next six years he was associated with the Westinghouse Electric Corporation, where he was engaged in advanced research and development programs pertaining to fire control systems, air-to-air and air-to-surface missile guidance, radar and infrared detection and space weapon systems. In 1959 he joined Nortronics, Palos Verdes Estates, Calif., where he presently serves as Director of Research; he is responsible for advanced research in various aspects of space electronics. Currently he is very active in the space programs of both NASA and the USAF. He has also been a regular lecturer at UCLA on graduate courses on infrared, space optics, etc.

Dr. Satyendra is a Fellow of the British Interplanetary Society and a member of the American Rocket Society, American Astronautical Society, Institute of Aerospace Sciences, AIEE, Sigma Xi, Tau Beta Pi, and Eta Kappa Nu.

❖

Kurt Schlesinger (A'41–SM'51–F'54) was born in Berlin, Germany, on April 20, 1906. He received the Dipl.-Ing. and Doctor of Engineering degrees, in 1928 and 1929, from the Technical University of Berlin.

From 1941 to 1944 he was with RCA Laboratories, Purdue University, Lafayette, Ind., and from 1944 through 1947, he was Consulting Engineer in Color Television for CBS, New York, N. Y. He was Head of a TV Research Department at Motorola, Inc., Chicago, Ill., from 1947 to 1958. In 1958 he joined the General Electric Company as a Consulting Engineer to the Cathode Ray Tube Department, Syracuse, N. Y.

Dr. Schlesinger has numerous U. S. Patents and has written over 30 technical articles. He has served on NTSC committees charged with studies of dot-interlace and color. He is a contributor to the "Handbook of Television Engineering," edited by Don Fink and published by McGraw-Hill Book Company. In 1957 he was Visiting Lecturer at Northwestern University, Evanston, Ill.

He is a member of the Television Society of London and the Research Society of America.

❖

Mortimer A. Schultz (SM'48) was born in Portland, Me., on March 21, 1918. He received the B.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1939.

From 1939 to 1945 he worked at the Westhinghouse Industrial Electronics Division, Baltimore, Md., where he was engaged in television and low-frequency radar projects. After a year at Photoswitch, Inc., Cambridge, Mass., as a Project Engineer for military radar attachments, he returned to Westinghouse Research Laboratories, as Section Manager in charge of radar, sonar, and industrial control. With the start of the nuclear submarine project, he transferred to the Westinghouse Bettis Field, Pittsburgh, Pa., where he was engaged in the development of the instrumentation and control system for the *Nautilus*. In 1955 he was on loan to the Sherwood fusion project at Princeton, N. J., and on his return to Pittsburgh was responsible for the design and operation of the Westinghouse Testing Reactor. In March, 1961 he opened his own nuclear consulting offices in Pittsburgh.

Mr. Schultz is a member of the American Nuclear Society and the AIEE.

❖

Oliver G. Selfridge was born in London, England, on May 10, 1926. He attended the Massachusetts Institute of Technology, Cambridge, from 1942 to 1950, receiving the S.B. degree, in 1945.

He worked for two years at the Signal Corps Laboratories, Fort Monmouth, N. J. Since 1952 he has been with the M.I.T. Lincoln Laboratory, Lexington, Mass., where he is a member of Division V, Information Processing. For the past few years he has been working on pattern recognition and other areas in artificial intelligence.

❖

Robert Serrell (A'35–SM'44) was born near Lyon, France, on October 16, 1904. He was a radio amateur in France in the twenties. He received the E.E. degree from the Institut Normal Electrotechnique, Paris, in 1926.

He was a Design Engineer at General Electric, Schenectady, N. Y., and at RCA, Camden, N. J., from 1927 to 1935. He was engaged in television engineering at RCA in Camden from 1935 to 1940, and was in charge of television transmitter development with the Columbia Broadcasting System in New York, in 1940 and 1941. He was Group Leader and Consulting Engineer at the Radio Research Laboratory of Harvard University, Cambridge, Mass., from 1942 to 1945, and in charge of television operations at CBS, 1944 to 1947. From 1947 to 1953 he was engaged in research in applied mathematics and in electronic computer theory at the Princeton Laboratories of RCA. He was Head of the RCA Computation Laboratory Princeton, N. J., from 1953 to 1957. Since September, 1957, he has been a Consultant engaged in the design of special-purpose electronic computers and data-processing machines and in the development of mathematical inference techniques. He is also a Research Associate of Haskins Laboratories, New York, N. Y.

Mr. Serrell is a Past Chairman of the IRE Electronic Computers Committee; a member of Sigma Xi, the Association for Computing Machinery (past Council member), the American Mathematical Society and the New York Academy of Sciences.

Ralph M. Showers (S'39–A'42–M'47–SM'55) was born in Plainfield, N. J., on August 7, 1918. He received the B.S.E.E. degree, in 1939, the M.S.E.E. degree, in 1941, and the Ph.D. degree, in 1950, all from the University of Pennsylvania, Philadelphia.

He is Professor of Electrical Engineering and Section Head of Communications Research in The Moore School of Electrical Engineering, University of Pennsylvania. He is also an Instructor of graduate courses in physical electronics, electronic circuits, communications systems, and of a seminar in operations research and systems engineering.

Dr. Showers was Chief U. S. delegate to the CISPR Conference (International Special Committee on Radio Interference) at The Hague, in 1958. He has been national Chairman of IRE-PGRFI, in 1960–1961, and is at present Chairman of the Philadelphia Section of IRE. He is Chairman of the American Standards Association C63 Ad Hoc Committee for the 1961 CISPR Conference.

Dr. Showers is a member of the AIEE, the American Standards Association, the Operations Research Society of America and Sigma Xi.

❖

Keeve M. Siegel (SM'57) was born in New York, N. Y., on January 9, 1923. He received the B.S. and M.S. degrees in physics from Rensselaer Polytechnic Institute, Troy, N. Y., in 1948 and 1950, respectively.

He has served with The University of Michigan, Ann Arbor, since 1948, first as a Research Associate, then as a Research Engineer. He became Head of the Upper Atmosphere Physics Section, in 1949 and Head of the Theory and Analysis Department, in 1952. Since 1957 he has been Head of the University's Radiation Laboratory and Professor of Electrical Engineering. His work at The University of Michigan has been in the fields of electromagnetic theory (*e.g.*, scattering and diffraction), high-altitude research and work in thermodynamics and hydrodynamics (*e.g.*, incompressible ideal subsonic flow), and passage of plane waves of sound in air. In 1961, he was appointed President and Director of the Ann Arbor Division of the Conductron.

Prof. Siegel is a member of the American Physical Society, American Institute of Physics, American Mathematical Society, and Sigma Xi. He is an Associate Fellow of the Institute of Aeronautical Sciences, a member of the USAF Scientific Advisory Board, and Consultant to the Advanced Research Projects Agency and several major corporations. He is a member of the Editorial Boards of the *Journal of Research of the National Bureau of Standards* and the *Journal of Mathematical Physics.* He is a member of Commission VI of URSI. He is listed in *Who's Who in America, American Men of Science. Who's Who in World Aviation,* and *World Directory of Mathematicians.*

❖

Samuel Silver (M'46–SM'50–F'54) was born in Philadelphia, Pa., on February 25, 1915. He received the B.A. and M.A. degrees in physics, in 1935 and 1937, respectively, from Temple University, Philadelphia, and the Ph.D. degree in physics, in 1940, from the Massachusetts Institute of Technology, Cambridge.

Following one year as a Postdoctoral Research Assistant in the Mendenhall Laboratory of Physics at Ohio State University, Columbus, he joined the faculty of the University of Oklahoma, Norman, where, as Instructor and Assistant Professor of Physics, he served until 1943. During that period he conducted research on molecular structures. He then joined the Radiation Laboratory at M.I.T. as a staff member in the Antenna Group where he carried on research on theory of microwave antennas and networks.

He joined the newly-formed Antenna Research Branch of the Naval Research Laboratory, Washington, D. C., in 1946, as head of a basic research group. He came to the Department of Electrical Engineering of the University of California at Berkeley, in 1947 and was made Professor of Electrical Engineering, in 1950 and subsequently Professor of Engineering Science. He spent the year 1953–54 at the Royal Technical University of Denmark and the Cavendish Laboratory, Cambridge University, under a Guggenheim Fellowship. In 1956 he was appointed Director of the Electronics Research Laboratory of the Electrical Engineering Department. He developed the research program on microwave antennas and applied electromagnetic theory and related areas of microwave electronics. He is currently directly engaged in upper atmosphere studies and solar phenomena in the microwave region. He was awarded a Guggenheim Fellowship for 1960–1961 for research in upper atmospheric physics. He was appointed Director of the Space Sciences Laboratory of the University of California on January 1, 1960.

Dr. Silver was Chairman of the International Commission VI (Radio Waves and Circuits), 1953–1960, and is currently Secretary of the USA National Committee of URSI. He is a Fellow of the American Physical Society and a member of the American Geophysical Society, the New York Academy of Sciences, and Sigma Xi.

❖

Donald B. Sinclair (J'30–A'33–M'38–SM'43–F'43) was born in Winnipeg, Manitoba, Canada, on May 23, 1910. He received the S.B., S.M., and Sc.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1931, 1932 and 1935, respectively.

Since 1936 he has been employed by the General Radio Company, West Concord, Mass. He has held the positions of Assistant Chief Engineer, Chief Engineer, Vice President and Director, and in 1960 he was appointed Executive Vice President and Technical Director. During World War II he was in charge of search-receiver work for radar countermeasures at the Radio Research Laboratory of Harvard University, Cambridge, Mass., and he was a member of the National Defense Research Committee on guided missiles. In 1943 he went to North Africa with the first Ferret plane, which was sent to the European theater of operations. For his work on countermeasures and guided missiles he received the President's Certificate of Merit in 1948. From 1954 to 1958 he was a member of the Technical Advisory Panel on Electronics of the Department of Defense, and he is presently serving as Consultant to the Defense Department.

Dr. Sinclair is a Fellow of the AIEE, and a member of Sigma Xi, Eta Kappa Nu, the AAAS, the American Physical Society and the National Association of Manufacturers Research Committee. He was President of the IRE in 1952 and has also served as Treasurer and Vice President. At various times he has been on the Executive Committee and on the Board of Directors.

❖

S. Fred Singer was born in Vienna, Austria, on September 27, 1924. He received the B.E.E. degree from Ohio State University, Columbus, in 1943, and the A.M. and Ph.D. degrees from Princeton University, N. J., in 1944 and 1948, respectively.

After service in the U. S. Navy from 1944 to 1946, he joined the Applied Physics Laboratory of The Johns Hopkins University, Silver Spring, Md. There he was a member of one of the earliest high altitude research groups using rockets. He and his collaborators discovered the electrojet current flowing in the equatorial ionosphere. From 1950 to 1952 he was Scientific Liaison Officer, Office of Naval Research, at the U. S. Embassy, London. In 1953 he joined the University of Maryland, College Park, where he is Professor of Physics, and Director of the experimental and theoretical programs of the Space Research Group. He has been intensely interested in and has published many original papers on satellites and their scientific applications.

Dr. Singer is a member of Commission IV of the International Scientific Radio Union, Director of the American Astronautical Society and a Fellow of numerous scientific societies in the U. S. and abroad.

❖

Hector R. Skifter (A'31–M'36–SM'43–F'51) was born in Austin, Minn., on March 4, 1901. He received the B.A. degree from St. Olaf College, Northfield, Minn., in 1922, and an Honorary D.Sc. degree, in 1945, also from St. Olaf College.

From 1922 to 1929 he was an Instructor in Physics and Mathematics at St. Olaf College. He was Chief Engineer of the Western Radio Company, St. Paul, Minn., from 1929 to 1932. During the years 1932 to 1942 he was a Consulting Radio Engineer, and from 1934 to 1942 he was Technical Supervisor of the National Battery Broadcasting Company, St. Paul. He was Associate Director of the Airborne Instruments Laboratory (AIL) of the Columbia University Division of War Research under a contract with the Office of Scientific Research and Development (OSRD), from 1942 to 1945. He was responsible for the development of the airborne magnetometer for submarine detection and the development of electronic countermeasures against German guided missiles. In August, 1945 when the OSRD contract was terminated, AIL was incorporated as an independent laboratory with Dr. Skifter as President. In June, 1958 he was elected to a Vice Presidency of Cutler-Hammer, when AIL merged with Cutler-Hammer.

In 1957, he served the Department of Defense as a part-time Consultant. In February, 1959, he took a leave of absence from his position as President of Airborne Instruments Laboratory to accept the full-time position of Assistant Director of Defense Research and Engineering. During his period of office he was responsible for technical evaluation and integration of defense weapons systems, planning and supervising the development of new and improved systems and their control environment, including anti-aircraft and anti-missile missiles and interceptor aircraft. As part of his duties he served as Chairman of the Ballistic Missile Defense Steering Group, spearheaded the installation of the BMEWS (Ballistic Missile Early Warning System) program, and coordinated the planning for the development of the NIKE-ZEUS and the ARPA DEFENDER programs. He served as a member of the Department of Defense Panel for the evaluation program of the BOMARC Weapon System and was the Chairman of the U. S. Delegation to the Tripartite Conference (United States, United Kingdom, and Canada) Technical Subgroup on Defense Against Ballistic Missiles. In April, 1960, he resigned from his position with the Department of Defense and rejoined AIL as President.

Dr. Skifter is a Vice-President and a Director of Cutler-Hammer, Inc., Chairman of the Board of Directors of the Research Analysis Corporation, and a Director of the American Research and Development Corporation and of Rabinow Engineering Company. He was formerly a member of the Board of Directors of Intercontinental Electronics Corporation and Cramer Controls Inc. He was also a member of the Gaither Security Resources Panel of the President's Science Advisory Committee. He is also Consultant to the President's Science Advisory Committee, Executive Office of the President, and a member of the Army Scientific Advisory Panel, the New York State Scientific Advisory Council and the Advisory Committee, Department of Ordnance, U. S. Army.

Dr. Skifter is a member of the Cosmos Club, the Wings Club, and the University Clubs of Washington and Milwaukee. He is a Director of Nassau Hospital, N. Y., a Trustee of Hofstra College, and a Director of the Long Island Fund.

❖

David Slepian (A'52–M'57–SM'60–F'62) was born in Pittsburgh, Pa., on June 30, 1923. He attended The University of Michigan, Ann Arbor, from 1941 to 1943, at which time he entered the U. S. Army and served in the Signal Corps until 1946. From 1946 to 1949 he attended Harvard University, Cambridge, Mass., where he received the M.A. degree, in 1947, and the Ph.D. degree in physics, in 1949. He spent the academic year 1949–1950 at the University of Cambridge, England, and at the Sorbonne, Paris, as a Parker Fellow in physics from Harvard.

In 1950 he joined the Bell Telephone Laboratories, Murray Hill, N. J., as a Research Mathematician and has remained in this position since. His interests here have been in communication theory and applied mathematics, especially applications of probability theory. During the academic year 1958–1959 he was a Visiting McKay Professor of Electrical Engineering at the University of California at Berkeley.

Dr. Slepian is a member of AAAS, the American Mathematical Society, Institute of Mathematical Statistics, Society for Industrial and Applied Mathematics, Commission VI of URSI, and Sigma Xi. He served on the Administrative Committee of PGIT from 1958–1960 and is presently on the Council of SIAM.

❖

George C. Southworth (M'26–F'41–L) was born in Little Cooley, Pa., on August 24, 1890. He received the B.S. degree from Grove City College, Pa., in 1914, and did graduate work both at Columbia University, New York, N. Y., and Yale University, New Haven, Conn., where he received the Doctorate degree, in 1923.

Prior to World War I he began experimental work at Grove City College, Pa., and during the war continued with research work at the National Bureau of Standards and at Yale University. Since 1923 he has been with the Bell System, where he has been associated with the transoceanic radio telephony and later with the early development of microwave techniques.

For his work in waveguides, he received the Morris N. Liebmann Prize of the IRE, in 1938 and the Stuart Ballantine Medal of the Franklin Institute, in 1947; for his work on microwave radiation from the sun, he received the Louis Levy Medal of the Franklin Institute, in 1946.

Dr. Southworth is a Fellow of the American Physical Society and the AAAS.

❖

Karl R. Spangenberg (A'34–SM'45–F'49) was born in Cleveland, Ohio, on April 9, 1910. He received the B.S. degree and the M.S. degree in electrical engineering from the Case Institute of Technology, Cleveland, in 1932 and 1933, respectively, and the Ph.D. degree from Ohio State University, Columbus, in 1937.

For some twenty years he has been on the staff of the Electrical Engineering Department of Stanford University, Calif. This period has been punctuated by technical work in the European Theater during World War II, a period as Head of the Electronics Division of the Office of Naval Research, 1948 to 1949, and a period as Head of the Electronic Engineering Department of the Instituto Technologica de Aeronautica in the state of Sao Paulo, Brazil, 1952 to 1954. During the past several years he has been a Consultant on engineering, management, and education, to various companies, mostly in the San Francisco Bay Area.

Mr. Spangenberg is the author of "Vacuum Tubes" and "Fundamentals of Electron Devices." He is a member of the American Physical Society.

❖

Malcolm L. Stitch (SM'58) was born in Elizabeth, N. J., on April 23, 1923. He received the B.A. degree in French and the B.S. degree in physics from the Southern Methodist University, Dallas, Tex., both in 1947, and the Ph.D. degree in physics from Columbia University, New York, N. Y., in 1953.

From 1949 to 1951 he was an Instructor in Physics at Sarah Lawrence College, Bronxville, N. Y., and the following year at Cooper Union College of Engineering, New York, N. Y. From 1953 to 1956 he was a Research Physicist at Varian Associates, Palo Alto, Calif., and since 1956 he has been with the Hughes Aircraft Company, Culver City, Calif., where he is currently Manager of Laser Development in the Radar and Missile Electronics Laboratory.

He designed the first high-temperature microwave spectrometer and with colleagues examined the microwave spectra of the alkali halides in the neighborhood of 1000°C. He also led the effort to exploit the systems application of lasers by being responsible for the design of the COLIDAR (COherent LIght Detecting And Ranging) ranging system at Hughes. He has also made contributions to the fields of frequency control of microwave tubes, maser theory and technology, atomic clocks, and quantum electronics.

Dr. Stitch is a member of the Physical Society, the AAAS, American Association of Physics Teachers, Physical Society of Japan, the New York Academy of Sciences, and Sigma Xi.

❖

Ernst Stuhlinger was born in Niederrimbach, Germany, on December 19, 1913. He received his Doctorate in Physics at the University of Tuebingen, Germany, in 1936.

He was appointed Assistant Professor of the Physics Department of the Berlin Institute of Technology and was a member of the faculty there from 1936 to 1941. He worked closely with Dr. Hans Geiger, developer of the Geiger counter, for seven years. From 1939 to 1941, he was a member of a research group conducting studies in nuclear physics. In the Spring of 1943, he joined the Rocket Development Center at Peenemuende, Germany, which was under the technical

supervision of Dr. Wernher von Braun, presently Director of George C. Marshall Space Flight Center, Huntsville, Ala. At Peenemuende, Dr. Stuhlinger carried on research in connection with the development of guidance and control systems for the V-2 guided missile. Dr. Stuhlinger, who became an American citizen on April 14, 1955, came to the United States in 1946 under the auspices of the Ordnance Corps, U. S. Army. He conducted research and development work with guided missiles at Ft. Bliss, Tex., and assisted in high altitude research firings of captured V-2's at White Sands Proving Ground, N. Mex. From 1956 until July 1, 1960, he was Director of the Research Projects Laboratory, Army Ballistic Missile Agency, U. S. Army Ordnance Missile Command at Redstone Arsenal, Huntsville, Ala. He is now the Director, Research Projects Division of the George C. Marshall Space Flight Center.

Dr. Stuhlinger, who has gained recognition for his feasibility and design studies of electrical propulsion systems for space ships, was awarded the American Rocket Society's Propulsion Award on December 7, 1960. He is a member of a number of professional societies including the ARS, AAS, and the International Astronautical Academy.

❖

William O. Swinyard (A'37–M'39–SM'43–F'45) was born in Logan, Utah, on July 17, 1904. He received the B.S. degree in mathematics and physics from Utah State University, Logan, in 1917, and has done graduate work at Columbia University, New York, N. Y., and Northwestern University, Evanston, Ill.

In 1930 he joined the Hazeltine Corporation, Bayside, N. Y.; in 1937 he was transferred to Chicago, Ill. He has been Chief Engineer of Hazeltine Research, Inc., since 1942, and Vice President and Director, since 1958. He is one of the founders of the National Electronics Conference, and has served the NEC as President and Chairman of the Board.

Mr. Swinyard is a Fellow of the Radio Club of America and the AAAS, a professional member of Eta Kappa Nu; and Past President of the Chicago Radio Engineers Club. He is a registered professional engineer in the state of Illinois, and a member of the National Society of Professional Engineers and the Illinois Society of Professional Engineers.

❖

Samuel A. Talbot (SM'61) was born in Holderness, N. H., on April 19, 1903. He received the A.B. degree, in 1925, from Cornell University, Ithaca, N. Y., the M.S. degree, in 1931, from Trinity College, Hartford, Conn., and the M.A. and the Ph.D. degrees in physics from Harvard University, Cambridge, Mass., in 1932 and 1938, respectively.

He was a Master at the Loomis School, Windsor, Conn., from 1927 to 1930, and an Instructor and Tutor in Physics at Harvard, from 1930 to 1935. In that year he joined the staff of the Department of Ophthalmology, The Johns Hopkins University, Baltimore, Md., working on physiological optics; he was appointed Associate Professor of Medicine (biophysics) in 1946. From 1942 to 1944 he worked for the Office of Science Research and Development. His main fields of interest are electro- and ballistocardiography, circulatory biophysics, cardiovascular sound, and biophysical methods and measurements. He is currently director of the P.H.S. Training Program in Bio-Medical Engineering at The Johns Hopkins Medical School.

Dr. Talbot is a member of the Physiological Society, the Optical Society, the Society of Experimental Biology, the American Heart Association, the Hospital Physicists Society, the Biophysical Society (Secretary), and the Visual Neurophysiology Society.

❖

Frederick E. Terman (A'25–F'37) was born in English, Ind., on June 7, 1900. He received the B.A. degree in chemical engineering, in 1920, and the E.E. degree, in 1922, both from Stanford University, Calif., and the D.Sc. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1924.

He has served on the faculty of Stanford University since his appointment as an Instructor in Electrical Engineering in 1925. In 1937 he became Professor and Head of the Electrical Engineering Department; in 1945 he was appointed Dean of the School of Engineering, and in 1955 became Provost, and since 1958 has been Vice President and Provost of Stanford.

On leave during World War II, he organized and directed the Radio Research Laboratory set up at Harvard University by the Office of Scientific Research and Development. The laboratory was the chief United States agency developing radar countermeasures.

In 1946 he was decorated by the British government, and in 1948 received the United States' highest civilian award, the Medal of Merit. Harvard University, the University of British Columbia, and Syracuse University have awarded him honorary doctoral degrees.

A Past President of the IRE, Dr. Terman received the Institute's Medal of Honor, in 1950. He is also a member of the AIEE, the American Physical Society, and the ASEE, and, in 1946, was elected to the National Academy of Sciences.

❖

George R. Town (A'37–SM'44–F'50) was born in Poultney, Vt., on May 26, 1905. He received the E.E. and the D.Engr. degrees from Rensselaer Polytechnic Institute, Troy, N. Y., in 1926 and 1929, respectively.

From 1929 to 1933 he did research at the Leeds and Northrup Company, Philadelphia, Pa. From 1933 to 1936 he taught at Rensselaer. In 1936 he joined the Stromberg-Carlson Company, Rochester, N. Y., where he remained until 1949; during the last five years of this time, he was Manager of Engineering and Research.

In 1949 he came to Iowa State University, Ames, as Associate Director of the Engineering Experiment Station and Professor of Electrical Engineering. For 20 months, in 1957 and 1958, he was on leave of absence from Iowa State University serving as Executive Director of the Television Allocations Study Organization, Washington, D. C. In 1959 he was named Dean of the College of Engineering at Iowa State University.

Dr. Town is a fellow of the AIEE and a member of ASEE, NSPE, Tau Beta Pi, Sigma Xi, Eta Kappa Nu and Phi Kappa Phi.

❖

John G. Truxal (S'47–A'48–SM'54–F'59) was born in Lancaster, Pa., on February 19, 1924. He received the B.A. degree from Dartmouth College, Hanover, N. H., in 1944, and the D.Sc. degree from Massachusetts Institute of Technology, Cambridge, in 1950.

He taught at Purdue University, Lafayette, Ind., and then at the Polytechnic Institute of Brooklyn, N. Y., where he has been Head of the Electrical Engineering Department, and is currently Vice President for educational development.

Dr. Truxal is Chairman of the American Automatic Control Council Theory Committee and Past Chairman of the IRE Professional Group on Education. He is a member of the Administrative Committee of PGAC and PGE, and a member of the IRE Editorial Board.

❖

Lester C. Van Atta (M'42–SM'43–F'52) was born in Portland, Ore., on April 18, 1905. He received the B.A. degree in physics from Reed College, Portland, in 1927, and the Ph.D. degree from Washington University, St. Louis, Mo., in 1931.

He served in the Physics Department of Princeton University, N. J., from 1931 to 1932, in the Physics Department of the Massachusetts Institute of Technology, Cambridge, from 1932 to 1940, and in the M.I.T. Radiation Laboratory as Head of the Antenna Group from 1940 to 1945. From 1945 to 1950 he was Head of the Antenna Research Branch of the Naval Research Laboratory, Washington, D. C. Since 1950 he has been at Hughes Aircraft Company, Culver City, Calif., originally as Head of the Microwave Laboratory, then as Associate Director of the Research Laboratories, and later as Director of Technical Information and Education. From May, 1960 to June, 1961, he was on professional leave from Hughes as Special Assistant for Arms Control to the Director of Defense, Research and Engineering, Office of the Secretary of Defense, Washington, D. C. In July, 1961 he returned to Hughes Aircraft Company as Technical Director of the Hughes Research Laboratories.

Since 1940 he has been active in military research and development at the M.I.T. Radiation Laboratory, the Naval Research Laboratory, and Hughes Aircraft Company. In addition, he served on the Research and Development Board as Chairman of the Panel on Antennas and Propagation from 1946 to 1949. More recently, he participated in a series of WSEG studies of continental air defense from 1956 to 1958, a series of PISGAH and other research and development conferences for the Army from 1957 to 1959, two ARGUS studies at Lawrence Radiation Laboratory, Livermore, Calif., from 1948 to 1959, and on the Gaither Committee in 1959.

At Hughes Aircraft Company he directed such systems studies as the Hemispherical Scan System for the Navy and the ICBM Passive Defense System for the Air Force, both systems based on electronic beam-scanning techniques. In connection with his responsibility for education at Hughes, he was active in the creation of the Southern California Industry-Education Council, and served as its Vice President during 1959–1960. His technical fields of specialization have included high resistors, electron scattering in gases, high-voltage and high-vacuum engineering, production and scattering of hard X rays, nuclear reactions by bombardment, microwave antennas and microwave components and systems.

Dr. Van Atta is a Fellow of the American Physical Society and has been active in the Research Society of America and the International Scientific Radio Union, especially Commission VI. He has been very active in the IRE, including chairmanship of the Antenna and Waveguide Technical Committee, the PGAP, and the Los Angeles Section. He is presently serving on the Board of Directors of the IRE. He is a member of the AAAS, the American Association of Physics Teachers, and the American Geophysical Union.

❖

Joseph H. Vogelman (M'46–SM'49–F'59) was born in New York, N. Y., on August 18, 1920. He received the B.S. degree, in 1940, from the College of the City of New York, and the M.S. and Ph.D. degrees, in 1948 and 1957, respectively, from the Polytechnic Institute of Brooklyn, N. Y., both in electrical engineering.

In 1945 after several years at the Signal Corps Radar Laboratory at Fort Hancock and Belmar, N. J., he joined the staff of the Watson Laboratories, Red Bank, N. J., where he served as Chief of the Development Branch until 1951, responsible for research and development of test equipment and microwave components and techniques. In 1951 he moved to the Rome Air Development Center, Griffiss Air Force Base, N. Y. From 1951 to 1953 he was Chief Scientist in the General Engineering Laboratory and Consultant on UHF and SHF theory and techniques to the U. S. Air Force. He was Chief of the Electronic Warfare Laboratory, directing all research and development in ground-based electronic warfare for the USAF, 1953 to 1956. From 1956 through June, 1959, he was Technical Director of the Communications Directorate with responsibility for the Air Force research and development effort in ground-based and ground-to-air communications. He is now Vice President of Research, Development and Engineering, Capehart Corporation, Richmond Hill, N. Y.

Dr. Vogelman is a Fellow of the AAAS and a member of Eta Kappa Nu, Sigma Xi, the AIEE, and the Armed Forces Communications and Electronics Association.

Willis H. Ware (A'43–SM'49–F'62) was born in Atlantic City, N. J., on August 31, 1920. He received the B.S. degree from the University of Pennsylvania, Philadelphia, in 1941, and, as a Tau Beta Pi Fellow, the S.M. degree from the Massachusetts Institute of Technology, Cambridge, in 1942, both in electrical engineering, and the Ph.D. degree, in 1951, from Princeton University, N. J.

From 1942 to 1946 he was employed by the Hazeltine Electronics Corporation, Little Neck, N. Y., where he did research and development in radar and IFF. In 1946 he became one of the original members of the staff of the Electronic Computer Project at the Institute for Advanced Study, Princeton, N. J., where he worked on the design and development of the large-scale general-purpose electronic digital computer, which later was to set the pattern for the construction of several other "Princeton-class" machines. In 1952 he joined The RAND Corporation, Santa Monica, Calif., where presently he is the Associate Head of the Computer Sciences Department, concerned principally with the development of large machines and their application to military and scientific problems.

Dr. Ware is a member of the ACM, the AAAS, Tau Beta Pi, Sigma Xi, Eta Kappa Nu, Pi Mu Epsilon, and Sigma Tau. He has been very active in the PGEC, having served as Chairman, in 1958–1959. He has been a member of the IRE Technical Committee 8 on Computers since 1951, and has served as IRE representative to the National Joint Computer Committee (1959 to 1961). In 1958 he served as Chairman of the Western Joint Computer Conference. In 1959 he was a member of an eight-man delegation of American scientists who visited the U.S.S.R. to discuss computers and related matters. In 1961 he was elected first Chairman of the American Federation of Information Processing Societies. He received the Achievement Award of the Los Angeles Section of the IRE in 1957.

❖

Ernst Weber (M'41–SM'43—F'51) was born in Vienna, Austria, on September 6, 1901. He received the Ph.D. degree, in 1926, from the University of Vienna, having studied philosophy, physics, and mathematics, and the D.Sc. degree, in 1927, from the Technical University of Vienna, having studied engineering, particularly field theory as applied to machinery.

After working as Research Engineer for the Austrian Siemens-Schuckert Company, he was transferred to the company's Berlin office, in 1929, and was also appointed Lecturer at the Technical University of Berlin. In 1930 he became Visiting Professor at the Polytechnic Institute of Brooklyn, N. Y.; the next year he was named a permanent Research Professor of Electrical Engineering in charge of graduate study. From 1942 to 1945 he was Professor of Graduate Electrical Engineering and Head of graduate study and research in that field. Under his direction the graduate enrollment increased from less than 200 to over 900 students.

Early in World War II he organized a microwave research group which developed the precision microwave attenuator, sorely needed for the accurate calibration of radar. In recognition of the contributions of the research group, he was awarded the Presidential Certificate of Merit. Out of this wartime research grew the Microwave Research Institute, which expends each year more than $1,000,000 on research projects for the military services, and the Polytechnic Research and Development Company, which was then owned by the Polytechnic Institute of Brooklyn.

In 1945 he was appointed Head of the Department of Electrical Engineering, and Director of the Microwave Research Institute. When a vice presidency for research was created by the Polytechnic Institute of Brooklyn in 1957, he was named to the position; following the death of the Institute's president, in June, 1957, he became Acting President and then President. He also served as President of the Polytechnic Research and Development Company until its sale in 1959. In January, 1959, he was elected President of the IRE.

A pioneer in the high-frequency electronic research, he holds more than 50 American, Canadian, and British patents in the field of microwave techniques. His published works include many scientific papers on electromagnetic fields, linear and nonlinear circuits, and microwave measurements. He has contributed to several books, and published "Mapping of Fields" and "Linear Transient Analysis." In 1959 he was appointed to the Army Scientific Advisory Panel by the Secretary of the Army.

Dr. Weber is a Fellow of the AIEE and the American Physical Society.

❖

Vernon I. Weihe (SM'45–F'58) was born in Louisville, Ky., on March 3, 1909. He received the B.S.E.E. degree from the Speed Scientific School of the University of Louisville, in 1931, and did graduate work at the University of Louisville and the University of Cincinnati, Ohio.

He has been engaged in air navigation, communications, and traffic control development since 1935. During World War II, as Chief Engineer of the Communications and Navigation Laboratory at Wright Field, Dayton, Ohio, he had cognizance over the development and procurement of many types of communication, navigation, and specialized equipments used during the war. Since 1946 he has been concerned with air navigation and traffic-control systems planning as a member of the Air Transport Association, as a Development Engineer, and currently as Director of Advanced Systems Planning for General Precision, Inc., Washington, D. C.

Mr. Weihe is a member of the Institute of Aeronautical Sciences, and a Past President of the Institute of Navigation.

❖

James O. Weldon (A'26–SM'46–F'54) was born in Canton, Mo., on March 15, 1905.

He was first engaged in radio engineering in 1927. From 1933 to 1952 he was a Consulting Engineer and manufacturer of transmitting equipment. He was appointed Chief Engineer of the Bureau of Communication Facilities, Overseas Branch, Office of War Information in August, 1942, and a few months later became Chief of that Bureau. He organized Continental Electronics Manufacturing Company, Dallas, Tex., in 1946. As President and Director of Engineering, he has closely directed the engineering design and development of most of Continental's high-power transmitters which are now in use by the Army, Navy, Air Force, the Bureau of Standards and the Voice of America.

❖

John W. Wentworth (S'48–A'50–M'55) was born in Greenville, Me., on November 3, 1925. He received the B.S.E.E. degree from the University of Maine, Orono, in 1949.

He then joined RCA and spent several years doing design and development work on broadcast studio equipment for color television. During this period, he also organized and taught a course in color television engineering, which later resulted in the publication of a textbook on the same subject. From 1954 to 1959 he served as the Manager of a group of design engineers engaged in projects related to television studio equipment. Since July, 1959 he has been engaged in the field of educational electronics, and in his present capacity of Manager, Educational Electronics, Broadcast and Television Division, RCA, Camden, N. J., he is in charge of a service which coordinates many of RCA's activities in the educational field.

Mr. Wentworth is a member of SMPTE, Tau Beta Pi, Phi Kappa Phi, and Sigma Xi.

Harold A. Wheeler (A'27–M'28–F'35) was born in St. Paul, Minn., on May 10, 1903. He received the B.S. degree in physics, in 1925, from George Washington University, Washington, D. C., and continued post-graduate studies until 1928 at The Johns Hopkins University, Baltimore, Md.

He was employed by the Hazeltine Corporation, Little Neck, N. Y., from 1924 to 1946, advancing to Vice-President and Chief Consulting Engineer. In 1959, he resumed some activity with this company as a Vice-President and a Director. Since 1947 his principal occupation has been as President of Wheeler Laboratories, Inc., Great Neck, N. Y., now a subsidiary of Hazeltine Corporation. In this capacity he is directing their Great Neck and Smithtown Laboratories, specializing in microwaves and antennas.

Mr. Wheeler has served the IRE in such positions as Director (1934, 1940 to 1945) and Chairman of the Standards Committee; he has contributed many papers to IRE periodicals and received the Morris N. Liebmann Memorial Prize from IRE, in 1940. He is a Fellow of AIEE and Radio Club of America, an Associate Member of IEE (British), and a member of Sigma Xi and Tau Beta Pi.

❖

William C. White (A'15–M'25–F'40) was born in Brooklyn, N. Y., on March 24, 1890. He received the E.E. degree from Columbia University, New York, N. Y., in 1912.

He entered the General Electric Research Laboratory in the summer of 1912 as an Assistant to Dr. Irving Langmuir during the years of pioneer development of the high vacuum tube. When this research reached the commercial stage, he was appointed engineer in charge of a development and design section for their manufacturer. Shortly before retirement in 1955, he returned to the Research Laboratory at the General Electric Company.

Mr. White served on the Board of Directors of the IRE, 1943 to 1947, and as its Treasurer during 1946.

❖

Laurens E. Whittemore (A'16–M'25–F'27) was born in Topeka, Kan., on August 20, 1892. He graduated from Washburn College, Topeka, in 1914, and received the M.A. degree from the University of Kansas, Lawrence, in 1915, where his major was physics.

After three years of graduate study and teaching at the University of Kansas, he was employed in the Radio Laboratory of the National Bureau of Standards, Washington, D.C., from 1917 to 1923. In 1924 he became a member of the staff of the U.S. Department of Commerce and served as Secretary of the Interdepartment Radio Advisory Committee. He was Assistant Secretary or Secretary of the four annual National Radio Conferences called by the Secretary of Commerce, Herbert Hoover, 1922–1925.

From 1925 to 1957 he was a member of the headquarters staff of the American Telephone and Telegraph Company, New York, N. Y., where he was concerned with, among other things, relations with Federal regulatory agencies. He attended a number of International Radio and Communication Conferences either on behalf of the U. S. Government or the American Telephone and Telegraph Company, and was Secretary of the International Radio Conference, Washington, D. C., in 1927.

Mr. Whittemore was Vice-President of the IRE, in 1928, and a member of the Board of Directors, from 1926 to 1929 and from 1935 to 1937. He has held assignments on a number of IRE Committees, including the Standardization Committee of which he was Chairman, from 1926 to 1929, and the Annual Review Committee of which he was Chairman, from 1940 to 1949. He served on the Board of Editors, from 1929 to 1953. He is a member of Sigma Xi.

❖

Lofti A. Zadeh (S'45–A'47–M'50–SM'56–F'58) was born in Baku, Russia, on February 4, 1921. He attended the American College, Teheran, Iran, and received the B.S.E.E. degree from the University of Teheran, in 1942. He came to the U. S. in 1944 and entered the Massachusetts Institute of Technology, Cambridge, where he received the M.S. degree, in 1946.

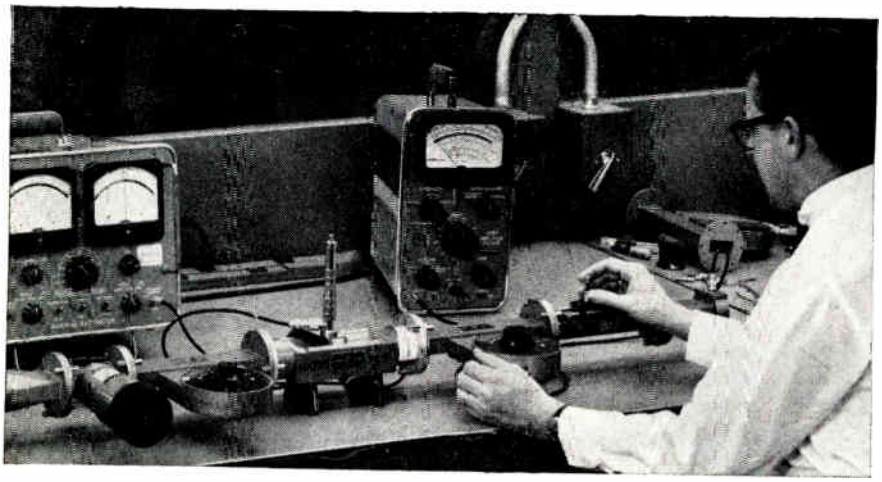
He joined the staff of Columbia University, New York, N. Y., as an Instructor in Electrical Engineering and received the Ph.D. degree, in 1949. He was promoted to Assistant Professor in 1950, Associate Professor in 1953, and full Professor in 1957. In 1956 he was a member of the Institute for Advanced Study, Princeton, N. J. At present, he is Professor of Electrical Engineering at the University of California at Berkeley.

He has published papers on the subjects of time-varying networks, optimal filters, and nonlinear systems, with emphasis on the general principles underlying the transformation of signals and the characterization of input-output relationships. His main interest at present is in system theory and sequential control processes.

Prof. Zadeh is Chairman of U. S. Commission VI of the International Scientific Radio Union (URSI), and a member of the AIEE, the American Mathematical Society, the American Physical Society, the Institute of Mathematical Statistics, the ACM, the Society for Industrial and Applied Mathematics, Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.

Index to Authors

☐ VSWR measurements of extreme precision
☐ FXR coax switch cuts Norden Laboratory test time
☐ How to reduce your cable inventory



VSWR measurements of extreme precision

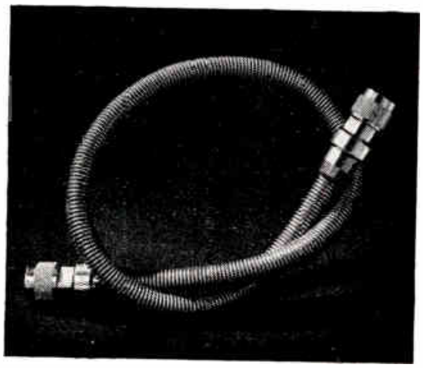
FXR's Model B813T VSWR Amplifier is a fully transistorized portable standing-wave amplifier with a full-scale maximum error at 5 db of ±0.05 db. We think it's the most accurate on the market.

The unit is designed for use on battery power for applications in the field where power is not available or in the laboratory where line noise might cause inaccuracies in measurement. Where power is available, however, the unit can be operated from the line in the normal manner.

Calibrated range of the B813T is 75 db. Normal, expanded and compressed meter scales are provided and can be used interchangeably without the need for readjustment of the gain control. The unit has special circuitry for bolometer protection and a meter display for bolometer resistance checking and current adjustment.

Other features of the B813T include controls and circuitry for selective meter damping, bandwidth selection and frequency peaking, range selection in 5-db steps, battery voltage checking and self-contained charging. Price: $285.00.

Need a high temperature cable for nuclear application?



. AMPHENOL ultra-high temperature flexible rf cable is capable of continuous operation at 1000°F with short excursions to higher temperatures.

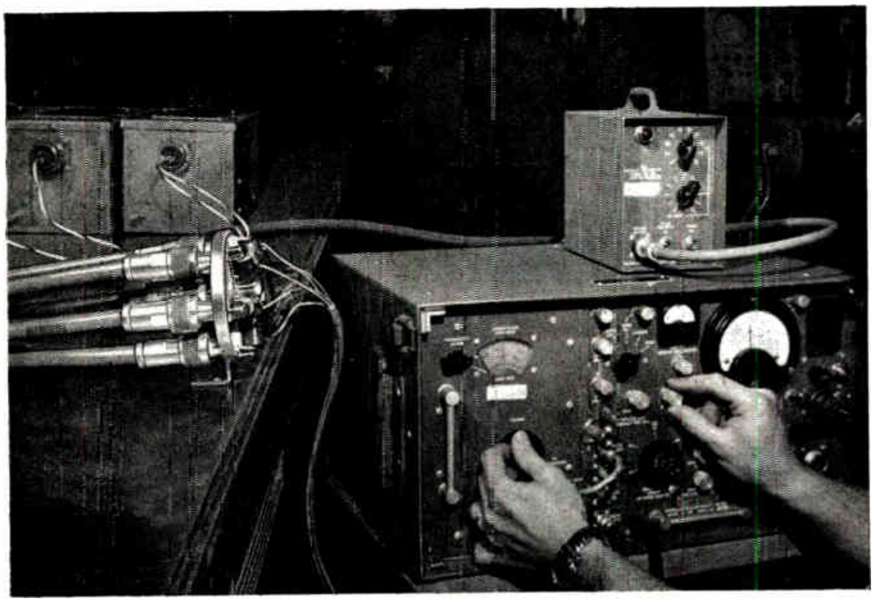
Application engineered for aircraft and missile temperature environments, the cable is a completely sealed rf transmission system consisting of inner and outer conductors separated by a dielectric of modified semi-solid silica.

Over all is a protective-sealed metallic-convoluted jacket. Cable ends are terminated at the factory with hermetically sealed 1000°F Series N plugs. System is resistant to nuclear radiation and is ideal for reactor use.

DATA

Altitude Insensitive-Moisture Resistant	Capacitance: 30.0 f/ft.
Resistant to Shock and Vibration	Velocity of Propagation: 69.0%
Resistant to Nuclear Radiation	Voltage Breakdown: 3500 Volts RMS
Connectors: Series N Plugs	Maximum Operating Voltage: 1000 VRMS
Impedance: 50 Ohms	Weight: Cable, 17½ lbs pr 100 ft.
	Connectors, 2½ ounces each

1000°F flexible rf cable is available in standard lengths up to 200 feet. Part number is 777-502. Specify length desired in order or quotation request. Length is measured from connector mating end to connector mating end.



FXR coax switch cuts Norden Laboratory test time

Norden, a division of United Aircraft Corporation, operates an environmental test laboratory at its Norwalk, Conn., plant.

Faced with an ever-increasing number of systems requiring noise and interference tests under military specifications, Norden engineers asked FXR to design a 50-ohm resistor-terminated coaxial switch that would have identical electrical characteristics through each of six outputs. Now, to check different channels, instead of disconnecting and connecting cable, Norden engineers just flick a small knob. All unused channels are grounded through 50-ohm terminations, with no significant crosstalk to interfere with the measurements being recorded.

The result is a faster, more convenient test procedure. If you're running noise tests under MIL-I 6181B, 6181D, 26600 or 19610 you'll find that this new coaxial switch will save hours of tedious work.

Maybe you have some other unusual rf switching problem. Why not ask us about it?

How to reduce your cable inventory

Most people we know cut up the cable they purchase, then put connectors on each end.

Why not let us make your cable assemblies? As the only major manufacturer of coaxial cable and coaxial connectors, FXR can probably save you money. We can certainly save you time.

How? Well, for one thing, you don't have to inventory so much cable or so many connectors. You don't have to set up production facilities for making these cable assemblies. And, you don't have to worry about quality control. We can produce sophisticated electrical assemblies to your rigid specifications. We have the people, facilities, and know-how to guarantee quality.

FXR manufactures Amphenol coaxial cable, Amphenol and "ipc" coaxial connectors and DK coaxial switches. Together, they constitute the industry's broadest line of coaxial components. Why not let us put them together for you?

The RF Products and Microwave Division Amphenol-Borg Electronics Corporation; 33 East Franklin Street, Danbury, Connecticut.

FXR™

Membership

The following transfers and admissions have been approved and are now effective:

Transfer to Senior Member

Agalides, E., Rochester, N. Y.
Belchis, S., Hartsdale, N. Y.
Benson, K. B., South Norwalk, Conn.
Bentivegna, M. J., Phoenix, Ariz.
Brown, R. B., 3rd, Bedford, Mass.
Caspers, J. W., San Diego, Calif.
Churchill, L. S., Jr., Plainfield, N. J.
Cohen, H. A., Baltimore, Md.
Cooney, J. J., Mishawaka, Ind.
Djordjevic, M. Z., Chicago, Ill.
Evans, A. P., New York, N. Y.
Gay, J. W., Columbus, Ohio
Henning, E. W., Phoenix, Ariz.
Jaffe, G., District Heights, Md.
Johnson, R. E., Maitland, Fla.
Knowles, C. H., Phoenix, Ariz.
Kreer, J. B., Morgantown, W. Va.
Parsons, L. W., Seattle, Wash.
Pinasco, S. F., Buenos Aires, Argentina
Shank, R. R., New Haven, Conn.
Zitovsky, S. A., Paramus, N. J.

Admission to Senior Member

Brown, J. H. U., Bethesda, Md.
Brustle, H. H., Adams, Mass.
Eckstein, A., Shrewsbury, N. J.
Edgar, R. B., Newport Beach, Calif.
Finke, E. L., Springfield, Va.
Hatoyama, G. M., Yokohama, Japan
Jaggi, G. C., Libertyville, Ill.
Nuzzo, D. R., Springdale, Conn.
O'Brien, J. S., Santa Ana, Calif.
Parszik, P. J., Pleasantville, N. Y.
Pease, H. G., Oriskany, N. Y.
Shallcross, J. S., Selma, N. C.
Staley, J. H., Anaheim, Calif.
Thompson, W. J., Mountain Lakes, N. J.
Ungar, S. H., St. Laurent, Que, Canada
Vollmer, J., Philadelphia, Pa.
Walton, R. C., San Jose, Calif.
Williams, E. C., The Hague, Netherlands
Wu, T. Y., Cambridge, Mass.

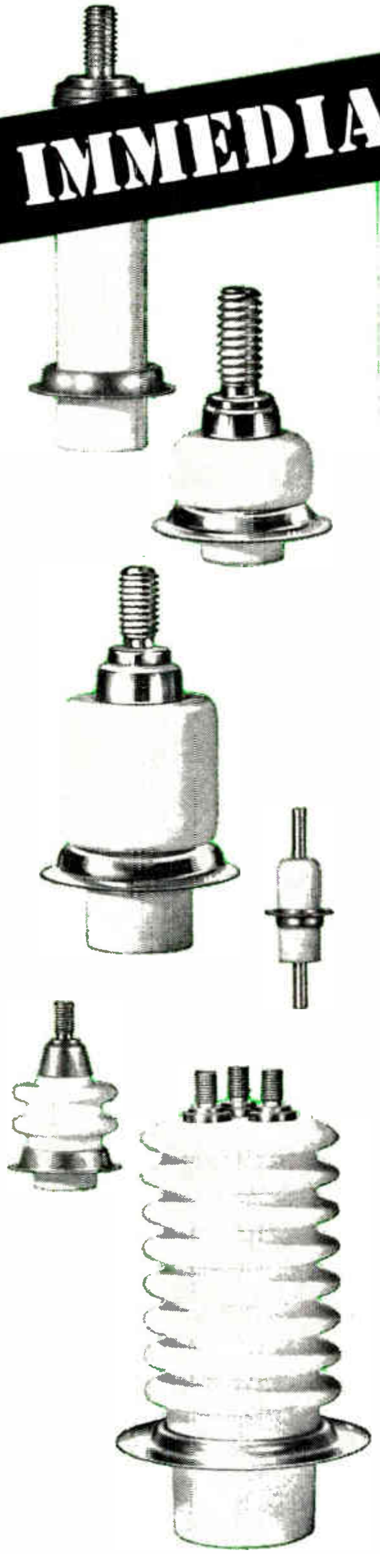
Transfer to Member

Bar-Lev, A., Kyriat Motzkin, Israel
Bean, E. W., Jr., Gainesville, Fla.
Brekhus, E. H., Minneapolis, Minn.
Carman, E. B., Derby, Conn.
Charfole, R. O., La Salina Edo Zulia, Venezuela
Collins, W. F., New Westminster, B. C., Canada
Davidson, R. C., Rochester, N. Y.
Findlay, D. A., Bayside, N. Y.
Fox, M. L., Charlotte, Mich.
Hatton, G. H. T., Perth, West Australia
Jezierski, E. C. J., Hythe N/R Southampton, Hampshire, England
Miller, C. J., San Diego, Calif.
Riggan, J. S., Anchorage, Alaska
Ryan, J. W., Methuen, Mass.
Sudbay, P. S., Gloucester, Mass.
Taagepera, A., Montreal, Que., Canada
Trofatter, K. F., Murray Hill, N. J.
Valentine, D. T., Palo Alto, Calif.
Vint, J. L., Richfield, Minn.
Vogel, S., New York, N. Y.

Admission to Member

Abbott, J. D., Sherborn, Mass.
Abodeely, E. F., Cedar Rapids, Iowa
Aho, J. K., APO, New York, N. Y.
Aldrich, J. E., Arcadia, Calif.
Amery, J. G., Lee's Summit, Mo.
Andrews, D. R., Yorktown Heights, N. Y.
Aufmuth, R. H., Mentor, Ohio

(Continued on page 138A)

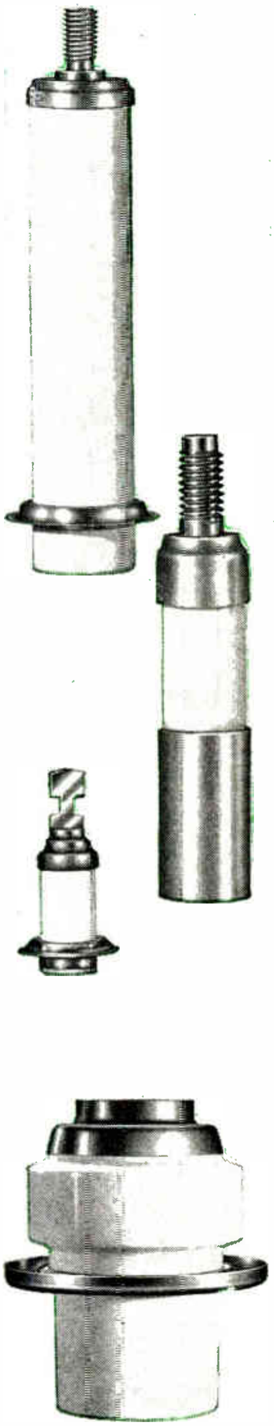


IMMEDIATE DELIVERY

of ALITE®

STANDARD HIGH ALUMINA BUSHINGS

In Over 100 Types and Sizes



More than 100 sizes of standard Alite high voltage terminals, feed-throughs, cable end seals and high amperage bushings are stocked for immediate, "off-the-shelf" delivery to simplify design problems, save time and help reduce costs.

One of these standard Alite high alumina ceramic-to-metal hermetic seals may be just the answer to your problems of reliability and maintenance where service conditions are extremely severe or critical.

- Alite hermetic seals — tested for vacuum-tightness with a high sensitivity helium mass spectrometer — are adaptable to virtually any assembly procedure, such as welding, brazing or soldering.
- Alite's mechanical strength and thermal-shock properties far exceed those of glass or porcelain.
- Alite retains its dielectric properties at elevated operating temperatures.
- Alite's smooth, white glaze resists build-up of contaminants on surfaces — easy to see, easy to clean.

From ceramic formulation to finished part, every manufacturing step is handled within our own plant. Strict quality control assures absolute adherence to specifications, utmost uniformity and reliability.

* * * *

When special bushings are called for, it is often possible to save time and money by modifying one of the standard units. However, when necessary a team of Alite engineers stands ready to help you by designing new bushings.

Write for these free helpful brochures. Bulletin A-40R illustrates, describes and gives complete specifications of standard Alite bushings, both high voltage and high amperage. Bulletin A-8 provides detailed descriptions of Alite formulations.
235-H

ALITE DIVISION

U. S. STONEWARE
BOX 119 ORRVILLE, OHIO

Membership

(*Continued from page 136A*)

Austin, D. F., Memphis, Tenn.
Baker, A. R., Buffalo, N. Y.
Baker, R. H., Offutt, AFB, Neb.
Barber, W. W., Catonsville, Md.
Barrick, D. E., San Diego, Calif.
Barron, C. E., Deansboro, N. Y.
Bayer, B. I., Syracuse, N. Y.
Bayliss, H. J., Duarte, Calif.
Beisty, J. M., Flushing, L. I., N. Y.
Bennett, A., Jr., Bountiful, Utah
Berndsen, P. J., Canoga Park, Calif.
Bilton, D. E., Trenton, N. J.
Binsley, D. L., Syracuse, N. Y.
Bordewieck, W. W., Wellesley Hills, Mass.
Bowers, E. J., Jr., Baltimore, Md.
Bowers, F. X., Mineola, L. I., N. Y.
Bremner, G. D., Victoria, B. C., Canada
Brennan, K. B., Chicago, Ill.
Brock-Nannestad, L. V., La Spezia, Italy
Brooks, D. W., Dallas, Tex.
Brown, M. F., St. Petersburg, Fla.
Burlas, T. C., Utica, N. Y.
Busenkell, L. A., Philadelphia, Pa.
Canale, S., Jr., Clay, N. Y.
Carr, E., Melbourne, Fla.
Castleman, J. I., No. White Plains, N. Y.
Chatelon, A. E., Montrouge (Seine), France
Chhabra, K. C., Pilani, Rajasthan, India
Chu, M. J., Chicago, Ill.
Churchill, F. E., Huntington Station, L. I., N. Y.
Churchill, J. G., Hyattsville, Md.
Claflin, M. L., Silver Creek, N. Y.
Clark, H. W., El Cajon, Calif.
Clow, R. G., Palo Alto, Calif.
Cole, H. S., Los Angeles, Calif.
Collins, E. L., Orlando, Fla.
Colodner, A. D., New Rochelle, N. Y.
Connell, R. A., Mount Kisco, N. Y.
Cooper, A. O., Fairfax, Va.
Cooper, D. F., Woodland Hills, Calif.
Cormia, K. C., Milwaukee, Wis.
Cormier, A. R., Lanham, Md.
Corry, J. R., Houston, Tex.
Cowgill, T. J., Cherry Hill, N. J.
Cox, T. E., Short Hills, N. J.
Crook, W. A., Oklahoma City, Okla.
Dabadghao, S. V., Bangalore, India
Daly, T. P., Redwood City, Calif.
Day, K. F., El Paso, Tex.
Debs, L. J., Van Nuys, Calif.
De Maeyer, L. C. M., Göttingen, Germany
Derry, J. H., Terre Haute, Ind.
Dial, R. E., Grand Island, N. Y.
Doss, D. W., Sunnyvale, Calif.
Eck, G. F., Johnsville, Pa.
Ehlers, J. F., Wichita, Kan.
Ehrlich, M., Somerset, N. J.
Eisenhauer, E. R., Union, N. J.
Erickson, R. W., Pacific Palisades, Calif.
Falotico, A. L., Woodmere, L. I., N. Y.
Feeley, T. E., Bethpage, L. I., N. Y.
Ferlan, A P., East Hartford, Conn.
Ferrandino, P. A., Huntington Station, L. I., N. Y.
Finch, E. B., Charlotte, N. C.
Fink, R. I., Northridge, Calif.
Finkel, A. G., Plainfield, N. J.
Fishbein, W., Shrewsbury, N. J.
Fishman, J. J., Waltham, Mass.
Flynn, H. G., Rochester, N. Y.
Forrester, J. E., Jr., Claremont, Calif.
Foshee, C. E., Alamogordo, N. M.
Franklin, D. L., Seattle, Wash.
Fransen, C. A., Broomall, Pa.
Freeze, J. P., Burlington, N. C.
Fuller, E. B., Webster, N. Y.
Funderburk, V. V., Fort Worth, Tex.
Fusco, A., Port Washington, L. I., N. Y.
Gadino, J. W., So. Hackensack, N. J.
Gerace, G. B., Pisa, Italy
Gilmore, C. P., Salt Lake City, Utah
Goldklang, S., Anaheim, Calif.
Grabl, S., Pasadena, Calif.
Gralnick, P. E. L., Philadelphia, Pa.
Grant, J. L., Eau Gallie, Fla.
Gray, D. C., Wheaton, Md.
Griffin, J. R., Stamford, Conn.
Grinnell, G. H., Menlo Park, Calif.
Gupta, K. C., Bangalore, India
Gurkewitz, A. M., New York, N. Y.
Hamilton, R. J., Halifax, N. S., Canada
Haag, H. D., State College, Pa.
Hanna, A., Bloomfield, N. J.
Hanna, W. J., III, Wayne, N. J.
Hastings, A. C., Santa Monica, Calif.
Hastings, R. H., Westwood, Mass.
Haynes, W. D., DeWitt, N.Y.
Haywood, V. E., New Orleans, La.
Hedlund, L. E., Omaha, Neb.
Heller, S. S., Haddonfield, N. J.
Helms, J. F., Flushing, L. I., N. Y.
Henderson, H. T., Arlington, Tex.
Henry, E. F., Orlando, Fla.
Heuer, E. C., Hometown, Ill.
Hicks, R. M., Anderson, Ind.
Hillock, R. T., Jr., Morristown, N. J.
Holbrook, W. A., Eau Gallie, Fla.
Holcomb, L. A., Rancho Cordova, Calif.
Horne, J. F., Orlando, Fla.
Humphrey, H. J., Framingham, Mass.
Iadicola, A. D., Riverton, N. J.
Iannuzzi, W. P., Cherry Hill, N. J.
Ingham, S. A., Moorestown, N. J.
Jagoda, N. H., Bedford, Mass.
Jamison, J. K., Syracuse, N. Y.
Jaselli, P., Rome, Italy
Jennings, F. W., Franklin, Mass.
John, W. W., Berlin, Zuehlendorf, Germany
Johnston, W. H., Baltimore, Md.
Kaito, Y., Kawasaki, Japan
Kane, D. L., Akron, Ohio
Kerr, J. R., Beaverton, Ore.
Kerr, M. J., Kirkland, Wash.
Kimura, T., Flushing, L. I., N. Y.
Kirilloff, V. V., Houston, Tex.
Koenig, R. L., Detroit, Mich.
Kopycienski, M., Linfield, Pa.
Kovanic, E. F., Livingston, N. J.
Kyle, J. C., Jr., Mill Valley AFS, Calif.
Lane, D. E., Clearwater, Fla.
Larson, B. G., Scotch Plains, N. J.
Larson, E. V., Bedford, Mass.
Latta, S. C., Jr., Philadelphia, Pa.
Lattanzi, E. W., Melrose, Mass.
Lavins, L. J., Ardmore, Pa.
Lee, G. W., Sudbury, Mass.
Le Gall, J. L., Cambridge, Mass.
Leong, K. O., Montreal, Canada
Levine, D. S., North Brunswick, N. J.
Levine, R. E., Newark, N. J.
Lilienthal, P. F., Hamburg, Pa.
Lohner, R. W., Golden, Colo.
Lubarsky, A. R., Utica, N. Y.
Macdonald, R. B., Lynbrook, L. I., N. Y.
Malarik, R. W., Grand Rapids, Mich.

(*Continued on page 140A*)

SEE IT

• PRESSURE
• STRAIN • DISPLACEMENT
• ACCELERATION • VIBRATION

any transducer-sensed variable shows up clearly on this

Sanborn® Amplifier-Indicator

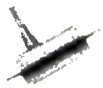
Here is a convenient, economical way to accurately measure variables from Sanborn and other makes of transducers. The Model 311 Amplifier-Indicator consists of an oscillator, high-gain carrier type amplifier, phase-sensitive demodulator, power supply and a large, multi-scale meter for direct-reading accuracy and ease. (A jack permits connection of an external recording or display device.) The instrument is entirely contained in a rugged 9″ x 6″ x 8″ high metal housing with leather handle. Weighing only 8 lbs., the Model 311 is an ideal portable test instrument for laboratory or field use.

Sensitivity . . . . ±250 μv full scale
Attenuator Ratios 2, 5, 10, 20, 50, 100, 200
Carrier Frequency . . . . . . 2400 cps internal, ±2% at 5 V rms
Output Monitor . . . . ±3 volts across load 1000 ohms or higher
Transducer Impedance . . . . 100 ohms min
Calibration . . . . 40 μv/volt of excitation
Zero Suppression . . . . . . Positive or negative (uncalibrated)
4-position switch; each position equal to approx. 5 turns of R BAL control (acts as coarse balance adjustment)
Power Requirements . . . . . 115 volts ±10%, 50-400 cps, 7 watts

Complete, Ready-To-Use Displacement Probe Transducers — Compact Model 580 (flange-mounted) and 581 (threaded) register 0.001″ linear displacement full-scale; 0.5% linearity; ±0.050″ stroke range; optional mounting block and carrier adapter.

Differential Transformers with Strokes up to 20″ — Rugged Linearsyn® series, shielded and unshielded; stroke lengths of ±0.005″ to ±1.0″ in stock, longer on special order; measures displacements as small as 0.000005″.

Pressure Transducers, Differential and Single-Ended — Series 267 and 268 for operation in carrier system have sensitivities of 40 μv/mm Hg/volt excitation and 40 μv/0.1 mm Hg/volt excitation; compact Monel case with standard Luer connectors; working ranges −100 to +400 mm Hg and −40 to +40 mm Hg.



Low Force Transducers — 5 standard ranges 0-1 gram (Model FTA-1) through 0-100 (FTA-100); displacement of 0.010″ full-scale.

OTHER SANBORN TRANSDUCERS (not carrier excited)

Miniature DC Differential Transformer — Built-in carrier system saves space, power and cost, eliminates phase and balancing problems; Model 7 DCDT-050 has ±0.050″ displacement range (other models, to ±3″); hermetically sealed, occupies less than ½ cu. in.

LVsyn® Velocity Transducers — Extremely durable, immersible units; 12 models with strokes 1.3″ to 22″, longer on special order; specials for high pressures, temperatures.

TRANSDUCER DIVISION
SANBORN COMPANY
175 Wyman Street, Waltham 54, Massachusetts
A SUBSIDIARY OF HEWLETT-PACKARD COMPANY

Membership

(Continued from page 138A)

Markiewicz, T. C., Merchantville, N. J.
Marsello, P., Haddonfield, N. J.
Marze, B. T., Pittsfield, Mass.
Massey, G. R., Los Angeles, Calif.
McClure, R. H., Wynnewood, Pa.
McCormick, D. E., State College, Pa.
McCown, R. A., Englewood, N. J.
McElroy, J. S., Medway, Mass.
McFarland, R. L., San Diego, Calif.
McLeod, A. D., Wakefield, Mass.
McMillen, G., Massapequa Park, L. I., N. Y.
McVay, J. L., Denver, Colo.
Meisels, M. W., New York, N. Y.
Menkes, B., White Plains, N. Y.
Merz, C. O., Deer Park, L. I., N. Y.
Mevel, J. Y., Rennes, France
Miley, F. P., Camarillo, Calif.
Miller, R., Orange, Calif.
Miller, W. C., Toronto, Ont., Canada
Milling, P. E., Houston, Tex.
Milner, I. Z., Lansdale, Pa.
Minners, C. J., St. Louis, Mo.
Molenda, J. T., Scranton, Pa.
Moon, M. F., Anaheim, Calif.
Morrow, G. E., Greenville, Tex.
Moses, L. F., New Hartford, N. Y.
Mowery, C. S., Jr., Winston-Salem, N. C.
Murphy, B. M., Cedar Rapids, Iowa
Murty, Y. S., Berhampur, Ganjam, India
Myers, B. L., Inglewood, Calif.
Nault, J. G., Montreal, Que., Canada
Nederostek, C. J., Philadelphia, Pa.
Neuman, J. B., Northridge, Calif.
Newell, B. M., Chelmsford, Mass.
Noll, P. C., Pennsauken, N. J.
Noveske, T. J., Seven Hills, Cleveland, Ohio
Ohlon, R. R., Stockholm, Sweden
Oishi, K., New York, N. Y.
Oliveto, F. E., Philadelphia, Pa.
Olmstead, G. A., Ottawa, Ont., Canada
Olsen, S. E., Newark, N. J.
Orr, C. B., Jr., Annandale, Va.
Paczynski, J. S., Montreal, P. Q., Canada
Padgett, C. L., Alexandria, Va.
Paljug, J. W., Falls Church, Va.
Paoli, T. L., Providence, R. I.
Parker, D. C., Seattle, Wash.
Peirce, R. W., LaGrange, Ill.
Pittman, C. W., Los Angeles, Calif.
Polkinghorn, F. A., Jr., Washington, D. C.
Pouncey, B. B., Jr., Westport, Conn.
Price, G. W., Huntsville, Ala.
Prochniak, E. J., Las Vegas, N. M.
Ralph, E. L., Sylmar, Calif.
Rao, G. M. S. V., Ithaca, N. Y.
Ratto, J. J., Bedford, Mass.
Reid, G. R., Norristown, Pa.
Renn, C. W., San Pedro, Calif.
Rhodes, F. L., Pomona, Calif.
Rice, R. W., Huntington Park, Calif.
Richards, C. A., Houston, Tex.
Richardson, G. A., Ossining, N. Y.
Richmond, R. G., Tiburon, Calif.

(Continued on page 142A)

introducing
a new Corporation
for precision production
of measuring and
testing equipment...

Western Electrodynamics

In 1955, Claude M. Hathaway — designer of many leading electronic instruments for science, industry and public utilities — sold his company and retired to his working ranch at Whitewater, Colorado. Five years later, in the seclusion of a complete engineering laboratory which he built at the ranch for the purpose, he worked on new instruments, improved their performance and made startling new discoveries. Repeated testing and checking proved the reliability of a broad new line of products.

Now a new corporation — Western Electrodynamics — has been formed to manufacture and market this wide product line. The temporary manufacturing facility is in production now. A completely modern plant is under construction and will be in operation by summer of 1962.

Western Electrodynamics is your assurance of the matchless Hathaway reputation in a well-founded organization of capable and dedicated men of experience.

New and improved products from Western Electrodynamics include:

Automatic fault-recording systems — To the electric power industry, Western Electrodynamics offers new and advanced equipment for automatically recording faults and abnormal conditions and for monitoring the performance of systems and components under fault conditions. Advances include simplified designs, higher accuracy, faster starting, and immediate record access.

New transient recording oscillograph — For general laboratory use, Model OSD provides increased resolutions of short, fast transients.

Scopograph — A laboratory-quality cathode ray oscilloscope with built-in recording features for standing waves and transient phenomena. Every laboratory needs this most versatile and convenient instrument.

Telemetry and remote-control systems — To industry Western Electrodynamics brings the telemetry developments of the space age. **Digitor 300** monitors an indefinite number of on-off functions over a single radio or metallic channel. **Digitor 500** transmits an indefinite number of measurements and control commands over a single radio or metallic channel. **Digitor** systems find an almost unlimited application in industry for remote measurement and control.

Electronic weighing equipment for industry — Products include electronic weighing components which reduce cost and complexity of weight-batching plants, improved load cells, and built-in motor truck scale which provides an indication of cargo weight at all times.

Strain measuring instruments and systems — Improved and miniaturized equipment for the recording of dynamics strain in machine parts of all kinds.

Measuring and testing equipment — This category includes electrical insulation measuring instruments, vibration and strain measuring devices, and electrical standards equipment.

Write today for 24 page color Facilities Brochure on the plant, its equipment, and key personnel.

Western Electrodynamics
INCORPORATED

P. O. BOX 98R, COLORADO SPRINGS, COLORADO

Membership

(*Continued from page 140A*)

Riddle, H. S., Trenton, N. J.
Roelse, A. L., Dallas, Tex.
Rogers, P. M., Greensboro, N. C.
Rollinson, W., Little Hulton, Manchester, England
Rosenow, A. C., Pomona, Calif.
Rosner, A., Great Neck, L. I., N. Y.
Rush, C. H., Jr., El Paso, Tex.
Sakamoto, Y., Princeton, N. J.
Sakuraba, M., New York, N. Y.
Sakus, P. J., Los Angeles, Calif.
Scales, J. R., Pasadena, Calif.
Schmitz, P. M., Jr., Riverdale, Md.
Schrader, E. N., Utica, N. Y.
Schultze, R. W., Levittown, Pa.
Schulz, R. F., Levittown, N. J.
Schwartz, A. A., Renton, Wash.
Schwartz, H., Long Branch, N. J.
Schwartz, J. L., New York, N. Y.
Schwarz, A. F., Delft, Holland
Scruggs, D. H., Hazlet, N. J.
Seckelmann, R., Schwelm/Westfalen, Germany
Sekyra, E., Rye, N. Y.
Selwyn, D., Oakland, N. J.
Sharek, H. A., Wayland, Mass.
Shaw, W. K., Hollywood, Calif.
Sherman, P. E., Winston-Salem, N. C.
Short, M. E., North Highlands, Calif.
Shuster, R. F., Mentor, Ohio
Siegert, R. H., Schenectady, N. Y.
Sieron, R. L., Centrebrook, Conn.
Silverberg, M., Riverton, N. J.
Simon, R., Floral Park, L. I., N. Y.
Simpson, J. E., Jr., Pomona, Calif.
Singer, A., Washington, D. C.
Skurnowicz, B. A., Gouldsboro, Pa.
Slattery, D. C., Garden City, L. I., N. Y.
Snow, E. J., Herkimer, N. Y.
Spencer, T. H., Camillus, N. Y.
Staas, G. L., Fort Worth, Tex.
Stahl, D. N., Hicksville, L. I., N. Y.
Stapp, F. E., Richardson, Tex.
Starks, K. H., Tallmadge, Ohio
Streifel, R. D., Los Angeles, Calif.
Stucki, F. F., Palo Alto, Calif.
Sullivan, G. A., Middlesex, N. J.
Swanson, R. W., Washington, D. C.
Szostak, J. J., Colonia, N. J.
Tankersley, J. C., Baltimore, Md.
Taylor, J. V., Jr., Livingston, N. J.
Teitelman, B., Waltham, Mass.
Tesar, V. R., Franklin Park, N. J.
Thomas, J. J., Princeton, N. J.
Thomas, J. S., Nashville, Tenn.
Thornton, P. W., Wichita, Kan.
Tischkewitsch, J. G., New York, N. Y.
Toepel, D. G., Seattle, Wash.
Toler, J. C., Huntsville, Ala.
Truffert, A., Montrouge, Seine, France
Tucker, S. S., Eau Gallie, Fla.
Upton, L. S., Sunnyvale, Calif.
Vigneron, H., Brussels, Belgium
Vlismas, T., Bronx, N. Y.
Volertas, V. F., Riverside, N. J.
Vyska, R. M., Sudbury, Mass.
Wakeman, B. F., Auburn, Wash.
Wallmeyer, E. W., Cheektowaga, N. Y.
Webb, D. C., Baltimore, Md.
Webster, R. B., Hempstead, L. I., N. Y.
Weiss, M., Mount Vernon, N. Y.
Weisz, T., Cranford, N. J.
Wenzel, J. H., Ithaca, N. Y.
Wheeler, J. W., Greensboro, N. C.
Whisler, N. L., Fort Wayne, Ind.
White, J. B., Huntsville, Ala.
Whitelock, H. E., Laurel, Md.
Williams, D. A., Jr., San Diego, Calif.
Wilson, R., McLean, Va.
Windham, T. W., Jr., Bloomingdale, N. J.

(*Continued on page 144A*)

for your mass-production wire stripping...

WORLD LEADERS IN

automatic

WIRE PROCESSING

EQUIPMENT

ARTOS MACHINES WILL

★ *MEASURE LENGTHS*

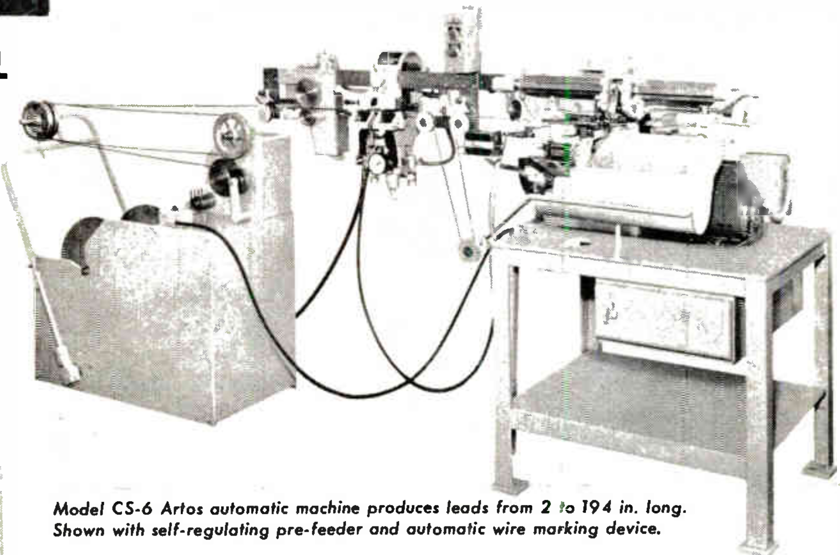
★ *CUT OFF LEADS*

★ *STRIP INSULATION*

...at speeds up to

8,000

pieces per hour!



Model CS-6 Artos automatic machine produces leads from 2 to 194 in. long. Shown with self-regulating pre-feeder and automatic wire marking device.

OPTIONAL AUXILIARY EQUIPMENT

This automatic equipment profitably adapts Artos machines to all of your individual job conditions:

DE-REELING DEVICE—Maintains even tension under all conditions.

MOTOR DRIVEN PRE-FEEDER—For extra heavy reels; self-regulating.

2-PLANE WIRE STRAIGHTENER—For very stiff wires; simplifies collection.

TWIST RETAINER—Saves costly handling.

WIRE MARKER—Identifies leads without color coding or tagging.

TERMINAL ATTACHING EQUIPMENT—Follows stripping in one continuous cycle.

PRE-DETERMINED COUNTER—For automatic machine shut-off.

YOU CAN ECONOMICALLY mass-produce finished wire leads ... to any specifications ... with unskilled labor ... on these Artos automatic processing machines.

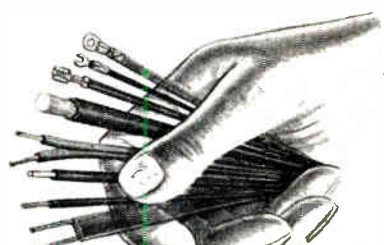
Leads are accurately measured, cut off and stripped without nicking the conductor—*at speeds up to 8,000 pieces per hour!* On some machine models, *terminal-attaching* can also be performed automatically in the same continuous cycle.

Artos machines handle almost any kind of wire without major set-up changes—solid or stranded, single or multiple conductor, with rubber, plastic, asbestos or braided insulation. Wire of 12-gauge and smaller can often be fed two at a time to double the output.

BROAD WORK RANGE

To fit your exact needs, the Artos line includes machines in a broad range of capacities, along with optional auxiliary equipment for every special job requirement. Here is the overall work range that can be handled:

WIRE SIZES: 30 to 000 gauge.
LEAD LENGTHS: 1 in. to 90 ft.
LENGTHS STRIPPED: ⅛ in. to 10½ in.



CUT HARNESS COST

Your wire harness needs can be handled faster and at lower cost by producing the leads on Artos automatic machines. Ask for engineering recommendations—submit details of your problem for expert analysis. No obligation.

WRITE

FOR ARTOS BULLETIN

Get complete data on Artos automatic machines and auxiliary equipment—learn how they can help you solve production problems. Write for your copy of Artos

BULLETIN 555

AGENTS THROUGHOUT THE WORLD

World Leaders in Automatic Machines for Finishing Wire Leads

ARTOS ENGINEERING CO.

2735 South 28th Street • Milwaukee 46, Wisconsin

Membership

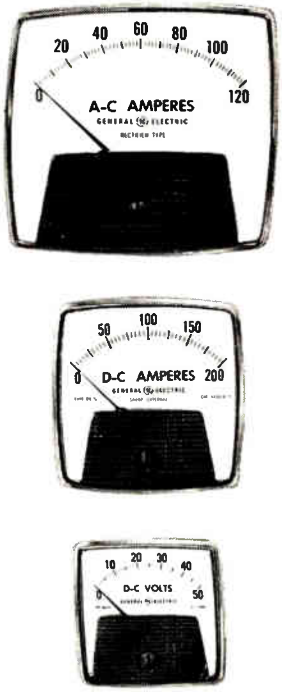
(Continued from page 142A)

Winje, R. L., Syracuse, N. Y.
Wolf, R. D., Bellevue, Neb.
Yeagle, E. W., Santa Clara, Calif.
Yee, G. N., Teaneck, N. J.
Zippel, G. A., Los Angeles, Calif.

Admission to Associate

Alexander, N., Mechanicsburg, Pa.
Allen, W. E., Villa Park, Ill.
Ammons, W., Jr., Redondo Beach, Calif.
Bardis, J. D., San Francisco, Calif.
Bates, J. H., Lexington, Ky.
Beringer, G. J., Chicago, Ill.
Berner, H. I. J., Brooklyn, N. Y.
Brennan, H. F., Cedar Grove, N. J.
Brennan, R. H., Cedar Grove, N. J.
Burk, E. G., Minneapolis, Minn.
Carlucci, R. J., Brooklyn, N. Y.
Casey, E. T., Garden City, L. I., N. Y.
Chatman, P. K., Lompoc, Calif.
Conroy, Q. H., West Hyattsville, Md.
Couchman, R. L., Malibu, Calif.
Cundiff, J. D., Sault Ste Marie, Mich.
Cvitan, R. J., Liverpool, N. Y.
De Prospo, M. V., Brooklyn, N. Y.
Depto, T. P., Patton, Pa.
Dierk, S. F., Palmyra, N. J.
East, R. J., Clearwater, Fla.
Edge, N. W., Brantford, Ont., Canada
Ervin, L. H., Shreveport, La.
Faley, R. L., Houston, Tex.
Fowler, F. A., Detroit, Mich.
Fuersr, B. W., Chicago, Ill.
Fullerton, R. T., La Jolla, Calif.
Gore, T. J., Roswell, N. M.
Gottlieb, L., Flushing, L. I., N. Y.
Graham, J. C., Washington, D. C.
Halverson, H. L., Ottawa, Ont., Canada
Harpole, T. H., Clinton, Ky.
Hicinbothem, W. A., Jr., Princeton, N. J.
Highstone, H. R., Titusville, Fla.
Hinckley, W. O., Redondo Beach, Calif.
Hollis, R. W., Philadelphia, Pa.
Horwitz, H., Havertown, Pa.
Ireland, W. E., San Francisco, Calif.
Isabelle, P. G., Shawinigan, P. Q., Canada
Jackson, F. D., Ewa Beach, Hawaii
Jetley, R. L., Delhi, India
Kmetz, R. B., Endicott, N. Y.
Kolen, O. G., Spenard, Alaska
Larson, S. J., Los Altos, Calif.
LeClerc, J. G., Ottawa, Ont., Canada
Litzenberger, M. L., West Keansburg, N. J.
MacKenzie, A., No. Vancouver, B. C., Canada
Madry, J. G., Jr., Melbourne, Fla.
Marshall, M., Brooklyn, N. Y.
Mazel, W., S/Bagneux, (Seine), France
McDonald, J. W., Jr., Pennsauken, N. J.
Meyer, W. J., Cleveland, Ohio
Miller, J. T., San Jose, Calif.
Mirabella, M. L., Rochester, N. Y.
Morris, E. C., New York, N. Y.
Morris, W. A., Dallas, Tex.
Nagle, E. M., Middletown, N. J.
Nelson, P. H., New York, N. Y.
Nichols, D. E., Jr., Dayton, Ohio
Nielsen, N. V., Copenhagen, Denmark
O'Connor, B. T., Indianapolis, Ind.
O'Connor, J. C., Arlington, Va.
Olsen, G., Mineola, L. I., N. Y.
Orlando, T. J., Brooklyn, N. Y.
Patella, F. W., Glen Rock, N. J.
Puntenney, F. D., Bonner Springs, Kan.
Rae, G. R., Cherry Hill, N. J.
Reyes, L. E., Santiago, Chile
Rheinhold, J. J., Jericho, L. I., N. Y.

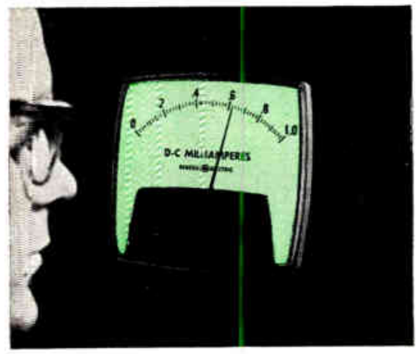
(Continued on page 146A)



BIG LOOK Panel Meters available in 2½, 3½ and 4½-inch sizes, AC, DC and rectifier-types. Custom scales and a variety of color windows.

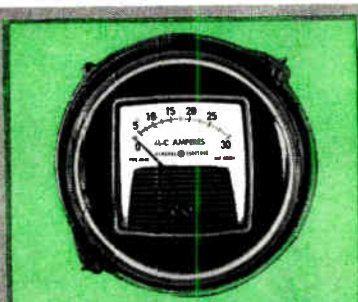
THE BIG LOOK*

PANEL AND TIME METERS FROM GENERAL ELECTRIC



New 4½-Inch AC and DC Panel Meters with mirror-backed or electroluminescent scales. Scale emits soft, even green light.

Full Line of BIG LOOK Elapsed Time Meters for match-and-lineup applications.

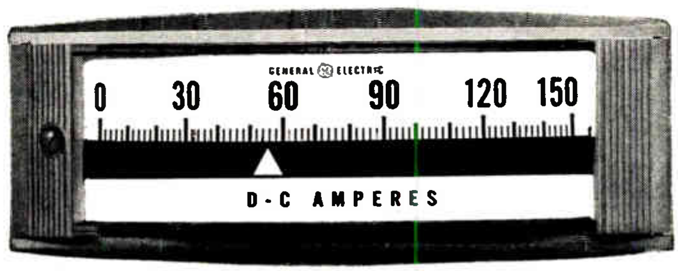


Conduit Case for 2½-inch panel and time meters. Portable test stand, too.

GENERAL ELECTRIC'S BIG LOOK PANEL AND TIME METERS, with compatible styling, bring distinctive appearance and unexcelled readability to your electrical and electronic products and equipment. ■ All BIG LOOK panel meters feature self-shielded d-c mechanism . . . sealed case on 2½- and 3½-inch models . . . snap-on, snap-off cover on 4½-inch model. Full line of a-c, d-c and rectifier-types available. ■ New BIG LOOK time meter family includes 2½- and 3½-inch reset and non-reset models with dust-proof construction . . . plus sealed models for military applications. Six extra-large digits read out from 0 to 99999.9 hours or minutes. ■ Immediate delivery from stock on all standard models and ratings . . . at low prices. Call your local G-E Sales Office or authorized Distributor today, or write for descriptive bulletins GEA-7034 (panel meters) or GEA-6710 (time meters), General Electric Company, Schenectady 5, N. Y. Canadian and foreign inquiries welcomed. *Trademark of General Electric Company.

595-06

NEW! EDGEWISE INSTRUMENTS permit single or cluster mounting . . . save up to 35% of panel space yet have big 4½-inch scale . . . require no panel drilling for stud holes. Vertical and horizontal models available.
Accuracy: ±1½% of full scale on all d-c and on a-c horizontal models, ±2% of full scale on a-c vertical model. Send for descriptive bulletin GEZ-3428.



Progress Is Our Most Important Product

GENERAL ELECTRIC

Membership

(Continued from page 144A)

Rhinebarger, L. G., Warren, Ohio
Rootenberg, J., Haifa, Israel
Rosenbaum, S., East Meadow, L. I., N. Y.
Roskelley, E. L., Odessa, Tex.
Ross, J., Beverly Hills, Calif.
Ruhlen, H. H., Anchorage, Alaska
Russo, J. R., So. Plainfield, N. J.
Rutkowski, G. B., Cleveland, Ohio
Saunders, S. G., Upper Montclair, N. J.
Schettini, L. J., Plainview, N. Y.
Schmahl, A. F., Dayton, Ohio
Segal, F., Cote St. Luc, (Mont.), Que., Canada
Shane, L. J., Philadelphia, Pa.
Shanley, J. F., Albany, N. Y.
Sheehan, L. J., Jr., Levittown, L. I., N. Y.
Smith, Don, Seattle, Wash.
Soled, M., Jersey City, N. J.
Sonnenberg, H. T., Galion, Ohio
Spani, T. L., Portland, Ore.
Steinberg, B. N., Philadelphia, Pa.
Strasser, J. A., New York, N. Y.
Stueber, F. S., Pine Bush, N. Y.
Takacs, J. M., Carlisle, Pa.
Tiersch, R. G., No. Massapequa, L. I., N. Y.
Tunwall, C. S., Jr., Marion, Iowa
Van Keuren, R. L., San Antonio, Tex.
Veroneau, W. J., Penacook, N. H.
Watson, J. L., Beaumont, Tex.
Whalley, L. E., Los Angeles, Calif.
Wheate, S. P., Biloxi, Miss.
Wrieley, B. E., Peoria, Ill.

Manufacturers' Cross-Reference Index

This index will serve two purposes:

1) It indicates corporate affiliations between firms, between a parent company and its divisions, etc.

2) It will help you in using the IRE DIRECTORY by suggesting where to look for a firm or division whose name you remember but whose listing in the Alphabetical Directory of Manufacturers may be under another name.

A

(Continued on page 148A)

TOGGLE SWITCH MODEL 20

PRECISION HIGH FREQUENCY ATTENUATORS

DC TO 500 MEGACYCLES

- 50, 70, or 90 ohm impedance
- High-Frequency, Precision Teflon & Silver Switches
- 1% Carbon Film Resistors ■ Fully Shielded Units
- Up to 101 db ■ Fixed 0 or 10 db insertion loss
- SWR: 1.2 : 1 max up to 250 mc
  1.4 : 1 max 250 to 500 mc
- Min. Insertion*: 0.1 db at 250 mc; 0.2 db at 500 mc
- Accuracy: At Full Attenuation: 0.5 db at 250 mc, 1.2 db from 250 to 500 mc
- Price: (Model 20) $79.00 f.o.b. factory ($87.00 F.A.S. N. Y.)

*Zero insertion loss (Model 20-0) $75.00 f.o.b. factory

KAY ELECTRIC COMPANY
DEPT. I-5 MAPLE AVENUE · PINE BROOK, MORRIS COUNTY, N. J. CAPITAL 6-4000

Write for Catalog Information

KAY
ELECTRONIC INSTRUMENTS

Sweeping Oscillators
Frequency Markers

Audio Spectrum Analyzers

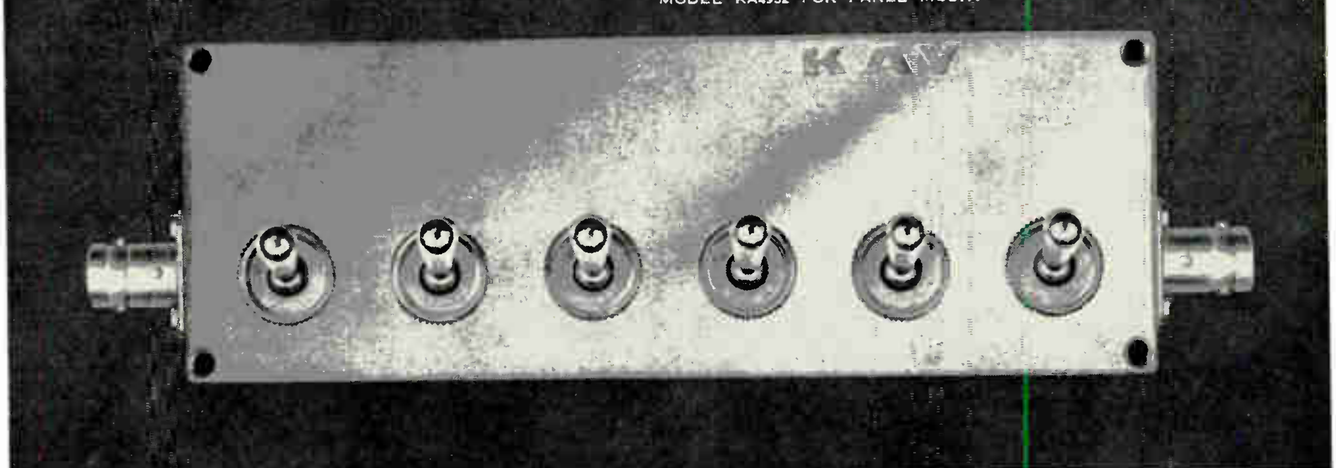
Pulse — Pulsed RF Generators

Precision Attenuators

Random Noise Generators

Kilo-Sweeps®
Rada-Sweeps®
Ligna-Sweeps®
Sona-Sweeps®
Marka-Sweeps®
Vari-Sweeps®
Mega-Sweeps®

MODEL KA4952 FOR PANEL MOUNT



Manufacturers' Cross-Reference Index

(Continued from page 146A)

(Continued on page 150A)

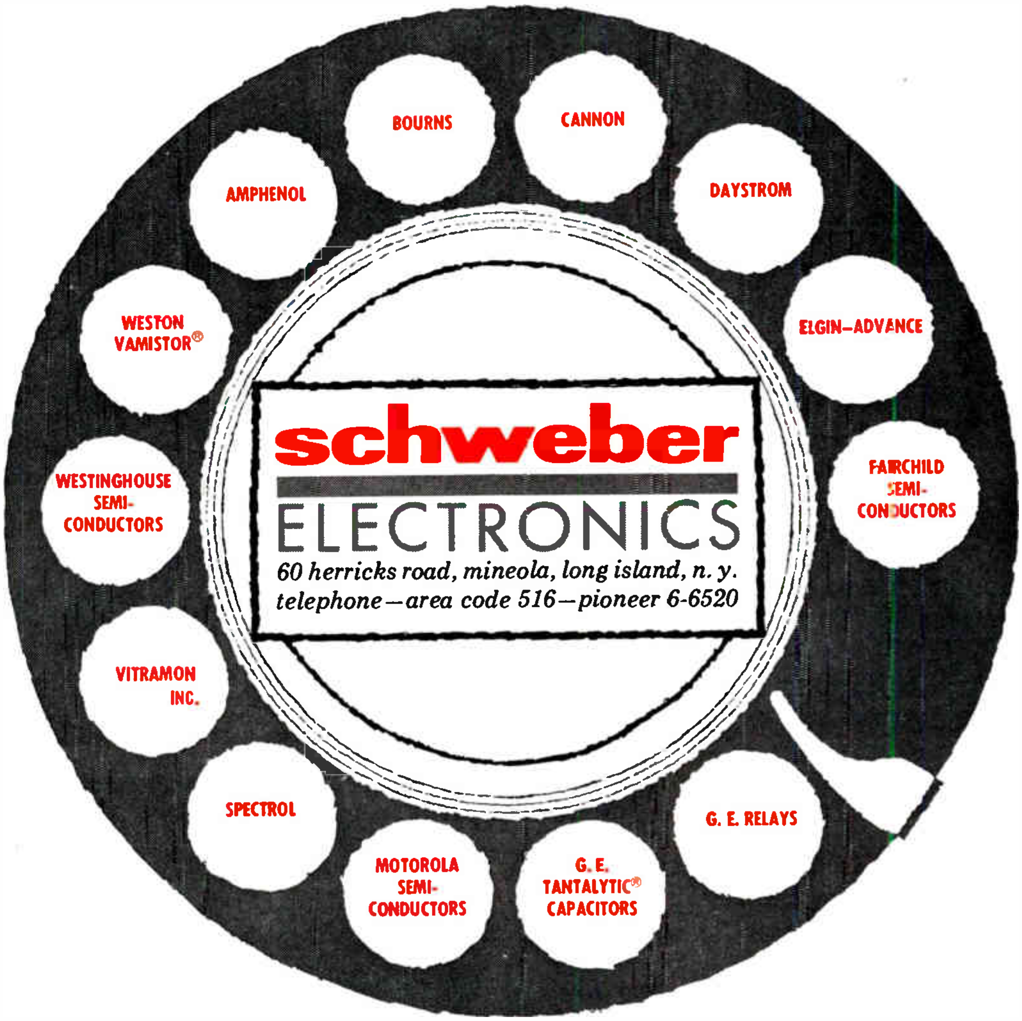
Manufacturers' Cross-Reference Index

(Continued from page 148A)

(Continued on page 153A)

Use your IRE DIRECTORY! It's valuable!

THIS AD WILL SAVE OUR CUSTOMERS OVER $100,000 A YEAR!



If you're in any of the states listed below, you can call SCHWEBER free! Just dial the special "enterprise" number shown and you will be connected directly to our New York Office, at no charge to you (that's even less expensive than a local call!). Remember, Schweber Electronics is as near as the phone on your desk. Call us first for expert component purchasing guidance and the very latest product information.

STATE	FREE PHONE
ARIZONA Phoenix	ENterprise #2151
CONNECTICUT Essex, Hartford, Meriden, Norwalk, Stamford, Windsor Locks	ENterprise #2606
DISTRICT OF COLUMBIA	ENterprise #1-6520
FLORIDA St. Petersburg	ENterprise #6520
ILLINOIS Aurora, Barrington, Champagne Chicago, Elmhurst, Lemont, Maywood, Morton Grove, Oaklawn, River Forest, Skokie, St. Charles, Tinley Park, Urbana, Wheaton	ENterprise #4710
INDIANA Fort Wayne	ENterprise #6520
IOWA Cedar Rapids	ZEnith #6520
MARYLAND Baltimore	ENterprise #93685
MASSACHUSETTS Alpine, Andover, Bigelow, Boston, Brighton, Burlington, Cambridge, Decatur, Lasell, Lexington, Maynard, Needham, Newton, Sudbury, Taunton, Twinbrook, Ulysses, Waltham, Watertown, Wayland, Wellesley, Wilmington	ENterprise #1440
MASSACHUSETTS Pittsfield	ENterprise #6440
MINNESOTA Minneapolis, St. Paul	ZEnith #5266
MISSOURI St. Louis Suburban	ENterprise #8495
NEW HAMPSHIRE Nashua	ENterprise #8106
NEW JERSEY Asbury Park, Bloomfield Camden, Eatontown, Elizabeth, Essex (No. 1 & 7), Fairlawn, Glen Ridge, Hasbrouck Heights, Holmdel, Kearney, Long Branch, Metuchen, Montclair, Moorestown, Nutley, Passaic, Paterson, Plainfield, Plainsboro, Princeton, Rahway, Red Bank	WX #3838
NEW YORK Binghamton, Buffalo, Elmira, Endicott, Ilion, Johnson City, Kingston, Liverpool, Owego, Plattsburg, Poughkeepsie, Rochester, Schenectady, Syracuse, Utica	ENterprise #4185
NORTH CAROLINA Charlotte, Raleigh	WX #3838
OHIO Cambridge, Cleveland, Columbus, Middletown	ENterprise #6523
PENNSYLVANIA Philadelphia	ENterprise #6339
PENNSYLVANIA Clifton Heights, Fort Washington, Norristown, Paoli	ENterprise #10137
PENNSYLVANIA York	ENterprise #20640
RHODE ISLAND Newport	ENterprise #3032
TENNESSEE Oak Ridge	WX #6520
TEXAS Dallas, Fort Worth, Grand Prairie, Richardson	ENterprise #1035
VIRGINIA Lynchburg, Roanoke, Waynesboro	ENterprise #172
VIRGINIA Hampton	ENterprise #1-6520
WISCONSIN Janesville, Lake Geneva, Milwaukee	ENterprise #1743

P.S. IF YOUR AREA IS NOT LISTED PLEASE LET US KNOW.

FREQUENCY CONTROL COMPONENTS & SYSTEMS

Bulova Electronics is primarily concerned with research, design, development and manufacture of frequency control devices, components and systems to meet the highest military and industrial specifications. Bulova experience in mastering many of the most difficult problems in miniature design and missile reliability can prove of immense value in your particular program. For more information on the products listed below, design or engineering assistance, write Department 2370.

BULOVA ELECTRONICS
61-10 Woodside Avenue
Woodside 77, N. Y.

BULOVA QUARTZ CRYSTALS

Frequency Range	Cut	Mode of Vibration	Overtone Order
2kc to 50kc	PT	W-1 flexure	Fundamental
6kc to 300kc	NT	W-1 flexure	Fundamental
6kc to 88kc	+5°X	W-1 flexure	Fundamental
90kc to 250kc	+5°X	1 extensional	Fundamental
80kc to 500kc	DT	face shear	Fundamental
100kc to 400kc	GT	extensional	Fundamental
150kc to 600kc	CT	face shear	Fundamental
500kc to 21mc	AT	thickness shear	Fundamental
1.5mc to 61mc	AT	thickness shear	3rd
2.5mc to 100mc	AT	thickness shear	5th

BULOVA FILTERS

Type	Model	Frequency Range	Band Width
1 NS	AD-23A	7kcs to 200kcs	.01% to .2%
		70kcs to 200kcs	.01% to .5%
	TG 9A	190kcs to 10mcs	.01% to .2%
	TG 13C	190kcs to 400kcs	.2% to .4%
2 NS	RP 13B	10kcs to 200kcs	.02% to .4%
	TG 9B	190 kcs to 5mcs	.02% to .4%
	AK 7A	5mcs to 20mcs	.02% to .4%
Reject	1R	180kc to 10mc	.02% to .4%
	2R	180kc to 10mc	
	4R	180kc to 10mc	
	6R	180kc to 10mc	
4NS	SL22	70kc to 200kc	.02% to 0.8%
	ME 25B	60kc to 200kc	.02% to 0.8%
	SL 15	180kc to 2mc	.02% to 0.5%
	LT 13A	250kc to 5mc	.02% to 0.5%
	AS 15A	1mc to 5mc	.01% to 0.5%
	WC 10A	5mc to 20mc	.02% to 0.5%
8NS	CE 30	60kc to 80kc	.02% to 0.8%
	CE 20	80kc to 200kc	.02% to 0.8%
	ME 30	180kc to 5mc	.02% to 0.5%
	AS 15A	1mc to 5mc	.01% to 0.5%
	WC 10A	5mc to 20mc	.02% to 0.5%
Single Side Band	5WL and 5WU		
	6WL and 6WU		

BULOVA OSCILLATORS

Bulova crystal controlled oscillators provide the "formula" for precise frequency and time regulation. High precision, sophisticated units, Bulova oscillators satisfy the stringent requirements arising out of higher speeds, decreased instrument time, tight frequency and time control references, shrunken communication and guidance spectrums.

Bulova specializes in tailoring frequency and time control packages and systems to specific requirements: supply output frequencies as high as 100 mc with frequency multipliers; frequencies as low as .001 cps with binary dividers. Stabilities as high as 2 pps 10^9 with very low aging rates are feasible. Miniaturization, too.

BULOVA RF COILS & XFORMERS

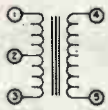
TYPE I

True L 4-6
20MH
22MH
3.4MH
14.8MH
70MH
15MH
11MH
20MH

TYPE II

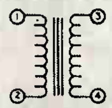
True L 4-6
65MH
16MH
20MH
65MH
150MH
30MH
15MH
5MH

TYPE III



True L 4-6	True L 2-3
80MH	307μH
15MH	250μH
15MH	940μH
5MH	424μH
400μH	100μH
20MH	1.25MH
6MH	480μH
18MH	4.5MH
37.5MH	2.34MH
3.6MH	18μH
70MH	17.5MH
18MH	88μH
15MH	3.75MH
50MH	12.5MH
2.5MH	57μH
3.5MH	61μH
35MH	8.75MH
5MH	78μH
7MH	100μH

TYPE IV



True L 4-6	True L 2-3
4.25MH	4.25MH
30μH	2.9μH
30μH	5.5μH
30μH	20μH
9MH	640μH

BULOVA OVENS

Model	Dimensions	Temperature Range*
BG 100	15/16" ID x 3/4"	−55°C to 5°C
BG 100 NLS	15/16" ID x 1 1/8"	−55°C to 5°C
BG 200	15/16" ID x 1 1/2"	−55°C to 5°C
BG 200 4N	15/16" ID x 2 1/4"	−55°C to 5°C
FC 100	3/8" x 13/16" x 3/4"	−55°C to 5°C
FC 200	3/8" x 13/16" x 1 1/2"	−55°C to 5°C
FC 200 SD	3/8" x 13/16" x 1 1/2"	−55°C to 5°C
AB 200	3/8" x 3/4" x 3/4"	−55°C to 5°C
AM 100	1 15/16" ID x 3 15/16"	−55°C to 5°C
AM 100 3S	1 15/16" ID x 3 15/16"	−55°C to 5°C
AM 100 4S	1 15/16" ID x 3 15/16"	−55°C to 5°C
AM 125	1 15/16" ID x 3 13/16"	−55°C to 5°C
AM 125 4S	1 15/16" ID x 3 13/16"	−55°C to 5°C
AM 150	1 15/16" ID x 3 15/16"	−55°C to 5°C
AM 175	1 15/16" ID x 3 13/16"	−55°C to 5°C
AM 175 3S	1 15/16" ID x 3 13/16"	−55°C to 5°C
AM 200	1 3/32" ID x 1 7/8"	−55°C to 5°C
AM 200 NL	1 3/32" ID x 2 1/4"	−55°C to 5°C
AM 200 SE	1 3/32" ID x 1"	−55°C to 5°C
AM 300	1 7/16" ID x 3 15/16"	−55°C to 5°C
AM 400	15/16" x 1 7/16" x 3 1/4"	−55°C to 5°C
AM 500	2 7/16" x 2 7/16" x 4 1/4"	−55°C to 5°C
AM 600	1 15/16" ID x 3 15/16"	−55°C to 5°C
AM 700	1 15/16" x 1 15/16" x 4 3/8"	−55°C to 5°C
BDX 10	Accommodates 4 diodes & 2 zener diodes	−55°C to 5°C
BDX 10 6	Accommodates 6 diodes	−55°C to 5°C
BDX PC	Accommodates 4 diodes	−55°C to 5°C
COM 175 7	5/8" x 13/16" x 13/16"	−55°C to 5°C

*Below nominal setting of oven temperature

Manufacturers' Cross-Reference Index

(Continued from page 150A)

American Phenolic Corp.—Amphenol-Borg Electronics Corp.
American Platinum & Silver Div.—Engelhard Industries, Inc.
American Precision Industries, Inc.—Delevan Electronics Corp.
American Pulley Co.—Van Norman Industries, Inc.
American Radiator & Standard Sanitary Corp.—American-Standard Controls Div.
American Research & Development—Textron Electronics, Inc.
American Research & Development Corp.—Intercontinental Electronics Corp.
American Seal Kap Corp.—Hardwick, Hindle, Inc.
American Semi-Conductor Co.—Spec-Tronics
American-Standard—Advanced Technology Labs.
American Steel & Wire—United States Steel Corp.
American Super-Temperature Wires, Inc.—Haveg Industries, Inc.
American Time Products, Inc.—Bulova Watch Co.
American Transformer Div.—Standard Electronics Div.; Dynamics Corp. of America
Amersil Quartz Div.—Engelhard Industries, Inc.
Amertran Div.—Dynamics Corp. of America
Ames Mfg. Corp.—Saxton Products, Inc.
Amherst Labs.—Sylvania Electric Products Inc.
Amperex Electronic Corp.—North American Philips; Price Electric Corp.
Ampex Corp.—Orr Industries; Telemeter Magnetics, Inc.
Amphenol-Borg Electronics Corp.—Borg Equipment Corp.; FXR Div.; Industrial Products-Danbury Knudsen Div.; RF Products Div.
Amphenol Electronics Corp.—Amphenol-Borg Electronics Corp.
Anaconda Wire & Cable Co.—American Brass Co.; Sequoia Wire & Cable Co.
Anatran Co.—Digitran Co.
Anchor Metals Div.—Kennedy & Co. (D. S.)
Anco Technical Writing Service, Inc.—Anelex Corp.
Andover Industries, Inc.—Erie Resistor Corp.
Andrew Antenna Corp.—Andrew Corp.
Anemostat Corp.—Standard Electronics Div.; Dynamics Corp. of America
Anodyne, Inc.—North Shore Nameplate Div.
Ansco Div.—General Aniline & Film Corp.
Antennex Co.—Clear Beam Antenna Co.
Anti-Friction Bearings, Inc.—Industrial Tectonics, Inc.
Anton Electronic Labs., Inc.—Lionel Electronic Labs.; Lionel Corp.
Anton Imco Electronic Corp.—Anton Electronic Labs., Inc.; Anton Machine Works
Anton-Imco Div.—Lionel Corp.
Antran Div.—International Electronics Mfg. Co.
Applied Dynamics, Inc.—Bowmar Instrument Corp.
Applied Electronics Corp.—Wayne-George Corp.
Applied Research Labs.—Bausch & Lomb, Inc
Applied Science Co. of Princeton—ASCOP Div.
Applied Science Corp.—Electro-Mechanical Research, Inc.
Arbame-Mallory, S. A.—Mallory & Co., Inc. (P. R.)
Arch Instrument Co., Inc.—Arch Gear Works, Inc.
Arco Capacitors, Inc.—Arco Electronics, Inc.
Arcolytic Div.—Arco Electronics, Inc.
Arde Associates—Arde Engineering Div.
Ardente Acoustic Labs., Ltd.—British Radio Electronics, Ltd.
Argus Camera Div.—Sylvania Electric Products Inc.
Argus Mfg. Co.—Cabtron Div.
Ark-Les Switch Corp.—Electric Terminal Corp.
Arkay International, Inc.—Radio Kits, Inc.
Arma Div.—American Bosch Arma Corp.
Armac Gear Div.—Franke Gear Works, Inc.
Armament Div.—Universal Match Corp.
Armstrong Furnace Co.—National Union Electric Corp.
Arnold Engineering Co.—Allegheny Ludlum Steel Corp.

(Continued on page 154A)

Manufacturers' Cross-Reference Index

(Continued from page 153A)

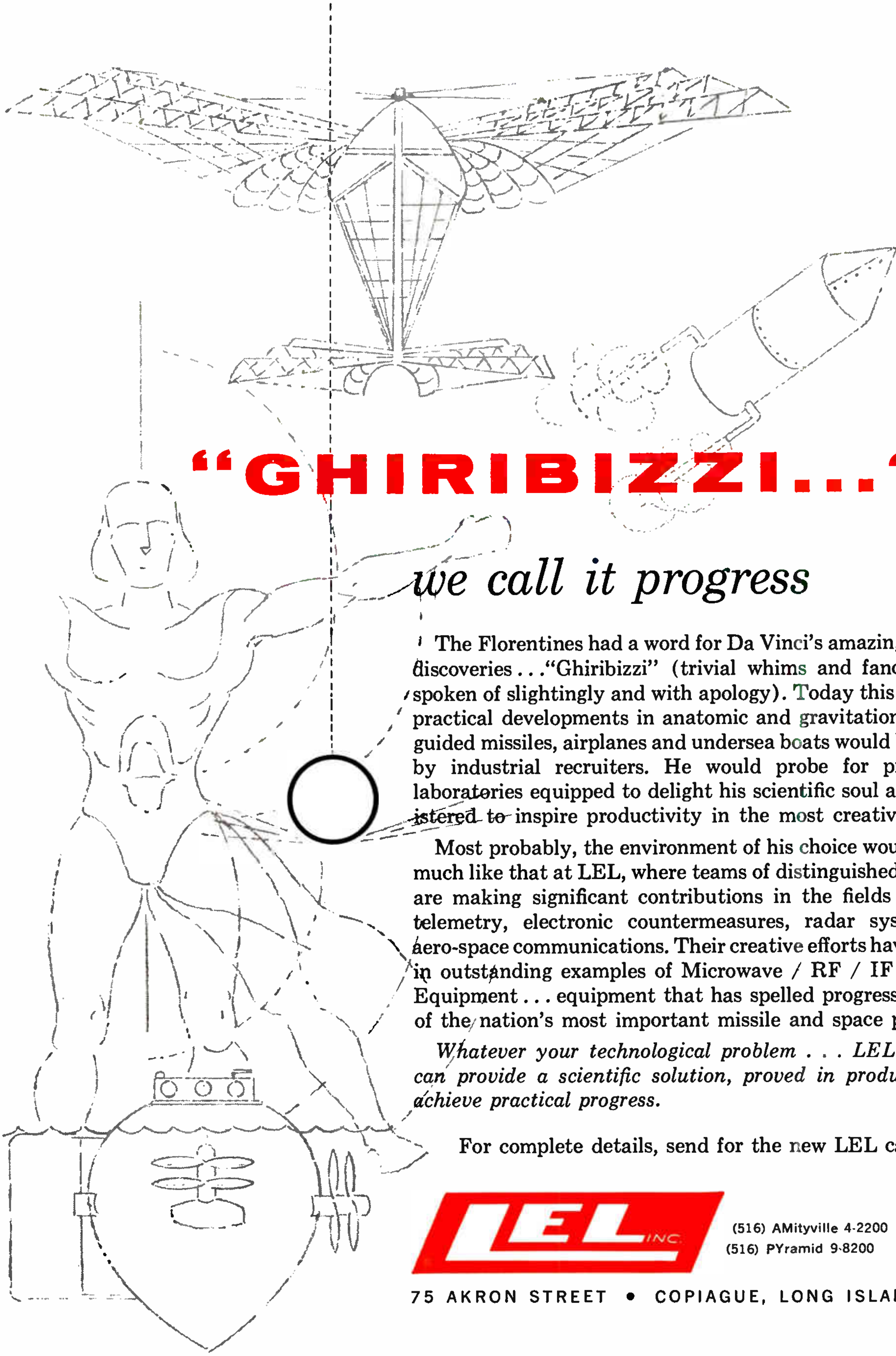
Arnold Mfg. Div.—Electronic Components Corp.
Arnoux Corp.—Automation Electronics, Inc.
Artisan Industries—Metal Fabricators Corp.
Artisan Metal Products, Inc.—Metal Fabricators Corp.
Assembly Products, Inc.—Metronix, Inc.
Associated American Trading Div.—Associated American Winding Machinery, Inc.
Associated Automation, Ltd.—Elliott-Automation, Ltd.
Associated Instrumentation & Controls, Ltd.—Automatic Timing & Controls, Inc.
Associated Missile Products Co.—Marquardt Corp., Pomona Div.
Associated Test Equipment Corp.—Associated Testing Labs., Inc.
Astra Technical Instrument Corp.—Arnoux Corp.; Automation Electronics, Inc.
Astrex, Inc.—U-Test-M Mfg. Co.
Astro-Electronics Div.—Radio Corp. of America
Astro-Science Corp.—American Avionics, Inc.
Astro-Space Labs.—Belock Instrument Corp.
Astrolab Testing Corp.—Crescent Engineering & Research Co.
Astrometrics, Inc.—Arnoux Corp.
Astronics Div.—Lear, Inc.
Astropower, Inc.—Douglas Aircraft Co.
Atchley Div. (Raymond)—American Brake Shoe Co.
Athenia Steel Div.—National-Standard Co.
Atlantic Research Corp.—Jansky & Bailey, Inc.; Northeastern Engineering, Inc.
Atlas Chain & Mfg. Co.—Atlas Precision Products Co.
Atlas Controls, Inc.—Atlas Engineering Co., Inc.
Atlas E-E Corp.—Atlee Corp.; Atlas Div.
Atlee Corp.—Applied Dynamics Corp.; Wesco Electrical Co., Inc.
Atomic Associates, Inc.—Baird-Atomic, Inc.
Atomic Development & Machine Corp.—Baird-Atomic, Inc.
Atomic Engineering Corp.—General Meters, Inc.
Atomic Labs., Inc.—Central Scientific Co.
Atomics International Div.—North American Aviation, Inc.
Atronic Products, Inc.—General Atronics Corp.
Audax, Inc.—Rek-O-Kut, Inc.
Audio Development Co.—ADC Products, Inc.
Audiotex Mfg. Co.—American Microphone Mfg. Co.; G-C Electronics Co.; Textron Electronics, Inc.
Augat Bros., Inc.—Darrill Products Corp.
Aurora Equipment Co.—Equipto Div.
Austin Co.—Austin Electronics
Auto-Timer Div.—Sloan Co.
Automark Marking Equip. Div.—Defiance Machine & Tool Co.
Automatic Coil Co.—Precise Electronics & Development Corp.
Automatic Coil Winder & Electrical Equipment Co., Ltd.—Avo, Ltd.
Automatic Electric Co.—Sylvania Electronic Systems Div.
Automatic Electric Labs., Inc.—Automatic Electric Co.
Automatic Electric Sales Corp.—General Telephone & Electronics Corp.
Automatic Mfg. Div.—General Instrument Corp.
Automatic & Precision Mfg. Corp.—A.P.M. Corp.
Automatic Seriograph Corp.—Litton Industries, Inc.
Automatic Temperature Control Co., Inc.—Automatic Timing & Controls, Inc.
Automation Corp. of America—Magnasync Corp.
Automation Electronics, Inc.—Arnoux Corp.
Automation Industries, Inc.—Jobbins Electronics Div.
Automation Instruments, Inc.—Automation Industries, Inc.
Avco Corp.—Crosley Div.
Aviation Electric, Ltd.—Bendix Corp.
Avien, Inc.—Colvin Labs., Inc.
Avion Division—ACF Electronics Div.
Avionic Div.—Oster Mfg. Co. (John)
Avis Industrial Corp.—Keystone Mfg. Co.
Axler Assoc.—Baird-Atomic, Inc.
Azoplate Div.—Engelhard Industries, Inc.

(Continued on page 156A)

Manufacturers' Cross-Reference Index

(Continued from page 154A)

(Continued on page 158A)



"GHIRIBIZZI...?"

we call it progress

The Florentines had a word for Da Vinci's amazing scientific discoveries . . ."Ghiribizzi" (trivial whims and fancies, to be spoken of slightingly and with apology). Today this pioneer of practical developments in anatomic and gravitational theory, guided missiles, airplanes and undersea boats would be courted by industrial recruiters. He would probe for progress in laboratories equipped to delight his scientific soul and administered to inspire productivity in the most creative sense.

Most probably, the environment of his choice would be very much like that at LEL, where teams of distinguished engineers are making significant contributions in the fields of missile telemetry, electronic countermeasures, radar systems and aero-space communications. Their creative efforts have resulted in outstanding examples of Microwave / RF / IF Receiving Equipment . . . equipment that has spelled progress for many of the nation's most important missile and space projects.

Whatever your technological problem . . . LEL engineers can provide a scientific solution, proved in products which achieve practical progress.

For complete details, send for the new LEL catalog.

LEL INC.

(516) AMityville 4-2200
(516) PYramid 9-8200

75 AKRON STREET • COPIAGUE, LONG ISLAND, N. Y.

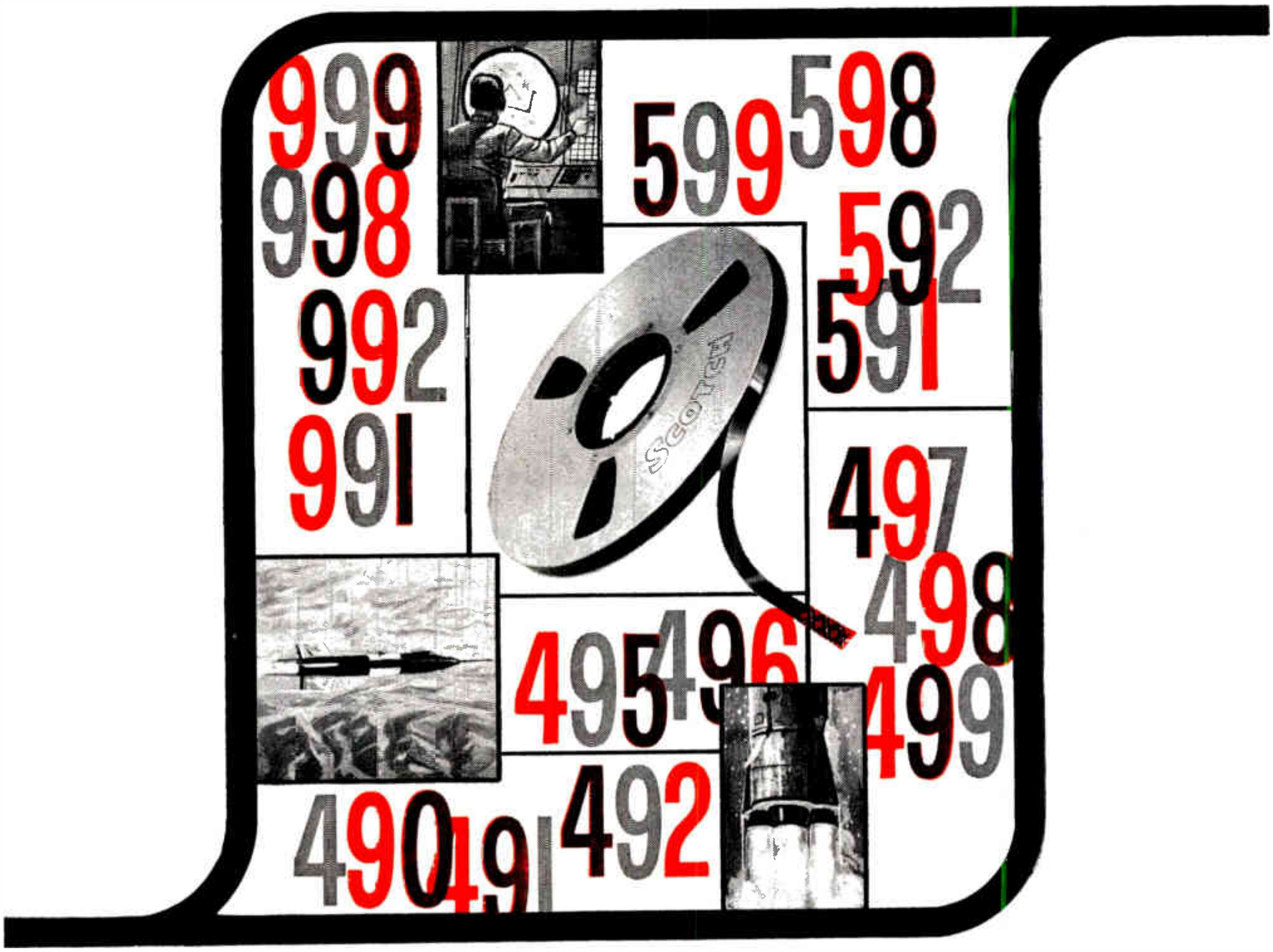
RESEARCH, DEVELOPMENT AND PRODUCTION OF LOW NOISE RECEIVING EQUIPMENT

Manufacturers' Cross-Reference Index

(Continued from page 156A)

(Continued on page 162A)

First as a matter of record . . . SCOTCH® BRAND Instrumentation Tapes



World's widest tape selection offers heavy-duty constructions for every instrumentation need!

Today's stepped up pace for data recording calls for magnetic instrumentation tapes that stay cool despite ever-increasing transport speeds, greater tensions, high heat build-up at recording heads. And the "SCOTCH" BRAND Instrumentation Tape line, with a tape for every instrumentation requirement, now includes 16 heavy-duty constructions that conquer difficult operating environments.

"SCOTCH" Heavy Duty Instrumentation Tapes are made with a special high-potency oxide and binder formulation that minimizes rub-off, withstands temperatures from —40°F to as high as 250° F! Field tests prove these tapes last a minimum of 15 times longer than ordinary tapes—capture signal with certainty despite high pressures and speeds.

The oxide coating affords nearly 1000 times greater conductivity than conventional tapes—drains off dust-attracting static charges to assure a clean tape pass every time! And exclusive Silicone lubrication protects against head wear, extends tape life! These three series of heavy duty tapes are available in a variety of widths and lengths . . .

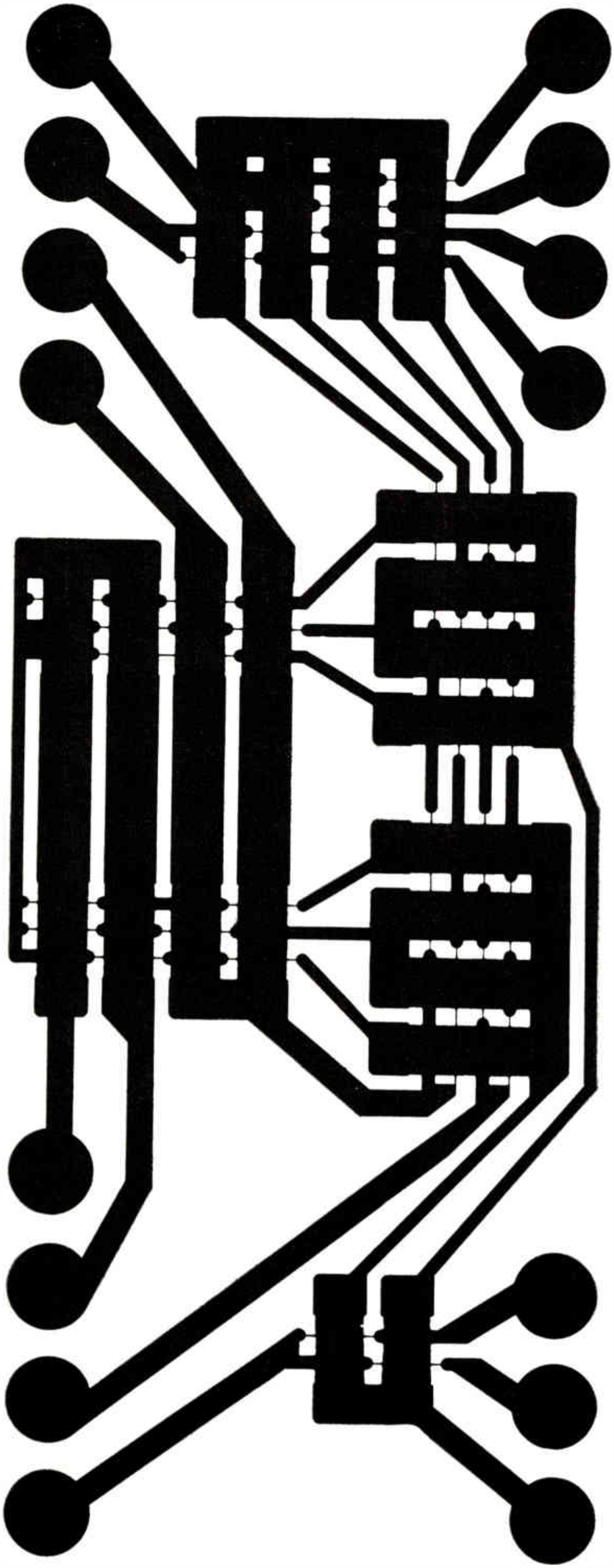
"400" series heavy duty tapes feature exceptionally long wear, excellent high and low frequency resolution. 8 constructions—.18 and .43 mil oxide coatings on .65, 1, 1.5 mil polyester backings; .56 mil coatings on 1 and 1.5 mil backings.

"500" series heavy duty tapes combine long wear, outstanding smoothness to assure sharp resolution of extremely high frequencies. 4 constructions—1 or 1.5 mil polyester backings; .18 or .43 mil oxide coatings.

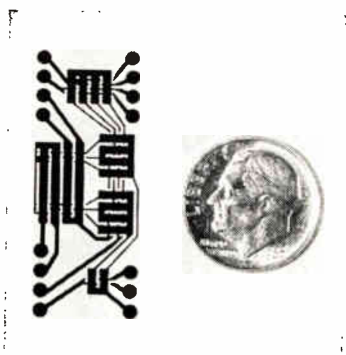
"900" series heavy duty tapes, designed especially for Mincom CM-100 series Recorder/Reproducers, provide ultra-smooth recording surfaces for critical short wave length requirements. 4 constructions—1 or 1.5 mil polyester backings; .18 or .43 mil oxide coatings.

Whatever your tape requirements—standard, high output, high resolution, sandwich or heavy duty—there's a right "SCOTCH" Instrumentation Tape. Consult your nearby 3M representative for helpful technical details. Or write Magnetic Products Division, Dept. MDA-52, 3M Company, St. Paul 1, Minn.

Magnetic Products Division 3M COMPANY



NEWS



CRYOGENIC PROGRESS in inertial guidance is typified by this 32-gate logic element for smaller, faster, more reliable computers. Other current cryogenic work includes high-precision gyros and accelerometers.

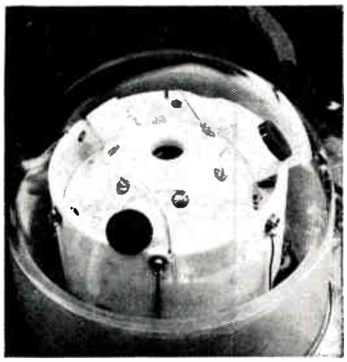
...OF DEFENSE TECHNOLOGIES

GUIDANCE

There are vast differences between the techniques required to guide missile flights and those needed for space ventures. Entirely different problems are presented in the tasks of guiding an air-to-air missile having a few thousand yards range, sending an ICBM to a target thousands of miles away, and boosting men and instruments to the frontiers of space. Experience in each is invaluable in satisfying new mission requirements.

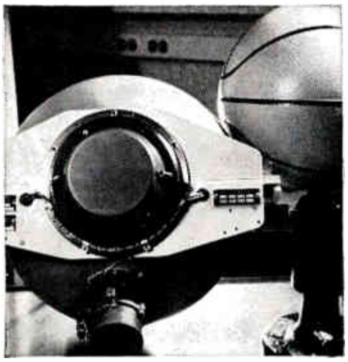
General Electric is producing infrared guidance for the Sidewinder, inertial guidance for the Polaris, and the radio-command system for Atlas ICBM's. It was this phenomenally accurate radio-command guidance that helped put the Mercury-Atlas into earth-orbit.

This wide range of guidance experience is enabling General Electric to apply new devices and techniques (such as electrostatic gyros and cryogenics) directly to existing problems, thus providing continuous and significant upgrading of guidance technologies and capabilities. Current activities include development of reference navigational systems for space and exploration of the best combinations of basic guidance forms for the sophisticated missiles and space vehicles of the future.

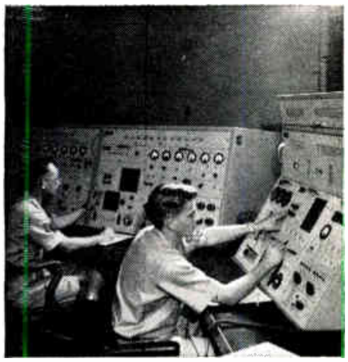


ELECTROSTATIC GYROS are being developed to meet future requirements for accuracies up to 1000 times greater than those attainable with present gyros. Other advantages will be simplicity and long life.

STAR-FIELD TRACKER correlates a viewed star pattern against a stored reference map for spacecraft attitude control. This device may replace the inertial table as an attitude sensor because of its zero drift.



INERTIAL GUIDANCE for Polaris has been extremely successful. Designed by M.I.T. and produced by General Electric, the small MK 2 system will help to extend the range of the advanced Polaris missile to 2500 miles.



RADIO-COMMAND GUIDANCE, already operational and proved accurate and reliable in ICBM and manned orbital flights, is currently being applied to booster programs and missile range instrumentation.

Progress Is Our Most Important Product

GENERAL ELECTRIC

DEFENSE ELECTRONICS DIVISION

Manufacturers' Cross-Reference Index

(Continued from page 158A)

Burlington Industries—Hess, Goldsmith & Co.
Burroughs Corp.—Control Instrument Co., Inc.
Burton Mfg. Co.—Trans Electronics, Inc.
Burton-Rogers Co.—Hoyt Electrical Instrument Works, Inc.
Bush Radio, Ltd.—Rank Cintel, Inc.
Bussmann Mfg. Div.—Edison Industries (Thomas A.); McGraw-Edison Co.
Butterworth & Sons Co. (H. W.)—Van Norman Industries, Inc.
Byron Jackson Div.—Borg-Warner Controls

C

C.B.M. Corp.—Bellandi Co., Inc. (M.)
CBS—Hytron—CBS Electronics Div.; Columbia Broadcasting System, Inc.
C&D Batteries Div.—Electric Autolite Co.
CDC Control Services, Inc.—CompuDyne Corp.
CG Electronics—Gulton Industries, Inc.
CGS Labs., Inc.—Trak Electronics Div.
CTS Corp.—Chicago Telephone of Calif., Inc.; Chicago Telephone Supply Corp.
Cadre Electric & Power Equipment Corp.—Cadre Industries Corp.
Caledonia Electronics & Transformer Corp.—Electro Networks, Inc.
Calidyne Co.—Ling Electronics; Ling-Temco-Vought, Inc.
Califone Corp.—Rheem-Califone Corp.; Rheem Mfg. Co.
California Capacitors Corp.—Electro-Capacitors Co.
California Eastern Aviation, Inc.—Chicago Electronic Div.; Dynalectron Corp.; Land-Air, Inc.
Calinoy Div.—Illinois Tool Works
Callahan Mining Corp.—Flexaust Co.
Calmag Div.—California Magnetic Control Corp.
Cambridge Div.—Airpax Electronics, Inc.
Camden Electronics, Inc.—Electro Networks, Inc.
Camerafiex Co.—Federal Mfg. & Engineering Corp.
Campbell Industries, Inc.—Clarostat Mfg. Co., Inc.
Canadian Controllers, Ltd.—Clark Controller Co.
Canadian Diaphlex Co.—Cook Electric Co.
Canadian Pacific Railway—Consolidated Mining & Smelting Co. of Canada, Ltd.
Canton Electronics—Technology Instrument Corp.
Capital Electric Products Co.—Magnatran, Inc.
Capitron Div.—AMP, Inc.
Carlisle Corp.—Tensolite Insulated Wire Co., Inc.
Carnf Co.—Potter Co.
Carol Electronics Corp.—Mohawk Business Machines Corp.
Carolina Industrial Plastics Div.—Essex Wire Corp.
Carr-Fastener Co.—Graphik-Circuits Div.; United Carr-Fastener Corp.
Carrier Corp.—Spectrol Electronics Corp.
Carter Div.—Carter Parts Co.
Cascade Div.—Monogram Precision Industries, Inc.
Casco Products Corp.—Standard Kollsman Industries, Inc.
Cass Div.—Hudson Wire Co.
Castle Electrical Instruments, Ltd.—Painton & Co., Ltd.
Cedar Engineering Div.—Control Data Corp.
Cenco Instrumenten Maatschappij, N.V.—Central Scientific Co.; Cenco Instruments Corp.
Central Electronic Mfrs.—Isotopes Specialties Co.; Nuclear Corp. of America
Central Electronics Corp.—Rauland Corp.
Central Scientific Co.—Cenco Instrument Corp.
Centroid Transformer Corp.—Tensor Electric Development Co., Inc.
Centronix, Inc.—Delta Coils, Inc., Jetronic Industries, Inc.; Winston Electronics Div.
Century Controls—Consolidated Controls Corp.
Century Geophysical Corp.—Century Electronics & Instruments, Inc.
Ceramic-Metal Assemblies Corp.—Shaw Instrument Corp.
Ceramics For Industry—CFI Corp.
Cerium Metals & Alloys Div.—Ronson Metals Corp.

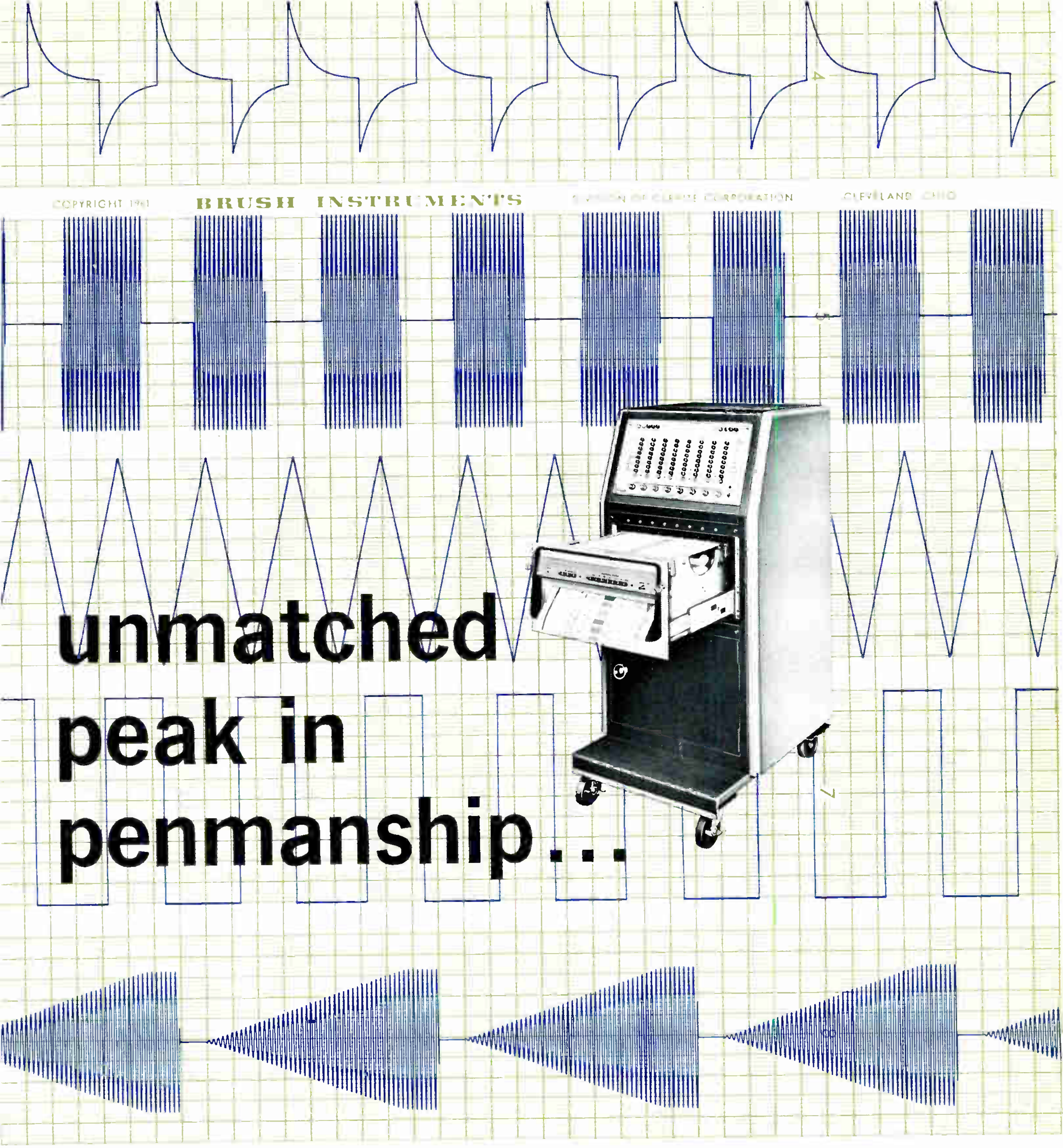
(Continued on page 164A)

Manufacturers' Cross-Reference Index

(Continued from page 162A)

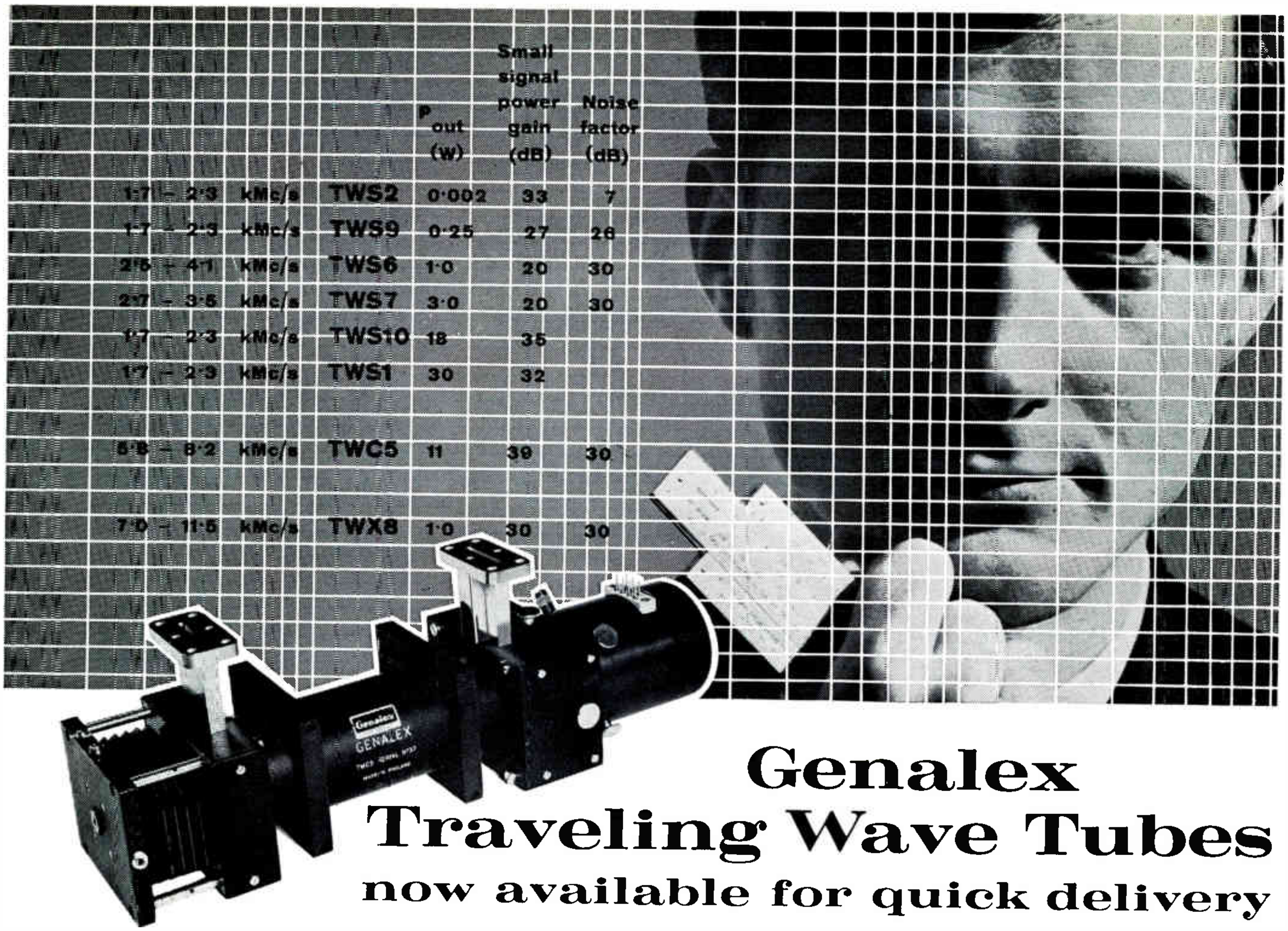
Cerl-Dale, Ltd.—Dale Products, Inc.
Cerramoplastic Molders, Inc.—Tech-Art Plastics Co.
Cerro de Pasco Corp.—Rockbestos Wire & Cable Co.; Titan Metal Mfg. Co.
Cesco Plant—Ferro Corp.
Cessna Aircraft Co.—Aircraft Radio Corp.
Chance Co. (A. B.)—Porcelain Products Co.
Chance Vought Aircraft, Inc.—Chance Vought Corp.; Ling-Temco-Vought, Inc.
Chance Vought Corp.—Genesy's Corp.; Information Systems, Inc.; Panellit Div.; Vought Electronics Div.
Chassis Trak Corp.—Western Devices, Inc.
Chatham Electronics Div.—Tung-Sol Electric, Inc.
Chemo Products, Inc.—Chemo-Textiles, Inc.
Cherry-Channer Corp.—Channer Corp.
Cherry Rivet Div.—Townsend Co.
Chevron Miter Joint Div.—Everard Tap & Die Corp.
Chicago Aerial Industries—Chicago Aerial Survey; Kintronic Div.
Chicago Electrical & Surgical—Hudson Bay Co.; Labline, Inc.
Chicago Electronic Engineering Co.—Chicago Magnetic Control
Chicago Electronics Div.—Land-Air, Inc.
Chicago Magnet Wire Corp.—Haydon Co. (The A. W.)
Chicago Metal Hose Corp.—Flexonics Corp.
Chicago Screw Co.—Standard Screw Co.
Chicago Standard Transformer Corp.—Stancor Electronics, Inc.
Chicago Telephone Supply Corp.—CTS Corp.; Chicago Telephone of Calif., Inc.
Chicopee Mills, Inc.—Johnson & Johnson; Lumite Div.
Chisholm-Ryder Co., Inc.—Premax Products Div.
Chromalox, Inc.—Wiegand Co. (Edwin L.)
Cinch Mfg. Co.—Graphik-Circuits Div.; Jones Div. (Howard B.); Ucinite Co.; United-Carr Fastener Corp.
Cinefonics Div.—Cook Technological Center, Cook Electric Co.
Cinema Engineering Div.—Aerovox Corp.
Cinema Television, Ltd.—Rank Cintel, Ltd.
Cinesone Div.—Federal Mfg. & Engineering Corp.
Circuit Engineering, Inc.—Stereatronics, Inc.
Circuitron, Inc.—Radio Engineering Products
Citation Div.—Harman-Kardon, Inc.
Clare Transistor Corp. (C. P.)—Universal Controls, Inc.
Clarkstan Corp.—Pacific Transducer Corp.
Clary Multiplier Corp.—Clary Corp.
Clemco Aero Products—Hathaway Instruments, Inc.
Cletron, Inc.—Cleveland Electronics, Inc.
Cleveland Cap Screw Co.—Standard Pressed Steel Co.
Cleveland Electronics, Inc.—Cletron, Inc.
Cleveland Graphite Bronze—Clevite Transistor Products Div.
Cleveland Instrument Co.—Bendix Corp.
Cleveland Powdered Metal Co.—International Powder Metallurgy Co.
Clevite Corp.—Brush Instruments; Clevite Transistor Products; Shockley Transistor Div.
Clifton Conduit Corp.—General Cable Corp.
Climax Molybdenum—American Metal Climax, Inc.
Cohu Electronics, Inc.—Kin Tel Div.; Massa Div.
Coleman Electronics, Inc.—Beattie-Coleman, Inc.
Coleman Engineering Co., Inc.—Beattie-Coleman, Inc.
Coleman-Kramer, Inc.—Coleman Electronics, Inc.
Collaro, Ltd.—Magnavox Co.
Collectron Corp.—Thermax Wire Corp.
Collins Electronics Mfg. Corp.—Collins Electronics, Inc.
Collins-Powell Co.—Elastic Stop Nut Corp. of America
Collins Radio Co.—Alpha Corp.; Communication Accessories Co.

(Continued on page 167A)



The totally new Brush Recorder Mark 200 made these incredibly crisp tracings. No other recorder in existence can match them. Note the line width. It never varies . . . regardless of writing velocity, regardless of chart speed. The writing mechanism is electrically signaled by the position-seeking "Metrisite" transducer . . . no parts to wear, infinite resolution, verifiable dynamic ½% accuracy. Traces are permanent, high-contrast, reproducible . . . on low cost chart paper. The Mark 200 has but three standard controls . . . attenuator, pen position, chart speed. Such fidelity, simplicity and economy are possible with no other direct writing recorder. Write for details . . . they'll speak for themselves.

brush INSTRUMENTS
DIVISION OF
37TH AND PERKINS **CLEVITE** CORPORATION CLEVELAND 14, OHIO



		P out (W)	Small signal power gain (dB)	Noise factor (dB)
1·7 – 2·3 kMc/s	TWS2	0·002	33	7
1·7 – 2·3 kMc/s	TWS9	0·25	27	26
2·5 – 4·1 kMc/s	TWS6	1·0	20	30
2·7 – 3·5 kMc/s	TWS7	3·0	20	30
1·7 – 2·3 kMc/s	TWS10	18	35	
1·7 – 2·3 kMc/s	TWS1	30	32	
5·8 – 8·2 kMc/s	TWC5	11	39	30
7·0 – 11·5 kMc/s	TWX8	1·0	30	30

Genalex Traveling Wave Tubes

now available for quick delivery

The Traveling Wave Tubes listed above are manufactured by General Electric Company Limited of England. They are available for quick delivery from the sole U.S. source, The Avnet System.

The TWC5, with 11 watts output, small signal gain of 39 db (5.8 to 8.2 kmcs) offers the ultimate in present-day performance: plug-in replaceability in the periodic magnet stack while meeting stringent matching and focusing requirements, *plus full life guarantees.* In quantity production and backed by extensive field experience, these tubes insure trouble-free systems.

These TWTs are established types, as are BWOs for all common bands. Avnet/Genalex can provide systems-proven tubes to fill all sockets in any microwave system.

Any tube bearing the Genalex label has long been recognized as embodying the highest technical advances in the Industry. Genalex has been designed into military systems because of compatibility with certain requirements and high performance standards.

For further information about how Avnet/Genalex can reduce your tube costs, or deliver such hard-to-get Traveling Wave Tubes as the TWC5, phone your Avnet Applications Engineer or mail the coupon.

The Avnet System

MEN / METHODS / MATERIALS / MANAGEMENT

AVNET INSTRUMENT CORP. *subsidiary* — AVNET ELECTRONICS CORP.

LOS ANGELES, CAL.; SUNNYVALE, CAL.; SEATTLE, WASH.; SALT LAKE CITY, UTAH · PHOENIX, ARIZ.; CHICAGO, ILL.; WESTBURY, L. I.; BURLINGTON, MASS.

AVNET markets from its stocking facilities:

BENDIX SCINTILLA CONNECTORS
MECHANICAL PRODUCTS
SPERRY SEMICONDUCTORS
BABCOCK RELAYS
SILICON TRANSISTORS
CREMAR CONNECTORS
SPRAGUE CAPACITORS
U. S. SEMCOR SEMICONDUCTORS
SCOPES OSCILLOSCOPES
CLARE RELAYS
MICRODOT CONNECTORS LERCO HARDWARE
AVNET AUTO. CONNECTOR AND CABLE TESTER
AVO MULTIRANGE METERS
WIDNEY DORLEC CONSTRUCTION SYSTEM
GENALEX TUBES
SULLIVAN PRECISION MEASURING APPARATUS
SERVO DESIGN AND TESTING EQUIPMENT

Rush me further information on the ☐ *Genalex Traveling Wave Tubes and/or send me the* ☐ *Genalex Tube Catalog*

NAME ____________

TITLE ____________

Clip this section to your letterhead and mail to:
East of Mississippi:
THE AVNET SYSTEM, Publications Section
70 State Street, Westbury, Long Island
West of Mississippi:
THE AVNET SYSTEM, Publications Section
5877 Rodeo Rd., Los Angeles 16, Cal.
Your request will be expedited within 90 minutes of receipt.

Manufacturers' Cross-Reference Index

(Continued from page 164A)

(Continued on page 184A)

Positions Open

The following positions of interest to IRE members have been reported as open. Apply in writing, addressing reply to company mentioned or to Box No.

The Institute reserves the right to refuse any announcement without giving a reason for the refusal.

Proceedings of the IRE
1 East 79th St., New York 21, N.Y.

THERMIONIC PHYSICIST or ENGINEER

Work will involve the design and development of prototype solar energy thermionic conversion systems for use as the source of electrical power for lunar and planetary spacecraft. Applied research and development directed toward improving reliability, life and efficiency of thermionic energy converters and generators. BS or MS with 1 to 2 years experience desired. Please send inquiries to Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, Calif. Att. J. I. Adams. An equal opportunity employer.

ASSISTANT or ASSOCIATE PROFESSOR

Applications are invited for an Assistant or Associate Professor of Electrical Engineering. Candidates should be well qualified academically, preferably to the doctorate level, and should have some research, industrial design, or teaching experience. Specialization may be in Control Systems, Circuit Theory, or other fields of Electronics. Write to: Chairman, Department of Electrical Engineering, Sacramento State College, 6000 J Street, Sacramento 19, California.

ELECTRONIC ENGINEER

Graduate Engineer with 8 yrs. of experience in design and development of electrical and electronic instruments, controls and transistor amplifiers. Exceptional opportunity for technical growth and administrative development. Write to: A R & D A—Div. of All American Engineering Co., 135 Main Street, Belleville 9, N.J.

INSTRUMENTAL SCIENCES ASSISTANTSHIPS

Leading to M.S. Degree. Engineering or Physics majors with strong background in instrumentation. Course work in fundamental principles and theory of basic instrumentation and process control. Research work in many fields of laboratory and process instrumentation. Stipend $2200 to $2500 per year. Send transcript and vitae to: Head, Electronics and Instrumentation, University of Arkansas, Graduate Institute of Technology, P.O. Box 3017, Little Rock, Arkansas.

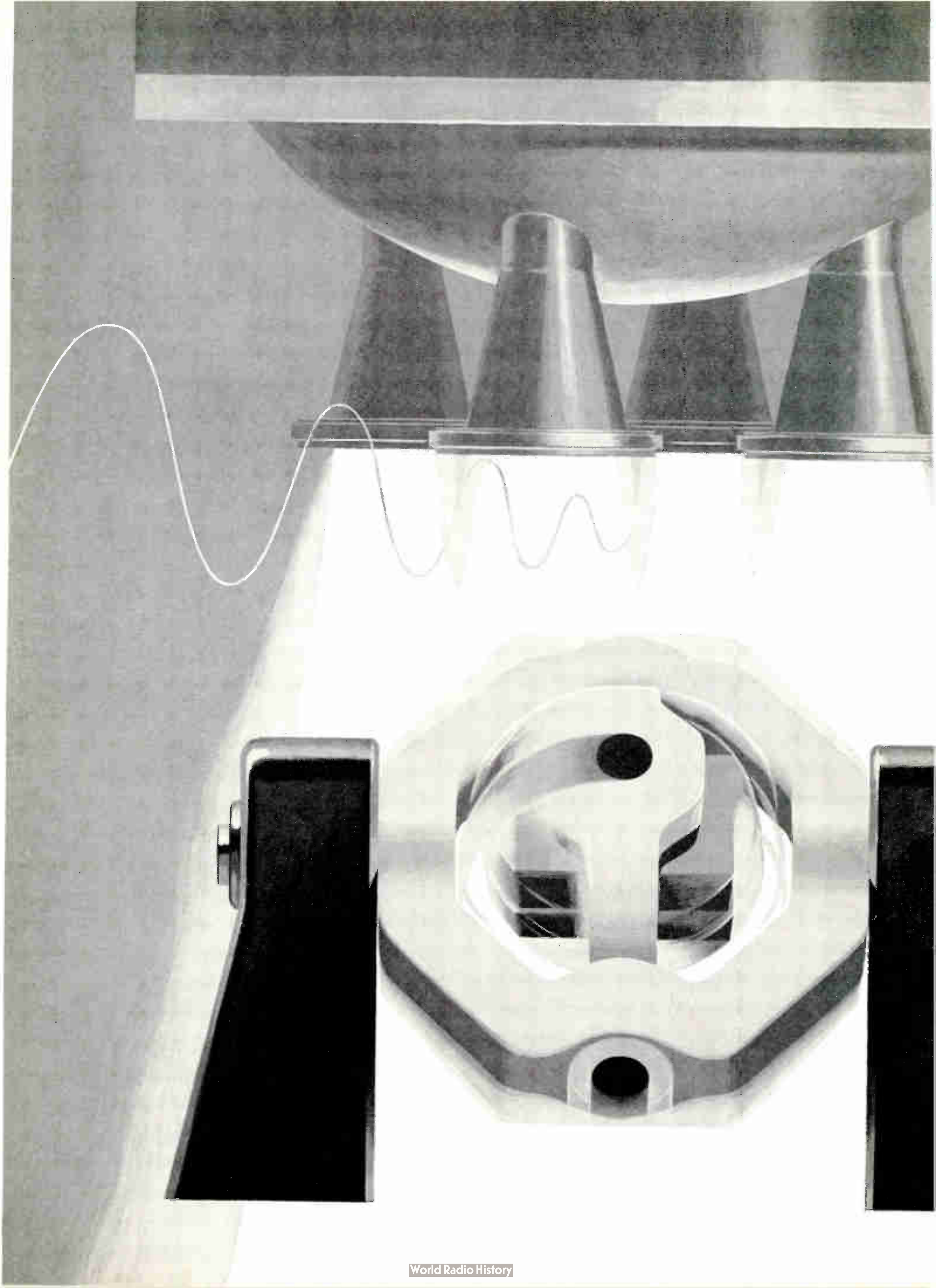
ASSISTANT PROFESSOR OF E. E.

Ph.D. required and competence in Circuit Theory and/or Solid State phenomena desirable. Should have a definite interest in research and some teaching experience. Position available Sept. 1962. Write Head of E. E. Dept., State Univ. of Iowa, Iowa City, Iowa.

ENGINEER-MATHEMATICIAN

Senior position open for analytical engineer or applied mathematician. Salary, pension, and other benefits fully competitive with industry. Ph.D. essential. Course work will be at the graduate and advanced under-graduate level. Ample time and encouragement for personal research work, including participation in activities of University Computation Center with 1620 computer. Pleasant living and working conditions in

(Continued on page 172A)



WHERE ENGINEERING AND PRODUCTION JOIN FORCES

HRB-SINGER MANUFACTURING SERVICES DIVISION*

Exciting careers in technical production are attracting engineers with the highest qualifications to HRB's Manufacturing Services Division. This unusual production facility is fabricating state-of-the-art components, sub-assemblies of major defense systems and complete infrared reconnaissance systems.

The Division maintains the following modern facilities: Production Department, Mechanical Design and Drafting, Etched Circuitry, Machine Shop, and Quality Control and Testing. These facilities offer the engineer the opportunity to choose his work from all lines of technical production, from conception and design to fabrication, check out and field test. Close liaison with HRB's R and D departments helps the Division engineer keep abreast of new concepts and developments in the industry.

In addition to this excellent technical environment, HRB offers unusual employee benefits. These include company-paid graduate study at The Pennsylvania State University, generous hospitalization, life insurance and vacation programs as well as a company-paid retirement program.

If you are interested in learning more about career opportunities at HRB-Singer, write George H. Rimbach, Supervisor of Personnel, Dept. R-3.

**for technical data write Dept. MSD.*

An equal opportunity employer

HRB

HRB-SINGER, INC.

A SUBSIDIARY OF THE SINGER MANUFACTURING COMPANY

Science Park, P.O. Box 60, State College, Pa.

Positions Open

(Continued from page 168A)

attractive uncrowded city. Address inquiries to Robert A. Chipman, Chairman E. E. Dept., University of Toledo, Toledo 6, Ohio.

TEACHING POSITION

Expanding work in the Department of Electrical Engineering, Air Force Institute of Technology, requires an added staff member. Ph.D. desired, but M.S. with appropriate experience may be satisfactory. Beginning salary approximately $9000-$12,000 depending on qualifications. Employment effected according to Civil Service Regulations. Address inquiries to Head, Department of Electrical Engineering, Air Force Institute of Technology, Wright-Patterson AF Base, Ohio.

TEACHING AT ALASKA

The University of Alaska will have an opening for an instructor or assistant or associate professor. We are teaching electronics and communications and are ready to expand into power system engineering. Research might be on the ionosphere or in northern problems in communications or power. Write to J. G. Tryon, Head, Department of Electrical Engineering, University of Alaska, College, Alaska. Airmail please.

ENGINEER

E.E.—Audio Circuitry. Design of commercial high-fidelity, phonograph, and AM-FM tuner circuitry. Responsible for overall equipment performance, components, evaluation and follow-through. Transistor circuitry experience desirable. Generous salary and benefits. Write Mr. R. D. Sommer, Westinghouse Television-Radio Division, Metuchen, N.J. An equal opportunity employer.

TEACHER

Teaching position in electrical engineering. M.S. or Ph.D. with interest in teaching and developing courses in an electrical engineering undergraduate program. Openings are immediately available in Electronics, Energy Conversion, Electromagnetic Fields and Power Systems. Salary commensurate with background and experience. Write to Harley Anderson, Chairman, Electricity & Electronics Department, General Motors Institute, Flint 2, Michigan.

SYSTEM ENGINEERING—EARTH SCIENCES

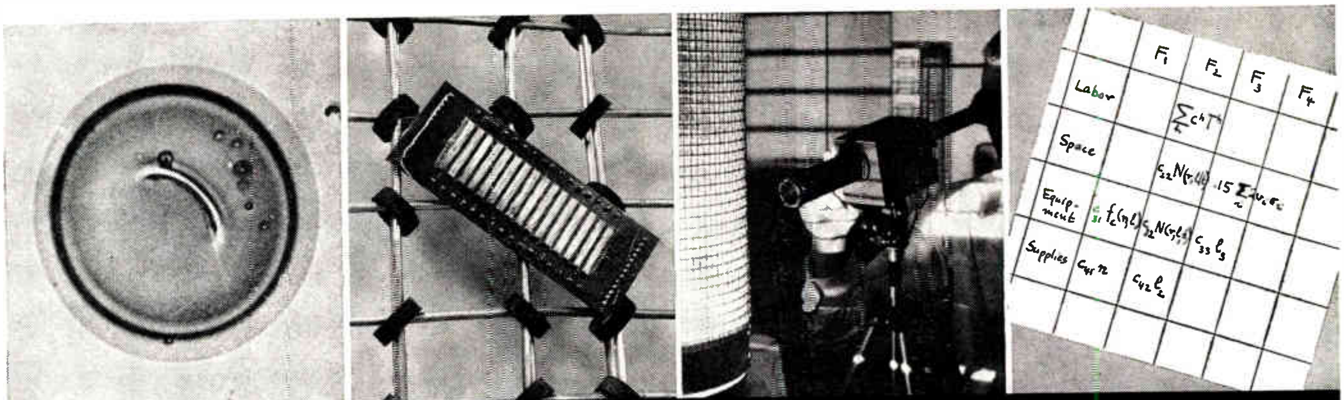
Creative electrical/electronic engineers to conduct research and development program on automation in data accumulation, reduction and processing in geology—geophysics, water resources investigations, geodesy, photogrammetry and allied fields including development of new techniques and equipment for sensing, transmitting, logging, processing and storing of mass data. Contact Placement Officer, U.S. Geological Survey, Washington 25, D.C.

ELECTRONICS ENGINEER—SALES

We are manufacturer's representatives selling high-quality instruments and OEM components to the electronic industry. The complexity of the product we handle requires a man having a BSEE degree and several years of experience. Fine, challenging opportunity with a fast growing organization. Compensation is by drawing against commission for field sales work. Contact Mr. George L. Boyden, Saunders & Company, 53 Prospect Street, Waltham 54, Mass.

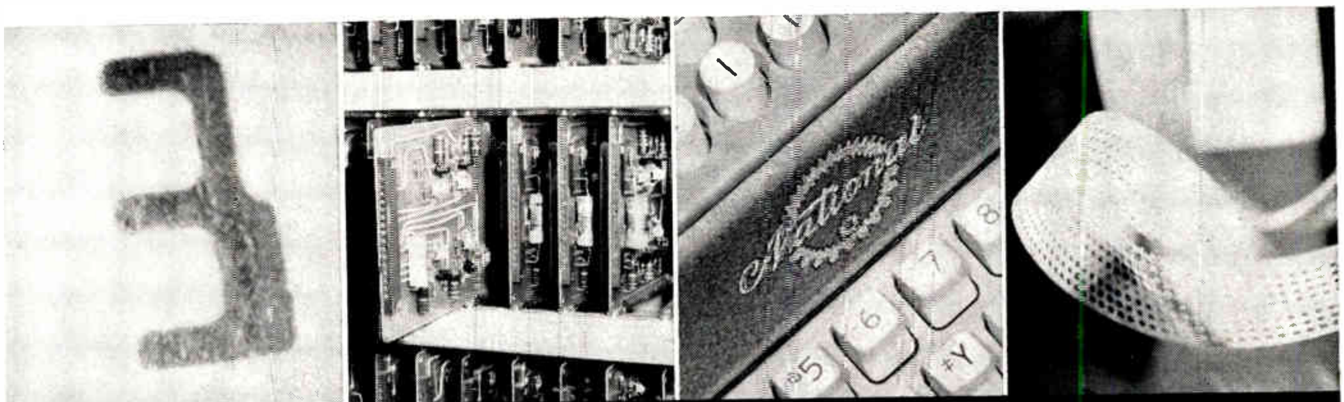
(Continued on page 174A)

NCR research



MICRO-ENCAPSULATION MAGNETICS THIN FILMS OPERATIONS RESEARCH

and development



OPTICS LOGIC & CIRCUITRY MECHANICS DIGITAL SYSTEMS

The National Cash Register Company presents above-average opportunity for the experienced scientist or engineer to apply his talents in a stable organization. Whether your interest lies in commercially oriented projects or advanced military electronic development, you will find lasting satisfaction and encouragement for professional advancement.

If your interests lie in one of the many fields of endeavor being carried on by The National Cash Register Company, you are invited to submit your resumé to us for consideration for one of the fine opportunities now open in Dayton, Ohio, at our modern Research and Development facility.

SEMI CONDUCTOR RESEARCH
LOGIC AND CIRCUIT DESIGN
COMPONENT EVALUATION
TECHNICAL WRITING
ADVANCED ELECTRONIC SYSTEMS
SYSTEMS RELIABILITY
DIGITAL COMMUNICATIONS

Positions open are at all levels of experience, including Project Manager. Prompt and personal attention will be given to your inquiry. All qualified applicants will be considered regardless of race, creed, color or national origin. Send resumé and letter of application to: T. F. Wade, Technical Placement, EN, The National Cash Register Company. Main and K Streets, Dayton 9, Ohio.

NCR PROVIDES TOTAL SYSTEMS – FROM ORIGINAL ENTRY TO FINAL REPORT – THROUGH ACCOUNTING MACHINES, CASH REGISTERS OR ADDING MACHINES, AND DATA PROCESSING
The National Cash Register Co. · 1,039 offices in 121 countries · 78 years of helping business save money

NCR

Positions Open

(Continued from page 172A)

ELECTRONIC SALES—INTERNAL

Inside technical sales. The expanded scope of our activity as an electronic manufacturer's representative has created an opportunity for a man to assist field salesmen by conducting mailing and phoning programs for our Waltham office. Ideal opportunity for a capable, disabled veteran. Secretarial skills such as shorthand and typing would be preferred but not necessary. Contact Mr. George L. Boyden, Saunders & Company, 53 Prospect Street, Waltham 54, Mass.

ELECTRONICS ENGINEER—EXPERIMENTAL SURGERY

For research program involving electronic organ and muscle control. M.S. or equiv. exp. servomechanisms; digital computers; transistor circuitry. Industrial background desirable. N.Y. area. Write Box 2069.

ELECTRONIC ENGINEER

Inductive Devices. Well established and growing company located in Culver City, Calif. has an excellent opportunity for a capable engineer to organize and manage the design and manufacture of low pass, high pass and band pass filters and electro-magnetic delay lines. Salary will be commensurate with training, experience and ability to manage. Liberal company benefits. Box 2070.

PHYSICISTS AND MICROWAVE ENGINEERS

SRI's expanding research efforts in our Electromagnetics Laboratory have created unique opportunities for M.S. and Ph.D. physicists and engineers with backgrounds and interests in the following areas: Lasers; Microwave Components; Low- and High-Power Microwave Filters; Solid-State Filters; Parametric Amplifiers. Our work involves theoretical and experimental research. Located on the San Francisco Peninsula, SRI offers rich professional association, industrial-level salaries and advanced study opportunities. Please send resume or contact R. K. Ditmore, Stanford Research Institute, Menlo Park, California. An equal opportunity employer.

TEACHING POSITION

Excellent teaching opportunity will be available beginning September. 1962. Ph.D. preferred but will consider outstanding candidates with lesser qualifications. Attractive full-year contract available. Salary range is $7500-$9500. Location in the Midwest at a medium-size private university. Send complete resume to Box 2071.

ELECTRONIC ENGINEERS

Speidel Corporation's Industrial Division offers exceptional opportunities for electronics engineers with B.S., M.S. and Ph.D. degrees to assume intermediate, senior and project engineer responsibilities in new programs for aerospace and commercial fields. Speidel desires men with experience in one or more of the following areas: Field Theory, Transistor Circuits, Magnetic Tape Recording, Electro-Mechanical Instruments, and Circuit Analysis for work on development of Recorder-Reproducers and Magneto-hydrodynamic Gyros. Contact (call collect 8 a.m.–8 p.m.) Mr. Joseph Motherway, Chief Engineer, R & D, Speidel Corp., Industrial Div., Speidel Industrial Park, Warwick, Rhode Island. REgent 9-7000. An equal opportunity employer.

(Continued on page 176A)

A RARE OPPORTUNITY

Stanford University, in Palo Alto, California, is now forming the nucleus of the engineering team building the world's largest electron accelerator. This two-mile linear microwave device will enable physicists to explore deeper into the atom than ever before. Those engineers and scientists selected to participate in this challenging project will share in the pleasures of . . . the cultural environment of the beautiful San Francisco Peninsula . . . the opportunity to work in an exciting intellectual atmosphere . . . four weeks paid vacation . . . an excellent retirement plan.

Electronic engineers with several years' experience are needed for senior positions in the research and development of such disciplines as:

- *microwave systems and components*
- *pulse circuits and pulse modulators*
- *ultra high power klystron tubes*
- *data handling equipment for accelerator central control systems*

PLEASE SEND YOUR RESUME TO:

Mr. G. F. Renner, Placement Manager
Stanford Linear Accelerator Center
Stanford University, Stanford, California

STANFORD LINEAR ACCELERATOR CENTER

Positions Open

(Continued from page 174A)

CAMBRIDGE ELECTRON ACCELERATOR

Inquiries are invited from electronics engineers to join the staff of the M.I.T.-Harvard 6 Bev Electron Accelerator. Experience in one or more of the following specialties is desirable: design and development of dc to nanosecond circuits, data handling and computer techniques. Excellent qualifications and 3 to 5 years practical experience required. Resumes, including salary requirements, should be addressed to the Director, Cambridge Electron Accelerator, Harvard University, 42 Oxford Street, Cambridge 38, Mass.

DEPUTY TECHNICAL DIRECTOR—ISRAEL

A practical engineer with a minimum of 7 years working experience and at least 3 years in development or design and series production scheduling of work on electronic components, such as coilpacks, I.F. transformers, potentiometers, tuning condensers, etc. This position is available on a 3 year contractual basis. For information inquire COMOI, 515 Park Ave., New York 22, N.Y. Plaza 2-0600.

ELECTRONICS TECHNICIANS

Excellent opportunities for experienced electronic technicians at large university radio observatory in development and operation of low noise radiometers with digitized output operating at UHF and microwave frequencies. Give resume of experience, references, and salary desired. Address correspondence to: Director, Ohio State Univ. Radio Observatory, 2024 Neil Avenue, Columbus 10, Ohio.

ASSOCIATE PROFESSOR OF E. E.

Position open September 1962 to teach undergraduate and graduate courses. Must have Ph.D. degree, or M.S. with equivalent experience. Specialty in networks or electromagnetic theory. Submit resume to Chairman, Electrical Engineering Dept., University of North Dakota, Grand Forks, North Dakota.

PH.D. OPENING

Excellent opportunity for Ph.D. with background in microwaves and/or solid state. Some teaching and some research required. Private consulting encouraged. Department has Ph.D. program. Good salary and rank arrangements will be made for exceptional man from either the industrial or academic fields. Address replies to: Dr. Fred Schumann, Chairman, Electrical Engineering Department, Vanderbilt University, Nashville, Tennessee.

MANAGER ADVANCED SYSTEMS ENGINEERING

Responsible for performing systems design for proposals, improved systems, integrating inputs (technical design & program planning). Provide preproposal and proposal & technical integration with customer up to award of contract, provide advanced research studies. Should have experience in technical direction of Airborne Weapon Systems. Send resumes to: Jordan-White, Inc., 420 Lexington Avenue, New York, N.Y.

(Continued on page 178A)

Positions Open

(*Continued from page 176A*)

MANAGER, ADVANCED SURVEILLANCE & RECONNAISSANCE SYSTEMS ENGINEERING

Responsible for performing preliminary systems design for proposals of new, improved systems, integrating technical design and program planning. Provide preproposal and proposal integration with customer to award of contract, manage system study programs, perform advanced research studies, provide master planning and guidance for advanced engineering programs in surveillance & reconnaissance systems. Should have knowledge of guidance, control, processing, communications subsystems. Send resume to: Jordan-White, Inc., 420 Lexington Avenue, New York, N.Y.

PROFESSOR OF ELECTRICAL ENGINEERING

The Department of Electrical Engineering seeks a Ph.D. for research and teaching in an expanding graduate program with emphasis on circuit analysis and synthesis, and the application of solid state devices. Should have experience in securing and directing sponsored research. Applications and inquiries may be sent to John E. Dean, Head, Department of Electrical Engineering, Colorado State University, Fort Collins, Colorado.

ENGINEER-PHYSICIST

Senior level engineer-physicist—Mathematician needed to direct active and existing program directed toward a better understanding of the fundamental mechanisms involved in sputtering phenomena. Should have experience in the ion bombardment field, and should be familiar with x-ray and electron microscopy as well as ultra high vacuum techniques. Support facilities excellent, subordinate personnel well qualified, and a high degree of freedom to conduct basic research. Send complete resume to Box 2072.

BROADCAST ENGINEER

Long term European assignment with private organization engaged in construction and operation of large high-powered shortwave radio stations. Top qualifications, experience and executive ability desired. Good salary plus overseas benefits. Reply to Box 2073.

TEACHING AND RESEARCH PROFESSORSHIP

Unusual opportunity for electronic engineer or physicist to initiate research program in electronics in rapidly growing urban institution. Income, with research, competitive with industry. Ph.D. or equivalent required. Reply to Dr. T. S. Peterson, Chairman, Division of Science, Portland State College, Portland 1, Oregon.

TRANSFORMER ENGINEER

Minimum 2 yrs. experience design and audio, power, MIL and industrial transformers. Expanding to multi-plant operations, both coasts. Top salary, full benefits, profit and pension plans. West Coast telephone 213-HO 3-5694; East Coast telephone 516-LOcust 1-6050. Person to contact for interview arrangements, Mr. Richard K. Chaber, Vice President, Engineering, at 516-LOcust 1-6050, Microtran Company, Inc., Valley Stream, N.Y.

ELECTRONICS ENGINEER, B.S. or M.S.

Unique opportunity in small expanding company for a versatile and resourceful man trained

(*Continued on page 180A*)

ENGINEERS, EE:

Our "Universal" Approach to QC Means Much Broader Technical and Managerial Responsibilities for

EVALUATION ENGINEERS

Because of the cost and complexity of the systems we develop and produce, our engineers simply can't afford the luxury of making a prototype for evaluation purposes. In almost every case, the first piece made . . . is shipped!

The situation becomes even more formidable when you analyze the number and types of subsystem interfaces, rigorous performance and reliability standards involved, and the number of subcontracting vendors whose individual quality standards must be equated with our own.

Small wonder that our Evaluation Engineers operate over such a broad spectrum of technical and managerial responsibilities. In briefest terms, these include:

1. Functional evaluation of military systems.
2. Environmental testing of the systems.
3. Determination of the uses for test equipment and preparation of test equipment specs.
4. Customer liaison on system performances and quality characteristics.
5. Technical leadership of quality operations to control the system requirements.
6. Design review for quality characteristics.

Product scope encompasses inertial navigation systems, analog and digital computers, optical alignment systems, antennas, electronic torpedoes, servo systems. BSEE is required.

Our location is convenient, western Massachusetts offers a host of cultural and 4-seasons recreational advantages for the entire family.

For more information, or to apply, please forward your resume in strict confidence to Mr. D. F. Kline, Div. 53-ME, Ordnance Department, General Electric Co., 100 Plastics Avenue, Pittsfield, Mass.

GENERAL ELECTRIC

An Equal Opportunity Employer

Positions Open

(*Continued from page 178A*)

in RF techniques, field theory, general electronics, feedback amplifier design and servomechanisms. Development, prototype and initial production runs. Minimum 3 years experience desirable. Reply to Boonton Polytechnic Company, P.O. Box 125, Boonton, New Jersey.

ELECTRICAL ENGINEERING

Staff positions beginning September, 1962. Undergraduate-graduate teaching and research in networks, systems, microwaves, and materials. Write to Joseph C. Michalowicz, Head, Department of Electrical Engineering, The Catholic University of America, Washington 17, D.C.

PROFESSOR OF PHYSICS

Ph.D. with teaching experience to be Department Head. To handle advanced undergraduate courses and to initiate research. College of Liberal arts and sciences. Resume to: R. B. Winslow, Department of Physics, Wheeling College, Wheeling, W.Va.

CRITICAL VACANCIES AT GRIFFIS AIR FORCE BASE, ROME, N.Y.

Research Psychologist (Physiological, Experimental & Engineering)	GS-180-7	$5355 pa
Civil Engineer (Gen)	GS-810-9	$6435 pa
Electronic Engineer (2) (Electro Magnetics)	GS-855-11	$7560 pa
Electronic Engineer (Gen)	GS-855-11	$7560 pa
Industrial Engineer	GS-896-11	$7560 pa

Positions Wanted

By Armed Forces Veterans

In order to give a reasonably equal opportunity to all applicants and to avoid overcrowding of the corresponding column, the following rules have been adopted:

The IRE publishes free of charge notices of positions wanted by IRE members who are now in the Service or have received an honorable discharge. Such notices should not have more than five lines. They may be inserted only after a lapse of one month or more following a previous insertion and the maximum number of insertions is three per year. The IRE necessarily reserves the right to decline any announcement without assignment of reason.

Address replies to box number indicated, c/o IRE, 1 East 79th St., New York 21, N.Y.

ENGINEER

Military retired. Over 20 years training and experience in electronics maintenance and operations management for USAF. Some R&D associated with 8 years assignments on special projects. Interested in electronics or geophysical operations with possible assignments in Australia or New Zealand. Box 3978 W.

TEACHER

B.S. and M.S. in E.E.; present position Research and Development Engineer with six years

(*Continued on page 182A*)

CAN YOU TAKE IT?

Do you have what it takes to manage a research or development program in advanced technologies? We have several important opportunities for highly qualified scientists and engineers. The demands are exacting—but the rewards are high (up to $25,000). Your resume must show achievements in any of these areas:

- COMMUNICATIONS TECHNIQUES
- WEAPONS SYSTEMS ANALYSIS
- SOLID STATE PHYSICS—CIRCUITRY
- ANTENNA DESIGN
- DIGITAL EQUIPMENT ENGINEERING
- MICROWAVE COMPONENTS AND SYSTEMS
- RADAR SYSTEMS AND TECHNIQUES

All fees and relocation expenses paid by our client companies. Negotiations will be held in strict confidence.

ACT NOW!

SEND YOUR RESUME TO:

Mr. J. Edward Beauregard
Dept. C-1

ABBOTT'S

EMPLOYMENT SPECIALISTS
150 Tremont St., Boston 11, Massachusetts

GOODYEAR AIRCRAFT CORPORATION

Arizona Division

is expanding its Electronic facilities and has long range positions that will challenge your talent in the technical areas listed below:

SR. DEV. ENG'R.

Develop wideband i-f amplifiers, i-f oscillators, i-f isolation amplifiers & related circuits.

SR. DEV. ENG'R.

Develop X-band radar circuits & equipment including wave guide assemblies & related circuits.

SR. DEV. ENG'R.

Servomechanisms, design close loop control system.

SR. DEV. ENG'R.

Develop radar transmitters & modulators including pulse forming networks, transformers, protection circuits, X-band amplifiers & related circuits.

SR. PACKAGING ENG'RS.

Layout packaging & detailing layout of electronic sub-assemblies including i-f and r-f units.

All qualified applicants will receive consideration for employment without regard to race, creed, color, or national origin.

Request Application or Send Resume To:

Engineering Personnel
Goodyear Aircraft Corporation
Litchfield Park, Arizona

Similar Positions at Goodyear Aircraft Corporation, Akron, Ohio.

Open once every 19 months

We just can't send a Mariner spacecraft to Venus any time we feel like it. Because of the different orbits of Earth and her most brilliant neighbor, we have only about one month, every 19 months, to get a shot at it.

Even at that, the launch window opens only one to two hours each day of that month.

The men with their noses pressed hardest against the pane are the scientists and engineers at Caltech's Jet Propulsion Laboratory. These are the men who designed the Mariner and its payload of scientific instruments. These are the men who'll find out what the moon is made of and if there's life on other planets. They keep their fingers crossed long after the last stage is fired.

But they know luck's a luxury we can't afford in other-world exploration. JPL needs minds. The kind of minds that won't leave anything to chance...that know what to do when the window goes up. If you have that kind of mind, maybe JPL is your kind of place. Write us about yourself. If you have a mind to.

JET PROPULSION LABORATORY
4804 OAK GROVE DRIVE, PASADENA, CALIFORNIA
Operated by California Institute of Technology for the National Aeronautics & Space Administration

jpl

All qualified applicants will receive consideration for employment without regard to race, creed or national origin / U.S. citizenship or current security clearance required.

Positions Wanted

By Armed Forces Veterans

(Continued from page 180A)

experience. Desires teaching position in Fall of 1962. Box 3979 W.

SYSTEMS ENGINEERING

M.S.E.E., 5 years experience digital and analog systems. Interested in process analysis, simulation, feasibility study, and proposal preparation related to development and application of process control computers. Box 3980 W.

ELECTRONICS ENGINEER

Age 32, married. Desires challenging position offering security and advancement possibilities. Twelve years experience in ground radar and communications. Five years of experience has been in systems engineering, primarily microwave. Training and supervisory experience; Commercial pilot. Box 3981 W.

REPRESENTATIVE/SUPERVISOR

Age 35. B.S. Mathematics, *summa cum laude.* Engaged continuously since 1951 in electronics teaching and field engineering. Sound, comprehensive background in electronics, radar, and computers. Fluent German and French. Thorough training in business administration. Successful supervisory experience. Would welcome overseas assignment. Box 3984 W.

EDUCATIONAL TELEVISION

Experienced broadcaster desires position as operations manager or assistant manager. Some teaching experience. Presently employed in designing of closed circuit TV. Will forward detailed resume on request. Box 3985 W.

ENGINEER

B.E. in Electrical Engineering 1953; P.E.; experience in all phases of manufacturing of electronic gear. Single, age 30. Will complete B.F.T. at American Institute for Foreign Trade in June; speak Portuguese. Desire position in management or technical sales in Brazil. Box 3986 W.

BIO-MEDICAL ENGINEER

10 years experience, publications, MSE (EE), desires teaching and/or research in medical engineering. Would like opportunity to initiate and guide research projects. I am principally an idea man. Box 3987 W.

ELECTRICAL ENGINEER

Electrical Engineer, BSEE 1950, desiring position as field engineer for system installation and for customer liaison with manufacturers, or as test engineer to design production tooling or test and quality control. Prefer Texas area. Box 3988 W.

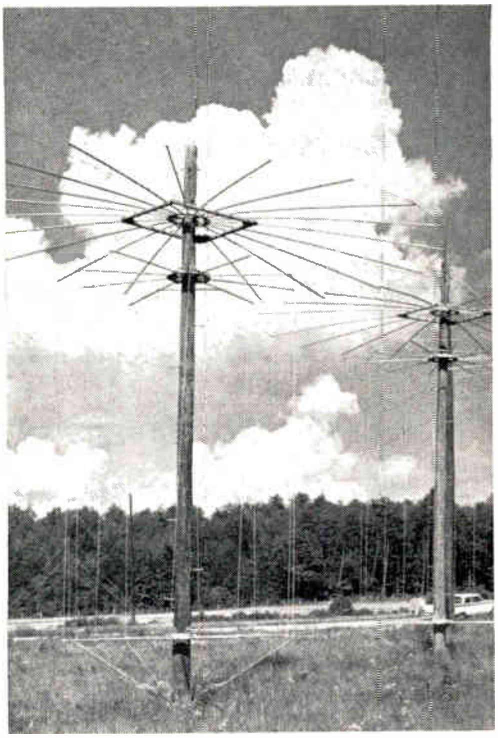
ENGINEER

B.E.E. with highest honor, Georgia Tech. 1950, 24 hours of graduate school. 2½ years design of two way mobile radio transmitters; 3 years manager of test laboratory; 5⅔ years manager of advance development group (automatic test and adjustment equipment). Desires challenging position with growing company. Box 3989 W.

SALES MANAGER

B.S. EE, Age 35, Married, 8 years experience in Electronic Component sales plus two years experience as project engineer for component manufacturer. Successful record in increasing sales. Tripled sales in present territory in three years. Box 3990 W.

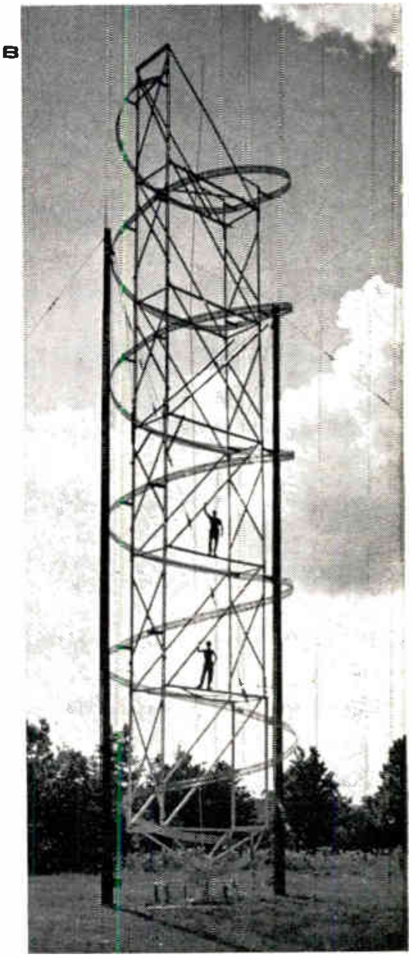
(Continued on page 184A)



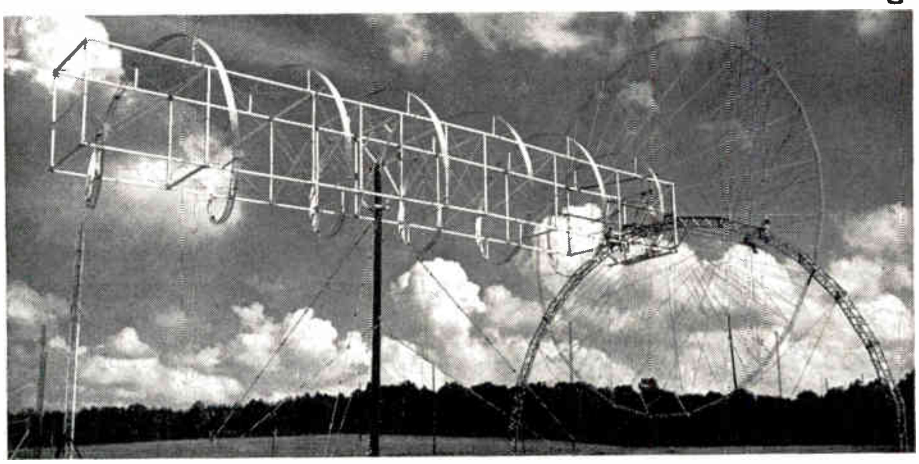
A

CLIMATE FOR CREATIVITY

B



C



A *Portion of a high gain broadband monopole, phased array, designed for electrical steering in azimuth and controlled elevation beamwidth.*

B *100 ft.-high, circularly polarized, axial mode broadband HF helix.*

C *Circularly polarized helical antenna for HF, trainable in azimuth.*

The creative scientist needs a climate where his ideas can flow freely . . . where he can find independence within a group dedicated to pursuits similar to his own. Because his creative work is the basic product of the Electro-Physics Laboratories, we have dedicated ourselves to the preservation of such a climate.

- Project responsibility is based on technical achievement.
- Programs are planned so that research personnel guide their projects from the concept stage, to applied research, to prototype development.
- Sympathetic and efficient administrative support is provided, which permits research personnel to focus their full talents on technical problems.
- The importance of the research effort instills in all participants a very real sense of contributing to our national purpose.

Staff appointments are available for research projects now under way and others still in the conceptual stages. Current problems include the investigation of ionospheric propagation in general and as applied to communications, the development of new and unique communications techniques, and the investigation of upper atmosphere phenomena by means of unusual rocket probe techniques. The available supporting laboratory program is being expanded to develop equipment and techniques for applications in these fields.

Security clearance helpful. For additional details, please write to: Mr. William T. Whelan, Director of Research and Development.

ELECTRO-PHYSICS LABORATORIES

ACF ELECTRONICS DIVISION, ACF INDUSTRIES

3355 52ND AVENUE, HYATTSVILLE, MARYLAND (A RESIDENTIAL SUBURB OF WASHINGTON, D. C.) AN EQUAL OPPORTUNITY EMPLOYER

Positions Wanted

By Armed Forces Veterans

(Continued from page 182A)

FLORIDA SALES REPRESENTATIVE

An aggressive engineer with M B A and established O.E.M. contacts is available to sell your electronic instruments or components. Age 30, married. Ultimate goal is to establish a representative organization. Box 3994 W.

PUBLICATIONS ENGINEER, SENIOR

Experience: Supervisor, computer manuals group; 18 years experience in electronics and writing. Graduate of Capitol Radio Eng. Inst., B.A. in English at Univ. of Md. 30 sem. units graduate study in English plus misc. training courses in electronics. Box 3995 W.

MICROMINIATURE COMPONENT ENGINEER

Graduate engineer, age 35, with ten years' experience in design, development and production of semiconductors, resistors, capacitors and microminiature circuits, seeks responsible position with manufacturer or user. Experience includes semiconductor device fabrication, thin films, photolithography, ceramic and glass-to-metal seals, ultramicrominiaturization techniques. Box 3996 W.

ELECTRONICS TECHNICIAN

Retiring from military service April 1963 with twenty years of communications-electronics experience. Graduate civilian and military schools. Resume upon request. Box 3997 W.

COMMUNICATIONS ENGINEER

Currently recalled as Captain in U.S. Army Security Agency. Industrial R & D experience in radio DF, communications satellites, and military communications systems. B.S. and M.S. in EE; PE; Ph.D. course work; languages; age 31; highest clearances. Desires long term assignment in Europe in communication systems engineering or management. Box 3998 W.

ELECTRICAL ENGINEER

Elect. Engr. Tech; RCA T3 Grad; Married; 28; Evening student at CCNY. 2½ yrs. with switching systems engineering dept. of R & D lab. Desires challenging position with room for advancement in N.Y.C. Box 3999 W.

Manufacturers' Cross-Reference Index

(Continued from page 167A)

(Continued on page 186A)

The Calibre of an Aerospace Organization Is Evidenced by the Calibre of the Problems Put To It by its Customers — and by the Company to its Engineers

Engineers, Scientists

what's your opinion:

1

ON HOW TO STORE AND PRESSURIZE CRYOGENIC PROPELLANTS IN SPACE, MAINTAINING INSTANT READINESS?

At Bell Aerosystems, Design Engineers are carrying on a concentrated attack on this problem, as it relates to space propulsion rocket engines, where weight and reliability are at a premium. This is an expanding activity at Bell. *Your* professional contributions will be welcomed if your experience and background equips you to analyze steady-state and transient heat transfer effects interrelated with thermodynamics and chemical kinetics of pressurization under varying space environments.

2

ON HOW TO DESIGN A DE-CRABBING SYSTEM TO AUTOMATICALLY ALIGN AIRCRAFT LONGITUDINALLY IN CROSS-WIND LANDINGS?

Correct alignment within 3 seconds of touchdown has been obtained by Bell engineers, in developing an all-weather landing system for conventional aircraft, successfully tested in 4000 landings of USN and USAF planes. Opportunities exist to participate in advanced studies of full spectrum of problems of automatic landing of very high-speed, advanced spacecraft.

3

ON HOW TO IMPROVE THRUST-TO-WEIGHT RATIO OF ELECTRO-STATIC PROPULSION DEVICES?

Applied research and advanced development in thermoelectrics for space propulsion is an important field of interest at Bell. As a qualified member of an R & D group in this area you would be concerned with finding practical answers to applications of electro-static propulsion for such missions as: interplanetary trips to Venus; interorbital transfers; deep space probes; satellite orientation control. Investigations are also in progress to develop light-weight ionization and neutralization systems; methods of creating, controlling and accelerating plasmas.

If you would like to pit your talents against the more significant and challenging problems of the Aerospace Age, why not look into the wide variety of professional openings at Bell, some of which are listed below. Bell Aerosystems originated the AGENA rocket engine, HIPERNAS inertial systems, the revolutionary Reaction Control Concept for space vehicle steering and many other advances in missilery, avionics and propulsion ... is now working on such diverse programs as the HYDROSKIMMER air cushion vehicle for the Navy, and a realistic space simulator for astronauts.

INERTIAL GUIDANCE

Various degree levels with 2 to 8 years specific experience in one of the following areas: logic design and magnetic memory development: low frequency design and development involving servos; systems and component testing, resolver chains, gyros, accelerometer integrators; inertial instrument development.

ELECTRONIC SYSTEMS

MS preferred, with minimum 5 years experience for synthesis and analysis of advanced systems such as air traffic control, target locators, visual simulators and feed-back controls. Openings also in Tucson, Arizona.

SYSTEMS ENGINEERING

BS in ME, Applied Math, AE or Physics. Experienced in the dynamics of closed-loop systems, guidance, mechanics of vehicle flight, space mechanics, analog and digital computing systems.

OPTICAL DEVELOPMENT

Advanced Engineering degree with minimum of 10 years experience in the design and construction of optical devices. For original design of optical equipment such as visual simulation apparatus, airborne sighting systems, and precision instrumentation.

DYNAMICS ENGINEERING

Degree in Math or Physics with experience in operation of analog computer equipment. To perform dynamic analysis of aircraft and missile systems, mechanization and operation of analog computer equipment in systems synthesis, design and evaluation.

Please address resumes to Mr. George Klock

BELL AEROSYSTEMS CO.

DIVISION OF BELL AEROSPACE CORPORATION—A textron COMPANY

An Equal Opportunity Employer P.O. Box #1, Buffalo 5, New York

Manufacturers' Cross-Reference Index

(Continued from page 184A)

(Continued on page 188A)

Manufacturers' Cross-Reference Index

(Continued from page 186A)

Deltron Co., Inc.—Servo-Tek Products Co.
Deluxe Coils, Inc.—Crown Tool, Inc.; Wabash Magnetics, Inc.
Dempa Shinbun, Inc.—Japan Electric Industry Div.
Designatronics, Inc.—Precise Electronics & Development Corp.; Sterling Instrument Div.; Automatic Coil Div.
Desomatic Products, Inc.—Atlantic Research Corp.
Deutschmann Corp. (Tobe)—Cornell-Dubilier Electronics Div.
Devoe Electric Switch Co.—Radio Engineering Products
Di-Met Corp.—Felker Mfg. Co.
Dialight Corp.—Dialtron Corp.
Diamond Electronics Div.—Diamond Power Specialty Corp.
Diamond H Switches, Ltd.—Hart Mfg. Co.
Diaphlex Div.—Cook Electric Co.
Diehl Div.—Singer Mfg. Co., Military Div.
Digital Instrument Labs.—Telemeter Magnetics, Inc.; Ampex Corp.
Discus U.S.A., Inc.—Birmingham Sound Reproducers, Ltd.
Divco-Wayne Electronics—Burton-Rodgers, Inc.
Djeco Div.—Djordjevic Engineering Co.
Documentation, Inc.—Benson-Lehner Corp.
Dodge Mfg. Corp.—Chicago Thrift-Etching Corp.
Dominion Diesel, Ltd.—Burke Electric & X-Ray Co., Ltd.
Don-Lan Electronics—R S Electronics Corp.
Donlen Mfg. Co., Inc.—ADC, Inc.
Donner Scientific Co.—Systron-Donner Corp.
Dot Products Co.—United-Carr Fastener Corp.
Douglas Microwave Co., Inc.—Spectra Electronics Corp.
Douglas-Randall, Ltd.—Bourns, Inc.
Dow Corning Corp.—Corning Glass Works
Drayer-Hanson Div.—Crane Co.
Dresser Industries, Inc.—Dresser-Ideco Co.; Dresser Electronics; Hermetic Seal Transformer Co. (HST); Southwestern Industrial Electronics Co. (SIE)
Driver Co. (Wilbur B.)—Philadelphia Insulated Wire Co.; Western Gold & Platinum Co.
Ducon Condenser, Ltd.—Mallory & Co., Inc. (P. R.)
Duncan Electric Co., Inc.—Davenport Mfg. Co.; Technique Associates
Dunleavey Electronics Corp.—Topic, Inc.
Dunn Steel Products Div.—Townsend Co.
Dupar Canada, Ltd.—Dominion Electrohome Industries, Ltd.
Duplex Electric Co., Inc.—Mosler Research Products, Inc.
Duramic Products, Inc.—Accurate Specialties Co., Inc.
Durez Plastics Div.—Hooker Chemical Corp.
Dutch Brand Div.—Johns-Manville Co.
Dymec Div.—Hewlett-Packard Co.
Dyna-Labs.—Dyna Magnetic Devices, Inc.
Dyna Magnetic Devices, Inc.—Guidance Controls Corp.
Dynac, Inc.—Dymec Div.
Dynacor, Inc.—Sprague Electric Co.
Dynacron Electronic Corp.—Eastern Precision Resistor Co.
Dynametrics Corp.—Western Union Telegraph Co.
Dynamic Electronics Div.—Capehart Corp.
Dynamic Instrument Co., Inc.—Dynisco, Inc.
Dynamics Corp. of America—Radio Engineering Labs., Inc.; Reeves-Hoffman Div.; Reeves Instrument Corp.; Standard Electronics Div.
Dynisco Div.—American Brake Shoe Co.

E

EMI Electronics, Ltd.—EMI-Cossor Electronics, Ltd.; Hoffman Electron Tube Corp
EMT Corp.—Lieco, Inc.
E-S Industries, Inc.—Time-O-Matic Div.
Eagle Lock & Screw Co.—National Scientific Labs., Inc.
Eagle Signal Co.—Gamewell Co.
East Coast Aeroproducts Div.—Federal Mfg. & Engineering Corp.
Eastern Air Devices, Inc.—Kurman Electric Co.
Eastern Fabricating Co.—Cleveland Fabricating Co.

(Continued on page 190A)

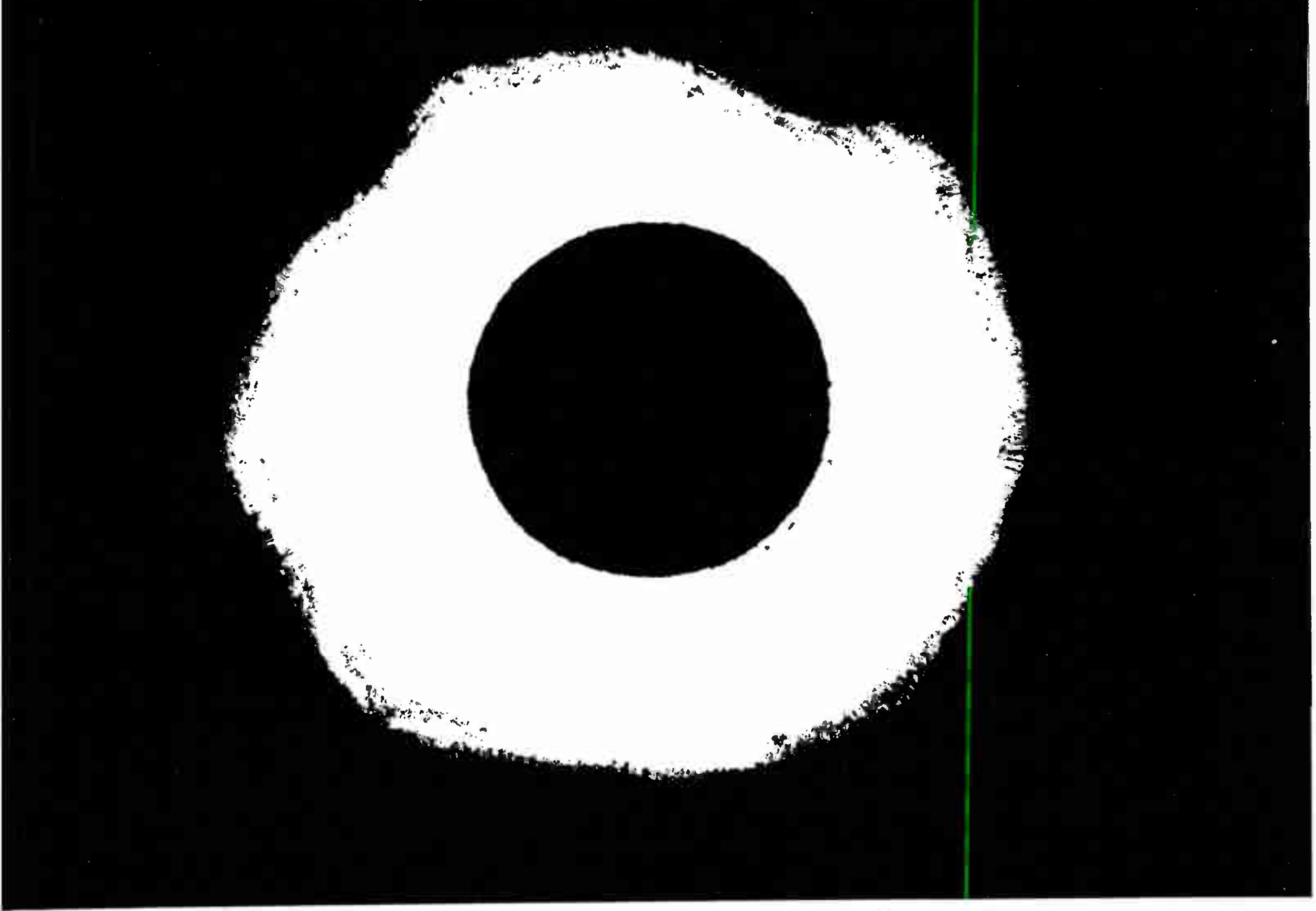
Manufacturers' Cross-Reference Index

(Continued from page 188A)

Eastern Numbering Machine Co.—International Eastern Co.
Eaton Paper Co.—Gorham Electronics Div.
Eclipse-Pioneer Div.—Bendix Corp.
Edcliff Instruments—Con-Elco Div.
Eddystone Div.—Baldwin-Lima-Hamilton Corp.
Edenfield Electric—Plug-In Instruments, Inc.
Edin Co.—Epsco, Inc.
Edison-Page—Page Communications Engineers, Inc.; Northrop Corp.
Educational Labs.—Reed Research, Inc.
Eemco Div.—Electronic Specialty Co.
Efcon, Inc.—General Instrument Corp.
Eico—EICO Electronic Instrument Co., Inc.
Eisler Transformer Co., Inc.—Eisler Engineering Co., Inc.
Eitel-McCullough, Inc.—Industrial Tubes, Inc.; National Electronics, Inc.
El-Tronics, Inc.—ALWAC Computer Div.; Robertson Electric Co., Inc.; Monarch Electric Div.; Warren Components Div.
Elasco, Inc.—Electronic Assembly Co., Inc.
Elastic Stop Nut Corp. of America—AGA Div.; Agastat Div.; Buchanan Electrical Products Corp.
Elastimold Div.—Elastic Stop Nut Corp. of America.
Elbeeco Div.—Aeroquip Corp.
Eldema Corp.—Genisco, Inc.
Eldico Electronics Div.—Dynamics Corp. of America
Eldon Industries, Inc.—Ungar Electric Tools Div.
Electra Screening Engineering, Inc.—Bellandi Co., Inc. (M.)
Electralab, Inc.—Electralab Printed Electronics Corp.
Electric Autolite Co.—C & D Batteries Div.
Electric Boat Div.—General Dynamics Corp.
Electric Cords & Supply Corp.—Royal Electric Corp.
Electric Distribution Products—Electric Machinery Mfg. Co.
Electric Eye Equipment Div.—Hurletron, Inc.
Electric Mfg. Co.—Gavitt Wire & Cable Co.
Electric & Musical Industries, Ltd.—E.M.I. Electronics, Ltd.
Electric Sorting Machine Co.—Mandrel Industries, Inc.
Electric Storage Battery Co.—Ray-O-Vac Co.
Electric Switch Corp.—Bernco Engineering Corp.
Electrical Communications, Inc.—Secode Corp.
Electrical Facilities, Inc.—Knopp, Inc.
Electrical Industries—North American Philips; Price Electric Corp.
Electrical Products Research & Development Co., Inc.—Eprad, Inc.
Electro Circuits, Inc.—Automation Industries, Inc.
Electro Devices, Inc.—Servospeed Div.
Electro Dynamics, Inc.—Electro Development Corp.
Electro-Fabricators—Blinn Co. (The Delbert)
Electro Instruments, Inc.—International Electronics Mfg. Co.; Antran Div.
Electro-Labs., Inc.—Hoffman Engineering Corp.
Electro-Machinery Div.—Design Tool Corp.
Electro Machinery Mfg. Co.—Mullenbach Div.
Electro-Measurements, Inc.—Electro Scientific Industries, Inc.
Electro-Mechanical Research, Inc.—ASCOP Div.; Computer Systems, Inc.
Electro-Mechanical Specialties Co., Inc.—Fidelitone, Inc.
Electro-Optical Systems, Inc.—Micro Systems, Inc.
Electro-Plex Div.—Nuclear Electronics Corp.
Electro-Precision Products, Inc.—Ecco Electronic Components Corp.
Electro-Pulse, Inc.—Servo Corp. of America
Electro Quip Control Div.—Filtors, Inc.
Electro-Security Corp.—Chapin Electronics, Inc.
Electro Snap Corp.—Controls Co. of America
Electro Switch Corp.—Electro Contacts, Inc.
Electro Tech Div.—Greenleaf Mfg. Div. of Systron-Donner Corp.
Electro-Technical Labs.—Mandrel Industries, Inc.

(Continued on page 192A)

An invitation to engineers and scientists interested in joining one of the most important developmental program areas of this generation—



RADIATION ELECTRONICS

is the unusually challenging area of NUDETS (477-L), a national network of nuclear detection and measuring devices now under development. Assignments for EE's include analysis, synthesis, integration, design, development of sensors, data processors and transfer equipment. Also damage assessment, hardening, equipment and component deterioration studies. BSEE or Physics; at least 2 years in solid state or experimental physics—instrumentation—circuit design. Also openings for ME's with electronic packaging experience. Write in confidence to Mr. P. W. Christos, Div. 53-ME.

DSD DEFENSE SYSTEMS DEPARTMENT

A Department of the Defense Electronics Division

GENERAL ELECTRIC

An Equal Opportunity Employer

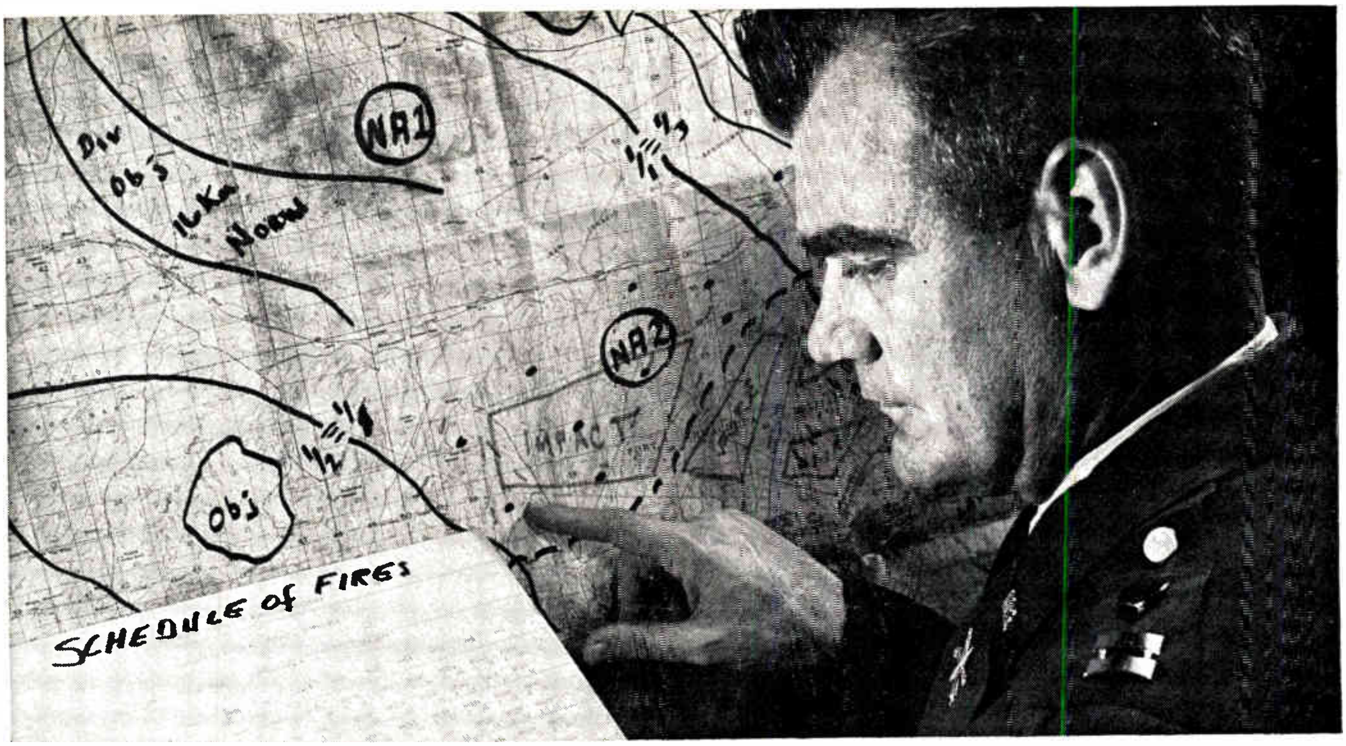
NORTHERN LIGHTS OFFICE BUILDING, SYRACUSE, NEW YORK

Manufacturers' Cross-Reference Index

(Continued from page 190A)

Electro-Therm, Inc.—American Instrument Co., Inc.
Electrocraft Co.—Electrocraft Components Div., GC-Textron Electronics, Inc.
Electrocraft Components Div.—GC Electronics Div.; Textron Electronics, Inc.; Richards Electrocraft
Electroflo Meters Co., Ltd.—Elliott-Automation, Ltd.
Electrol, Inc.—General Computers, Inc.
Electromatic Parts, Inc.—Electric Terminal Corp.
Electrometric, Inc.—Whitewater Electronics, Inc.
Electron Corp.—Ling-Temco-Vought, Inc.
Electron Div.—Controls Co. of America
Electron Products—Electro-Physics Labs.
Electron Research, Inc.—Eric Resistor Corp.
Electronic Apparatus Co.—Mercury Contacts, Inc.
Electronic Cable Corp.—Solartron, Inc.; Centroid Engineering & Mfg. Corp.
Electronic Counters, Inc.—Potter Instrument Co., Inc.
Electronic Devices Co., Ltd.—Nichols, Ltd. (R.H.)
Electronic Energy Conversion Corp.—Behlman Engineering Co.
Electronic Engineering Co. of California—Engineered Electronics Co.
Electronic Enterprises, Inc.—Keystone Electronics Co.
Electronic Fabricators, Inc.—Efcon, Inc.
Electronic Industries Mfg. Co.—Audio Electronic Corp.
Electronic Instrument Co.—EICO
Electronic & Marine Products, Inc.—General RF Fittings, Inc.
Electronic Materials Dept.—Cominco Products, Inc.
Electronic Noise Generator Co.—Elgenco, Inc.
Electronic Prototypes, Inc.—Pastoriza Associates, Inc. (James)
Electronic Research Assoc.—Era Electric Corp.
Electronic Researcb & Development Co.—Mandrel Industries, Inc.
Electronic Specialty Co.—Technicraft Div.
Electronic Stethoscope Corp.—Hamlin, Inc.
Electronics Corp. Pan America—Electronics Corp. of America
Electronics Development, Inc.—Sherman Industrial Electronics Co.
Electronics Guild—Bulova Watch Co.
Electronics International Co.—International Electronic Research Corp.
Electronized Chemicals Corp.—High Voltage Engineering Corp.
Electrons, Inc.—Compact Circuits Co.
Electropac, Inc.—Computer Control Co., Inc.
Electrosnap Switch Div.—Control Switch Div.
Electrospace Corp.—General Hermetic Sealing Corp.
Elektro-Serv Div.—Applied Technology Corp.
Elexso Corp.—Mandrel Industries, Inc.
Elgin Labs., Inc.—Erie Resistor Corp.
Elgin Metalformers Corp.—Emcor Ingersoll Products Div.; Borg-Warner Corp.
Elin Div.—International Electronic Research Corp.
Elk Co.—Magnetic Powders, Inc.
Elliott Bros. London, Ltd.—Elliott-Automation, Ltd.; Isotope Developments, Ltd.
Elliott & Evans—Telectron Div.
Elliott Nucleonics, Ltd.—Elliott-Automation, Ltd.
Elmeg Elektromechanik—Presin Co.
Elmenco Products Co.—Arco Electronics, Inc.
Elmet Div.—North American Philips Co., Inc.
Elsin Electronics Corp.—Specialty Electronics Development Corp.
Emarco Corp.—Mercury Contacts, Inc.
Emerson Radio & Phonograph Corp.—Granco Products, Inc.; Emertron, Inc.
Empresa Products de Aluminio, Ltda.—Southwire Co.
Encapsor Products Sales Corp.—Telechrome Mfg. Co.
Ender-Monarch Co.—Acoustica Assoc., Inc.
Enderes Div.—Muter Co.
Endevco Corp.—Digitran Co.

(Continued on page 194A)



AUTOMATED ARTILLERY FIRE PLANNING

TRADING HOURS FOR MINUTES WHEN IT COUNTS

Targets to destroy... troop movements to cover... a bridge to save... a gun emplacement to wipe out... a precisely coordinated schedule of firepower and manpower and brainpower.

The Divisional Fire Plan, prepared automatically by electronic computers in 12 minutes, tells the captain where every artillery battery will hit—and when. He checks his Schedule of Fires against the tactical map, applying human judgement, experience and intelligence to the information supplied by the computers. If he wants to, he can alter or refine the schedule—even scrap it completely and have a new one in his hands in a few minutes.

A simulation exercise today; a powerful system in the field tomorrow.

At recent exercises, an Army-Industry team successfully demonstrated that automatic data processing systems can help to reduce the time it takes to prepare a Divisional Fire Plan from a matter of hours to a matter of minutes. When computers process ballistic solutions at electronic speeds, they aid in achieving first-round target hits. This means saving ammunition—conventional or nuclear. It also means surprise—hitting the target while it is unprepared and vulnerable. The long-range program of developing Command Control Information Systems is now being extended to other Tactical Field Army functions, utilizing the FIELDATA family of equipment.

Development and testing of artillery fire planning and control systems are being done by the U.S. Army Electronic Proving Ground, Fort Huachuca, Arizona, with the cooperation and support of the U.S. Army Artillery and Missile School, Fort Sill, Oklahoma. Technical assistance is being provided by Ramo-Wooldridge, a Division of Thompson Ramo Wooldridge Inc.

Challenging openings now exist at Ramo-Wooldridge's Fort Huachuca office for experienced computer programmers, systems analysts and test and applications engineers. All qualified applicants will receive consideration for employment regardless of race, creed, color or national origin.

Contact Mr. R. J. Kremple

Thompson Ramo Wooldridge Inc. **RW** Division
8433 Fallbrook Ave., Canoga Park, Calif., Diamond 6-6000

Manufacturers' Cross-Reference Index

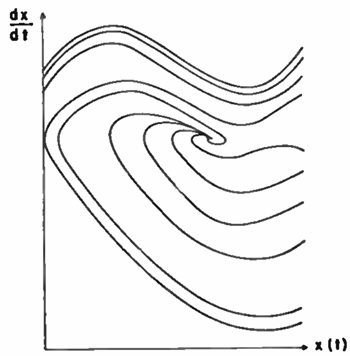
(Continued from page 192A)

(Continued on page 196A)

NON-LINEAR DIFFERENTIAL EQUATIONS

The non-linear differential equation shown above the phase-plane plot governs the behavior of one form of the well-known electronic circuit used for the reception of weak communication signals. The circuit, known as the phase-lock loop, has a mechanical analog consisting of a mass subjected simultaneously to non-linear restoring and resistive forces.

$$\frac{d^2x}{dt^2} + (\alpha + \kappa \cos x)\frac{dx}{dt} + \gamma \sin x = \beta$$



INTEGRAL CURVES IN THE PHASE PLANE

Simpler equations of this basic type have been studied for a number of years, and computers have produced many phase-plane plots for special values of the parameters, such as shown above. For fuller understanding of the circuit operation, however, new analytical methods are being investigated at Philco which may give an overall picture of the solutions of this, and even more general, equations.

This study is typical of the Philco Scientific Laboratory's activity in applied mathematics. Philco offers attractive staff positions in this field and in a number of others such as electrodynamic theory, optical communications, pattern recognition, and solid state electronics. Applications and inquiries may be addressed to: Mr. R. A. Ruth, Philco Scientific Laboratory, Blue Bell, Pennsylvania.

PHILCO SCIENTIFIC LABORATORY

PHILCO®

A SUBSIDIARY OF *Ford Motor Company*

an equal opportunity employer

if you have your eye on an exciting future......

■ ■ ■ ■ ■ ■ exciting things are happening at Eimac. Who's the world's foremost developer and manufacturer of electron power tubes? Eimac—with a long line of firsts in the field. The first modern beam magnetic optics. The first practical matrix cathode. The first metal-ceramic tubes. The first modulating anode.

Eimac tubes have powered the first satellite bounce communication. The longest command-response achieved by man. The first radar contact with the moon. Venus. And the sun.

Tomorrow: not even the sky's the limit.

You can make this exciting future **your** future. Each of our development laboratories has openings for microwave tube or power grid tube development engineers. If your forte is developing new higher power, higher frequency tubes, or cavities or test equipment. Or if you have a flair for solving tube production problems and advancing production techniques and equipment, there are exciting challenges to investigate at Eimac. You'll work in an area renowned for its concentration of electronics industry, sunny weather, fine schools and pleasant living conditions. If you can qualify, write: Mr. C. F. Gieseler, Department C, Eitel-McCullough, Inc., San Carlos, California. An equal opportunity employer.

KEEP YOUR EYE ON Eimac

Manufacturers' Cross-Reference Index

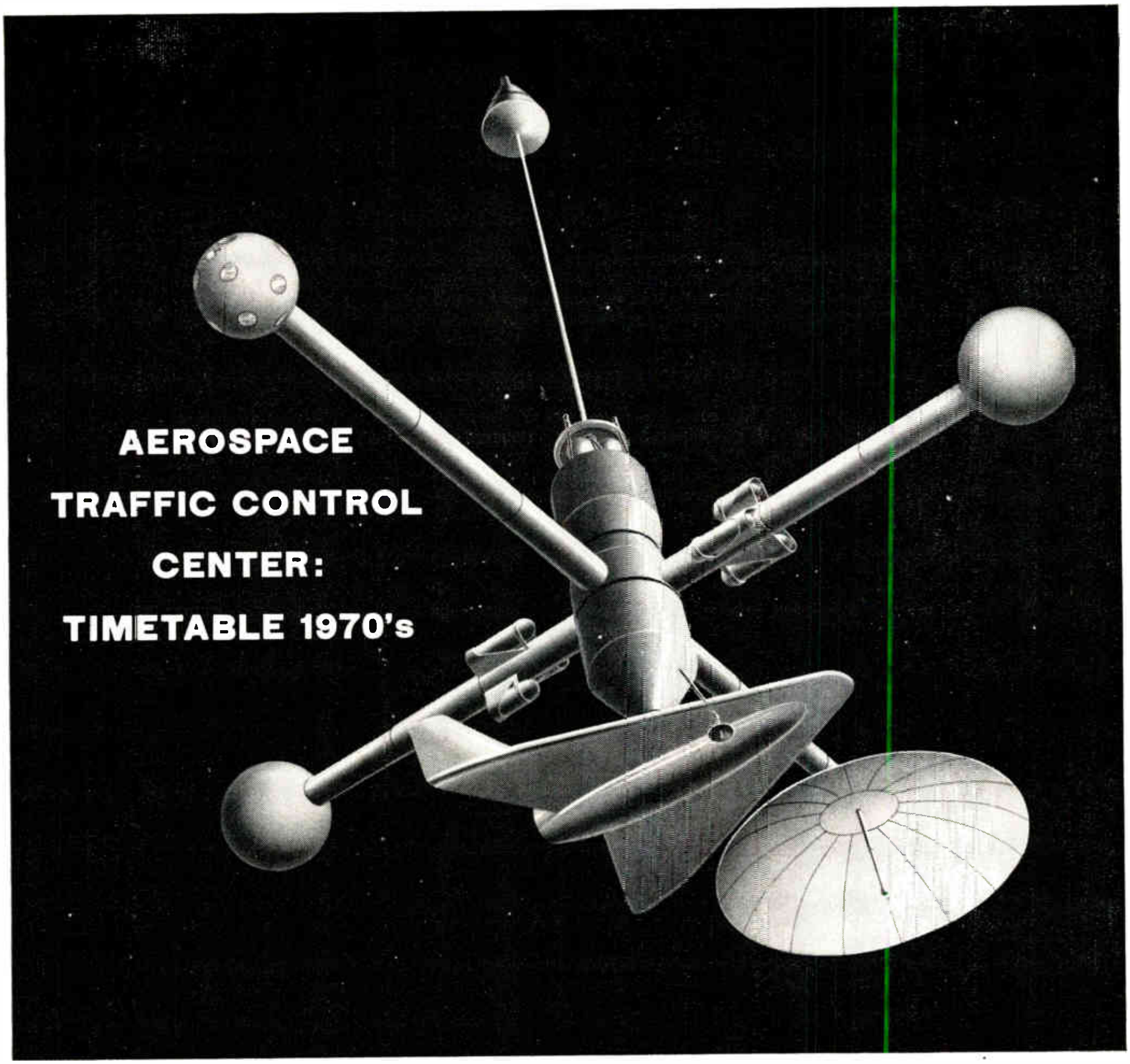
(Continued from page 194A)

(Continued on page 198A)

Manufacturers' Cross-Reference Index

(Continued from page 196A)

(Continued on page 200A)



A typical example of advanced design by Lockheed-California Spacecraft

Will command centers based in space be an outstanding development in the ten-year span from 1967 to 1977?

The answer at Lockheed-California Company's Spacecraft Organization is—Yes. Proof: the Spacecraft design pictured here. It reflects the maturity developed in our thinking about manned space systems.

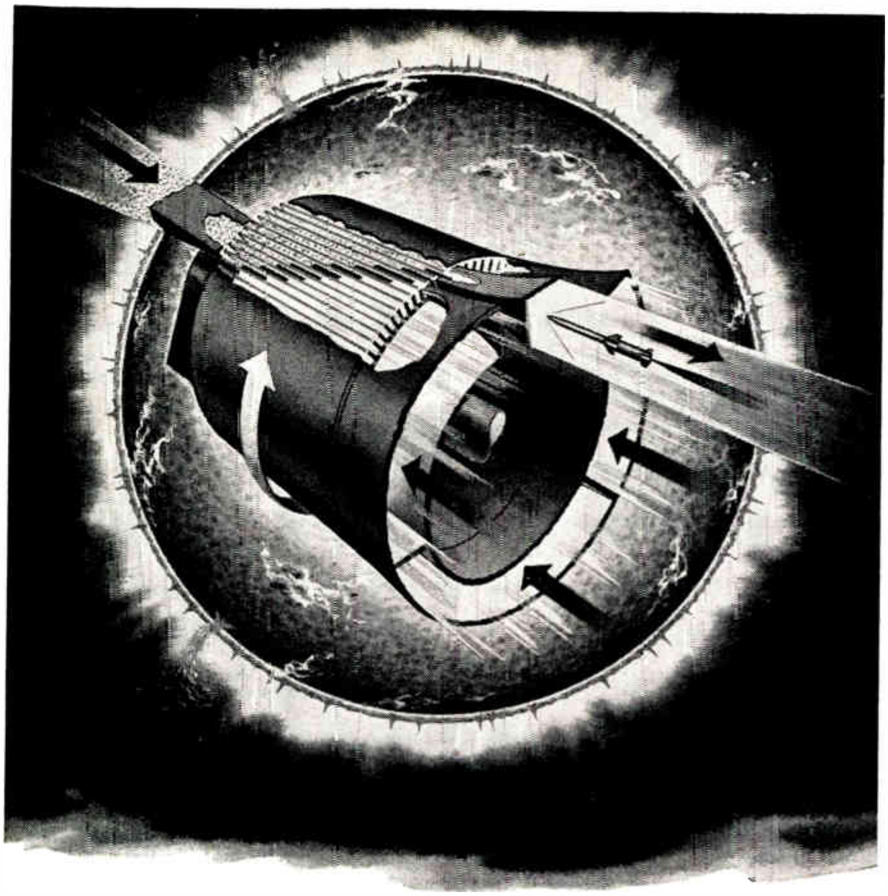
The station—aeroscope traffic control center of the 1970's—will be assembled in orbit in a series of pieces brought together by rendezvous techniques. Included: Command center; living quarters; maintenance station; radar and infrared sensing devices; nuclear power supply; communication links with the earth and other space vehicles. As now planned, 12 people will man the vehicle. Their tour of duty will be measured in weeks.

For four years Lockheed-California Spacecraft has concentrated on the needs of man in space. Activities embrace all fields pertaining to development of complex spacecraft as well as supporting technologies. An operation of such magnitude opens many doors of opportunity.

Scientists and Engineers of outstanding talent and training are needed to develop new Spacecraft, Aircraft, ASW concepts in: Human Factors; Physics (theoretical, plasma, high-energy, solid state, infrared, optics, nuclear); Thermodynamics; Servosystems; Reliability; Guidance and Control; Dynamics; Electronic Systems; Aerospace Ground Equipment; Bioastronautics; Systems Integration and Trade-Off; Space Mechanics; Sub-Systems Synthesis and Analysis; Nuclear, Electric and Liquid Rocket Propulsion; Electronics Research; Hydrodynamics. Send résumé to: Mr. E. W. Des Lauriers, Manager Professional Placement Staff, Dept. 1805, 2402 N. Hollywood Way, Burbank, California. An equal opportunity employer.

LOCKHEED *CALIFORNIA COMPANY*

A DIVISION OF LOCKHEED AIRCRAFT CORPORATION



Manufacturers' Cross-Reference Index

(Continued from page 198A)

General Precision, Inc.—GPL Div.; Kearfott Div.; Librascope Div.; Link Div.
General Precision Lab., Inc.—GPL Div.
General Railway Signal Co.—Electrons, Inc.
General Riveters, Inc.—General-Electro Mechanical Corp.
General Scientific Corp.—San Fernando Electric Mfg. Co.
General Steel Wares, Ltd.—Alliance Motor Div.
General Switch Co.—Kurman Electric Co.
General Telephone & Electronics Corp.—Automatic Electric Co.; Sylvania Electric Products Inc.; Lenkurt Electric Co., Inc.; Sylvania Electronic Systems
General Telephone & Electronics Labs., Inc.—Sylvania Electronic Systems
General Thermodynamics Corp.—Falcon Div.
General Time Corp.—Haydon Div.; Stromberg Time Corp.; Westclox Div.
General Tire & Rubber Co.—Bolta Products Div.
General Transistor Corp.—General Instrument Corp.
General Ultrasonics Co.—Acoustica Assoc., Inc.
Genisco, Inc.—Bekey Electric Div.; Eldema Corp.; Genistron, Inc.
Gentape Corp.—General Plastics Corp.
Geophysics Corp. of America—Itek Corp.
George Products Co., Inc.—Paramount Products Co., Inc.
Geosciences & Instrumentation Div.—Texas Instruments Incorporated
Gering Products, Inc.—Gering Plastics Div.; Studebaker-Packard Corp.
Getters Electronics, Inc.—Edwards High Vacuum (Canada), Ltd.
Giannini Co. (G.M.)—Giannini Controls Corp.
Giannini Controls Corp.—Datex Corp.
Giannini Scientific Corp.—Flight Research, Inc.; Wiley Electronic Products Co.
Gibson Div.—Associated Spring Corp.

(Continued on page 202A)

COMMUNICATIONS and RADAR & DATA-HANDLING ENGINEERS/SCIENTISTS

SADTC

The *SHAPE AIR DEFENCE TECHNICAL CENTRE* has a few vacancies for qualified engineers/scientists for technical investigations and planning studies concerning NATO Air Defence Systems. Candidates should be university graduates with at least 5 years' professional experience in a responsible position, and should have worked in one or more of the following fields:

COMMUNICATIONS

- —Line transmission techniques
- —Operation of communications networks
- —Digital techniques
- —Microwave systems and techniques
- —Scatter communications
- —HF radio voice or telegraph systems

RADAR & DATA-HANDLING

- —Military ground environment systems
- —Calibration, test and evaluation of military radar systems
- —Radar data extraction, IFF and data-handling systems
- —Radar electronic counter counter measure techniques
- —Air traffic control and guided weapon control systems
- —Digital data extraction, transmission and display techniques

Duties will include the writing of technical reports.

The SADTC is an international establishment located in The Hague, Netherlands, and has staff from many NATO nations. The work offers challenging problems and is of a unique international character.

The basic salary will depend on background and experience and will be supplemented by a foreign allowance of between 60 and 70%. Reimbursement of the cost of moving families and household effects to The Hague and back to the country of origin after termination of contract is made. Staff members benefit from a number of concessions such as freedom from income tax in the Netherlands on the total emoluments.

Applications, including detailed information on professional qualifications and experience, should be sent to: The Director, SHAPE Air Defence Technical Centre, P.O. Box 174, The Hague, Netherlands.

ELECTRONIC ENGINEERS

R & D

Outstanding career opportunities for engineers with at least a B.S. in Electrical Engineering or Physics to work in our Systems Engineering Group. Should have strong analytical capability with at least 3 to 5 years research experience in one or more of the following areas:

INFRARED DETECTION SYSTEMS
MEDICAL ELECTRONIC SYSTEMS
MAGNETIC TAPE RECORDERS

Duties entail project responsibility from the concept stage, through design, to model fabrication. The Mechanics Research Division provides scientific research and development services both for the government and industry.

Salaries commensurate with experience. Excellent opportunity for personal and professional growth contributing to diversified and challenging projects. Liberal company benefits including a full tuition reimbursement plan. Chicago suburb location. Kindly forward your resume in confidence to R. B. Martin, Dept. I-15.

AMF

AMERICAN MACHINE & FOUNDRY COMPANY

Mechanics Research Division
7501 North Natchez
Niles 48, Illinois

"An Equal Opportunity Employer"

Senior Physicist

Project Manager / Lunar Program

Well-established engineering firm now entering aerospace field offers participation in the early stages of broad new programs to a qualified physicist. Current programs in exotic aerospace research require Ph.D. with experience in diversified space exploration problems.

Strong theoretical and analytical background is desired, comprising some or all of the following: Optics and infra-red systems · Solar simulation · wave propagation · Kinetics · Quantum mechanics · Radiation · Detonations · Explosions.

SALARY: $15,000–$22,000
FEE PAID

Please submit resume in complete confidence to:
Engineering Dept.

Fanning
PERSONNEL AGENCY, INC.
"The Award-Winning Service"
180 Broadway, New York 38, N. Y.
DIgby 9-3800

Manufacturers' Cross-Reference Index

(Continued from page 200A)

(Continued on page 204A)

Manufacturers' Cross-Reference Index

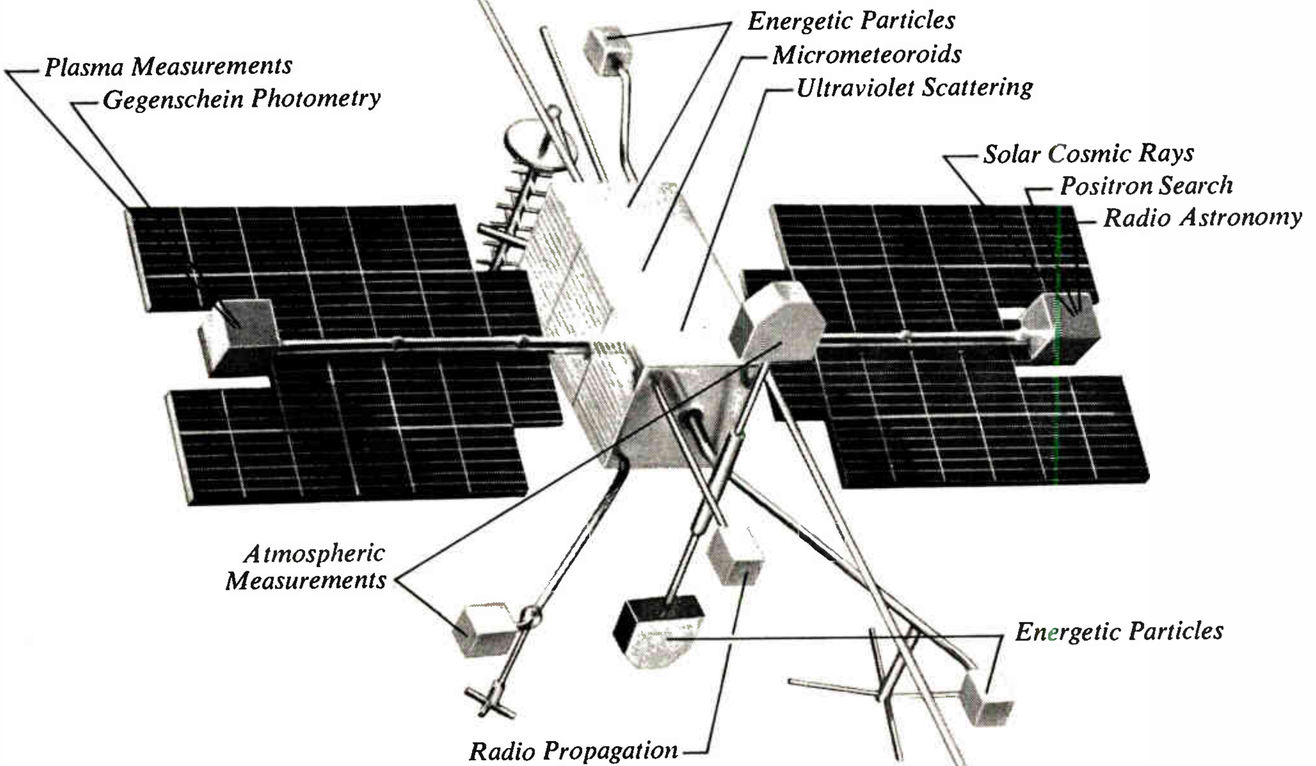
(Continued from page 202A)

Harris-Intertype Corp.—Gates Radio Co.; PRD Electronics, Inc.
Harrison Labs. (B.M.)—Harrison Electrosonics, Inc.
Hartech, Inc.—Diamond Tool Engineering Co.
Hartford Machine Screw Co.—Standard Screw Co.
Harvard Industries, Inc.—Frequency Standards Div.
Harvey-Wells Corp.—Hercon Electronics Corp., Harvey-Wells Electronics, Inc.
Hathaway Instruments, Inc.—Dale Products, Inc.; Dale Electronics, Inc.; Lionel Corp.; Hathaway-Denver Div.; Induction Heating Corp.
Haveg Industries, Inc.—American Super-Temperature Wires, Inc.
Haydon Co. (A.W.)—Haydon Instrument Co.; Haydon Switch, Inc.; North American Philips; Price Electric Corp.
Haydon Div.—General Time Corp.
Haydon Instrument Co.—Haydon Co. (The A.W.)
Haydon Switch, Inc.—Haydon Co. (The A.W.)
Haynes Stellite Co.—Union Carbide Corp.
Hazeltine Corp.—Wheeler Labs., Inc.
Heat-X, Inc.—Dunham-Bush, Inc.
Heavy Minerals Co.—Vitro Chemical Co.
Hedin Tele-Technical Corp.—General Automatic Corp.
Heeger, Inc.—LMB-Heeger, Inc.
Heiland Div.—Minneapolis-Honeywell Regulator Co.
Helipot Div.—Beckman Instruments, Inc.
Hemisphere Products, Inc.—Haveg Industries, Inc.
Henry Furnace Co.—Westinghouse Electric Corp.
Herbrand Div.—Van Norman Industries, Inc.
Hercon Electronics Corp.—Harvey-Wells Electronics Div.; Hermetic Connector Corp.
Hermes Electronics Co.—Hycon-Eastern, Inc.; Itek Corp.
Hermes Electronics Corp.—Western Union Telegraph Co.
Hermetic Connector Corp.—Hercon Electronics Corp.
Hermetic Pacific Corp.—Hermetic Seal Corp.
Hermetic Seal Corp.—Thermal Controls, Inc.
Hermetic Seal Transformer—Dresser/HST Electronics
Herron Optical Co.—Bausch & Lomb, Inc.
Hertner Electric Co.—General Precision, Inc.
Hetherington, Inc.—Controls Co. of America
Hetherington Switch Div.—Control Switch Div.
Hewlett-Packard Co.—Boonton Radio Corp.; Dymec Div.; Moseley Co. (F.L.); Palo Alto Engineering Co.; Sanborn Co.
Hi-G, Inc.—Control Indicating Corp.
Hi-G Relay, Inc.—Pendar, Inc.
Hi Products, Inc.—General Magnetics, Inc.
Hi-Q Div.—Aerovox Corp.
Hi-Speed Equipment, Inc.—Metal Fabricators Corp.
Hickok Electrical Instrument Co.—Supreme Electronics Corp.
Higbie Mfg. Co.—Avon Tube Div.
High Vacuum Equipment Corp.—Robinson Technical Products, Inc.
Highland Engineering Co.—Highland Design, Inc.
Hill Engineering & Mfg. Co.—Hill Electronics, Inc.
Hillburn Electronics Corp.—Loral Electronics Corp.
Himmel Engineering Co.—Universal Circuit Controls
Hirschmann Radiotecnisches Werk (Richard)—Rye Sound Corp.
Hobart Bros.—Motor Generator Corp.
Hobbs Corp. (John W.)—Stewart-Warner Corp.
Hodge Controls Div.—Ovitron Corp.
Hoffman & Co., Inc. (H.L.)—E.M.I. Electronics, Ltd.
Hoke, Inc.—Pamar Electronics, Inc.
Hol-Gar Mfg. Co.—Hollingsworth Co.
Holt & Co., Inc. (A.)—Composite Industrial Metals, Inc.

(Continued on page 206A)

Magnetic Fields

OGO: its first mission. Sometime in 1963, OGO (NASA's Orbiting Geophysical Observatory) will be launched into an elliptical orbit around the earth. It will gather, process and transmit data on the physics of near-earth and cislunar space. Here are some of the studies OGO may undertake in this initial flight: *Energetic particles,* with nine separate experiments on the flux and characteristics of these particles (including cosmic ray and plasma studies). *Radio propagation and astronomy,* through measurements of ambient radio energy not accessible from earth. *Micrometeoroids,* to determine the mass distribution and direction of interplanetary dust in the vicinity of earth. *Magnetic fields,* their intensity, direction and variation near earth and in space. *Atmospheric measurements,* to study the pressure, temperature and composition of earth and cislunar space. *Ultraviolet scattering,* from hydrogen in space. *Gegenschein photometry,* to study sunlight scattered by interplanetary matter. OGO will be launched into a wide range of orbits and may carry as many as 50 different experiments on each of its missions. This Orbiting Geophysical Observatory will be one of the most versatile earth satellites man has ever built.



**Captions indicate possible arrangement of instrumentation clusters which OGO may carry.*

OGO: its challenge. Today OGO demands advanced techniques in spacecraft design and development to meet its need for flexibility. It is a challenging responsibility to STL engineers, scientists and supporting personnel, who design it, fabricate it, integrate it, and test it. This versatile spacecraft will be manufactured at STL's vast Space Technology Center where expanding space projects (OGO, Vela Hotel and other programs) create immediate openings for engineers and scientists in fields such as Aerodynamics; Spacecraft Heat Transfer; Analog and Digital Computers; Applied Mathematics; Electronic Ground Systems; Power Systems; Instrumentation Systems; Propellant Utilization; Propulsion Controls; System Analysis; Thermal Radiation; Trajectory Analysis. For Southern California or Cape Canaveral positions, write Dr. R. C. Potter, One Space Park, Department —I, Redondo Beach, California, or P. O. Box 4277, Patrick AFB, Florida. STL is an equal opportunity employer.

VLF Radio Propagation
Magnetic Fields

SPACE TECHNOLOGY LABORATORIES, INC.
a subsidiary of Thompson Ramo Wooldridge Inc.

Los Angeles • Vandenberg AFB • Norton AFB, San Bernardino • Cape Canaveral • Washington, D.C. • Boston • Huntsville • Dayton.

Manufacturers' Cross-Reference Index

(Continued from page 204A)

(Continued on page 208A)

Manufacturers' Cross-Reference Index

(Continued from page 206A)

Hufford Corp.—Siegler Corp.
Huggins Labs., Inc.—Menlo Park Engineering
Hughes Aircraft Co.—Santa Barbara Research Center
Hughes Semiconductors—Hughes Aircraft Co.
Hughson Chemical Co.—Lord Mfg. Co.
Hull-Standard Div.—Hull Corp.
Huntington Alloy Products Div.—International Nickel Co., Inc.
Hupp Electronics Co.—JEM Electronics Corp.
Hurletron, Inc.—Wheaton Engineering Div.
Hussey Supply Div.—Aeroflex Corp.
Hycon Eastern, Inc.—Hermes Electronics Co.
Hycor Div.—International Resistance Co.
Hydraulic Research & Mfg. Co.—Bell Aerospace Corp.; Bell Aircraft Corp.
Hygrodynamics, Inc.—American Instrument Co., Inc.
Hymac Electronics Corp.—Minneapolis-Moline Co.; Molectronics Div.; Motec Industries
Hypro Tool Co.—Continental Screw Co.
Hyres Electrical Div.—Hydraulic Research Div.
Hytron Radio & Electronics, Inc.—CBS

I

IERC Div.—International Electronic Research Corp.
IMC Magnetics Corp.—Induction Motors of California; Induction Motors Corp.; PSP Engineering Co.
IMI Branch—Curtiss-Wright Corp.
IT-S, Ltd.—Illinois Tool Works
ITT Components Div.—International Tel. & Tel. Corp.
ITT Federal Div.—International Tel. & Tel. Corp.
ITT Industrial Products Div.—International Tel. & Tel. Corp.
ITT Labs.—International Tel. & Tel. Corp.
Illinois Precise Casting—Vap-Air Div.
Illinois Tool Works—Fastex Div.; Licon Switch & Control Div.; Paktron Div.; Pascol Div.; Shakeproof Div.
Imperial Radar & Wire, Inc.—Saxton Products, Inc.
Inca Mfg. Div.—Phelps Dodge Copper Products Corp.
Indar Corp.—Mallory & Co., Inc. (P.R.)
Indian Co.—Donnelly Mfg. Co.
Indiana General Corp.—Advanced Vacuum Products, Inc.; General Ceramics Corp.; Indiana Steel Products Co.
Indiana Products, Inc.—Standard Electrical Products Co.
Indiana Steel Products Co.—Indiana General Corp.
Indiana Steel & Wire Co., Inc.—General Cable Corp.
Induction Heating Corp.—Hathaway Instruments, Inc.
Induction Motors Corp.—IMC Magnetics Corp.
Induction Motors of California—IMC Magnetics Corp.
Industrial Camera & Instrument Corp.—Fairchild Camera & Instrument Corp.
Industrial Development Engineering Assoc., Inc.—I.D.E.A., Inc.
Industrial Development Labs., Inc.—Industrial Devices, Inc.
Industrial Engineering Co.—Vap-Air Div.
Industrial Hardware Mfg. Co., Inc.—Industrial Electronic Hardware Corp.
Industrial Precision Products Co.—IE Mfg.
Industrial Products Co.—Industrial Products-Danbury Knudsen Div.; RF Products Div. of Amphenol-Borg
Industrial Products-Danbury Knudsen Div.—Amphenol-Borg Electronics Corp.
Industrial Reactor Labs., Inc.—Corning Glass Works
Industrial Television, Inc.—ITI Electronics, Inc.
Industrial Timer Corp.—Line Electric Co.
Industrial Tubes, Inc.—Eitel-McCullough, Inc.
Industro, Inc.—Industro Transistor Corp.
Inet Div.—Leach Corp.
Information Science Center—Collins Radio Co.
Information Systems, Inc.—Chance Vought Electronics Div.; Panellit Div.; Ling-Temco-Vought, Inc.

(Continued on page 210A)

Manufacturers' Cross-Reference Index

(Continued from page 208A)

(Continued on page 212A)

Manufacturers' Cross-Reference Index

(Continued from page 210A)

Irvington-Baker Refining Div.—Engelhard Industries, Inc.
Irvington Div.—Minnesota Mining & Mfg. Co.
Isotopes Specialties Co., Inc.—Nuclear Corp. of America
Isotronics, Inc.—Constantin & Co. (L.L.)
Itek Corp.—Hermes Electronics Co.; Vectron, Inc.

J

J-B-T Instruments, Inc.—Shurite Meters Div.
James Vibrapowr Co.—James Electronics, Inc.
Jamieson Labs.—Thermo Cal, Inc.
Janline, Inc.—Mark Products Co.
Jansky & Bailey, Inc.—Atlantic Research Corp.
Jarco Services, Inc.—Weighing & Controls, Inc.
Jeffers Electronics Div.—Speer Carbon Co.
Jefferson Electric Co.—Jefferson Electronic Products
Jefferson, Inc. (Ray)—Jetronic Industries, Inc.
Jefferson Products Co.—Jefferson Wire & Cable Corp.
Jensen Mfg. Div.—Muter Co.
Jessall Plastics Div.—Electric Storage Battery Co.
Jet Vac Corp.—Metal Fabricators Corp.
Jetronic Industries, Inc.—Centronix, Inc.; Delta Coils, Inc.; Winston Electronics Div.
Jobbins Electronics Div.—Automation Industries, Inc.
Jobling & Co., Ltd. (J.A.)—Corning Glass Works
Johnson & Assoc. (C.M.)—Johnson Co., Inc. (Corydon M.)
Johnson & Johnson—Lumite Div.; Chicopee Mills; Permacel Div.
Johnson, Matthey & Co., Ltd.—Bishop & Co. (J.)
Johnson, Matthey & Mallory, Ltd.—Mallory & Co., Inc. (P.R.)
Johnson Research Corp.—Johnson Co., Inc. (Corydon M.)
Jones Div. (Howard B.)—Cinch Mfg. Co.
Jones Electronics Co. (M.C.)—Bendix Corp.
Jordan Co., Inc.—Jordan Controls, Inc.
Jordan Electronics Div.—Victoreen Instrument Co.
Jowil Electronics, Inc.—Avionics Corp. of America

K

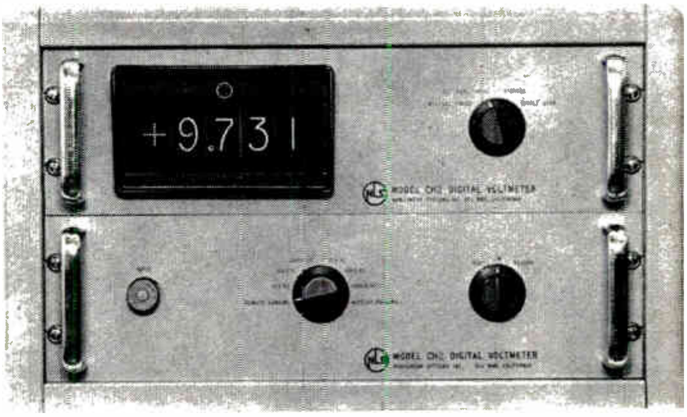
K-C-K Corp.—Texas Capacitor Co.
KVM Co.—Weighing & Controls, Inc.
Kaiser Industries Corp.—Kaiser Aircraft & Electronics
Kalbfell Electronix—San Diego Scientific Corp.; Instruments, Inc.
Kama Div.—Narda Microwave Corp.
Kaman Aircraft Corp.—Kaman Nuclear Div.
Kay Labs.—Kin Tel Div.
Kaylock Div.—Kaynar Mfg. Co., Inc.
Kearfott Div.—General Precision, Inc.
Kearfott Semiconductor Corp.—General Precision Equipment Corp.
Keleket X-Ray Corp.—Tracerlab, Inc.; Laboratory For Electronics, Inc.
Keller Tool Div.—Gardner-Denver Co.
Kellogg Switchboard & Supply Co.—International Tel. & Tel. Corp.
Kelsey-Hayes Co.—Utica Drop Forge & Tool Div.
Kemet Co.—Union Carbide Corp.
Ken-Tron Corp.—Waltham Horological Corp.
Ken-Wel, Inc.—Webster Electric Co.
Kenmark Corp.—Master Engraving Studios
Kennecott Copper Corp.—Chase Brass & Copper Co.; Okonite Co.
Kensico Tube Corp.—Robinson Technical Products, Inc.
Kent Mfg.—Thomas & Betts Co., Inc.
Kent Mill Div.—Beach-Russ Co.
Ketay Dept.—United Aircraft Corp.; Norden Div.
Key Electric Corp.—Digitronics Corp.
Keystone Electronics Co.—Electronic Enterprises, Inc.
Kidde & Co., Inc. (Walter)—Kidde Ultrasonic & Detection Alarms, Inc.

(Continued on page 216A)

NEW!

4 MAJOR DIGITAL VOLTMETER ADVANCEMENTS FROM NLS

Now Measure Varying Voltages With Great Accuracy . . . Make 15,000 Highly Accurate DC Measurements/Sec . . . Measure VDC-Ratio-Ohms to 5 Digits . . . Buy a Quality DVM for $1,460

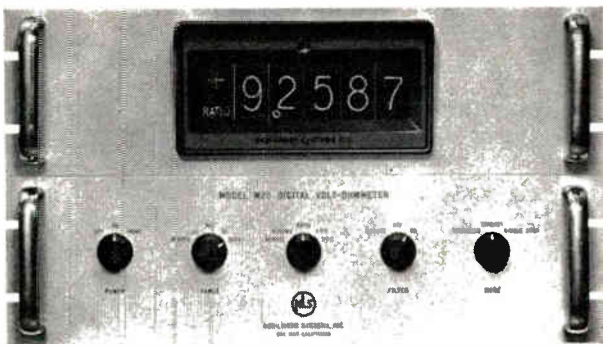


CH2 DIGITAL VOLTMETER

CH2 SAMPLE-AND-HOLD DIGITAL VOLTMETER — permits you to measure varying voltages at a definite point in time — with accuracy never before approached . . . starts to measure within 0.5 μsec., and is totally unaffected by voltage change during measurement . . . features accuracy of ±0.01% ±1 digit, constant input impedance of 10 megs, wide measuring range (up to 1,000 v.), and extremely high slewing rates up to 300,000 v./sec. (rate at which CH2 can "catch up to" and "follow" input changes). Price: $6,985.

MODEL 15 A/D CONVERTER

MODEL 15 A/D CONVERTER — 4-digit instrument bringing high accuracy to high-speed measuring and data logging . . . 15,000 measurements/sec . . . accuracy: ±0.01% ±1 digit from 0 to full scale from 0 to 40°C . . . any range from ±1 to ±100 v. full scale . . . true bipolar digital output, high output current, uses internal or external clock . . . constant input impedance. Price: $4,985.



M25 VOLT-RATIO-OHMMETER

M25 5-DIGIT VOLT-RATIO-OHMMETER — ultra-reliable instrument measuring DC volts, DC ratio and ohms with full 5-digit resolution of 0.001% and accuracy of ±0.01% of reading ± 1 digit over entire range of ±0.0001 to ±999.99 v. and .1 ohm to 1 meg . . . 10 to 1000 meg input z . . . twice speed of fastest stepping switch DVMs . . . advanced transistor circuitry with ultra-reliable mercury relays with 171 years life expectancy . . . input filter . . . remotely programmable . . . fully automatic . . . data logging output . . . AC or low-level DC with accessories. Price: $5,985

484A VOLTMETER-RATIOMETER

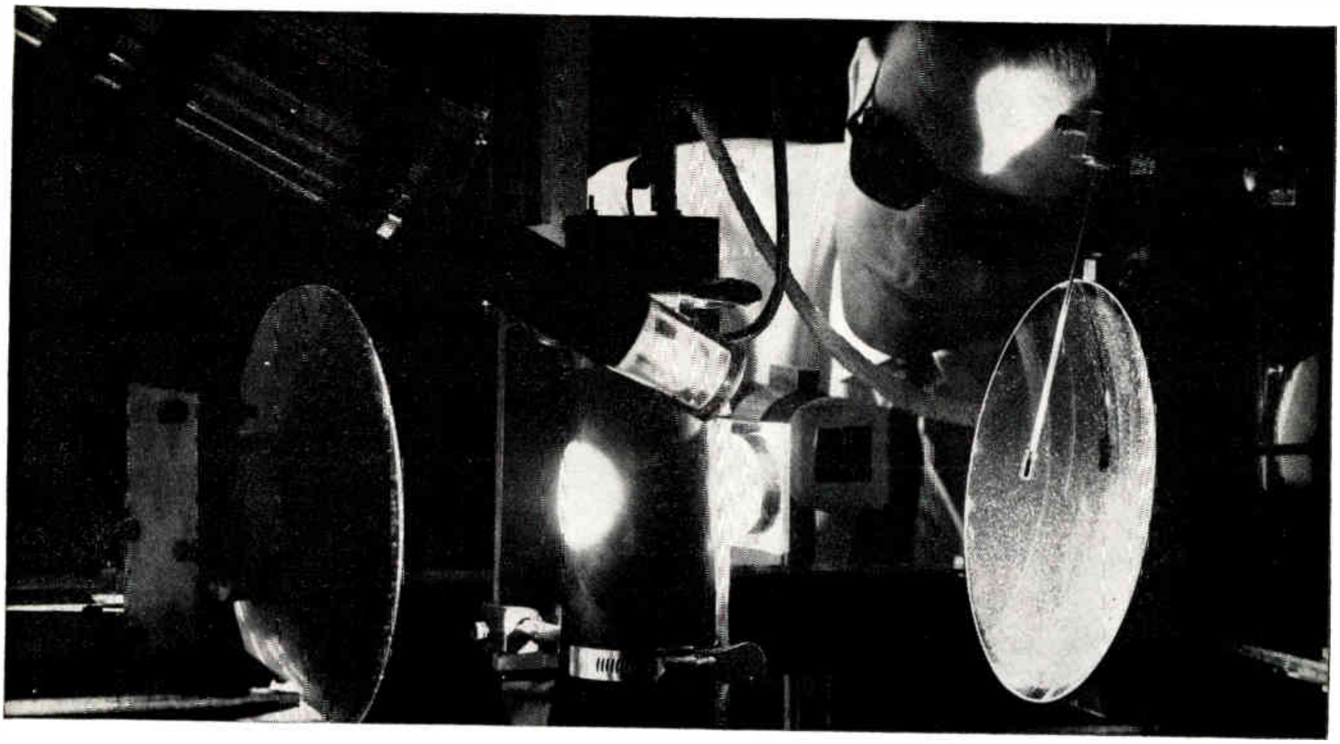
484A DIGITAL VOLTMETER-RATIOMETER — most versatile and highest quality instrument at low cost . . . measures ±0.001 to ±999.9 VDC and DC ratio up to ±99.99% . . . ±0.01% accuracy . . . 10 meg input . . . auto range and polarity . . . input filter . . . built-in auto print control . . . only a few dollars more than cheapest DVMs without such quality features as plug-in stepping switches that can be replaced in seconds for troubleshooting, snap-out readout, wire-wound resistors, epoxy fiberglass circuit boards . . . measures AC or low-level DC with accessories. Price: $1,460. **784 DIGITAL OHMMETER** — low-cost companion to 484A with same high quality features: $1,460.

LOOK TO NLS FOR DIGITAL VOLTMETERS — These instruments represent additional measuring achievements from the company that offers you the world's most complete line of digital voltmeters and associated instruments — most complete by purpose and by price. Prices are F.O.B. destination in U.S.A.

Originator of the Digital Voltmeter

non-linear systems, inc.

DEL MAR, CALIFORNIA



Exploring the possibilities in Coherent Light

At Bell Laboratories, Donald F. Nelson studies a beam of coherent red light produced by a continuously operating ruby optical maser. The heart of the device is a uniquely shaped ruby crystal immersed in liquid nitrogen in the tubular glass dewar extending from upper left to center. Light from the mercury arc lamp (lower center) is reflected by round mirror at left to mirror at right and then is focused on the ruby crystal to produce maser action. Coherent light emerging from end of dewar is picked up by a detector.

Is it feasible to take advantage of the enormous bandwidth available at optical frequencies? Could coherent light, for example, be sent through protecting pipes to provide high-capacity communication channels between cities?

To study such possibilities it is, first of all, necessary to have a source of continuous coherent radiation at optical frequencies. Such a source was first produced when Bell Laboratories scientists developed the gaseous optical maser.

Recently, our scientists demonstrated the generation of continuous coherent light by solid materials. Using a crystal of neodymium-doped calcium tungstate, a material developed at Bell Laboratories, continuous optical maser action was obtained in the near infrared. It has also been attained with visible light, using a new optical "pumping" arrangement to excite a ruby crystal. (See illustration above.)

Multichannel light highways for communications are still far from realization. But with continuous sources of coherent light available, it becomes possible to explore the problems of modulating, transmitting, detecting, amplifying and, in general, controlling light for possible communications applications.

BELL TELEPHONE LABORATORIES

World center of communications research and development

From G-V® the world's most experienced time delay engineers!

INSTANT RESET THERMAL TIMING ELEMENT, TYPE DM

- Used in industrial and communications equipment
- Instant reset during or after delay
- Delay intervals from 2½ seconds to 6 minutes
- Heater voltages from 6.3 to 230 volts
- Full temperature compensation

RED LINE THERMAL TIMING RELAY

- The low cost timing relay
- Rugged, steel sheathed heater
- Shatter-proof — no glass
- Dust tight metal enclosure
- Tamper proof
- Factory preset
- Octal size
- Available from stock

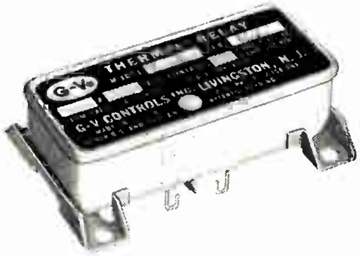
Stock Time Delays: 2, 5, 10, 20, 30, 45, 60, 90, 120 and 180 seconds

Stock Heater Voltages: 6.3, 26.5, 115 and 230 volts

VOLTAGE & CURRENT SENSING RELAYS

- For overload protection, voltage regulation, overvoltage cut off, under voltage alarm, low frequency cut off, battery charge control
- Available in R, T, and H Series, Miniature Size — Flanged or Plug-in
- Heater voltages from 2 to 230 volts
- Heater currents from .015 to 5 amps
- Heater resistances from 1 to 12,000 ohms

PT THERMAL RELAY



- The most precise, sturdiest thermal relay ever built
- Operates up to 2,000 cps, 20 g vibration and 50 g, 11 ms shock
- Time delay: 3 to 120 seconds. Tolerance ±5%.
- Heater voltages: 6.3 to 115 v. for delays up to 12 sec; 6.3 to 230 v. for longer delays.

ADJUSTABLE TRANSISTORIZED TIME DELAYS

- Time Delays—1 millisecond to 10 minutes
- Superior recycling characteristics both during and after timing cycles
- Protected against premature operation due to transient voltage surges
- Protected against lengthening or shortening delay due to change in voltage during timing
- Protected against damage through applying reverse polarity

H SERIES TIME DELAY RELAYS

- Hot wire heaters
- Short delays — rapid recovery
- Hermetically sealed, still adjustable
- Meet military requirements
- Miniature Size, Flanged or Plug-in

For time delays of 1/10 to 5 seconds.

Heater voltages up to 32 volts AC or DC.

ELECTRICAL THERMOSTATS

- Hermetically sealed, still adjustable
- Fast response to temperature change
- Bracket, flange or pipe plug mounting
- Vibration up to 2,000 cycles per second
- Accurate repeatability

Operating temperatures from —65° to +300°F.

Contact ratings to 1,000 watts.

R SERIES TIME DELAY RELAYS

- Metal sheathed heaters
- Hermetically sealed, still adjustable
- Meet military requirements
- Miniature Flanged, Miniature Plug-in or Octal Size

For time delays of 12 to 240 seconds.

Heater voltages up to 230 volts AC or DC.

T SERIES TIME DELAY RELAYS

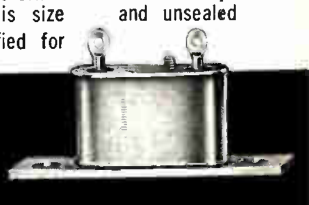
- Ribbon heaters
- Medium delays–faster recovery
- Hermetically sealed, still adjustable
- Meet military requirements
- Miniature Flanged, Miniature Plug-in or Octal Size

For time delays of 3 to 12 seconds.

Heater voltages up to 160 volts AC or DC.

VE SERIES MINIATURE ELECTRICAL THERMOSTATS

- Temperature operated switches in crystal can size
- Twice the contact motion and twice the contact moving force ordinarily found in thermostats of this size
- Fully qualified for military applications under vibration up to 2,000 cps, yet economical enough for commercial use
- Available factory-set and hermetically sealed or adjustable and unsealed



For complete information, write today or call (Area Code 201) 992-6200, Mr. George Compton.

G-V®

G-V CONTROLS INC.
Livingston, New Jersey

Manufacturers' Cross-Reference Index

(Continued from page 212A)

Kimberly-Clark Corp.—Schweitzer Div. (Peter J.)
Kin Tel Div.—Cohu Electronics
Kindt Design Service—Trought Associates, Inc.
Kinetics Corp.—High Vacuum Equipment Corp.
Kingston Industries, Inc.—Kingston Electronics Div.
Kinney Vacuum Div.—N.Y. Air Brake Co.; Kinney Mfg. Div.
Kintronic Div.—Chicago Aerial Industries, Inc.
Kip Electronics Div.—York Research Corp.
Klar & Beilschmidt—Electronic Applications, Inc.
Knapic Electro-Physics—Columbus Electronics Corp.
Knob Corp. of America—TV Development Corp.
Koiled Kords, Inc.—Whitney Blake Co.
Kolcast Industries Div.—Thompson Ramo Wooldridge Inc.
Kollmorgen Corp.—Inland Motor Corp. of Virginia
Kollsman Instrument Corp.—Kollsman Motor Corp.; Richardson-Allen Corp.; Standard Coil Products Co., Inc.; Standard Kollsman Industries, Inc.
Kollstan Semiconductor Elements, Inc.—Kollsman Instrument Corp.
Kolux Corp.—Victoreen Instrument Co.
Kontro Co.—Metal Fabricators Corp.
Kraftware Corp.—Henry & Miller Industries, Inc.
Kriser Corp.—Lektra Labs.
Kruger Instruments (Harold)—Beckman Instruments, Inc.
Kuehne Mfg. Co.—KTV Tower & Communication Equipment Co.
Kulite-Bytrex Corp.—Bytrex Corp.; Kulite Semiconductor Products, Inc.; Kulite Tungsten Co.

Kuthe Labs., Inc.—International Tel. & Tel. Corp.

L

LEAD Div.—Lockheed Electronics & Avionics
L.G. Electronic Co., Inc.—Great Eastern Mfg. Co.
LME Dept.—General Electric Co
LMK Mfg. Co.—Dawe Instruments, Ltd.
L-R Mfg. Co.—Ripley Co., Inc.
Labline, Inc.—Hudson Bay Co.
Laboratory For Electronics, Inc.—Eastern Industries Div.
Lacey Mfg. Co., Inc.—Barden Corp.
Laible Mfg. Co.—Daven Co.; General Mills, Inc.
Lake Erie Machinery Co.—Bell Aircraft Corp.
Lakeshore Div.—Bendix Corp.
Lamco Div.—Vocaline Co. of America, Inc.
Land-Air, Inc.—Chicago Electronic Div.; Stepper Motors Div.
Langevin Div.—Maxson Corp. (W.L.)
Lansdale Div.—Philco Corp.
Latrobe Plant—Carborundum Co.
Leach & Garner Co.—General Findings & Supply Co.
LearCal Div.—Motorola Aviation Electronics, Inc.
Leesona Corp.—Patterson Moos Research, Inc.; Leesona Moos Labs. Div.
Leetronics, Inc.—Lee Spring Co.
Lehigh Engineering Assoc.—Lehigh Design Co., Inc.
Leich Electric Co.—Sylvania Electronic Systems Div.
Leland Airborne Products—American Machine & Foundry Co.
Leland, Inc. (G.H.)—Ledex Corp.
Lenkurt Electric Co., Inc.—Sylvania Electronic Systems Div.
Lerco Electronics, Inc.—Micro Gee Products, Inc.
LeRoi Div.—Union Switch & Signal Div.
LeTourneau-Westinghouse Co.—Union Switch & Signal Div.
Levinthal Electronic Products—Radiation, Inc. (Radiation at Stanford)
Lewyt Mfg. Corp.—Budd Electronics, Inc.
Leybold Hoch Vakuum Anlagen GmbH—NRC Equipment Corp.
Librascope Div.—General Precision, Inc.
Licon Div.—Illinois Tool Works
Lignacraft Design & Drafting Service—Wyrco Projects, Inc.
Lima Electric Motor Co., Inc.—Consolidated Diesel Electric Corp.
Linde Co.—Union Carbide Corp.
Lindsay Structures, Inc.—Lindsay Div.
Line Electric Co.—Industrial Timer Corp.
Line Scale Co., Inc.—Quick Charge Div.
Ling-Altec Electronics, Inc.—Ling Electronics
Ling Systems, Inc.—Missile Systems Corp.
Ling-Temco-Vought, Inc.—Altec Lansing Corp.; Altec Service Co.; Calidyne Co., Inc.; Chance-Vought Div.; Continental Electronics Mfg. Co.; Electron Corp.; Fenske, Fedrick & Miller, Inc.; Information Systems, Inc.; Ling Electronics Div.; Peerless Electrical Products Div.; Temco Electronics; Ling-Altec Div.; United Electronics Co.; University Loudspeakers, Inc.
Link Aviation, Inc.—Link Div.; General Precision, Inc.
Link-Belt Co.—Syntron Co.
Linlar, Inc.—Permoflux Products Co.
Lion Fastener Co., Inc.—Southco Div.; South Chester Corp.
Lionel Corp.—Anton Electronic Labs., Inc.; Anton Machine Works; Telerad Div.
Lionel Electronic Labs.—Anton Electronic Labs.
Liquid Carbonic Div.—General Dynamics Corp.
Lisle Instrument Systems—Automatic Timing & Controls, Inc.
Little Giant Pump Co.—Quick Charge Div.
Litton Industries, Inc.—Airtron, Inc.; Monroe Calculating Machine Co.; Triad Transformer Corp.; Utrad Corp.; Westrex Co.
Litton Systems, Inc.—Westrex Co.
Logistics Research—Alwac Computer Div.
Lorain Tool Enterprises, Inc.—Telkor, Inc.
Loral Electronics Corp.—Alpha Wire Corp.
Los Angeles Coil Co.—LMB-Heeger, Inc.
Los Angeles Resistor, Inc.—Pendar, Inc.
Lourdes Instrument Corp.—Hudson Bay Co.
Louthan Mfg. Co.—Ferro Corp.
Lumite Div.—Chicopee Mills, Inc.
Lund Aviation, Inc.—Land-Air, Inc.
Luther Mfg. Co.—Giannini Controls Corp.
Lycoming Div.—Avco Corp.

Lynchburg Transformer Co., Inc.—Toroids Unlimited, Inc.

M

MB Mfg. Co.—MB Electronics Div.
MCM Machine Works—Quick Charge Div.
M & C Nuclear, Inc.—Texas Instruments Incorporated
MFR Electronic Instruments—Analab Instruments
MRC Coating Corp.—Materials Research Corp.
MSI Telecommunications, Inc.—Microwave Services, Inc.
M.V.M., Inc.—American Elite, Inc.
Macallen Co., Inc.—Insulation Mfrs. Corp.
Machlett Labs., Inc.—Raytheon Co.
Mack Electronics—Airtronics International Corp.
Mack Electronics Div.—Electronic Assistance Corp.
Mackenzie Products Co.—Deitz Co., Inc.
Magna Products, Inc.—Thompson Ramo Wooldridge Inc.
Magnaflux Div.—General Mills, Inc.
Magnavox Co.—Magnavox Research Labs.
Magne-Head Div.—General Transistor Corp.; General Instrument Corp.
Magnecord, Inc.—Midwestern Instruments, Inc.
Magnetic Powders, Inc.—National-U.S. Radiator Corp.
Magnetic Recording Industries, Inc.—Thompson Ramo Wooldridge Inc.
Magnetic Shield Div.—Perfection Mica Co.
Magnetic Specialties, Inc.—Nothelfer Winding Labs., Inc.
Magnetic Windings Div.—Essex Wire Corp.
Magnetics, Inc.—Control Div.
Magnodyne Corp.—Cornwell Electronics Corp.
Majestic—Philips Electronics Industries, Ltd.
Makepeace Div. (D.E.)—Union Plate & Wire Co.; Engelhard Industries, Inc.
Mallory Batteries, Ltd.—Mallory & Co., Inc. (P.R.)
Mallory & Co., Inc. (P.R.)—Radio Materials Co.; Mallory Controls Co.
Mallory Metallurgical Products, Ltd.—Mallory & Co., Inc. (P.R.)
Mallory Sharon Metals Corp.—Mallory & Co., Inc. (P.R.)
Manning, Maxwell & Moore, Aircraft Products Div.—Consolidated Controls Corp.
Mansol Ceramics, Inc.—Faradyne Electronics Corp.
Mantec Div.—Wyle Mfg. Corp.
Mapico Iron Oxides Unit—Columbian Carbon Co.
Marathon Special Products Corp.—Marathon Electric Mfg. Corp.
Marconi's Wireless Telegraph Co., Ltd.—Marconi Instruments
Marcy Spraying & Finishing Corp.—Dolin Metal Products, Inc.
Marine Electric Corp.—Galbraith-Pilot Marine Corp.
Marion Instrument Div.—Minneapolis-Honeywell Regulator Co.; Precision Meter Div.
Mark Mobile, Inc.—Mark Products Co.
Marma Electronic Co.—Godfrey Mfg. Co.
Marman Div.—Aeroquip Corp.
Marquardt Corp.—Cooper Development
Marshall Co. (G.S.)—Wahlgren Magnetics; Marshall Industries; Electro-Physics Labs.
Marshall-Eclipse Div.—Bendix Corp.
Martin Co. (A.M.)—Atlantic Research Corp.
Martin Furniture Div.—Electro-Voice, Inc.
Martin-Marietta Corp.—Martin Co.; American-Marietta Co.
Maryland Ceramic & Steatite Co.—Maryland Lava Co.
Maryland Electronic Mfg. Co.—Litton Industries, Inc.
Massa Labs., Inc.—Massa Div.
Massachusetts Gear & Tool Co.—Geartronics Corp.
Master Craft Trailers, Inc.—Cadre Industries Corp.
Mastermark, Inc.—Master Engraving Studios
Mattern X-Ray Div.—Land-Air, Inc.
Maxson Instruments Div.—Maxson Corp. (W.L.)
Mayhew Machine Co.—Gardner-Denver Co.
McAleer Mfg. Co.—Avon Tube Div.
McCauley, Inc. (L.D.)—Robertson Electric Co., Inc.
McColpin-Christie Corp.—Christie Electric Co.
McElroy Corp.—Warren Mfg. Co., Inc.
McGraw-Edison Co.—Bussmann Mfg. Div.; Edison Industries (Thomas A.); Measurements Div.; Pittsburgh Lectrodryer Div.

(Continued on page 220A)

WE SHARE YOUR PRIDE

on this, the 50th anniversary of the Institute of Radio Engineers, because one of the original founders of the I.R.E. also founded our company, the HOGAN FAXIMILE CORPORATION.

The man who helped found us both: the late Dr. John V. L. Hogan, pioneer in radio, television, and facsimile development.

We here at Hogan Faximile try to continue Dr. Hogan's lifelong tradition of attacking new frontiers in the communication and allied fields. And we've been fortunate enough to have made some valuable gains. For instance, through engineering and chemical R & D teamwork, Hogan Faximile has developed and produced...

...*High-speed plotters and printer plotters* for Polaris, B-70, Nike Zeus, and other programs. These plotters are used in data logging, countdown recording, flight test data reduction, and other applications.

...*Event-analyzing recorders,* which can also provide instantaneous spectrum analysis.

...*A high-speed data retrieval system* to print out copy from microfilm or hard copy by remote.

...*A microdensitometer* to measure both the line and background density of microfilm, so as to determine proper settings for reproducing microfilm of a predetermined quality.

...*Fully transistorized commercial facsimile equipment* to operate at speeds of a page in six minutes, three minutes, or one minute. This equipment contains such advanced features as automatic background control, automatic stack feed, and automatic copy cut-off.

...*Electrolytic recording paper* capable of high marking speeds and with desired characteristics for special applications.

Hogan Faximile know-how in the graphic communications and allied fields can assist you in obtaining your project goals.

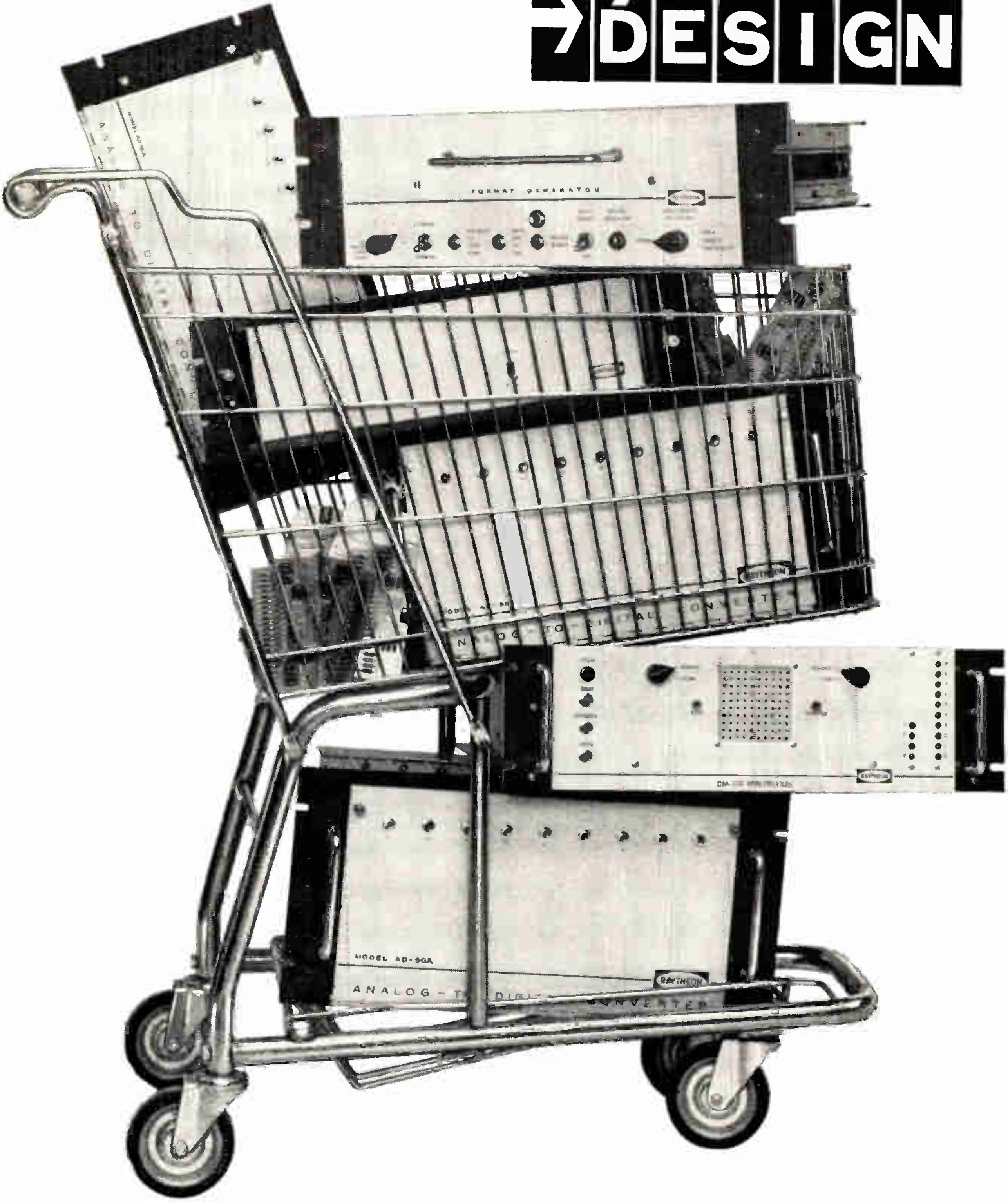
HOGAN FAXIMILE CORPORATION

635 GREENWICH STREET • NEW YORK 14, NEW YORK

Off the Shelf!

Raytheon Data-Design's High Speed Digital Handling Units

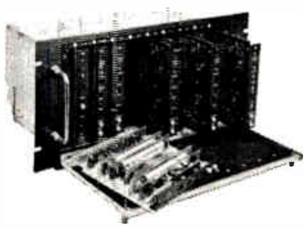
DATA-DESIGN



Complete Line . . .

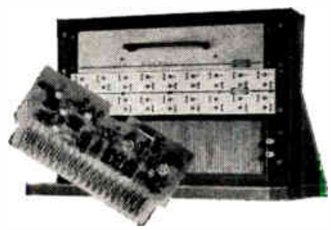
The Raytheon DATA-DESIGN units described below are unmatched in the digital field today for speed, accuracy and expandability. Together they form the fastest, most expandable data acquisition and conversion systems ever built. Each of these units is now available off-the-shelf from Raytheon, for system integration by the user.

Also available from Raytheon's unique DATA-DESIGN Service is the design and manufacture of complete digital systems or units based on your individual requirements and specifications.



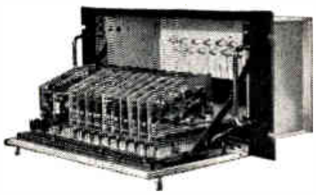
AD-50A Ultra-High Speed
ANALOG-DIGITAL CONVERTER

Up to 5 million samples/second. Up to 8-bit accuracy output, with lower bit outputs available at reduced cost. Digitizes pulses less than 0.3 microseconds wide. Direct, useable readout every 0.2 microseconds. 25 nanosecond aperture time. Every output valid and useable immediately. No need for sample and hold circuitry in most applications.



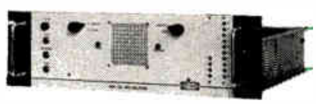
High Speed
MEGACYCLE BUILDING BLOCKS

Adaptive circuitry adjusts power requirements to load for minimum power drain. In quiescent state, draw 1/7th the power of conventional blocks. Temp. Specs. –30° to –65°C. "Worst Case Analysis" criteria used in all block designs. Interchangeable wiring panels mean only one step from logical design to finished equipment.



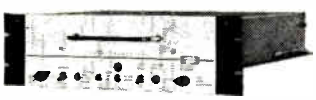
AD-10A High Speed
ANALOG-DIGITAL CONVERTER

Up to 500,000 complete 10-bit conversions/second. Serial readout at 5-mc bit rate; parallel readout at 500-kc word rate. Plug-in module construction. Synchronous or asynchronous operation.



DM-120 High Speed
MULTIPLEXER

Up to 1 million analog switchings/second. Expandable from 2 to 24 channels. As many as 576 channels obtained by stacking up to 24 Raytheon DM-120's. Compatible with both Raytheon Converters and with Raytheon Power Supply Units. For Multiplexer-Converter System, one Raytheon DPS-1A will power both units.



FORMAT GENERATOR

Arranges and prepares binary data from the conversion equipment in proper format for suitable entry into a computer or a tape storage unit. Also provides record identification information and parity check pulses on each character to be entered. Available in formats compatible with most major computers.

RAYTHEON

MAIL TODAY!

RAYTHEON COMPANY — PI-5
1415 Providence Turnpike
Norwood, Massachusetts

Please send DATA-DESIGN literature on following:

____

____

Name______________________

Title______________________

Company___________________

Address____________________

City and State______________

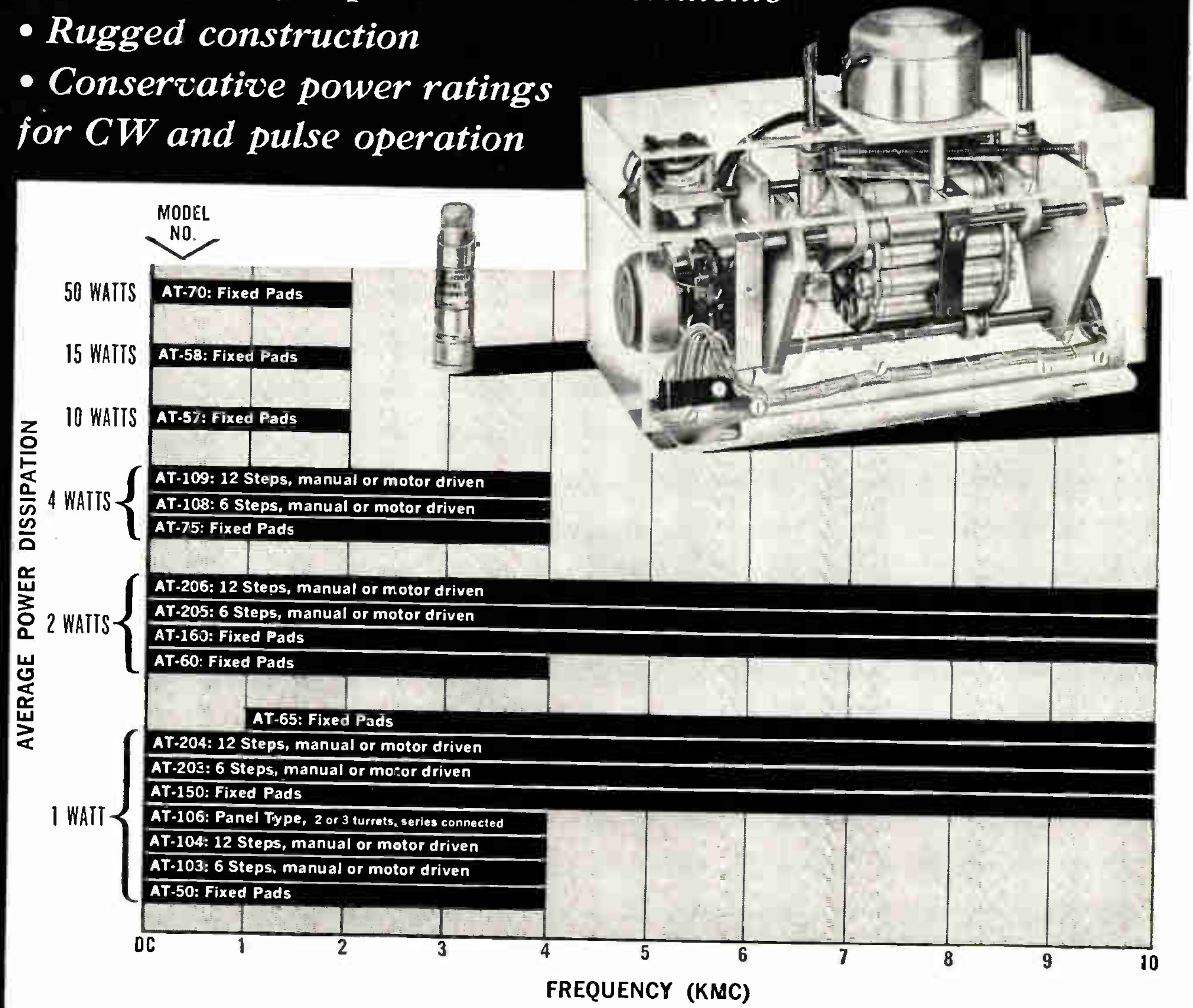
Manufacturers' Cross-Reference Index

(Continued from page 216A)

(Continued on page 224A)

FROM A SINGLE FIXED PAD to a 12-POSITION AUTOMATIC STEP ATTENUATOR...EMPIRE CAN MEET YOUR COAXIAL ATTENUATOR NEEDS...RATINGS FROM 1-50 WATTS...FOR OPERATION FROM DC TO 10 KMC

- *Low VSWR • Excellent accuracy • Optimum reliability*
- *High stability deposited carbon elements*
- *Rugged construction*
- *Conservative power ratings for CW and pulse operation*



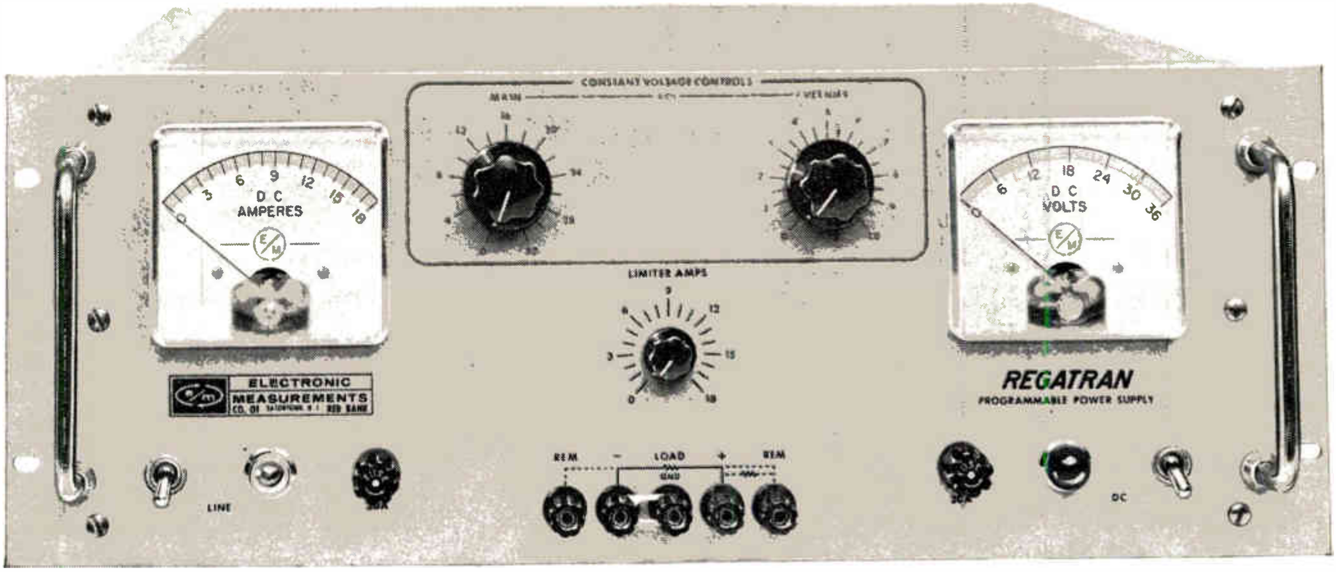
For further technical information, send for Catalog 614 to:

ed

EMPIRE DEVICES, Inc.

AMSTERDAM, NEW YORK • VICTOR 2-8400

MANUFACTURERS OF: FIELD INTENSITY METERS • POWER DENSITY METERS • IMPULSE GENERATORS • COAXIAL ATTENUATORS • CRYSTAL MIXERS



another first from Electronic Measurements

NEW "PV" Series Power Supplies

BRIEF SPECIFICATIONS

BASIC MODEL NO.	DC OUTPUT VOLTS	DC OUTPUT AMPERES	DIMENSIONS IN INCHES PANEL HEIGHT	DIMENSIONS IN INCHES PANEL WIDTH	DIMENSIONS IN INCHES DEPTH BEHIND PANEL
PV32-5	0-32	0-5	3½	19	17¼
PV32-10	0-32	0-10	5¼	19	16½
PV32-15	0-32	0-15	7	19	15⁷⁄₁₆
PV32-30	0-32	0-30	8¾	19	16¼
PV36-5	0-36	0-5	3½	19	17½
PV36-10	0-36	0-10	5¼	19	16½
PV36-15	0-36	0-15	7	19	15⁷⁄₁₆
PV36-30	0-36	0-30	8¾	19	16¼
PV60-2.5	0-60	0-2.5	3½	19	17¼
PV60-5	0-60	0-5	5¼	19	16½
PV60-7.5	0-60	0-7.5	7	19	15⁷⁄₁₆
PV60-15	0-60	0-15	8¾	19	16¼

- 0.01% or 2 millivolts regulation
- All solid-state with SCR input
- Programmable over the entire voltage and current range
- Long-line remote sensing
- Continuously variable current limiting
- Slaved series or parallel operation
- Up to 44% reduction in panel height

request specification sheet 2020

ELECTRONIC MEASUREMENTS
COMPANY OF RED BANK
EATONTOWN • NEW JERSEY

Manufacturers' Cross-Reference Index

(Continued from page 220A)

National Semiconductor Prods.—Hoffman Electronics Corp.
National-Standard Co.—Reynolds Wire Div.
National Tube Co.—United States Steel Corp.
National-U.S. Radiator Corp.—Crane Co.
National Vendors, Inc.—Universal Match Corp.
National Wire Corp.—Hatfield Wire & Cable
Naugatuck Chemicals Div.—U.S. Rubber Co.
Nautec Corp.—Motor Products Corp.; Trionics Corp.; Aermoton, Inc.; Air Motor Div.
Nems-Clarke Co.—Vitro Electronics Div.
Neptune Meter Co.—Revere Corp. of America
New Castle Products—Pyre-Electric, Inc.
New Departure Div.—General Motors Corp.
New England Tape Co.—Graphik-Circuits Div.
New Haven Clock & Watch Co.—Condenser Products Div.
New Hermes Engraving Machine Corp.—Hermes Plastics, Inc.
New London Instrument Co.—Waltham Electronics Corp.
New Process Metals, Inc.—Ronson Metals Corp.
New Rochelle Tool Corp.—New Rochelle Thermatool Corp.
New York Air Brake Co.—Kinney Vacuum Div.
New York Transformer—Essex Electronics
Newton Screw Machine Products Co.—Peck Spring Co.
Niagara Electric Sales—Bradford Components, Inc.
Niagara Frontier Div.—Bell Aerospace Corp.
Nicad Div.—Gould-National Batteries, Inc.
Nichols & Roe Electrical Labs.—Nichols, Ltd. (R.H.)
Nickel Cadmium Battery Corp.—Gould-National Batteries, Inc.
Nickel-Cadmium Battery Div.—Nife, Inc.
Nielsen Tool & Die Co.—Nielsen Hardware Corp.
Niforge Corp.—General Communication Co.
Nikon, Inc.—Nippon Kogoku K.K.
No Strip Co.—Thordarson-Meissner
Noma Lites, Inc.—American Screw Co.
Norbute Corp.—Kurman Electric Co.
Nord Photocopy & Business Equipment Corp.—Television Utilites Corp.
Norden Div.—Hamilton Standard Div.; United Aircraft Corp.
Norden-Ketay Corp.—United Aircraft Corp.
Norfolk Precision Machine Corp.—Precision Microwave Corp.
North American Aviation, Inc.—Autonetics Div.
North American Electronics Div.—International Resistance Co.
North American Geophysical—Mandrel Industries, Inc.
North American Philips Co.—Ferroxcube Corp. of America; Alliance Mfg. Co.; Mepco, Inc.; Consolidated Electronics Industries, Inc.; Amperex Electronic Corp.; Electrical Industries; Haydon Co. (The A.W.); Philips Electronics & Pharmaceutical Industries Corp.; Philips Electronic Instruments; Elmet Div.; Philips Labs.; Price Electric Corp.
North Electric Co.—Ericsson Telephone Co., Ltd. (L.M.)
North Hills Electric Co., Inc.—North Hills Electronics Co., Inc.
North & Judd Mfg. Co.—Con-Torq, Inc.
North Shore Nameplate Div.—Anodyne, Inc.
Northern Radio Mfg. Co., Ltd.—Northern Radio Co., Inc.
Northrop Corp.—Page Communications Engineers, Inc.; Nortronics Div.
Norton Labs., Inc.—Auburn Plastics, Inc.
Nortronics Div.—Northrop Corp.
Nostrand Metal Fabricating Corp.—Dolin Metal Products, Inc.
Nuclear Corp. of America—Central Electronic Mfrs.; Isotopes Specialties Co.; Research Chemicals Div.; Semcor Solid State Div.
Nuclear Development Corp. of America—Ray Proof Corp.
Nuclear Instrument & Chemical Corp.—Nuclear-Chicago Corp.
Nuclear Power & Equipment Corp.—Ultek Corp.
Nuclear Products-Erco Div.—ACF Electronics
Nucledyne Div.—Cook Electric Co.
Nytronics, Inc.—Essex Electronics

O

OK Electronics Corp.—Hermetic Seal Corp.; Thermal Controls, Inc.
Oak Electronics Corp.—Oak Mfg. Co.
Oberline, Inc.—Ultraudio Products Div.
Ohio-Apex Div.—FMC Corp.
Oliver-Shepherd Industries, Inc.—Shepherd Industries, Inc.; S-I Electronics, Inc.
Olsen Mfg. Co. (C.A.)—Westinghouse Electric Corp.
Olympic Development Corp.—Barnes Engineering Co.
Olympic Div.—Siegler Corp.
Olympic Products—Essex Electronics
Olympic Radio & Television—Siegler Corp.
Omaton Div.—Burndy Corp.
Omega Lighting, Inc.—Kent Lighting Corp.
Omnitronics, Inc.—Borg-Warner Corp.
Omohundro Div.—United States Chemical Milling Corp.
Onan Div.—Studebaker-Packard Corp.
Oneida Electronics, Inc.—Canadian Aviation Electronics, Ltd.
Onondaga Electronics Div.—Speer Carbon Co.
Onondaga Pottery Co.—Speer Carbon Co.
Operadio—DuKane Corp.
Optical Film Engineering Co.—New York Air Brake Co.
Original Enderes Co.—Jensen Mfg. Co.
Orr Industries—Ampex Corp.
Orradio Industries—Orr Industries
Ortholog Div.—Gulton Industries, Inc.
Ossining Div.—Hudson Wire Co.
Ostlund, E.M.—Telecontrol Corp.
Otis Radio & Electric—Coilcraft, Inc.
Owens-Corning Fiberglas Corp.—Corning Glass Works
Owens-Illinois Glass Co.—Kimble Glass Co.
Oxford Components Div.—Oxford Electric Corp.
Oxford Electric Corp.—Hudson Lamp Co.
Oxley Developments Co., Ltd.—British Radio Electronics, Ltd.
Ozalid Div.—General Aniline & Film Corp.

(Continued on page 228A)

For the vital telemetry link of the Agena Launch system, Lockheed chose HELIAX. The requirements of this launch complex link, typical of careful missile range instrumentation, illustrate the advantages of HELIAX.

No RF Noise: The unique Andrew connector design captivates inner conductor to outer conductor, eliminating RF noise from sliding connector contacts.

No RF Leakage: The continuous copper outer conductor protects the circuit from other RF signals.

Low Loss, Low VSWR: Insures maximum circuit reliability, minimum signal distortion.

Flexible, Rugged: The corrugated outer conductor of HELIAX makes it the only flexible air dielectric cable. This corrugated outer conductor withstands high crushing forces and heavy vibration.

Easy to Install: Andrew connectors are easily attached in the field, even by inexperienced personnel. HELIAX flexibility makes it much easier to pull into place on the tower and through ducts.

No Corrosion Problem: HELIAX, made of copper, is less affected by corrosive fumes and fluids present in launch areas.

These advantages make HELIAX the preferred coaxial cable for all RF transmission.

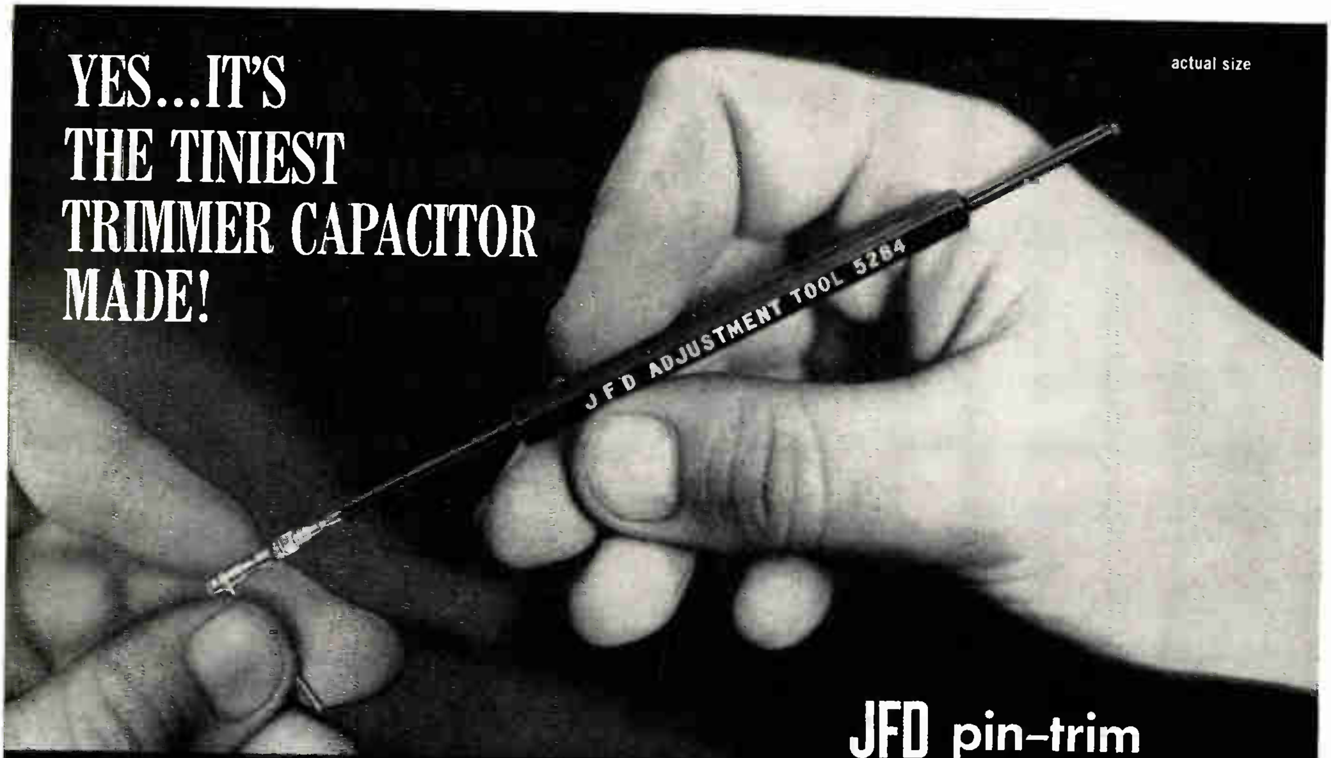
CHICAGO:	LOS ANGELES:	TORONTO:	NEW YORK:	WASH., D. C.:	BOSTON:
P. O. Box 807	941 E. Marylind	606 Beech Street	P. O. Box 416	P. O. Box 5897	P. O. Box 296
Chicago 42, Illinois	Claremont, Calif.	Whitby, Ontario	Ridgewood, N. J.	Wash. 14, D. C.	Westwood, Mass.
Fieldbrook 9-3300	NAtional 6-3505	MOhawk 8-3348	Gilbert 5-2500	OLiver 4-2540	DAvis 6-6500

Andrew

Write for Catalog H,
P. O. Box 807,
Chicago 42, Illinois

ANTENNAS • ANTENNA SYSTEMS • TRANSMISSION LINES

YES...IT'S THE TINIEST TRIMMER CAPACITOR MADE!



JFD pin-trim

If you are designing for missile environments or applications requiring more capacity in less space, look in to the Pin-Trim. It provides a practicable solution to the challenge of end-product miniaturization with high operational stability.

The new Pin-Trim delivers: (1) more capacitance per cubic centimeter than any other conventional piston trimmer; (2) 75 per cent less weight and 50 per cent less volume than JFD's own miniature trimmers; (3) greater sensitivity; (4) finer adjustment.

If you are looking for maximum compactness between stacked circuit boards, or less stray capacitance in a given area, check the JFD Pin-Trim specifications for your subminiature trimmer applications.

For further data, call your local JFD Field office or your JFD franchised Industrial Distributor.

- Overall diameter: 1/8 inch. Overall length above panel: 3/8 inch to 1 inch.
- Double the sensitivity of JFD standard trimmers. Special adjust mechanism provides 102 turns per inch for extra fine adjustment.
- Increased maximum to minimum capacitance ratio per unit (minimum: 0.5 pf.).
- Operating temperature −55° to +125°C.
- Low temperature coefficient of capacitance.
- Anti-backlash design for precise tuning resolution.
- Low inductance for high frequency use.
- Ultra linear tuning assures accurate alignment—absolute repeatability. Standard slotted end for screwdriver adjustment.
- Rugged shock and vibration resistance.
- 500 V. DC working voltage.
- 10^6 megohms insulation resistance.
- Q factor of 500 (measured as per JFD #5178).
- 0.5 inch ounce tuning torque.
- Meet or exceed applicable performance requirements of MIL-C-14409A.

JFD Adjustment Tool No. 5284 (illustrated) available at 85¢.

Model*	Capacitance Range MMF Measured Per JFD #5177		D.C. Working Volts	Dielectric Strength Measured For 5 Seconds at 50% R.H. at Max. Rated Cap.	Insulation Resistance Measured After One Minute at 500V D.C. and 50% R.H.	Q Factor Measured Per JFD #5178	Unit Weight Grams	Dimen.** Max. ±1/32
	Min.	Max.						
PT901	0.5	2.0	500	1000	10^6 Megohms	500	0.62	3/8"
PT902	0.5	3.0	500	1000	10^6 Megohms	500	0.64	1/2"
PT903	0.5	5.0	500	1000	10^6 Megohms	500	0.79	3/4"
PT904	0.5	7.0	500	1000	10^6 Megohms	500	0.94	1"

* These units are also available in the same capacitance values for printed circuit boards in models PT911, PT912, PT913 and PT914.
** Length front of panel.

U.S. Patent No: 2,922,093 Canadian Patent No: 604,810

JFD

JFD ELECTRONICS CORPORATION

Components Division • 6101 16th Avenue, Brooklyn, New York • Phone DEwey 1-1000 • TWX-NY25040

JFD WESTERN
P. O. Box 3416
Van Nuys, Calif.
Phone: EMpire 4-4131

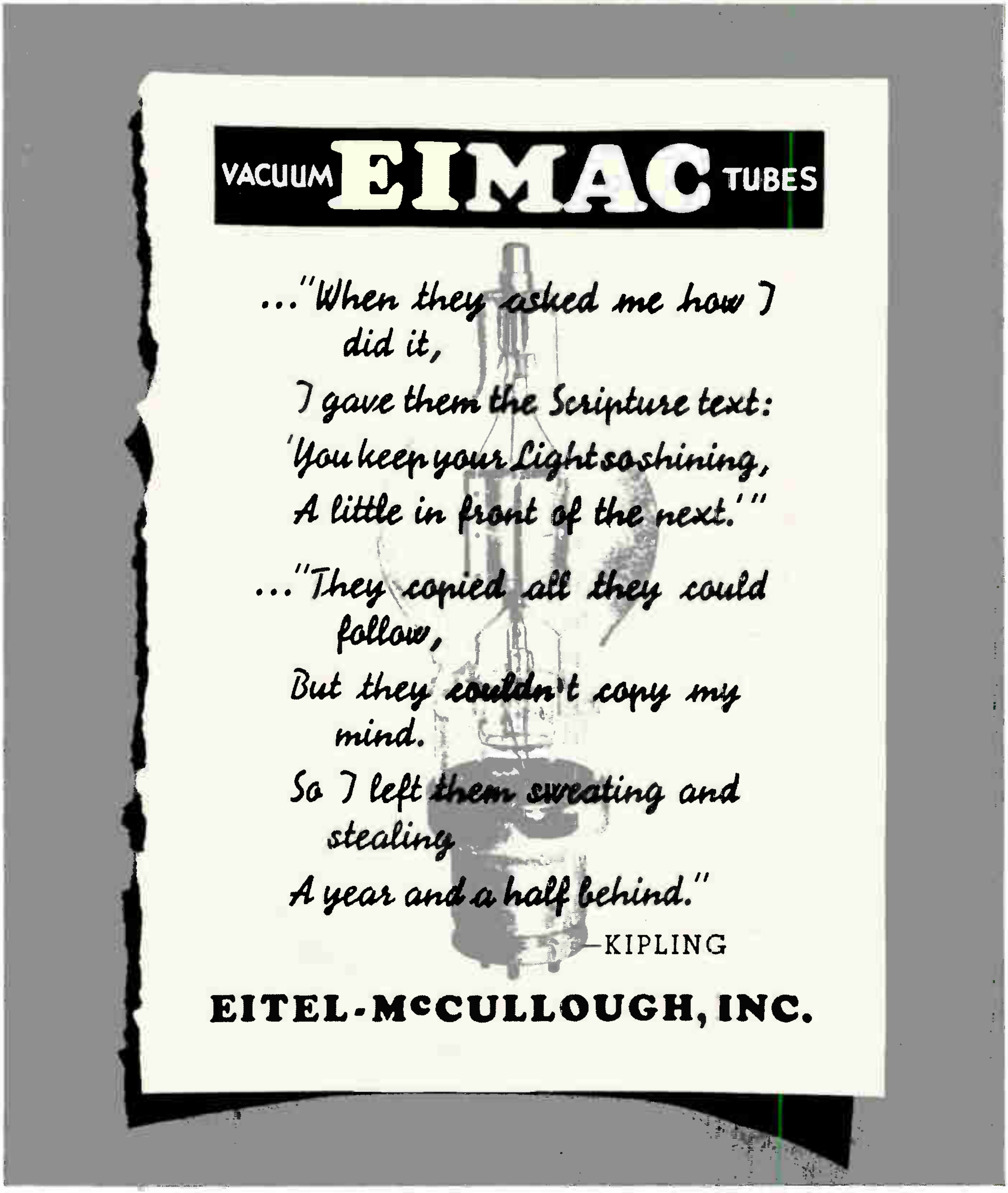
JFD MIDWESTERN
P. O. Box 588
Skokie, Illinois
Phone: 675-1140

JFD NORTHEASTERN
Ruth Drive, P. O. Box 228
Marlboro, Mass.
Phone: HUntley 5-7311

JFD CANADA
51 McCormack Street
Toronto, Ontario, Canada
Phone: ROger 9-1129

VARIABLE TRIMMER PISTON CAPACITORS • FIXED METALIZED INDUCTORS • LC TUNERS • DIPLEXERS
FIXED AND VARIABLE DISTRIBUTED AND LUMPED CONSTANT DELAY LINES • PULSE FORMING NETWORKS

VISIT JFD INDUSTRIAL CONFERENCE BOOTH T-118 AT THE 1962 ELECTRONIC PARTS SHOW



25 years ago we ran this ad.

Technology has changed since then. But not Eimac's position as industry leader.

Today, as 25 years ago, Eimac leads the way in electron power tube technology. With its research, engineering and manufacturing capabilities, Eimac can always be counted on to meet tomorrow's tube needs today. Another reason to keep your eye on Eimac for advanced power grid tubes, high power klystrons, microwave devices, accessory products. Eitel-McCullough, Inc., San Carlos, California. Subsidiaries: Eitel-McCullough, S.A., Geneva, Switzerland; National Electronics, Geneva, Illinois.

KEEP YOUR EYE ON Eimac

Manufacturers' Cross-Reference Index

(Continued from page 221A)

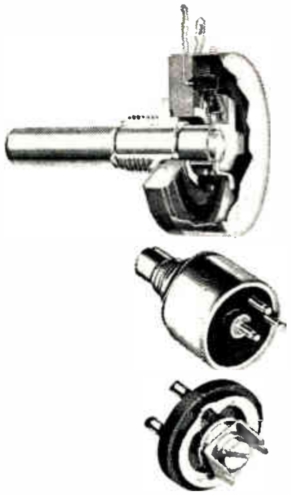
P

(Continued on page 232A)

Centralab: a RELIABLE source for RELIABLE PRODUCTS

FOR MILITARY & COMMERCIAL APPLICATIONS

Most CENTRALAB products are available for immediate delivery in industrial quantities at factory prices, from CENTRALAB stocking distributors.



HOT MOLDED COMPOSITION VARIABLE RESISTORS

Meet MIL-R-94 environmental and test requirements. Provide smooth noise-free operation and high stability for which hot molded units are well known.

2 WATT @ 500V—Raised resistance track reduces surface contamination. Wide clearance between bushing and collector track for increased high voltage stability. Carbon collector and pick-off brushes.

1-3/32" diam., 5/8" deep. 500 ohms—5 meg, linear taper; 10K-2.5 meg log taper.

3/4 WATT @ 350V—Metal enclosed units, available with triple "O" ring shaft seal. 23/32" diam., 1/2" deep. Same resistances as 2 watt unit.

1/2 WATT @ 350V—Similar construction as 3/4 watt. 1/2" diam., 15/32" deep. 100 ohms—5 meg, lin. taper.

1/3 WATT at 350V TRIMMER—locking type variable fixed resistor. Extremely resistant to shock vibration, acceleration, can be encapsulated. 19/32" diam., 11/32" deep. 50 ohms—5 meg, linear taper.

Write for Group G bulletins.

VARIABLE RESISTORS

1/4W. Linear Motion, wirewound 100 to 20K ohms and composition 10K to 2.5 meg. .250" x .325" x 1.250".

1/10W. Microminiature, 0.286" diam: **Ultraminiature**, 0.502" diam; **Subminiature**, 5/8" diam; 500 ohms to 10 megs. **1/4W. and 1/2W. Subminiature**—For instruments and military. 43/64" diam. 500 ohms to 2.5 megs.

1/4W. Multiple Miniature—Up to 4 variable and 9 fixed resistors on a 3/4" x 2-1/4" steatite plate. 1000 ohms to 5 megs.

1/2W. Standard—For radio, TV. Single, Twin or dual-concentric. 15/16" diam. 250 ohms to 10 megs.

1-1/2W. Wirewound—for military and instruments. 11/16" diam., 5/16" deep. 4 ohms to 30K ohms.

2-5W. Wirewound—For instruments and TV. Single or dual-concentric. 1-5/32" diam. 1 ohm thru 100K ohms. Also available as 20 watt audio L and Bridged T Pad attenuators.

Write for Group B bulletins.

ELECTRONIC SWITCHES

Subminiature rotary—15/16" diam. for military and high reliability applications. Rating 0.5 amp. at 6VDC, 100 ma. at 110VAC. Laminated phenolic, steatite, single or multiple sections.

Miniature rotary—1-1/4" diam. for military and high reliability applications. Rated 2.0 amps. at 15VDC and 250 ma. at 110VAC. Available single or multiple ceramic, Mycalex and phenolic sections.

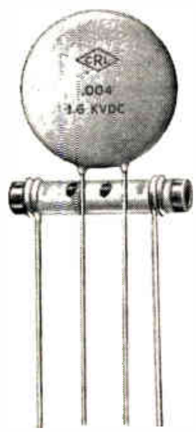
Standard rotary—1-5/16" diam. laminated phenolic or steatite insulation, single or multiple sections. Rating 2 amps. at 15-VDC. 150 ma. at 110VAC. (Resistive load.)

Heavy-duty power—For transmitter, industrial control, laboratory testing, military, commercial. 2-13/16" diam. 7-1/2 amp. rating at 110VAC. 25,000 cycles minimum.

Spring return—1-5/16" diam., coil or C-type springs, phenolic, steatite, single or multiple sections.

Specialized—Lever, slide, tone, tuner sections, others.

Write for Group P bulletins.



CERAMIC CAPACITORS

Ultraminiature—3 to 10VDC, 0.22—2.2 mfd.—for low power factor transistor applications.

Temperature compensating—Discs, tubulars, 150 v to 6,000 VDCW. 1 mmf.—0.1 mf. Capacitance +100 to —5250 ppm.

Bypass—Coupling—Discs, tubulars, 150 v to 6,000 VDCW, 1 mmf.—0.1 mf.

High Voltage—High Accuracy—High Voltage types, up to 30,000 VDCW, High Accuracy types, ± 1% tolerance, 500 VDCW, up to 2,500 mmf.

Trimmer—Tubular or flat. Meet MIL-C-81A specifications. 1 mmf. to 400 mmf. 500 VDCW.

Feed-thru—10—5000 mmf., 500—1,500 VDCW, bushing, shoulder, ring, eyelet, resistor-capacitor combinations.

Specialties—Stand-off, button-shape, potted, other capacitors.

Write for Group D bulletins.

PACKAGED CIRCUITS

Complete computer and Radio-TV circuits, amplifiers, oscillators, detector networks, resistor networks—including transistors, capacitors, resistors, wiring and inductance, manufactured to your specific performance limits.

PEC* circuits result in substantial savings in assembly costs. These high reliability packaged circuits can be supplied in a wide variety of terminals for printed circuit board applications.

1/2W. resistors meet applicable MIL-R-11 specifications, 50 ohms to 50 megs. Capacitors up to .01 mmf.

Write for Group Y bulletins.

*Trade Mark

ENGINEERED CERAMICS

High alumina—85%, 95%, 99%—for high frequency, high temperature applications close tolerance (±.00025") designs. L6A Jan-I-10 grade.

Grade Jan-I-10-L5A steatite, Grade L2A Corderite, and Grade L2A Electrical Porcelain. For applications where high dielectric and compressive strength, high dimensional stability, low loss and low power factor are required, there is a CENTRALAB ceramic material for the job.

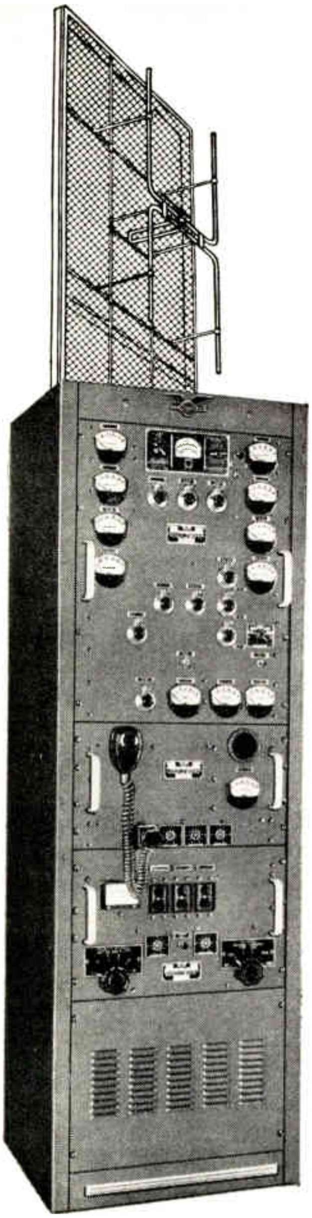
CENTRALAB also specializes in metalizing of ceramics, for hermetic seals or mechanical attachment of other ceramic or metal parts.

Write for Group X bulletins.

G-6140 **Centralab**

THE ELECTRONICS DIVISION OF GLOBE-UNION INC.
920E EAST KEEFE AVENUE • MILWAUKEE 1, WISCONSIN
In Canada: Centralab Canada Ltd., P. O. Box 400, Ajax, Ontario

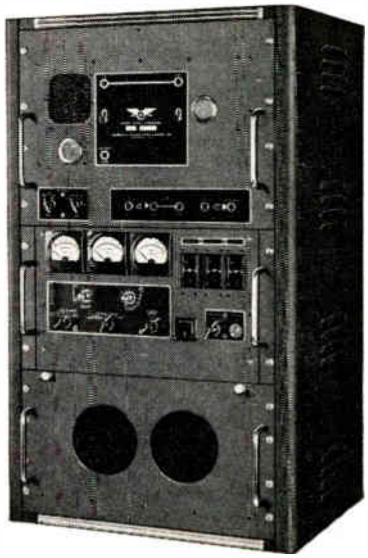
ELECTRONIC SWITCHES • VARIABLE RESISTORS • CERAMIC CAPACITORS • PACKAGED ELECTRONIC CIRCUITS • ENGINEERED CERAMICS



AEROCOM PRESENTS VHF AM TRANSMITTERS and RECEIVERS

AEROCOM communications equipment is designed with both performance and reliability in mind, and is produced by experienced personnel using high-quality materials. The following features are found in all three transmitters: Single crystal controlled frequency (plus an additional frequency ½% away from main frequency): stability ± .003% or ± .001% over temperature range of 0° C to + 55° C, any humidity up to 95%; audio system incorporates high level plate modulation, with compression; forced ventilation with air filter is employed. Welded steel cabinets.

Model 10V1-A—1000 Watts output—Successfully being used in Troposcat service for communications with aircraft beyond the optical horizon. Frequency range 118-153 mc. Can be completely remote controlled by using AEROCOM's remote control equipment. All tuning from front panel by means of dials. Power requirements 210-250 V 50/60 cycles, single phase.



Model VH-200—200 Watts output in range 118-132 mc. Excellent for both point-to-point and ground-to-air communications. Press-to-talk and audio input may be remoted using single pair of telephone lines. Power requirements 105-120V 50/60 cycles. Also available for use above 132 mc; output drops gradually to 150 watts at 165 mc.

Model VH-50—50 Watts output. Frequency range 118-153 mc. Outstanding low power transmitter for ground-to-air service. With remote control provisions; main power control with front panel switch. Convection cooling for press-to-talk service—otherwise forced air cooling. Power requirements 115/230 V 50/60 cycles.

Model 85 VHF Receiver. A high performance, low noise, single channel crystal controlled, single conversion VHF receiver. Stability normally ± .001% (with oven crystal ± .0005%) over temperature range 0° C to + 55° C. Sensitivity ½ microvolt or better for 1 watt output with 6 db signal to noise ratio. Standard selectivity bandwidth 30 kc; other widths available. Spurious response down 90 db. Frequency range 118-154 mc. Power requirements either 115 V or 230 V 50/60 cycles. Made for standard rack panel mounting.

As in all AEROCOM products, the quality and workmanship of this VHF equipment is of the highest. All components are conservatively rated. Replacements parts are always available for all AEROCOM equipment.

Complete technical data available on request

AER-O-COM

FCC Type Accepted for Aviation Service

3090 S. W. 37th Avenue—Miami 33, Florida

MICRO-RADIONICS, INC.

MRI

VAN NUYS, CALIFORNIA

A NEW NAME

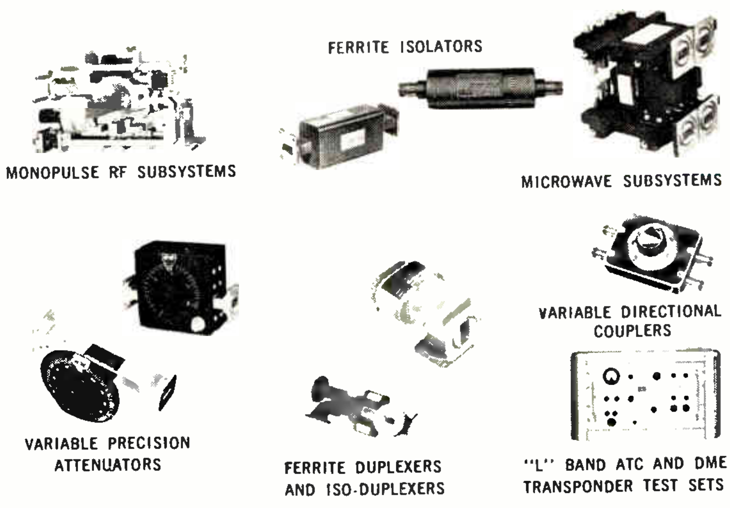
for recognized quality

MICROWAVE EQUIPMENT

FORMERLY MICROWAVE Kearfott DIVISION

KEARFOTT MICROWAVE DIVISION BECOMES

MICRO-RADIONICS, INC.



Acquired by the aggressive, young management team who formerly guided Kearfott Microwave Division, M-R-I will extend its line of high quality products, long recognized as the finest in the microwave field. Now integrated into one complete facility, M-R-I can offer a quick reaction capability to fit any development, engineering, or production requirement.

M-R-I engineering experience, acquired over the many years under the Kearfott banner, is being supplemented by new state-of-the-art advancements, a value analysis team, and improved production methods. Representatives located in major cities, coast-to-coast, are ready to assist with your microwave design needs.

Creative engineers are invited to investigate the rewarding opportunities at M-R-I.

MRI

MICRO-RADIONICS, INC.

Formerly Kearfott Microwave Division

14844 OXNARD STREET, VAN NUYS, CALIFORNIA

STate 6-1760 TWX: VNYS 5451

M-R-I IS AN INDEPENDENT ORGANIZATION, NOT AFFILIATED IN ANY WAY WITH KEARFOTT DIVISION OR ANY PART OF GENERAL PRECISION, INC.

Manufacturers' Cross-Reference Index

(Continued from page 228A)

Phillips Doup Co.—Ray, Daisley Co., Inc.
Phillips-Eckardt Electronic Corp.—Phillips Control Co.
Phoenix Engineering & Mfg. Co.—Telecomputing Corp.
Phonolo Co., Ltd.—Dominion Electrohome Industries, Ltd.
Photographic Products, Inc.—Beattie-Coleman, Inc.
Photomechanisms, Inc.—MacLaren & Co., Inc. (F.B.)
Photoswitch Div.—Electronics Corp. of America
Photstat Corp.—Itek Corp.
Physical Sciences Corp.—Packard-Bell Electronics
Pi-Square Engineering Co.—Philbrick Researches, Inc. (George A.)
Pic Automation Controls Div.—General Controls Co.
Pickard & Burns, Inc.—Gorham Electronics Div.
Pioneer-Central Div.—Bendix Corp.
Pioneer Scientific Corp.—Bausch & Lomb, Inc.
Pittsburgh Corning Corp.—Corning Glass Works
Pittsburgh Lectrodryer Div.—McGraw-Edison Co.
Plastic Process Co.—United-Carr Fastener Corp.
Platronics Div.—Palumbo Bros., Inc.
Pluta Mfg. Corp.—General Circuits, Inc.
Point Mugu Div.—Land-Air, Inc.
Polarad Electronics Corp.—Sterling Transformer Corp.; Telewave Labs., Inc.
Pollock Mfg. Co., Ltd.—Dominion Electrohome Industries, Ltd.
Polyken Sales Div.—Kendall Co.
Polymer Corp.—Polymer Corp. of Pennsylvania
Polymer Processes, Inc.—Polymer Corp. of Pennsylvania
Polypenco, Inc.—Polymer Corp. of Pennsylvania
Polytechnic Research & Development Co., Inc.—PRD Electronics, Inc.
Pomona Div.—Marquardt Corp.
Pop Rivet Div.—United Shoe Machinery Corp.
Port Electric Corp.—Rollan Electric Co.
Portchester Instrument Corp.—Premier Instrument Corp.
Porter Metal Products—Kent Lighting Corp.
Posen & Kline Tube Co., Inc.—Handy & Harman
Potter Instrument Co.—Electronic Counters, Inc.
Poughkeepsie Coil Co., Inc.—Endicott Coil Co., Inc.
Power Controls Div.—Ovitron Corp.
Power Magnetics, Inc.—Varo Inc.
Power Sources, Inc.—Microwave Assoc., Inc.
Powers Chemco, Inc.—Photocircuits Corp.
Pratt & Whitney Co.—United Aircraft Corp.
Prazision-Edelstein GmbH—Fidelitone, Inc.
Precise Die & Stamping Co.—Cinaudagraph, Inc.
Precision Apparatus Co., Inc.—Pacotronics, Inc.
Precision Film Labs., Inc.—Maurer, Inc. (J.A.)
Precision Metalcraft, Inc.—Precision, Inc.
Precision Paper Tube Co.—Resinite Corp.
Precision Products Div.—Chicago Dynamic Industries, Inc.
Premier Crystal Labs., Inc.—Premier Research Labs., Inc.
Premier Instrument Corp.—Portchester Instrument Corp.; Premier Microwave Corp.
Prentiss Wire Mills—Porter Co., Inc. (H.K.)
Pres Industrial Products, Ltd.—Universal Instruments Corp.
Pres-To-Line Corp.—Tamar Electronics, Inc.
Prescott Television Co.—Vue-Tronics, Inc.
Pressure Elements, Inc.—Colvin Labs., Inc.
Prest-O-Lite Batteries Div.—Electric Autolite Co.
Presto Recording Corp.—Bogen-Presto Div.
Prewitt Aircraft—Atlantic Research Corp.
Price Electric Co.—Haydon Co. (The A.W.); North American Philips; Mepco, Inc.
Princeton Div.—Curtiss-Wright Corp.
Printed Electronics Corp.—Electralab Printed Electronics Corp.
Production Research Corp.—Radio Condenser Co.
Profile, Inc.—Coil Co. of America, Inc.
Progress Mfg. Co.—Chester-Morton Electronics Corp.
Progressive Mfg. Co.—Torrington Co.
Prudential Industries, Inc.—Atlas Precision Products Co.
Pure Carbon Co.—Stackpole Carbon Co.
Pure Carbonic Co.—Air Reduction Co., Inc.
Purnell Electronics—Coils-Electronics Co.
Pye, Ltd.—Cathodeon, Ltd.
Pye Telecommunications, Ltd.—Pye Corp. of America
Pyramid Electric Co.—General Instrument Corp. (Capacitor Div.)
Pyroferric Co., Inc.—Q.O.S. Corp.
Pyrometer Co. of America, Inc.—Magnetic Instruments Co., Inc.

Q

Quality Components, Inc.—Vitroseal Corp.

R

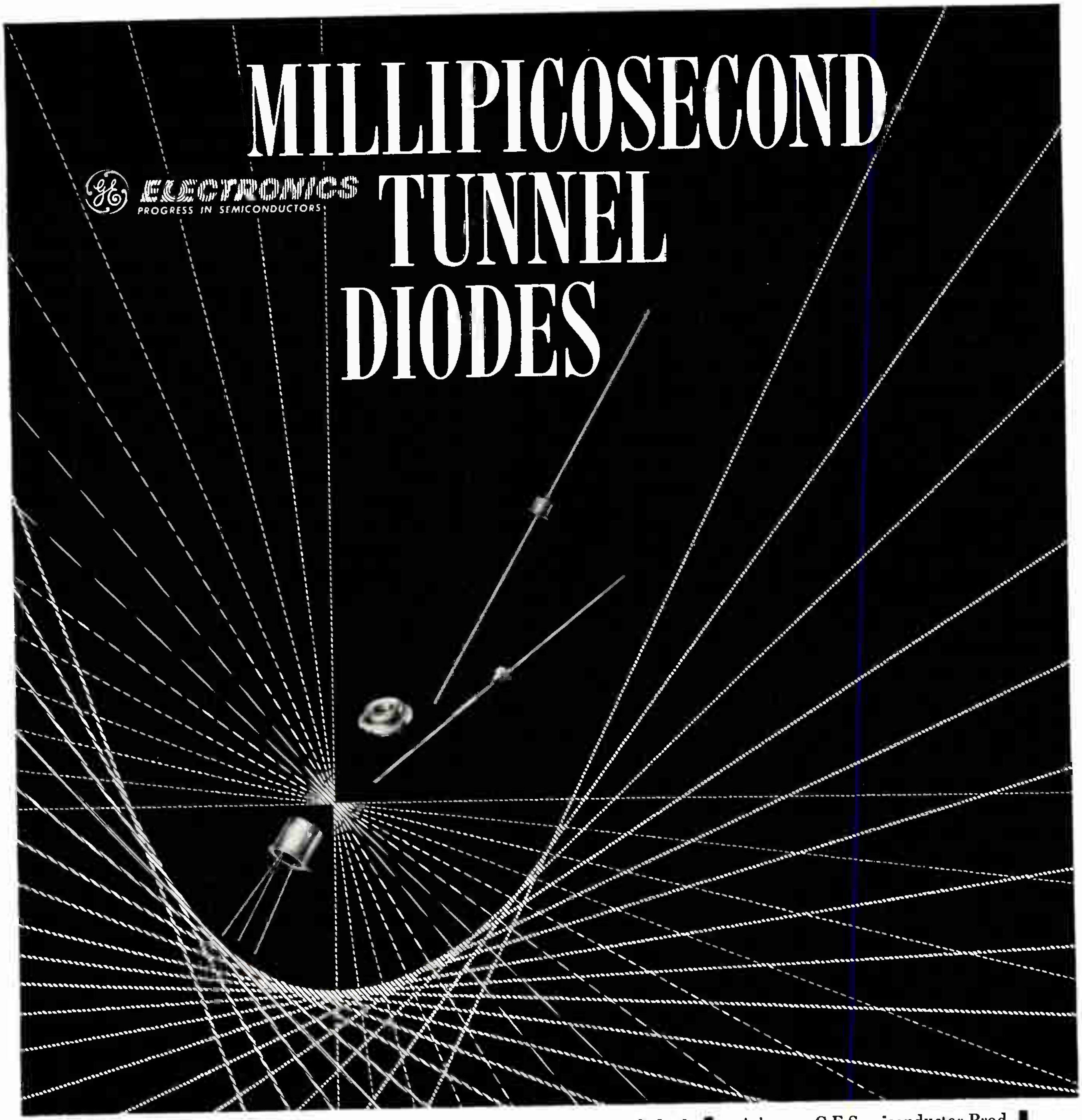
RBM Controls Div.—Essex Wire Corp.
R-F Electronics, Inc.—Electro Contacts, Inc.
RFI Shielded Enclosures—Ace Engineering & Machine Co.
R-M-I Div.—Ramage & Miller, Inc.
Rad Mfg. Corp.—Douglas Microwave Co., Inc.
Radclin, Ltd.—United States Radium Corp.
Radcor, Inc.—Wiegand Co. (Edwin L.)
Radiart Corp.—Cornell-Dubilier Electronics Div.
Radiating Systems Div.—Electronic Specialty Co.
Radiation Electronics Co.—Comptometer Corp.
Radiation, Inc.—Levinthal Electronic Products, Inc.
Radio City Products Co., Inc.—General Metal Products Co.
Radio Coil Winding Specialties—Smallwood, Ltd. (S.G.)
Radio Condenser Co.—Production Research Corp.; Thompson Ramo Wooldridge Inc.
Radio Engineering Labs., Inc.—Dynamics Corp. of America; Standard Electronics Div.
Radio Industries, Inc.—Thompson Ramo Wooldridge Inc.
Radio Mfg. Engineers, Inc.—Electro-Voice, Inc.
Radio Materials Div.—Mallory & Co. (P.R.)
Radio Music Corp.—Cornwell Electronics Corp.
Radio Receptor Co.—General Instrument Corp.
Radio Specialties Div.—Lapp Insulator Co., Inc.
Radionic Labs., Inc.—Mobile Electronics, Inc.
Ramo-Wooldridge Corp.—Thompson Ramo Wooldridge Inc.
Ramsey Corp.—Thompson Ramo Wooldridge Inc.
Ranco, Inc.—Wilcolator Co.
Rank-Xerox, Ltd.—Xerox, Inc.
Rapid Specialties Div.—Lapp Insulator Co.
Rattray Co. (George)—Instruments For Industry, Inc.
Ravenna Metal Products—Standard Screw Co.
Ray-O-Vac Div.—Electric Storage Battery Co.
Raybun Foundry—Gorham Electronics Div.
Rayclad Tubes, Inc.—Raychem Corp.
Raymond Mfg. Div.—Associated Spring Corp.
Raytheon Co.—Applied Electronics Co.; Machlett Labs., Inc.; Sorensen & Co., Inc.; Webster Mfg. Div.
Raytherm Corp.—Raychem Corp.
Rea Magnet Wire—Aluminum Co. of America
Rebat Batteries Div.—Electric Autolite Co.
Red Bank Div.—Bendix Corp.
Redmond Co.—Controls Co. of America
Reed Instrument Co.—Reed Research, Inc.
Reed Roller Bit Co.—Cleco Air Tools Div.
Reeves Equipment Corp.—Tandberg of America, Inc.
Reeves-Hoffmann Div.—Dynamics Corp. of America
Reeves Instrument Corp.—Dynamics Corp. of America
Refractories Div.—Carborundum Co.
Regan Industries, Inc.—R S Electronics Corp.
Regulators, Inc.—Electric Specialty Co.
Reiner Electronics Co.—Radio City Products Co., Inc.; General Metal Products Co.
Reinhold Engineering & Plastics Co.—Haveg Industries, Inc.
Rek-O-Kut Co., Inc.—Audax, Inc.
Relay Exchanges, Ltd.—Goodman's Industries, Ltd.
Relcoil Products Corp.—Hi-G, Inc.

(Continued on page 234A)

Manufacturers' Cross-Reference Index

(Continued from page 232A)

(Continued on page 238A)



General Electric is currently supplying to customers doing advanced circuit work special tunnel diodes with switching speeds expressed in millipicoseconds (thousandths of a trillionth part of a second....) faster than can be measured by present day instruments. The best of these devices switches in the time it takes light to travel 6/1000 of an inch! We are also supplying the TD-100 series "off the shelf," with nominal switching speeds of 15 to 50 picoseconds. TD-101, 102, 103 and 104 have peak currents of 100, 50, 22 and 10 ma, with maximum capacitances of 6, 5, 4, and 2 pf respectively.

Recently introduced tunnel diode matched pairs (TDP-1 through 5) guarantee peak currents to be DC and AC matched within ± 1% over the entire temperature range from 15 to 55°C, making possible high performance logic circuits, memory circuits, and level detectors. Also available: low level switching tunnel diodes in both axial or TO-18 package, with peak forward current ratings from 1 to 22 ma; high speed microminiature switches from 1 to 10 ma; miniature microwave communication tunnel diodes from 1 to 10 ma; and computer back diodes with forward current ratings from 5 to 30 ma. Prices now range as low as $1.35 in 1000-up quantities.

Ask your G-E Semiconductor Products District Sales Manager about these and other production tunnel diodes *now available* to meet your particular requirements in applications such as switching circuits, oscillators, converters, and data and signal processing. For your free copy of "Applications for Tunnel Diodes" by John Phelps, G-E Manager-Application Engineering, write to Semiconductor Products Department, Section 23E126, General Electric Company, Electronics Park, Syracuse, New York. In Canada: Canadian General Electric, 189 Dufferin Street, Toronto, Ont. Export: International General Electric, 159 Madison Ave., N.Y. 16, N.Y.

GENERAL ELECTRIC

This is TELONIC Versatility...

A Sweep/Signal Generator for Audio to 3000 MC

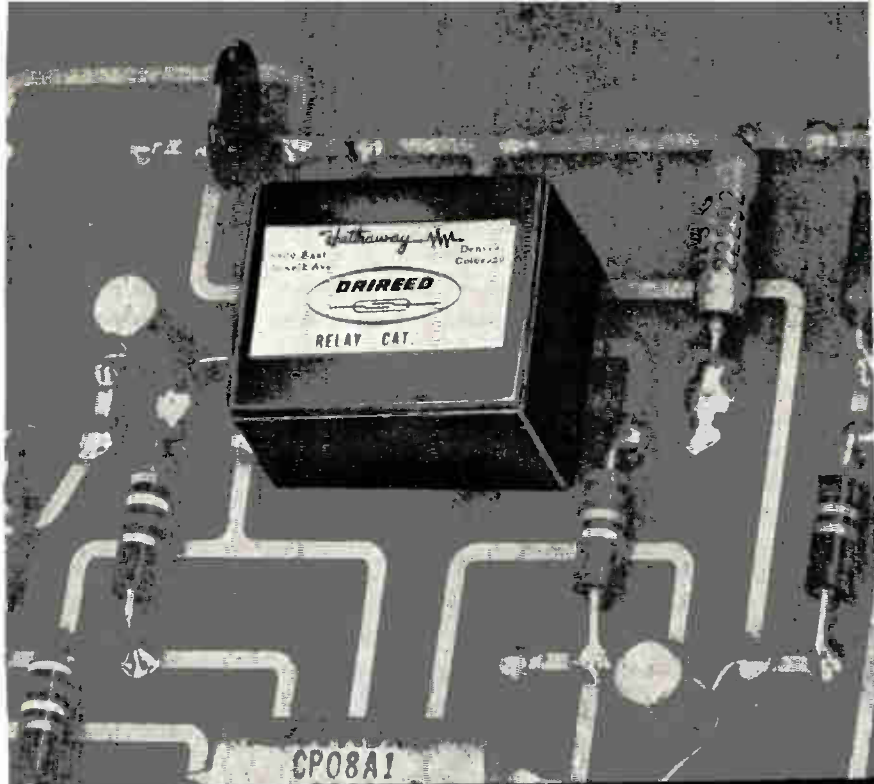
As a major designer and manufacturer of RF instruments and components, Telonic once again leads the field with the introduction of the SM-2000 Sweep and CW Signal Generator. New from every standpoint, the SM-2000 provides unmatched versatility for laboratory or production operations. Now, with one instrument and several, interchangeable plug-in oscillators, an engineer can cover a frequency range from audio to 3000 mc.

Telonic has designed 19 different oscillator heads for specific and general purposes that enable the user to change range of the SM-2000 in a matter of seconds. For general applications, only two plug-in units are necessary to cover frequencies from .5 to 2000 mc. And, in addition, the operator may select four different functional modes with the SM-2000—swept RF, modulated swept RF, CW, and modulated CW. He can set attenuation from 0 to 60 db in 1 db steps with the two built-in attenuators. He also has provisions in the instrument for use of an external marker, or for adding up to eight fixed, plug-in markers if desired.

All these features are combined with the fine basic performance that has made the name Telonic synonymous with the best in RF instrumentation—low VSWR, high display linearity and excellent workmanship. If you would like more complete details on this new sweep generator please write for Technical Bulletin T-233.

Telonic Industries, Inc.

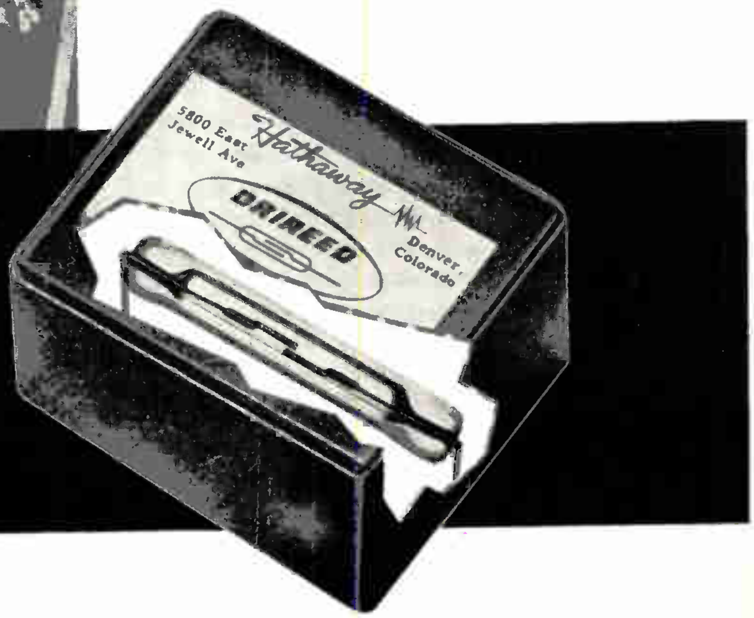
BEECH GROVE, INDIANA—PHONE STATE 7-7241



Hathaway

HATHAWAY DRIREED RELAYS SAVE UP TO 75% DESIGN SPACE

...and give the reliability
of solid state units
at a fraction of their cost



Heart of Hathaway Relays is the Drireed—industry's most nearly perfect switch.

- **Vibration:** 20 G's to 2,000 cps, energized or non-energized
- **Shock:** 50 G's
- **Operate time typical:** 800 microseconds, including bounce, for 28 vdc unit at nominal voltage
- **Closed contact resistance:** .05 ohms nominal
- **Natural contact frequency:** 2,700 to 3,200 cps
- **Contact rating:** 10,000,000 operations at 100 milliamperes, 115 V 60 cps resistive. 50,000,000 operations at 100 milliamperes, 28 vdc resistive. 1,000,000,000 operations at 2 milliamperes, 28 vdc

Miniaturization with solid state reliability and the advantages of low cost and drift free operation is accomplished in the Hathaway Drireed Relay. The Form A type pictured above is designed for either printed circuit or on the chassis mounting. An ideal application is scanning very low level signals at a rapid rate without the use of rotating parts or brushes.

Hathaway relays are ideally applicable to satellite components, ground support equipment and anti-submarine detection equipment. The same operational superiority is available to all applications requiring economy of weight and space, long life, fast, trouble-free action, resistance to shock and vibration, and immunity to the effects of acceleration.

Relays engineered and manufactured to your requirements are a Hathaway specialty... we will furnish a modular assembly of any number of contacts in a completed form.

What can these relays do for your design? For information, contact your Hathaway representative, call or write us at the address below.

Hathaway INSTRUMENTS, INC.
A SUBSIDIARY OF THE LIONEL CORPORATION
5800 EAST JEWELL AVENUE, DENVER 22, COLORADO

SKyline 6-8301
TWX DN 656

FOREIGN REPRESENTATION BY TERMINAL RADIO INTERNATIONAL, LTD., 3 WEST 61ST ST., NEW YORK 23, N. Y.

Manufacturers' Cross-Reference Index

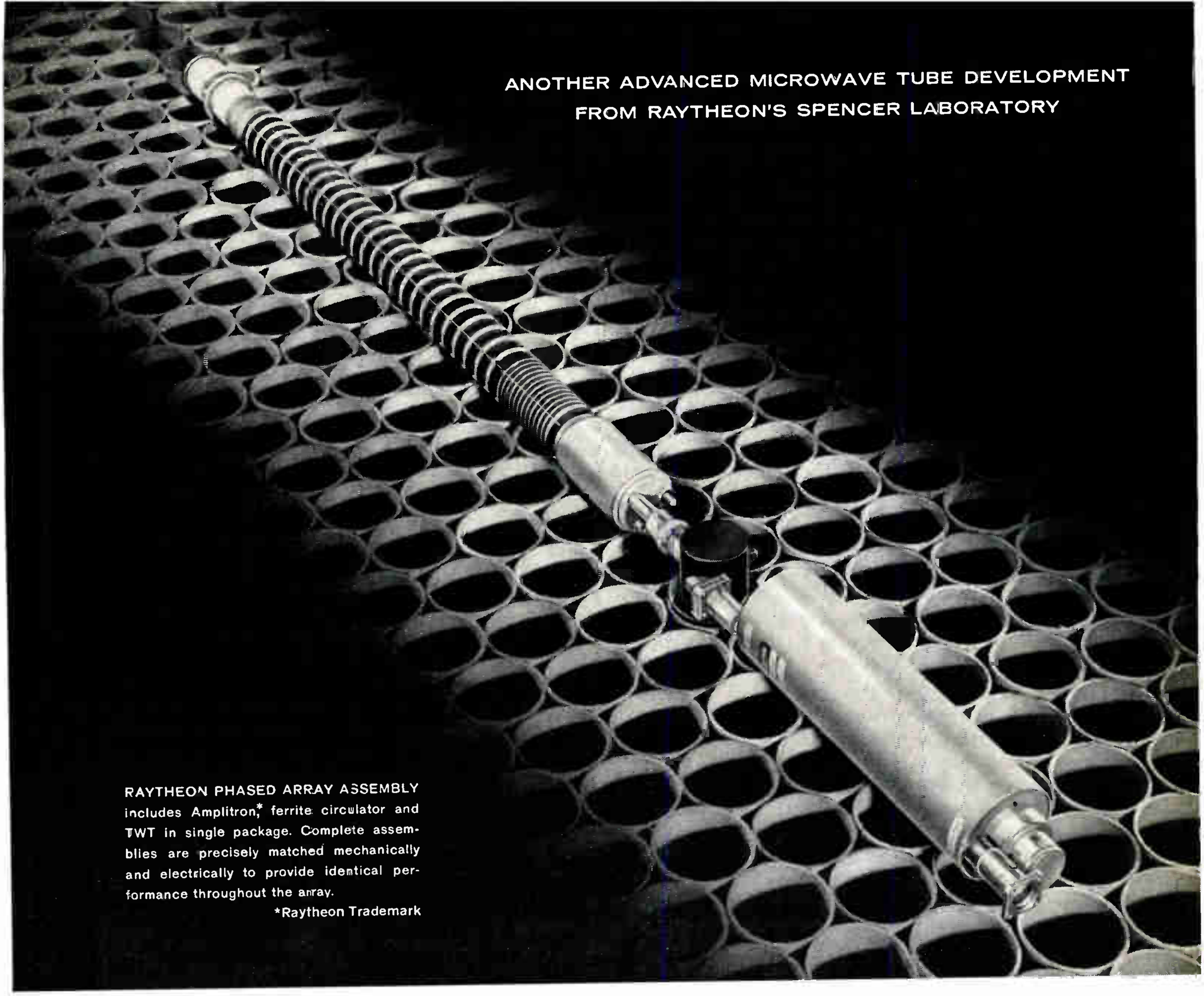
(Continued from page 234A)

Robertson Co. (H.H.)—International Research & Development Corp.
Robertson Technical Products, Inc.—Kupfrian Mfg. Div.; Robinson Vibrashock Div.
Robinson Aerial Surveys Div.—Aeroflex Corp.
Robinson Aviation, Inc.—Robinson Technical Products, Inc.
Robinson Machine Works, Inc.—Nugent Electronics Co., Inc.
Robinson Plastics, Inc.—Robinson Technical Products, Inc.
Robinson Technical Products, Inc.—High Vacuum Equipment Corp.
Rockbestos Products Corp.—Rockbestos Wire & Cable Co.
Rockwell-Standard Corp.—Air-Maze Div.
Rockwood Sprinkler Co.—Gamewell Co.
Rogers Electronic Tubes & Components Div.—Philips Electronics Industries, Ltd.
Rogers Industries, Inc.—Rogers Electronic Corp.
Rohnco, Inc.—Rohn Mfg. Co.
Rola Div.—Muter Co.
Roller-Smith Div.—Federal Pacific Electric Corp.
Rollins Corp.—Inductor Engineering, Inc.
Rome Cable Corp.—Aluminum Co. of America
Rome Mfg. Co.—Revere Copper & Brass, Inc.
Romec Div.—Lear, Inc.
Ronar, Inc.—Hamlin, Inc.
Ronson Corp.—Ronson Metals Corp.
Rosenthal-Isolatoren—United Mineral & Chemical Corp.
Ross Mfg. Co.—Advance Ross Electronics Corp.
Ross Tool Co.—Columbus Electric Mfg. Co.
Rotal Electric Co.—Electric Cords & Supply Corp.
Roth Rubber Co.—Vap-Air Div.
Roy Co. (Milton)—Serfass Corp.
Royal Electric Corp.—International Tel. & Tel. Corp.
Royal Industries, Inc.—Calibration Standards Corp.; Ideal-Aerosmith Div.
Royal McBee Corp.—Instrument Development Labs., Inc.
Rubel Corp.—International Radiant Corp.
Rubicon Instruments Div.—Minneapolis-Honeywell Regulator Co.
Rue Products—Electronet Div.
Russell Gasket Co.—Cleveland Fabricating Co.
Ryan Aeronautical Co.—Aerolab Development Co.; Ryan Electronics; Transdata, Inc.
Rypinski Co. (C.A.)—Secode Corp.

S

S.F.D. Labs., Inc.—Varian Associates
SIE Div.—Dresser Industries
SKF Industries—Reed Instrument Bearing Co.
S. M. Heat Treating Co.—Standard Metals Corp.
S-M-S Corp.—Mallory & Co., Inc. (P.R.)
STL Products Div.—Space Technology Labs.; Thompson Ramo Wooldridge Inc.
S & T Sales, Ltd.—Spilsbury & Tindall, Ltd.
Safe Lighting, Inc.—Inertia Switch, Inc.
Saint Joe Machines, Inc.—Electrend Products Corp.
Salem Div.—Cannon Electric Co.
Sampson Electric Products Co.—Oxford Electric Corp.
Sandia Corp.—American Telephone & Telegraph Corp.
Sandimar Recording Co.—Westinghouse Electric Corp.
Sanford Mfg. Corp.—Micromech Mfg. Corp.
Sanford Plastics Div.—Bonny Mfg. Corp.
Saratoga Industries Div.—Espey Mfg. & Electronics Corp.
Sargent & Greenleaf, Inc.—Security Devices Lab.
Savage Industries, Inc.—Wiley Electronics Co.
Savoy Electronics, Inc.—Bassett, Inc. (Rex)
Savoy Industries, Inc.—Savoy Electronics, Inc.
Saxon Electronics—G-V Controls, Inc.
Scatter Communications, Inc.—Datronics Engineers, Inc.; Developmental Engineering Corp.
Schafer Custom Engineering—Textron Electronics, Inc.
Schloer, Inc. (H.C.)—Modern Design Div.
Schlumberger Corp.—Computer Systems, Inc.
Schmidt & Co. (Hans)—Electromatic Equipment Co.
Schoenberger Co. (W.J.)—Bell Aircraft Corp.
Schomandl—Electronic Applications, Inc.
Schutter—CWS Waveguide Corp.
Schwien Engineering—Whittaker Gyro Div.
Science Electronics, Inc.—General Electronic Labs., Inc.
Scientific & Process Instrument Div.—Beckman Instruments, Inc.
Scientific Products, Ltd.—Painton & Co., Ltd.
Scintilla Div.—Bendix Corp.
Scranton Corp.—Eicor Div.
Seaboard Pacific Div.—Associated Spring Corp.
Sealectro Corp.—Deltime, Inc.
Sears Co. (M.J.)—Audioscars Corp.
Seco Mfg. Co.—Seco Electronics, Inc.
Seismograph Service Corp.—Seiscor Div.
Selcose Metals Corp.—Secon Metals Corp.
Semcor Solid State Div.—Nuclear Corp. of America
Semiconductor Devices, Inc.—Kaiser Electronics, Inc.
Semimetals, Inc.—General Instrument Corp.
Seminole Div.—Airpax Electronics, Inc.
Sentinel Co.—Magnavox Co.
Sequoia Wire & Cable Co.—Anaconda Wire & Cable Co.
Servel, Inc.—Burgess Battery Div.
Service Bureau Corp.—International Business Machines Corp.
Servo Corp. of America—Electro-Pulse, Inc.
Servo Dynamics—National Co.
Servo Electronic Switch & Signal Co.—Servo Corp. of America
Servocontrol Div.—Oilgear Co.
Servomechanisms, Inc.—Mechatrol Div.
Servotrol, Inc.—Technology Instrument Corp.
Sessions Clock Co.—Haydon Co. (The A.W.)
Seth Thomas Div.—General Time Corp.
Shakeproof Div.—Illinois Tool Works
Shakespeare Co.—Columbia Products Co.
Shasta Div.—Beckman Instruments, Inc.
Shasta Mfg. Co.—Shaw Insulator Div.
Sheffco Mfg. Corp.—Townsend Co.; Textron, Inc.
Sheffield Corp.—Bendix Corp.
Shell Development Co.—Hallikainen Instruments
Shelton Eyelet Div.—United Shoe Machinery Corp.
Shepard Associates, Inc.—Shepard Labs., Inc.
Shepherd Electrical Labs.—Nichols, Ltd. (R.H.)
Shepherd Industries, Inc.—S-I Electronics, Inc.
Sherman Classics, Inc.—United International Dynamics Corp.
Sherman Industries, Inc.—American Silver Co., Inc.
Shockley Transistor Div.—Clevite Corp.
Short Wave Plastic Forming Co.—Cavitron Electron Oscillator Co.
Shreve Mfg. Co.—Diamonite Products Mfg. Co.
Shurite Meters—J-B-T Instruments, Inc.
Sickles Div. (F.W.)—General Instrument Corp.
Siegler Corp.—Bogen-Presto Co.; Hallamore Electronics Co.; Magnetic Amplifiers, Inc.; Jack & Heintz Div.; Olympic Radio & Television
Sierra Electronic Corp.—Philco Corp.
Sightmaster of Calif. Co.—Chemalloy Electronics Corp.
Sigma Instruments, Inc.—Fisher-Pierce Co.
Sigmadyne of Conn.—Cornwell Electronics Corp.
Signal Indicator Corp.—Dialight Corp.
Sikorski Aircraft Div.—United Aircraft Corp.
Simmonds Aerocessories, Inc.—Simmonds Precision Products, Inc.
Simmons Machine Tool Corp.—Simmons Fastener Corp.
Simmons-Woodward, Inc.—Universal Match Corp.
Simplex Wire & Cable Co.—Hitemp Wires Co.
Singer-Bridgeport—Singer Mfg. Co., Military Div.
Singer Mfg. Co.—Diehl Mfg. Co.; HRB-Singer, Inc.
Single Crystal Corp. of America—Semi-Elements, Inc.
Skottie Electronics, Inc.—Astron Corp.
Slaco Div.—Hallikainen Instruments
Slepian & Co. (Arthur)—Asco Wire & Cable Co.
Slip Ring Co. of America—A-1 Precision Products
Smalley's Radio, Ltd.—Century Electronics & Instruments, Inc.

(Continued on page 243A)



RAYTHEON PHASED ARRAY ASSEMBLY includes Amplitron,* ferrite circulator and TWT in single package. Complete assemblies are precisely matched mechanically and electrically to provide identical performance throughout the array.

*Raytheon Trademark

Raytheon tubes bring new modular solution to phased array problems

TYPICAL OPERATING CHARACTERISTICS

	Model CLM Circulator*		QKW 1013 TWT	QKS 1012 Amplitron
Peak pwr.	15 kw	Pk. power out.	5 kw	100 kw
Average pwr.	800W	Av. power out.	400W	2.5 kw
Frequency (Mc)	1215-1400	Frequency (Mc)	1215-1400	1215-1400
Isolation	18 db min.	Gain	50 db	13 db
Insertion loss	0.4 db max.	Efficiency	20-25%	70%
VSWR	1.25	Pulse dur.	750μs	500μs

*3-port Wye-junction circulator with load.

New permanent magnet concept makes possible half wavelength mounting of transmitter tubes

Raytheon now has an advanced Amplitron, TWT and ferrite circulator for modular mounting in a permanent magnet lattice. The new assembly provides excellent efficiency, bandwidth and gain over a wide range of power output levels.

The TWT-circulator-Amplitron assemblies are expressly designed for minimum noise and uniformity of phase and gain characteristics from unit to unit. To provide greater operating flexibility, the TWT employs a modulating anode and isolated collector.

Raytheon can offer the new combination at low cost and in large quantities. Write today for complete technical details. Address Raytheon Company, Microwave and Power Tube Division, Waltham 54, Massachusetts.

RAYTHEON

RAYTHEON COMPANY

MICROWAVE AND POWER TUBE DIVISION

why Ampex uses NIKON OPTICAL COMPARATORS

The new Ampex AR-300, developed principally for military and scientific applications, is the only recorder capable of covering a 4-megacycle bandwidth. It can capture and record phenomena beyond the range of any equipment in use today.

The most vital component in the AR-300 is its rotating recording head, designed to achieve the high, head-to-tape velocity required for wide band response, yet maintaining a relatively low, reel-to-reel speed.

Nothing in the manufacture of the AR-300 proved as critical as the production of this head. Most of its parts did not lend themselves to conventional inspection methods. While .0001″ tolerances had to be maintained, Ampex standards call for inspection equipment capable of at least 10 times the required accuracy.

This inspection problem was finally solved when the task was assigned to a Nikon 6 Optical Comparator. Not only did the Nikon Comparator provide the exact measurements required during 'in-process' inspection, but it also permitted consistent duplication of these measurements, time after time.

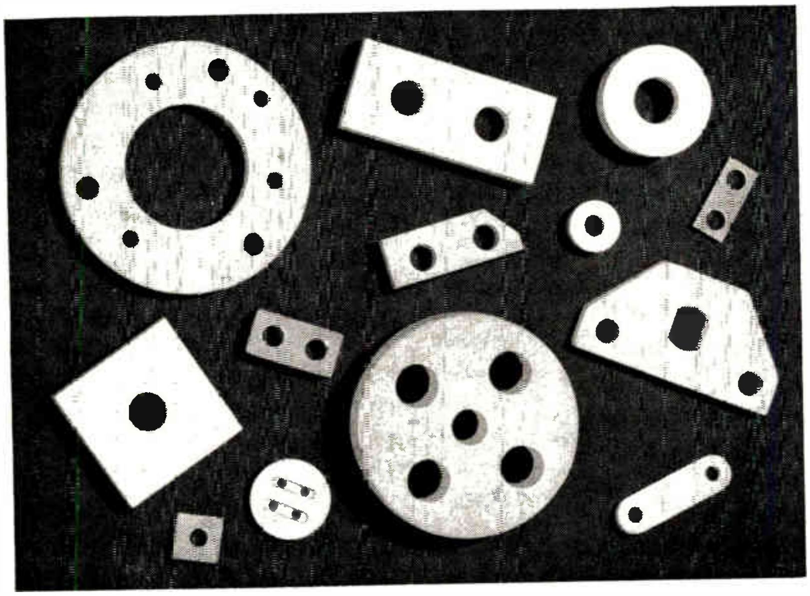
Today, the Nikon 6 has a permanent position on the AR-300 production team. Furthermore, there are now five additional Nikon comparators serving Ampex in other production projects, and in development—helping to maintain the high quality standards for which Ampex products are justly famous.

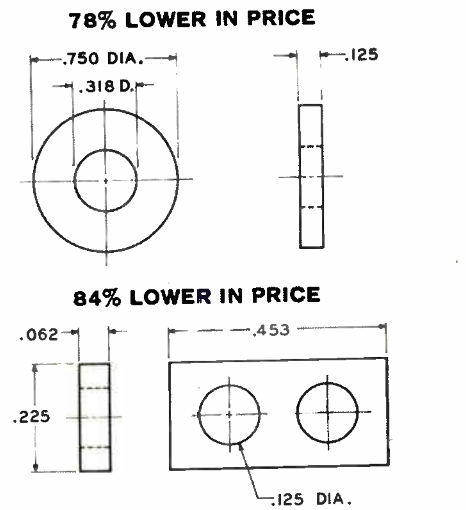
Why not investigate what a Nikon comparator can do for you? Write for complete, illustrated catalog to Dept. IRE-5.

Instrument Division • Nikon Incorporated • 111 Fifth Ave., N. Y. 3, N. Y.
subsidiary of Ehrenreich Photo-Optical Industries, Inc.

Ferride Coil Under Inspection Showing Size Relationship

New cost savings for users of dielectric materials





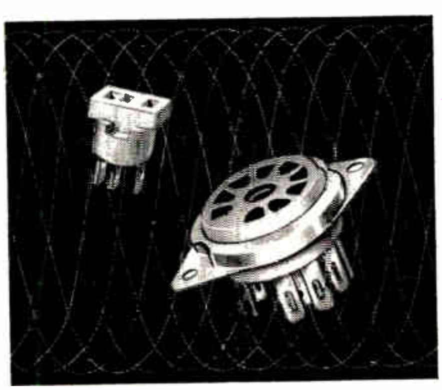
. . . a progress report on the MYCALEX METHOD from Jerome Taishoff, President, Mycalex Corporation of America

"We're proud to announce the MYCALEX METHOD: a unique molding and finishing technique that delivers higher-performance parts at lower cost. In line with our policy, we pass all savings on to our customers.

"The diagram above outlines typical savings –78% and 84%–both in addition to the temperature endurance, total dimensional stability, dielectric strength and low-loss SUPRAMICA® and MYCALEX® have delivered for years."

Your choice of SUPRAMICA 620 "BB", 560 and 555 ceramoplastics, as well as MYCALEX 410 glass-bonded mica.

So to ease your profit squeeze . . . get top-grade electronic-insulation materials for less than ever . . . look into the new MYCALEX METHOD. Send blueprints and drawings for specifics.

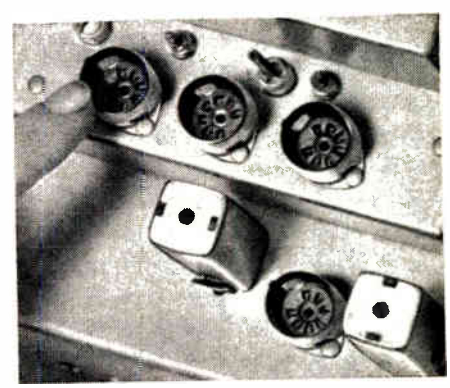


IN SUPPLY

Sockets of MYCALEX glass-bonded mica are yours in standard or custom-designed types. Also supplied in SUPRAMICA ceramoplastics on special order . . . sockets illustrate the complex configurations in which these versatile materials can be produced.

IN SERVICE

For 11 years, MYCALEX tube sockets in important AM, FM and low-level audio circuits have helped to make this custom-made Fisher TRI 6G the superb radio-phono-TV instrument it is.



General Offices and Plant:
126 Clifton Boulevard
Clifton, N. J.

SYNTHAMICA® SYNTHETIC MICA — SUPRAMICA® CERAMOPLASTIC — SINCE 1919 MYCALEX® THE INSULATOR

MYCALEX
CORPORATION OF AMERICA

Executive Offices:
30 Rockefeller Plaza
New York 20, N. Y.

World's largest manufacturer of ceramoplastics, glass-bonded mica and synthetic mica products

Manufacturers' Cross-Reference Index

(Continued from page 238A)

Smith Corp. (A.O.)—Crowley Div.
Smith Paper Co.—Schweitzer Div. (Peter J.)
Sodeco—Landis & Gyr, Inc.
Soiltest, Inc.—Central Scientific Co.
Sola Electric Co.—Hevi-Duty Electric Co.; Basic Products Corp.
Solar Aircraft Co.—Electronic Systems Development Corp.
Solartron Electronic Group, Ltd.—Solartron, Inc.
Solid State Instrument Labs.—Rowan Controller Co.
Sonic Energy Products—Bendix Corp.
Sonobond Corp.—Aeroprojects, Inc.
Sonotec, Inc.—Langevin Div.
Soreng Products Corp.—Controls Co. of America
Sound Equipment Corp. of Calif.—Kwikheat Mfg. Co.
South Chester Co.—Southco Div.; Lion Fastener Co.
Southwestern Industrial Electronics—Dresser/SIE
Sovirel—Corning Glass Works
Space Communications Div.—Radiation, Inc.
Space Electronics Corp.—Pacific Automation Products, Inc.
Space Recovery Systems, Inc.—CBS Labs.
Space Technology Labs., Inc.—Thompson Ramo Wooldridge Inc.; STL Products Div.
Sparta Mfg. Div.—Diamonite Products Mfg. Co.
Spartan Co.—Magnavox Co.
Specialloid, Inc.—Radio Engineering Products, Ltd.
Specialty Engineering & Electronics Co.—Specialty Electronics Development Corp.
Spectra Electronics Corp.—Douglas Microwave Co., Inc.
Spectrolab, Inc.—Textron Electronics, Inc.
Speedgrinder Mfg. Co.—Everard Tap & Die Corp.
Speer Carbon Co.—Jeffers Electronics Div.
Spellman Television Co., Inc.—Spellman High Voltage Co.
Sperry Rand Corp.—Remington Rand Univac; Ford Instrument Co.; Sperry Gyroscope Co.; Sperry Microwave Electronics Co.; Sperry Piedmont Co.; Vickers, Inc.; Wheeler Electronic Corp.; Wright Machinery Co.; Wright Div.
Spinco Div.—Beckman Instruments, Inc.
Spinform Div.—Antenna Systems, Inc.
Spiroid Div.—Licon Switch & Control Div. of Illinois Tool Works
Sprague of Wisconsin, Inc.—Sprague Electric Co.
Sprague Products Co.—Sprague Electric Co.
Stability Capacitors, Ltd.—British Radio Electronics, Ltd.
Standard Coil Products Co., Inc.—Kollsman Instrument Corp.; Kollsman Motor Corp.; Standard Kollsman Industries, Inc.
Standard Electrical Products Co.—General Electronic Control, Inc.
Standard Electronics Div.—Dynamics Corp. of America
Standard Electronics Research Corp.—National Standard Electronics, Inc.
Standard Felt Co.—Victoreen Instrument Co.
Standard Kollsman Industries, Inc.—Kollsman Instrument Corp.; Standard Coil Products Co., Inc.; Richardson-Allen Corp.
Standard Plastics—Essex Electronics
Standard Pressed Steel Co.—International Electronic Industries, Inc.
Standard Products Co., Inc.—Standard Electrical Products Co.
Standard Steel Works Div.—Baldwin-Lima-Hamilton Corp.
Standard Winding Co.—Ovitron Corp.
Stanray Corp.—Gillen Co. (John)
Stantron Div.—Wyco Metal Products
Star Expansion Industries Corp.—Star Precision Devices
States Electronics Corp.—Bludworth Marine Div.
Stavid Div.—Lockheed Electronics Co.
Stavid Engineering, Inc.—Lockheed Electronics Co.
Stearns Magnetic Products Div.—Indiana Steel Products Div.
Steel Co. (Herman D.)—Swiss Jewel Co.
Steel Improvement & Forge Co.—Sifco Metachemical, Inc.
Stepper Motors Div.—Land-Air, Inc.
Sterilshield Instruments, Inc.—Baker Co., Inc.
Sterling Precision Corp.—Wheelabrator Corp.
Steuben Glass, Inc.—Corning Glass Works
Stimsonite Div.—Elastic Stop Nut Corp. of America
Stokes Corp. (F.J.)—International Pump & Machine Works
Strand Labs., Inc.—Ferrotec, Inc.; Microwave Development Labs., Inc.
Stratford Electronics Div.—Bruno-N.Y. Industries Corp.
Stratos Div.—Fairchild-Stratos Corp.
Stratton & Co., Ltd.—British Radio Electronics, Ltd.
Stromberg-Carlson—General Dynamics/Electronics
Stromberg Time Corp.—General Time Corp.
Strong Electric Corp.—General Precision Equipment Corp.
Stupakoff Div.—Carborundum Co.
Sturrup, Larrabee & Wamers, Inc.—Sturrup, Inc.
Sun Chemical Corp.—Electro Technical Products
Sunnyvale Development Center—Sperry Gyroscope Co.
Sunnyvale Div.—Westinghouse Electric Corp.
Sunvic Controls, Ltd.—Associated Electrical Industries, Ltd.
Super Tool Co.—Van Norman Industries, Inc.
Superior Tube Co.—Johnson & Hoffman Mfg. Corp.
Supreme Electronics—Hickok Electrical Instrument Co.
Supreme Instruments Corp.—Supreme Electronics Corp.
Supreme Transformer Div.—Oxford Electric Corp.
Surf Chemical Corp.—Driver-Harris Co.
Susquehanna Corp.—Computer Engineering Assoc., Inc.; Susquehanna Sciences, Inc.
Sutton Electronic Co.—Essex Electronics
Swift & Sons, Inc. (M.)—Swift Textile Metalizing & Laminating Corp.
Swiss Jewel Co.—Steel Co. (Herman D.)
Switch-Lock, Inc.—Pendar, Inc.
Sylphon Div.—Robertshaw-Fulton Controls Co.
Sylvania Electric Products Inc.—Sylvania Electronic Systems; General Telephone & Electronics Corp.; Argus Cameras, Inc.
Symphonic Electronic Corp.—Lynch Corp.
Synchrosolve, Inc.—Rotating Components, Inc.
Synthetic Mica Co.—Mycalex Corp. of America
Syntorque Corp.—McLean Syntorque Corp.
Syntorque Motors, Inc.—Brevel Products Corp.
Syrkus & Guttman, Inc.—Associated American Winding Machinery, Inc.
Syscom Div.—AMP, Inc.
Systematics Div.—General Instrument Corp.
Systems Development Corp.—Electronic Systems Development Corp.
Systems Labs. Div.—Electronic Specialty Co.
Systron Corp.—Systron-Donner Corp.
Systron-Donner Corp.—Donner Scientific Co.; Greenleaf Mfg. Div.

T

T.E.M., Inc.—Microlab
TMC Canada, Ltd.—Technical Materiel Corp.
TMC Industries Corp.—Technical Materiel Corp.
TM Development Corp.—Technical Materiel Corp.
Taco West Corp.—West Instrument Corp.
Taller Cooper Div.—American Electronics, Inc.
Talley Industries, Inc.—Microtech, Inc.
Tammen & Denison—Chicago Electronic Div.
Tapco Div.—Thompson Ramo Wooldridge Inc.
Taylor Instrument Co.—Taylor-Emmett Controls, Inc.
TechRep Div.—Philco Corp.
Technical Electronics Co.—Price Electric Corp.
Technical Industries Corp.—Packard-Bell Electronics Corp.
Technical Materiel Corp.—TMC Canada, Ltd.
Technical Operations, Inc.—Microwave Assoc., Inc.; Power Sources, Inc.; Western Union Telegraph Co.
Technical Products International—Technical Materiel Corp.
Technical Products & Services Co.—Chemalloy Electronics Corp.
Technical Research Group—TRG, Inc.

(Continued on page 246A)

KLYSTRONS

400-2855 Mc
100 kw to
40 Mw Pulse

Here: L-3401
1254 to 1386 Mc
Tunable
5 Mw peak
(minimum)

Height: 9 ft.

Length: 13 in.

TRAVELING WAVE TUBES

400-11,000 Mc
20 mW to 5.5 kw

Here: L-3711
4000-8000 Mc
1 W CW (minimum)

CROSSED FIELD AMPLIFIERS

8500-11,000 Mc
1 kW CW to 500 kw Pulse

Diameter: 12 in.

Here: L-3650
8500-9600 Mc Tunable
1 kW CW (minimum)

2 in. Cube

MINIATURE MAGNETRONS

8600-16,500 Mc/30 W to 4.0 kw peak

Here: L-3719
15,000 ± 100 Mc Fixed
750 W (minimum)

Diameter: 6 in.

BARRATRON®

Transmitting Tube
UHF through X-band
Broad bandwidths of Near
White Noise

Here: One of L-3330 Series

1 METER TO 4 MILLIMETERS
MILLIWATTS TO MEGAWATTS
ELECTRON TUBES BY LITTON

Height: 28 in.

SWITCH TUBES

150-350 kVdc Collector Voltage
20-300 A Switching Current

Here: L-3620
350 kVdc Collector Voltage
30 A Collector Current

Diameter: 8 in.

M-BWO's

1000-11,000 Mc
125-200 W CW

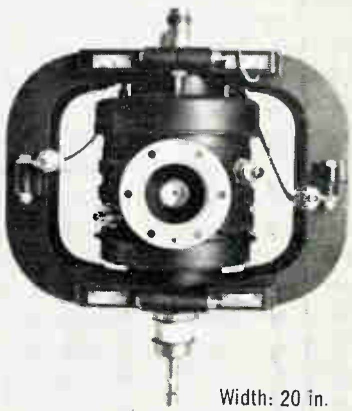
Here: L-3724
2500-3550 Mc
Tunable
180 W CW
(minimum)

Width: 8 in.

MILLIMETER WAVE TUBES

18-80 kMc
.03-50 W

Here: L-3689
68-80 kMc Tunable
0.5 W (nominal)



Width: 20 in.

PULSE & CW MAGNETRONS

406-17,000 Mc
30 W to 2 Mw peak
110 to 500 W CW

Here: L-3455
406-450 Mc Tunable
2 Mw peak (minimum)

Also latest developments in display devices and accessory equipment. Contact us at San Carlos, California, for complete information.

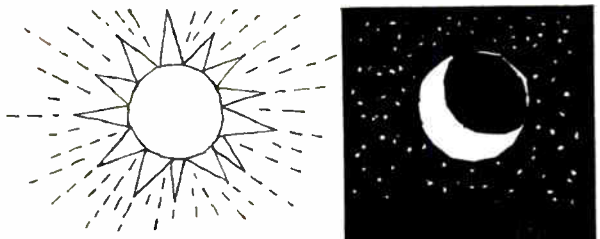
LITTON INDUSTRIES
Electron Tube Division
MICROWAVE TUBES AND DISPLAY DEVICES

Manufacturers' Cross-Reference Index

(Continued from page 243A)

Technicraft Div.—Electronic Specialty Co.
Technicraft Labs., Inc.—Electronic Specialty Co.
Technology Instrument Corp.—Acton Labs., Inc.; Technology Instrument Corp. of Calif.; Bowmar Instrument Corp.
Tecumseh Products Co.—Ohio Semiconductors Div.
Tee Corp.—Thermo Electron Engineering Corp.
Tele-Dynamics Div.—American Bosch Arma Corp.
Tele-Signal Corp.—GPL Div. of General Precision, Inc.
Telechrome Mfg. Corp.—Hammarlund Mfg. Co., Inc.
Telecom, Inc.—Webster Electric Co.
Telecomputing Corp.—Narmco Materials Div.; Whittaker Gyro Div.
Telecomputing Services, Inc.—Telecomputing Corp.
Telectrosonic Corp.—Telectro Industries Corp.
Telefunken—American Elite, Inc.
Telegraph Construction & Maintenance Co., Ltd.—Telcon, Inc.
Telegraphie sans Fil, Cie. Generale de—Steafix; American Radio Co., Inc.
Telemeter Magnetics, Inc.—Invar Electronics Corp.; Ampex Corp.
Telemetering Corp. of America—Pacific Mercury Electronics
Teleprinter Corp.—Western Union Telegraph Co.
Teleprompter Corp.—Western Union Telegraph Co.
Telerad Div.—Lionel Corp.
Teletronics Lab., Inc.—Crosby-Teletronics Corp.
Televex Co.—International Audio Stylus Corp.
Television & Radar Co.—Plastoid Corp.
Television Specialty Co.—Federal Mfg. & Engineering Corp.
Telewave Labs., Inc.—Polarad Electronics Corp.
Telex, Inc.—Ballastran Div.; Aemco Div.; Lumen Div.
Telonic Engineering Corp.—Telonic Industries, Inc.
Temco Aircraft Corp.—Fenske, Fedrick & Miller, Inc.; Ling-Temco-Vought, Inc.; Temco Electronics Div.
Tempel Mfg. Co.—Tempel Steel Co.
Tempo TV Sales Corp.—Clear Beam Antenna Co.
Tennessee Corp.—Chester Cable Corp.
Tenney Engineering—Pam Associates, Inc.
Test Institute Corp.—Benmar Heater Div.
Texas Electronic Products Corp.—Dorsett Electronics, Inc.
Texsteam Corp.—Vap-Air Div.
Textron Electronics, Inc.—Allegany Instrument Co., Inc.; Electronic Research Co.; G-C Electronics Div.; Globe Electronics Div.; MB Electronics Div.; MB Mfg. Co., Inc.; Schafer Custom Engineering Div.; Spectrolab, Inc.
Textron, Inc.—Bell Aerospace Corp.; Boots Aircraft Nut Co.; California Technical Industries; Dalmo Victor Co.; Homelite Corp.; Nuclear Metals, Inc.; Textron Electronics, Inc.; Townsend Co.; Hydraulic Research Div.
Thermal Dynamic Products, Inc.—Waltham Precision Instrument Co.
Thermal Syndicate, Ltd.—Thermal American Fused Quartz Co.
Thermocal Div.—Jamieson Electronic Labs., Inc.
Thomas Closure & Specialty Corp.—Thomas & Sons, Inc. (William)
Thomas Organ Co.—Pacific Mercury Electronics Div.
Thompson-Bremer Co.—American Machine & Foundry Co.
Thompson Products, Inc.—Thompson Ramo Wooldridge Inc.
Thompson Ramo Wooldridge Inc.—Good-All Electric Mfg. Co.; Thompson-Ramo-Wooldridge Products Co.; Pacific Semiconductors, Inc.
Thordarson - Meissner — Gramer - Halldorson Transformer Corp.
Three Point One Four Corp.—Chlorivolt Corp.
Ticket Office Equipment Div.—Toeco Metal Products, Inc.
Times Facsimile Co.—Westrex Co.; Litton Industries, Inc.
Titania Div.—American Lava Corp.
Titania Electric Co. of Canada, Ltd.—Gulton Industries, Inc.
Titanium Alloy Mfg. Div.—National Lead Co.
Tobe Deutschmann Corp.—Cornell-Dubilier Electronics Div.
Tocco Div.—Ohio Crankshaft Co.
Toms River-Cincinnati Chemical Corp.—Ciba Products Corp.
Tool-Lab, Inc.—Beckman Instruments, Inc.
Topp Industries—United Industrial Corp.
Topp Mfg. Co.—U. S. Science Corp.
Toroids Unlimited—New England Transformer Co., Inc.
Torque Industries, Inc.—Co Engineering Co.
Toshiba Co.—Tokyo Shibaura Electric Co.
Townsend Co.—Boots Aircraft Nut Co.; Textron, Inc.; Sheffco Mfg. Corp.
Trak Electronics Co.—CGS Labs., Inc.
Trans Digital Systems Div.—Cook Electric Co.
Trans-Electric Bahama—Savoy Electronics, Inc.
Trans-Sonics Pacific—Trans-Sonics, Inc.
Transcon Div.—Aeroflex Corp.
Transcontinental Systems—Entron, Inc.
Transdata Div.—Ryan Aeronautical Co.
Transformer Engineers—Electro Instruments, Inc.
Transformer Model Shop—New England Transformer Co., Inc.
Transicoil Div.—Daystrom, Inc.
Transistor Applications Co.—Glentronics, Inc.; New England Transformer Co., Inc.
Transograph—Chart-Pak, Inc.
Transport Products Corp.—Chemalloy Electronics Corp.
Transtech, Inc.—Avionics Corp. of America
Travers & Co., Inc.—Precision Microwave Corp.
Trent Tube Co.—Crucible Steel Co. of America
Tri-Point Mfg., Inc.—Tri-Point Plastics, Inc.

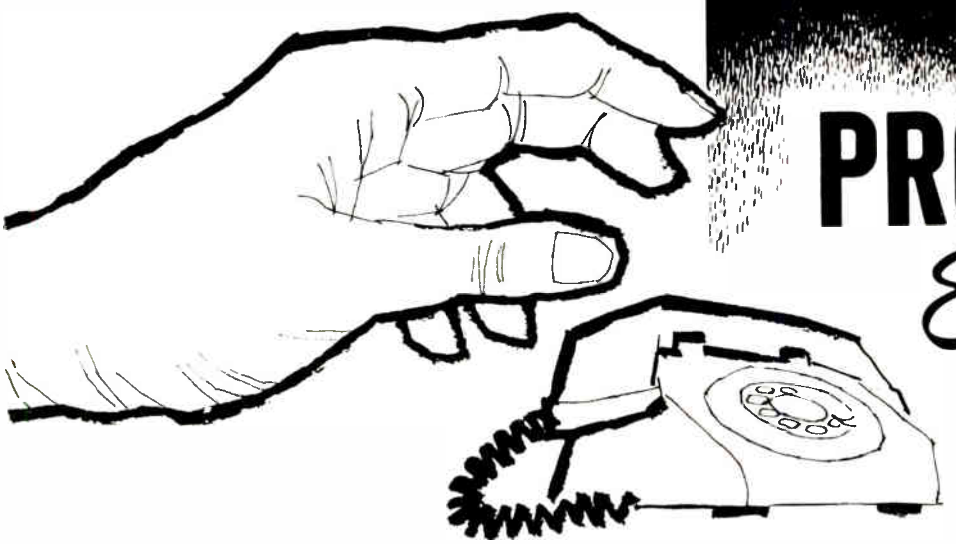
(Continued on page 248A)



WHAT IS A DAY SAVED WORTH TO YOU?

components & cable ordered now reach you *tomorrow*

when you call



PROGRESS *Electronics*

AUTHORIZED DISTRIBUTOR:

AMPHENOL · BORG · ipc · DAGE · CANNON PLUGS · GREMAR · HUGHES · ITT · S · Fenwal Electronics · ZT

Now OEM projects need not be held up for lack of essential parts. Whatever you might need — connectors, plugs, thermistors, transistors, cables, switches and potentiometers — Progress Electronics can supply them to you *from stock*. No fuss, no worry, no costly waiting — phone in your order today, you'll get them tomorrow!

We carry the *Amphenol-Borg* line. *IPC, Dage* — we have them. *Cannon, Fenwal, Gremar, Hughes, ITT, Silicon, Zippertubing* — just about every famous name in the book. And we continue to add lines for an even more complete service. We've also added warehouse space — 6,000 square feet, this year, in New York alone. This enlarged stock of high-precision components and cable is all ready for sameday shipment to those who pick up the phone and call.

How good is our service? It's tops, and here's the proof: electronics engineers and companies keep coming back to us, and recommend us to others. (You'll recommend us, too!) So we're growing. This year, we've added 10 men to our New York staff. And up in Canada, we've opened a new office in Montreal, for faster service to our Canadian clients.

Our speed should suit you fine! Write us for catalogs, or better still, call us for a specific item you need in a hurry.

Immediate delivery at factory prices guaranteed:

PROGRESS *Electronics* CO., INC.

MAIN OFFICE:
107 Franklin St., New York 13
212 - CAnal 6-5611 TWX New York 1-1377

BRANCHES:
11924 Santa Monica Blvd., Los Angeles 213 - GRanite 9-8884
480 Port Royal St. West Montreal, Canada 388-5746

RF & MULTI-PIN CONNECTORS, COAXIAL AND SPECIAL CABLES, THERMISTORS, DIODES, TRANSISTORS, RECTIFIERS, POWER DEVICES, ZIPPERTUBING®

Manufacturers' Cross-Reference Index

(Continued from page 246A)

Triad Transformer Corp.—Litton Industries, Inc.
Trimpot Div.—Bourns, Inc.
Trincor Corp.—Trinity Equipment Corp.
Triplett & Barton, Inc.—Sperry Products Co.
Trolex Corp.—CTS Corp.
Troy Electronic Devices, Inc.—Ram Meter, Inc.
Truscon Div.—Republic Steel Corp.
Tucson Instrument Corp.—Technology Instrument Corp.
Tuffline Div.—Whitney Blake Co
Tullamore Electronics Corp.—Victoreen Instrument Co.
Tweezer Weld Co.—Federal Electronic Products, Inc.
Twenty-First Century Electronics, Inc.—Osborne Electronics Sales Corp.
Twin Coach Co.—York Body & Equipment Co.

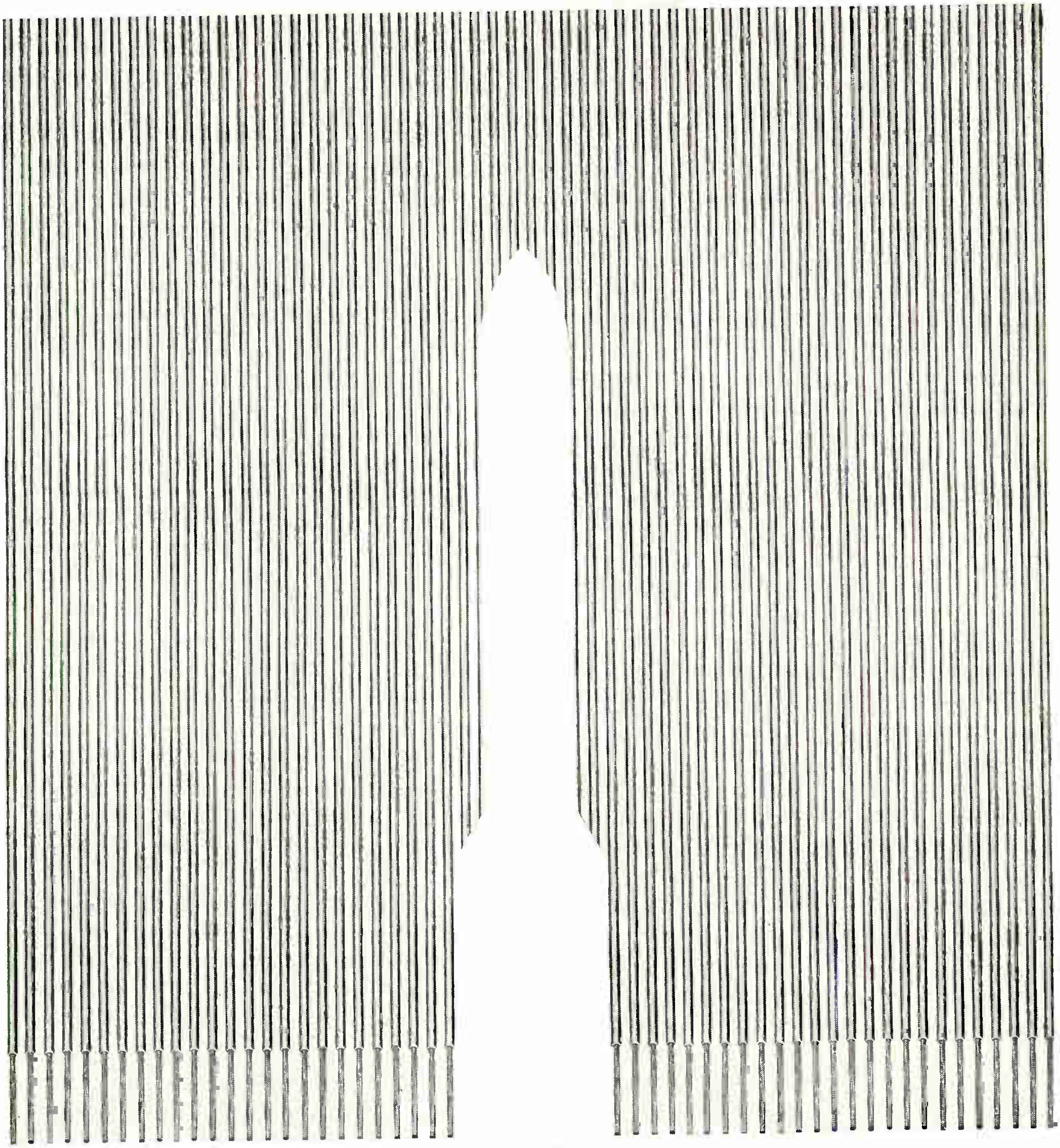
U

U.M.F. Mfg. Corp.—Rimak, Inc.
Ucinite Co.—United-Carr Fastener Corp.
Ultrasonic Corp. of America—Powertron Ultrasonics Corp.
Underwood Corp.—Canoga Div.
Unimax Switch Div.—Maxson Corp. (W.L.)
Union Carbide Corp.—Kemet Co.; National Carbon Co.; Linde Co.
Union Elect-Sonics Corp.—Union Ultra-Sonics Corp.
Union Gear & Sprocket Corp.—Union Ultra-Sonics Corp.
Unisco, Inc.—Eastern Foilcal Nameplate Corp.
United Aircraft Corp.—Hamilton Standard Div.; Norden Div.
United-Carr Fastener Corp.—Graphik-Circuits Div.; Cinch Mfg. Co.; Monadnock Mills; Palnut Co.; Ucinite Co.
United Condenser Corp.—Film Capacitors, Inc.
United Control Corp.—Electro Development Corp.; Palomar Scientific Corp.; Smith-Florence, Inc.
United Electric Controls Co.—United Sensor & Control Corp.
United Electronics Co.—Ling-Temco-Vought, Inc.
United Geophysical Corp.—United Electrodynamics, Inc.
United Industrial Corp.—Aircraft Armaments, Inc.; American Engineering Co.; Micro-Path, Inc.; Topp Industries; U. S. Science Corp.
United Insulator Div.—Telegraph Condenser Co., Ltd.
United Mfg. Co.—Maxson Corp. (W.L.)
United Pressed Products Co.—Uppco, Inc.
United Research Corp. of Menlo Park—Hamilton Standard Div.
United Scientific Labs.—De Wald Radio Div.
United Sensor & Control Corp.—United Electric Controls Co.
United Specialties Co.—Sibley Co.
U. S. Asbestos Div.—Raybestos-Manhattan, Inc.
U. S. Ceramic Tile Co.—Diamonite Products Mfg. Co.; Sparta Mfg. Co.
U. S. Electronics Development Corp.—U. S. Semiconductor Products
U. S. Engineering Co.—Litton Industries, Inc.
U. S. Flare Div.—Atlantic Research Corp.
U. S. Gasket Co.—Garlock Electronic Products
U. S. Industries, Inc.—USI Western Design; Western Design & Electronics Div.
U. S. Metals Refining Co.—American Metal Climax, Inc.
U. S. Relay Co.—U. S. Relay-Electronics
U. S. Science Corp.—Aircraft Armaments, Inc.
U. S. Semcor—U. S. Semiconductor Products; Nuclear Corp. of America
United Testing Labs.—United Electrodynamics, Inc.
Unitronics Co.—Siegler Corp.
Universal Controls, Inc.—Clare & Co. (C.P.)
Universal Dynamics—Acoustica Assoc., Inc.
Universal Fabricating Corp.—Missile Systems Corp.
Universal Mfg. Co., Inc.—Universal Toroid Coil Winding, Inc.
Universal Match Corp.—Reflectone Electronics, Inc.
Universal Metal Products Corp.—Universal Match Corp.
Universal Toroid Coil Winding, Inc.—Universal Mfg. Co., Inc.
Universal Transistor Products Corp.—Telechrome Mfg. Corp.
Universal Winding Co.—Leesona Corp.
Universal Wood Products—Electronic Teaching Labs., Inc.
University Loudspeakers, Inc.—Ling-Temco-Vought, Inc.
Useco—Litton Industries; U. S. Engineering Co.
Utah Radio Products—Carter Parts Co.
Utah Research & Development Co.—Interstate Electronics Corp.
Utica Drop Forge & Tool Div.—Kelsey-Hayes Co.
Utrad Corp.—Litton Industries, Inc.

V

Vacuum-Electronic Engineering Corp.—Veeco Vacuum Corp.
Vacuum Equipment Div.—Kinney Mfg. Div.; New York Air Brake Co.
Vacuum Furnace Div.—Brew & Co., Inc. (Richard D.)
Vacuum Tube Products Co.—Hughes Aircraft Co.
Valor Electronics, Inc.—Valor Instruments, Inc.
Value Engineered Products, Inc.—Telecomputing Corp.
Valvotechnia, Ltda.—Eletronica Industrial, Ltda.
Vanderbilt Tire & Rubber Corp.—Iuso Electronic Products, Inc.
Vapor Heating Corp.—Vap-Air Div.; Vapor Corp.
Varcum Chemical Div.—Reichhold Chemicals, Inc.
Vard Div.—Royal Industries, Inc.
Varian Associates—Bomac Labs., Inc.
Vascoloy-Ramet Div.—Fansteel Metallurgical Corp.
Vectrol Engineering, Inc.—Sprague Electric Co.
Vectron Corp.—Applied Dynamics Corp.

(Continued on page 250A)



Manufacturers' Cross-Reference Index

(Continued from page 218A)

Vectron, Inc.—Itek Corp.
Ventures, Ltd.—Metal Hydrides, Inc.
Vernistat Div.—Perkin-Elmer Corp.
Vibrapowr Co.—James Electronics, Inc.
Vibro-Ceramics Div.—Gulton Industries, Inc.
Vickers, Inc.—Sperry Rand Corp.
Victoreen Instrument Co.—Fast & Co. (John E.); Electronic Products Co.; Federal Mfg. & Engineering Corp.; Jordan Electronics Div.; Tullamore Electronics Corp.
Vidmar, Inc.—Volkert Stampings, Inc.
Vidya, Inc.—Itek Corp.
Viking Air Products Div.—Crane Co.
Visulite Co.—Keystone Electronics Corp.
Vitro Corp. of America—Nems-Clarke Co.; Vitro Chemical Co.
Vitro Electronics Div.—Nems-Clarke Co.
Vlier Engineering Corp.—Barry Wright Corp.
Vocaline Co. of America—Bristol Motors Div.
Vought Electronics Div.—Chance Vought Corp.; Ling-Temco-Vought, Inc.
Vue-Tronics, Inc.—Prescott Television Co.
Vultel, Inc.—Baso, Inc.

W

WRL Electronics—Globe Electronics
Wac Engineering Co.—WacLine, Inc.
Wagner Litho Machinery Div.—Reynolds Wire Div.
Wahlgren Electrical Mfg. Co.—Wahlgren Magnetics; Electro-Physics Labs.
Walden Engineering Corp.—Precision Microwave Corp.
Waldman & Sons (Joseph)—Epoxy Products Div.
Waldorf Studio of Engineering—Barta-Griffin Co.
Wales-Strippit, Inc.—Houdaille Industries, Inc.
Walkesville Corp.—Reed Research, Inc.
Walsco Electronics—G-C Electronics Co.; Textron Electronics, Inc.
Waltham Horological Corp.—Ken-Tron Corp.
Waltham Precision Instrument Co.—Boesch Mfg. Div.; Camblock Corp.; Electro-Mec Div.; Thermal Dynamic Products, Inc.
Walworth Co.—Belz Industries, Inc.
Waring Products Corp.—Dynamics Corp. of America
Warminster Fiberglass Co.—Fischer & Porter Co.
Warren Components Div.—Robertson Electric Co., Inc.
Warren Wire Co.—Wire Co. of America.
Waterbury Brass Goods—American Brass Co.
Waterbury Button Co.—Waterbury Companies, Inc.
Wave Particle Div.—Ramage & Miller, Inc.
Waveco Corp.—Microwave Associates, Inc.
Webb Wire Div.—Carpenter Steel Co.
Weber Showcase & Fixture Co., Inc.—Weber Aircraft Corp.
Webster-Chicago Corp.—Webcor, Inc.
Webster Mfg.—Applied Electronics Co.; Raytheon Co.
Weighing & Control Components, Inc.—CompuDyne Corp.
Welch Connectors—Control Dynamics Corp.
Welch Mfg. Co. (W.M.)—Welch Scientific Co.
Weldmatic Div.—Unitek Corp.
Well Instrument Developing Co.—Mandrel Industries, Inc.
Wells Mfg. Co., Inc.—Wells Electronics Co.
Welwyn Electric, Ltd.—Welwyn International, Inc.
Wesco Electrical Co., Inc.—Atlee Corp.
West Virginia Pulp & Paper Co.—Hinde & Dauch Div.
Westbury Electronics Corp.—Intercontinental Electronics Corp.
Westclox Div.—General Time Corp.
Western Alloy Products Co.—Mallory & Co., Inc. (P.R.)
Western Astronautics—Crescent Engineering & Research Co.
Western Automatic Machine Screw Co.—Standard Screw Co.
Western Development Labs.—Philco Corp.
Western Gold & Platinum Co.—Driver Co. (Wilbur B.)
Western Insulated Wire Co.—Pacific Resistor Co.
Western Lithograph Co.—Westline Products Div.
Western Molded Fibre Products, Inc.—Hawley Products Co.
Western Union Telegraph Co.—Gray Mfg. Co.
Westinghouse Air Brake Co.—Melpar, Inc.; Union Switch & Signal Div.
Weston Instruments Div.—Daystrom, Inc.
Westwood Div.—Houston Fearless Corp.
Wheaton Co. (T.C.)—Tronex, Inc.
Wheelabrator Corp.—Bell Aircraft Corp.
Wheeler Electronic Corp.—Sperry Rand Corp.
Wheeler Insulated Wire Co.—Wheeler Electronic Corp.
Wheeler Labs., Inc.—Hazeltine Electronics Div.
Whirlclad Div.—Polymer Corp. of Pennsylvania
White Electron Devices, Inc. (Roger)—Litton Industries, Inc.
White Industries—Airtronics International Corp.
Whitney & Co. (J.H.)—Summers & Mills, Inc.
Whitney Metrology Lab. (Eli)—Sheffield Corp.
Whittaker Controls—Telecomputing Corp.
Whittaker Gyro—Telecomputing Corp.
Whittin Machine Works—Harvey-Wells Electronics, Inc.
Wiancko Engineering Co.—Daystrom, Inc.
Wico Div.—Centralab Electronics
Wilcox Products Co.—Flo-Tronics, Inc.
Willard Machine—Robertson Electric Co., Inc.
Willmott-Taylor, Ltd.—Kassel Industries, Inc.
Willor Mfg. Corp.—Loral Electronics Corp.
Willrose Corp.—Horlick Co., Inc. (William I.)
Wilson Co. (K.R.)—Premax Products Div.
Wilson Div. (H.A.)—Engelhard Industries, Inc.
Wincharger Corp.—Rauland Corp.; Zenith Radio Corp.
Winco Electronics—Wintronics
Winsted Hardware Mfg. Div.—Dynamics Corp. of America
Winston Electronics—Jetronic Industries, Inc.
Winston Electronics, Ltd.—Dynamics Corp. of America
Wire Co. of America—Warren Wire Co.
Wire-Wrap Div.—Gardner-Denver Co.
Wirecom Div.—Cook Electric Co.
Wirt Co.—Continental Carbon, Inc.; Continental-Wirt Electronics Corp.
Wisco Div.—Electric Storage Battery Co.
Wolfe Electronics Corp.—Westinghouse Electric Corp.
Woodford Mfg. Corp.—Fusite Corp.
Woolfe Co. (Franklyn C.)—Parker Seal Co.
Worcester Wire Div.—National-Standard Co.
World Trade Corp.—International Business Machines Corp.
Worthington Corp.—Electric Machinery Mfg. Co.
Wright Electronics, Inc.—Midland Mfg. Co.
Wright Line, Inc.—Barry Wright Corp.
Wright Machinery Co.—Sperry Rand Corp.
Wuerth Tube-Saver Corp.—Wuerth Products Corp.
Wyle Labs.—Wyle Mfg. Corp.; Wyle-Burgoyne, Inc.
Wyle-Parameters, Inc.—Wyle Labs.

X

X-Acto, Inc.—Handicraft Tools, Inc.
X-Ray Monitoring Corp.—Atomic Film Badge Corp.

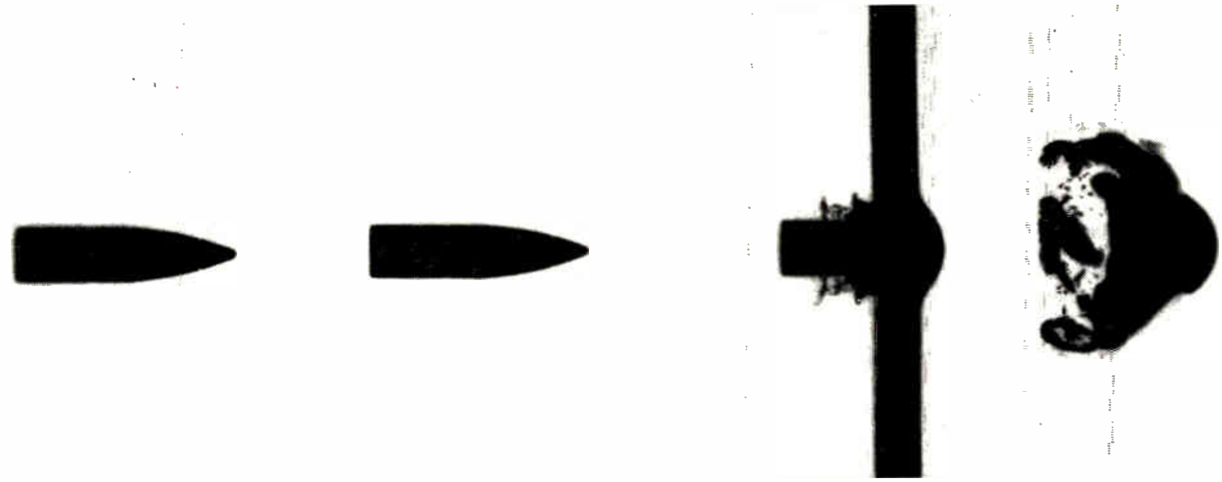
Y

York Div.—Bendix Corp.
Young Spring & Wire Corp.—Gonset Div.
Yuba Consolidated Industries, Inc.—Yuba-Dalmotor Div.

Z

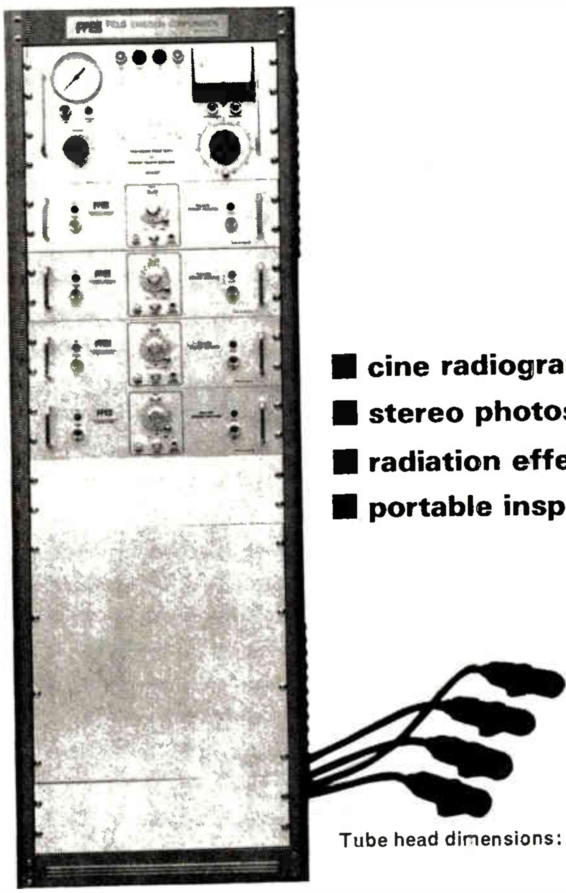
Zacharias Co. (Robert)—Precision Electroplating Co.
Zangerle & Peterson Co.—Wells-Gardner Electronics Corp.
Zell Products Corp.—Molecu-Wire Corp.
Zenith Radio Corp.—Central Electronics, Inc.; Rauland Corp.; Wincharger Corp.
Zenith Radio Research Corp.—Rauland Corp.

Use your IRE DIRECTORY! It's valuable!



X-ray of projectile penetrating steel plate.

New flash X-ray system



Tube head dimensions: 2" x 10"

fexitron model 730 series

- cine radiographs
- stereo photos
- radiation effects
- portable inspection

Performance data: Frame rates to 10^6/Sec ■ 30 to 100 nanoseconds ■ 75 KV to 300 KV ■ Dose rates to 10^8R/Sec ■ For additional information write or call: Field Emission Corporation, 611 Third Street, McMinnville, Oregon, Telephone: 472-5101 ■

Field Emission Corporation

CONHEX SUB-MINIATURE R. F. CONNECTORS

Designed and manufactured to meet the industry's demand for a better connector. Real gold-plated, not just flashed. Maximum impedance match, minimum VSWR. Available in Crimp-On or Clamp-On designs.

SCREW-ON
Clamp-On

SNAP-ON
Clamp-On, Crimp-On

PRESS-FIT® TEFLON TERMINALS

Sealectro and only Sealectro makes Press-Fit terminals. Ultra-quality provides optimum reliability and assembly ease. Available in all standard EIA colors.

SUBMINIATURE STANDOFFS
Teflon bushings in diameters from .148" to .218". Choice of lugs, including turrets.

SUBMINIATURE FEEDTHRUS
For connections through chassis or casing. Choice of lug designs including hollow-tube.

MINIATURE STAND-OFFS
For components and assemblies stressing higher voltages. Wide choice of lugs and turret designs.

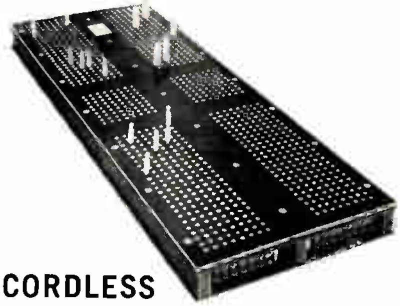
MINIATURE FEEDTHRUS
For handling voltages up to 4000 VDC. Wide choice of bushing diameters and lengths.

SEALECTRO COMPONENTS OFFER PROVEN RELIABILITY IN ALL INDUSTRIAL AND MILITARY APPLICATIONS

PRESS-FIT® TEST JACKS AND PLUGS

Male and female. Ideal for test instrument probes, jumper plugs and multiple pin plugs. Test jacks for printed circuitry.

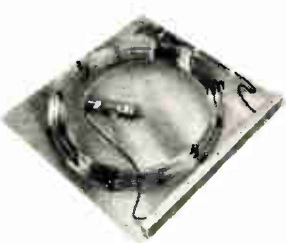
SEALECTOBOARD



CORDLESS PROGRAM BOARDS

The revolutionary program board. Simplifies multi-channel switching. Provides complete connections with insertion of single pin. No cord clutter. Accessories include component holders for inserting diodes, resistors, or other components at any circuit point. Available in any X- and Y- configuration. Standard Proto-Kits including board, shorting pins and component holders available from distributors.

DELTIME MAGNETOSTRICTIVE DELAY LINES



Longer delay times utilizing the magnetostrictive principle for maximum stability and mechanical ruggedness. Delay times up to 10 milliseconds at repetition rate of 655 KC with return-to-zero, or 5 millisecond delay at 1 MC with return-to-zero. Suitable for data storage. Completely humidity and magnetically shielded. Many standard fixed and variable models available.

WRITE FOR CATALOGS—
STATING PRODUCT INTEREST

Sealectro CORPORATION
139 HOYT STREET • MAMARONECK, N. Y.

IRE People

(Continued from page 130A)

The appointments of **Gerald Smith** (SM'56) as Director of Development Engineering, and **Ronald G. Oschmann** (M'56) as Section Manager for Electronic Systems in the company's Development Engineering Systems Department have been announced by Daystrom, Incorporated, Military Electronics Division, Archbald, Pa.

G. Smith

R. G. Oschmann

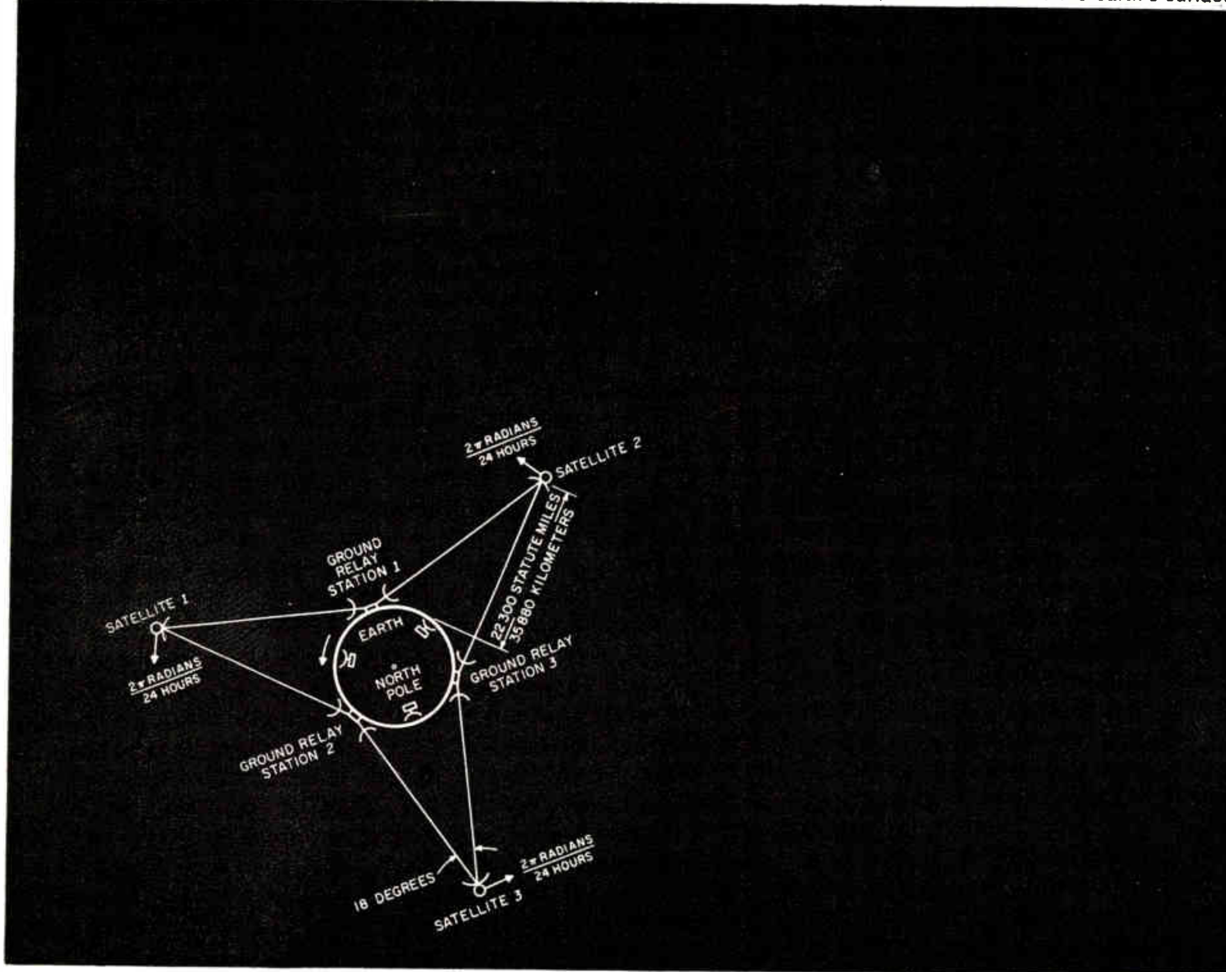
Mr. Smith joined Daystrom, Incorporated in 1952 as an electronic engineer in the Research and Development Department. His service also included the position of Senior Engineer, Development Engineer and Section Manager for Electronic Systems. Prior to joining Daystrom, Incorporated, he was associated with Marine Electric Co. and National Union Tube Co.

Mr. Smith attended City College of New York and graduated magna cum laude with a B.S. degree in electrical engineering. He received his Master's degree in electrical engineering from the Stevens Institute of Technology. He is a member of Tau Beta Pi and Eta Kappa Nu.

Mr. Oschmann joined Daystrom, Incorporated in 1952 as an electronic engineer. He accepted a position in 1954 with the Curtiss-Wright Corp., returning to Daystrom in 1955. Since then he has held a variety of engineering positions with the company.

He is a graduate of the Pennsylvania State University with a B.S. degree in electrical engineering.

The geometry of a satellite communication system in which each satellite maintains a fixed position relative to the earth's surface.*



THE "WHY" OF SATELLITE COMMUNICATION

In the past 60 years the world's population has nearly doubled. Demographers expect it to re-double in the next 20 years. This "population explosion" has triggered a "communication explosion" of such vast proportions that even planned cable and microwave capacity soon will be outstripped. Today, it is obvious that new communication channels must be opened./At ITT we have invested many thousands of man-hours and dollars in our conviction that a radio-relay satellite system is a feasible and economic way to expand world telephone capacity and provide a reliable data transmission and television link as well. In fact, at ITTFL we already have a satellite ground station terminal built and ready to function. We have become deeply involved in other pioneer space communication projects too, such as the RELAY and COURIER programs. Backed by the skills of scientists and engineers in 24 countries, ITT associates and divisions are actively engaged in creating communications systems and hardware to serve future generations./International Telephone and Telegraph Corporation, ITT Building, 320 Park Avenue, New York 22, N. Y.

*Louis Pollack, "Radio Communication Using Earth-Satellite Repeaters," Electrical Communication, (the technical journal of ITT) Vol. 36, No. 3, 1960

IRE ®
50th Anniversary

ITT

Advertising Index

DISPLAY ADVERTISERS

NEW KEITHLEY MILLIOHMMETER

fast, accurate, direct-reading

—permits low resistance measurements from 10 micro-ohms!

Accurate low resistance measurements can now be read directly with a maximum sample dissipation of only 10 microwatts. Exceptionally stable, the Keithley Model 503 requires no balancing—as encountered in Kelvin Bridges—and is designed for rapid measurements. The line-operated 503 supplies an output voltage usable either for chart recording or control functions.

The measurement technique involves an ammeter-voltmeter method using an ac test current. Four terminals are employed, two furnishing a known test current to the sample and two measuring the resultant voltage drop. The voltage is measured by a synchronous ac voltmeter sensitive only to the test current frequency.

The Model 503 lends itself to a wide variety of applications by combining laboratory precision with production line ruggedness. Typical uses include measurements of internal resistance of dry cells, resistivity profiles of thermoelectric materials and low value resistors; measurement of temperatures with thermistors and resistance changes in conductors due to temperature and humidity effects; as well as dry-circuit testing of relay contacts, semi-conductor resistivity measurements, contact resistance of vibrators, relays and choppers, and safe measurement of fuses and squibs.

RANGE: 0.001 to 1000 ohms full scale. The test current, the input voltage drop, and sample power dissipation for full scale readings are given below.

Range Ohms	Applied Current ma, rms	Voltage Drop µv, rms	Maximum Power in Sample Microwatts
0.001	100	100	10
0 003	33	100	3 3
0.010	10	100	1 0
0 030	3.3	100	0 33
0 10	1.0	100	0 10
0 30	0 33	100	0 033
1 0	3 0	3000	9
3.0	1.0	3000	3
10	0.3	3000	0 9
30	0.10	3000	0.3
100	0.03	3000	0 09
300	0.01	3000	0 03
1000	0 003	3000	0 009

ACCURACY: 1% of full scale on all ranges for meter indications. 0.5% of full scale on all ranges at output voltage terminals.

SPEED OF RESPONSE: 0.25 second to 90% full scale on all ranges.

STABILITY: No visible drift after 15 minute warmup.

REPEATABILITY: Within 0.25% of full scale range setting.

OUTPUT CHARACTERISTICS: +100 millivolts dc at full scale, output impedance 800 ohms.

CALIBRATION: Provision for verification and adjustment on front panel.

POWER REQUIREMENT: 105-125 volts, 50-1000 cps, 30 watts. May be wired for 210-250 volt line.

TWO-IN-ONE CONSTRUCTION — A new package design permits choice of bench or rack mounting by means of a conversion kit supplied with each unit at no extra cost.

PRICE:
Model 503 . . . . **$675.00**
Model 503C (Contact Meter Model) **825.00**

send for complete specifications in latest engineering note . . .

KEITHLEY INSTRUMENTS
12415 EUCLID AVENUE CLEVELAND 6, OHIO

electrometers • micro-microammeters • microvoltmeters • power supplies • ac amplifiers

PROFESSIONAL CONSULTING SERVICES

Advertising Rates

Professional Consulting Service Cards

1 insertion $35.00
12 insertions $300.00
50% discount to IRE members on 3 or more consecutive insertions.

In IRE DIRECTORY $35.00 flat rate

New rate effective January 1963
1 Insertion $40.00
12 Insertions $420.00
50% discount to IRE members on 3 or more consecutive insertions

In IRE DIRECTORY $40.00 flat rate

1 inch only, no cuts or display type.

Advertising restricted to professional engineering and consulting services by individuals only. No product may be offered for sale, and firm names may not be mentioned in cards.

Prof. D. S. BUGNOLO, Sc.D.
Consultant
Satellite & Space Communication
Propagation, Antennas, Information Theory
61 Lillian Place Red Bank, N.J.
SHadyside 7-1549

BASIC & EXPERIMENTAL PHYSICS
WILLIAM K. COBURN, P. E.
Studies, proposals, electronic and mechanical R&D, circuit design, prototype fabrication, tests, and reports for Communications, Detection, Instrumentation, and Real-time digital and analog control. Computation, Calibration, Field Engineering, Packaging, Publications Services.
Box 689, Falmouth (Cape Cod) Mass.
AC 617-KI 8-2175

SIDNEY FRANKEL, Ph.D.
& associates
Engineering Research • Technical Consultation
Feasibility Studies
Communication and Control Systems
Radio and Radar Techniques
1165 Saxon Way Menlo Park, Calif.
DAvenport 6-9166

Proposals Research Development
Reports Prototype Units
DR. A. W. FRIEND, P. E.
CONSULTING ENGINEER AND PHYSICIST
5903 City Avenue TEnnyson 9-3098
Philadelphia 31, Pa. MOhawk 4-0275

EARL GREENBERG & associates
Solid State Circuits Laboratory
Transistorized digital/analog calculators
Special devices and controls
Design—Development—Prototype Construction
11-16 154th Street, Whitestone, New York
HIckory 5-2459

LEON HILLMAN
Engineering Staff and Laboratory
• Instrumentation • Control Systems
• Electronic Research and Development
255 County Rd., Tenafly, N.J. LOwell 7-3232

LEONARD R. KAHN
Consultant in Communications and Electronics
Single-Sideband and Frequency-Shift Systems
Diversity Reception - Stereophonic Systems
Television Systems
81 South Bergen Place, Freeport, L.I., N.Y.
FReeport 9-8800

LEONARD J. LYONS
Consulting Mechanical Engineer
COOLING—HEATING—FLUID FLOW
Analysis and Testing
Heat Transfer & Thermodynamic Problems
3710 Crestway Dr. AXminster 1-5446
Los Angeles 43, California

Advertising Index

LEN MAYBERRY
ELECTRONICS CONSULTING ENGINEER
Laboratories and Staff
Transistorized Modular Equipment
Product Design • Test Equipment Design
111 South Oak Street ORegon 8-4847
Inglewood, California

EUGENE MITTELMANN, E.E., Ph.D.
Consulting Engineer, Physicist
ELECTRONICS FOR INDUSTRY
Analysis, Research and Development
549 West Washington Boulevard
CHICAGO 6, ILLINOIS
CEntral 6-2983

Telecommunications Consulting Engineers
"Worldwide Experience"
V. J. Nexon and staff
WIRE, CARRIER, RADIO, MICROWAVE, TROPO SCATTER
Feasibility Studies, Terrain Surveys, System Design, Job Supervision, Government Proposals, Market Research, Product Planning.
U.S. 46 & Cisco, Denville, N.J. OAkwood 7-7400

E. M. OSTLUND & associates
Electronic Engineers
Radio—Microwave—Carrier—
Communication—Control—
Systems and Equipment
Consulting—Research—Development
ANDOVER, NEW JERSEY
Tel: PArkway 9-6635 P.E.N.J.

NATHAN GRIER PARKE
and staff
Consulting Applied Mathematicians
Research • Analysis • Computation
Bedford Road • Carlisle, Massachusetts
Area Code 617 • EMerson 9-3818

MYRON M. ROSENTHAL
& staff
Microwave & Electronic Systems
19 Brookline Dr., North Massapequa, N.Y.
PErshing 5-2501

INDUSTRIAL AND MILITARY ELECTRONICS
EUGENE H. SHEFTELMAN
System & Component Analysis & Design
Radar • Communications • Computers
Advanced Signal Processing Techniques
10 BEVERLEE DR. NASHUA, N.H.
TUXEDO 2-7961

Analysis—Research—Design
Electronics and Automation
GARDNER H. SLOAN
R & D Consultant
Electrical Engineer Physicist
One Story Street Harvard Square
Cambridge 38, Mass. Eliot 4-0766

HAROLD A. WHEELER
Laboratories and Engineering Staff
Consultation—Research—Development
Radar and Communication Antennas
Microwave Assemblies and Components
Main office: Great Neck, N. Y. HUnter 2-7876
Antenna Laboratory: Smithtown, N. Y.

Advertising Index

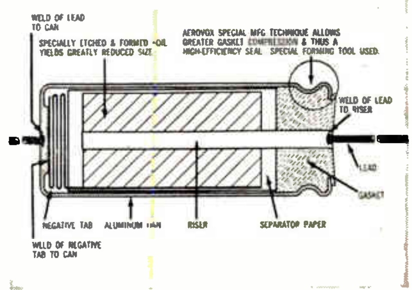
Advertising Index

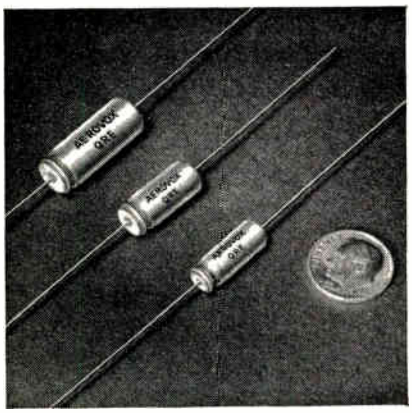
PRODUCT NEWS FROM AEROVOX

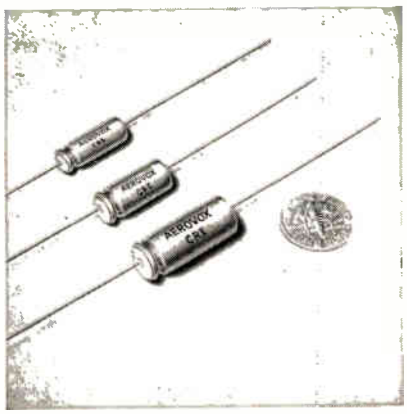
New Long-Life 85°C Electrolytic Capacitors in Ultra-Miniature Sizes a Product of New Design and Manufacturing Techniques

Aluminum-cased units permit circuit applications previously impossible

A new series of ultra-miniature tubular electrolytic capacitors is now available from Aerovox. A product of the continuing Aerovox program of advanced research and development, the greatly reduced sizes of these high-quality units have been made possible by the use of a revolutionary etching and formation process. All critical terminations are welded, thus eliminating the danger of open circuits with the passage of time in service. A unique, high-efficiency seal has been produced by specially designed forming tools which were developed by Aerovox after months of intensive engineering effort to improve on conventional sealing methods.







TYPE QRE SPECIFICATIONS

Operating Temperature: −40°C to +85°C
Capacitance Tolerance: standard capacitance is −10% to +100% of rated capacitance.

DC Leakage Current:

Volts DC	Current – Microamperes
3 to 6	1.0
10 to 15	2.5
25 to 50	5.0
100 to 150	15.0

Surge Voltage:

Rated DC Working Voltage	Surge Voltage (Max.)
3	5
5	8
6	10
10	14
12	15
15	20
25	45
50	70
100	125
150	175

TYPE CRE SPECIFICATIONS

Operating Temperature: −30°C to +85°C
Capacitance Tolerance: standard capacitance is −10% to +100% of rated capacitance.

DC Leakage Current:

Volts DC	Current–Microamperes
3 to 6	1.0
10 to 15	2.5
25 to 50	5.0
100 to 150	15.0

Surge Voltage:

Rated DC Working Voltage	Surge Voltage (Max.)
3	4
6	8
12	15
15	18
25	40
50	65
100	125
150	175

High-Reliability Type QRE
Due to a totally new design concept, Aerovox has achieved a new industry high for capacitors of this type. Type QRE capacitors offer a useful life expectancy of more than 10 years when operated within ratings.
The combination of long life, ultra-miniature size, and outstanding temperature characteristics now makes available an aluminum electrolytic capacitor which can be used in many circuit applications heretofore not considered possible with capacitors of this type. Design engineers in the computer and communications fields in particular can benefit from the extraordinary advantages offered by Type QRE. These units are manufactured in specially constructed super-clean "White Room" production areas where only the most experienced operators are employed. An exhaustive 100% testing program assures you high-reliability performance.

Commercial Type CRE
Type CRE ultra-miniature units are ideally suited for use in bypass, filter, and coupling applications in low voltage, compact, miniaturized equipments. This is especially true where assembly space is at a premium, such as personal radios, hearing aids, microphones, and wire receivers.

Availability
Aerovox Type QRE and CRE Ultra-Miniature 85°C Aluminum Cased Electrolytic Tubular Capacitors are available in prototype quantities for immediate delivery from the factory. See your Aerovox Representative for delivery information on production quantities.

Complete Technical Data
Call your nearest Aerovox Field Representative or write today for a free copy of Bulletin 201B7 (Type QRE) and 201B6 (Type CRE).

AEROVOX CORPORATION
NEW BEDFORD DIVISION — NEW BEDFORD, MASS.

Technical Leadership — Manufacturing Excellence

Advertising Index

exploring moon craters or detecting ocean-bottom seismic whispers!

On the 50th anniversary of the IRE, electronics boundaries extend from outer space to ocean depths. Scientists and engineers exploring these once inaccessible realms require an ever-increasing flow of new materials, products and services. ■ To aid these space-age pioneers, Texas Instruments has developed telemetry and command-control equipment for space and lunar vehicles, nuclear fuel cores for naval reactors, and undersea seismic stations that will function unattended at 1600 fathoms. These are typical accomplishments of TI's unique combination of technologies—ranging from basic materials; through the manufacture of electronic components, assemblies and systems; to the world-wide application of geoscience. ■ Wherever you work in electronics, you can rely on TI products and services. Write today for brochure describing how TI capabilities can help you.

RELY ON TI
MATERIALS
COMPONENTS
ASSEMBLIES
SYSTEMS
SERVICES

TEXAS INSTRUMENTS
INCORPORATED
P. O. BOX 5474 • DALLAS 22, TEXAS
14 PLANTS IN 8 COUNTRIES • OFFICES IN 85 PRINCIPAL CITIES OF THE WORLD