ROBERT E. MATKINS

VOLUME 4 □ NUMBER 7 JULY 1967

IEEE spectrum

features

Departments: *please turn to the next page*

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

departments

the cover

On-line handprinting input transducers are paving the way for greater computer flexibility. The systems discussed in the article beginning on page 72 recognize handprinted alphanumeric characters with a high degree of accuracy. The character extremes shown on the cover were fed into the system to evaluate recognition capability of distorted figures; the tinted areas were not recognized.

Spectral lines

IEEE's educational role. In order to evaluate the effectiveness of the IEEE in fulfilling its constitutional objective of being an "educational organization," the Executive Committee recently appointed an *ad hoc* committee, under the chairmanship of Dr. Bernard M. Oliver, which is studying ways to strengthen our educational activities.

The present activities of the Institute include many that are related to educational institutions or are directed toward educational ends. Examples are:

1. Section sponsorship of lecture series to assist the practicing engineer to orient himself in new fields.

2. Group sponsorship of organized workshops directed toward teaching subjects of emerging importance.

3. Publication of articles of a tutorial nature in technical journals of the Institute.

4. Publication of the STUDENT JOURNAL six times a year for undergraduate Student Members in the United States and Canada. (Students in Regions 8, 9, and 10 and graduate students in Regions 1–7 receive SPECTRUM instead.)

5. Operation of Student Branches at 230 universities and colleges and 59 technical institutes for the approximately 20 000 Student Members of the IEEE.

6. Publication of career guidance pamphlets.

7. Active participation in ECPD, whose major activity is accrediting engineering curricula in U.S. colleges, universities, and technical institutes.

I have omitted from this list the regular publications and numerous conferences and meetings that are the IEEE's major means for disseminating technical information. The term "educational activities" is being used to designate those activities of an instructional rather than an informative nature, activities designed to develop for the participant the background material needed to understand and utilize the new information reported in the regular publications and at meetings and conferences. Thus, educational activities are close in organization and content to a university course, although the format may be different.

Although it would be difficult to draw a sharp line of distinction between IEEE's educational and informative activities—and little reason to do so—the committee's concern has been focused on the educational aspect.

The committee identified three areas in which the Institute might well offer improved services to its members: (1) in continuing education, i.e., programs for the engineer and scientist who wishes to further his education without interrupting his employment; (2) in the provision of greater assistance to both students and faculty of colleges, universities, and technical institutes; and (3) in a more effective program at the secondary school level for informing students of the challenges and excitement of careers in the electrical/electronics area and for providing teachers with material to enrich the curriculum. In each of these areas, the attempt was made to identify services that the Institute is uniquely qualified to provide. It is not the intent to compete with universities or private publishers but to supplement their activities in ways a professional society is specially equipped to do.

The Institute has some unique assets with which it can develop its educational program. In its 196 Sections, it has a geographically widely dispersed organizational structure; in its award structure it has a means of recognizing professional achievement; in its membership it has the leading professionals in their fields. But, most important, there is a tradition of voluntary service that enables the Institute to develop programs by calling on the best talent available regardless of whether the individuals are employed by industry, government, or universities. Thus the task is to determine the educational services that our present and future members need and then to develop useful programs that utilize these important assets.

The committee has discussed a number of possibilities that warrant further study. These include:

1. The publication of monographs and course material in new technological areas as soon as possible after the technical advances are made. There is now a serious time delay between developments in industrial and governmental laboratories and the appearance of this material in such a published form that the nonspecialist can absorb it quickly and efficiently. This possibility is also being studied by the Publications Board.

2. The sponsorship of lecture series recorded on video tape or in slide-tape format for circulation to Sections, Student Branches, or other appropriate units.

3. Development of guides to the literature to assist the neophyte in a specific field. These guides would list various textbooks, review and tutorial articles, films, etc., that are available, and include critical comments.

4. The organization of educational materials in important new technological fields into ordered sequences that would include problem sets and some sort of examination system. An individual would undertake a systematic program and upon its completion receive an appropriate certificate. An accrediting agency such as ECPD in the United States, might certify the contents of the program so that a man completing it is given individual recognition that has meaning to employers and possibly universities.

In order to develop and support a program including ideas of this type, it may well be necessary to make some organizational rearrangements within the Institute. This question is also under study by the *ad hoc* committee.

An *ad hoc* committee study cannot hope to develop in detail a complete program for as complex an organization as the IEEE. However, it is clear that the Institute has significant assets for service in the educational area. I hope we are wise enough to exploit these potentialities effectively.

F. Karl Willenbrock

Authors

Commercial satellite communications experience (page 63)

E. J. Martin (M) is presently serving as manager of the Advanced Systems Analysis Department for Communications Satellite Corporation (Comsat), Washington, D.C., where he has responsibility for the generation, analysis, and evaluation of concepts for new applications of satellite communications.

He was graduated from Fordham University, Bronx, N.Y., with an A.B. degree in mathematics in 1954 and, three years later, he received the Exceptional Service Award of the U.S. Air Force. Mr. Martin received the M.S. degree in mathematics and physics from Northeastern University in 1959. He was with the Air Force Cambridge Research Laboratories and was a project officer in the Communications Satellite Project Office of the Defense Communications Agency prior to joining Comsat in 1964.

He has published many papers on radio propagation and communications.

W. S. McKee, who is a member of the technical staff of the Space Segment Implementation Division of Communications Satellite Corporation, received the B.S. degree in electrical engineering in 1958 from the University of Maryland and attended graduate school at George Washington University. Between 1959 and 1964 he was a technical representative for the Hughes Aircraft Company. In the following year, while serving as a senior project engineer with the Western Union Telegraph Company, he was responsible for the implementation of portions of the Autodin Overseas program, including contract monitoring and preparation of manning documents. He is presently engaged in monitoring the orbital position and condition of Early Bird, and in the coordination of satellite launch and positioning operations for the Intelsat II program.

The human side of engineering (page 70)

E. M. Turner (SM) presently serves as technical manager of the Antenna Radome Group in the Research and Technology Division at Wright-Patterson Air Force Base, Ohio. He received the E.E. and M.S.E.E. degrees in 1934 and 1948, respectively, from the University of Cincinnati. Between 1936 and 1941 he was a power plant engineer with the Appalachian Electric Power Company. For the next five years he served the U.S. Army as a radar officer and, subsequently, he joined the Erie Resistor Corporation as a development engineer in ceramics and plastics. In 1948 he joined the Wright-Patterson AFB as a project engineer in electronic systems, assuming his present post 11 years later.

Mr. Turner is the inventor of the spiral slot and the scimitar antennas, the antennaverter, the space-frequency filter circuit, and the radant supergain structure. In addition, he has made substantial contributions in the areas of pulse analyzers, panoramic receivers, low-frequency direction finding, unfurable, and electrically small antennas, and the polarization control of antenna systems.

Handprinting input device for computer systems (page 72)

J. G. Simek (M) is on the staff of the Systems Development Division of the International Business Machines Corporation, Endicott, N.Y. He received the bachelor of science degree in 1957 and the master of science degree in 1963, both in electrical engineering, and both from Cornell University, Ithaca, N.Y. While an undergraduate at Cornell, he was elected to Eta Kappa Nu.

Mr. Simek joined the International Business Machines Corporation in 1957 and, until he returned to Cornell under an IBM Graduate Study Grant, he was engaged in the design of various transistor and magnetic circuits for use in data collection and communications systems. Since 1963 he has been involved with studies in handprinting recognition; he is presently the manager of a group responsible for the design of special small-signal circuits.

C. J. Tunis (M), who is spending the academic year as visiting professor of electrical engineering at Stanford University, Stanford, Calif., was serving as a senior engineer at the International Business Machines Corporation Product Development Laboratory, Endicott, N.Y., during the time in which he was involved in the writing of the article that appears in this issue.

Dr. Tunis received the B.Eng. (Engineering Physics) and M.Sc. (Physics) degrees from McGill University, Montreal, Canada, and the Ph.D. degree in electrical engineering from the University of Manchester, England, in 1958.

He joined International Business Machines in Endicott in that year, and his most recent assignment was as manager of Advanced Electrical Technology. His work has been in the fields of digital subsystem design, adaptive systems, and pattern recognition.

New horizons in quantum electronics (page 82)

N. Bloembergen (F), Gordon McKay Professor of Applied Physics at Harvard University, received the Ph.D. degree from the University of Leiden in 1948. For the next three years he was a member of the Harvard Society of Fellows. He was an associate professor of applied physics before assuming his present post in 1957. He also was a visiting professor at the University of California, Berkeley, and at Ecole Normale Superieure, Paris, France. The latter post was held in 1957 as a Guggenheim Fellow.

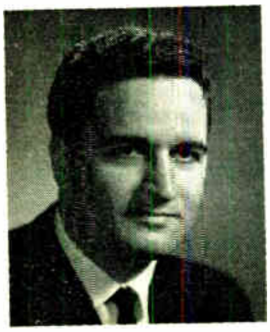
He is a fellow of the American Academy of Arts and Sciences and of the American and Dutch Physical Societies, and he has been an associate editor of the *Physical Review*, the *Journal of Chemical Physics*, and the *Journal of Applied Physics*, and of the IEEE Journal of Quantum Electronics. He is a recipient of the Morris Liebmann Award.

Cryogenic random-access memories (page 91)

A. R. Sass is group leader of cryoelectric research at the Computer Research Laboratory, David Sarnoff Research Center, RCA, Princeton, N.J. He received the B.S.E.E. degree from the Massachusetts Institute of Technology in 1958, and the M.S.E.E. and doctoral degrees from Purdue University in 1960 and 1962, respectively. During the summers between 1959 and 1961 he was engaged in research on cryoelectric memories and cryoelectric logic components. Between 1960 and 1962 he served as an instructor of electrical engineering at Purdue, where he worked on cryoelectric computer components. For the next two years, while stationed at the U.S. Army Lewis Research Center, he also investigated cryoelectric devices. Dr. Sass joined the Computer Research Laboratory in 1964, and assumed his present post in 1966. He has been researching the feasibility of cryoelectric memories.

W. C. Stewart (M), who is with the RCA David Sarnoff Research Center, was graduated magna cum laude with departmental honors from Duke University, where he was awarded the B.S.E.E., M.S., and Ph.D. degrees in 1958, 1961, and 1964, respectively. Prior to his graduation, he worked on magnetic amplifier applications for Sperry Rand Corporation and with the General Electric Company, and on research instrumentation in collaboration with the Duke Parapsychology Laboratory. As a staff member of the school's Department of Electrical Engineering, he taught field theory and did research on superconducting circuits, the subject of his graduate thesis. He was also a consultant on a thin-film capacitor program with the Research Triangle Institute. Dr. Stewart joined the RCA Laboratories in 1964, where his primary interest continues to be in research on superconducting devices.

He is a member of Sigma Xi, Tau Beta Pi, Eta Kappa Nu, and Phi Beta Kappa.



L. S. Cosentino (M) is a member of the David Sarnoff Research Center, Radio Corporation of America, Princeton, N.J. He was graduated from the City College of New York with a bachelor of electrical engineering degree in 1960 and, two years later, he received the master of science degree from Princeton University.

He has been a member of the technical staff of the David Sarnoff Research Center since 1960. His past endeavors have included work on the design and development of tunnel diode circuitry for ultrahigh-speed computers. For the last five years, he has been engaged in research in the cryoelectrics area, including switching devices, memories, and associated electronics.

Mr. Cosentino is a member of Eta Kappa Nu and the American Association for the Advancement of Science.

Diagnostic engineering (page 99)

John Dent is manager of diagnostic engineering in the Systems Development Division of IBM Corporation, Kingston, N.Y. He was graduated from the University of Massachusetts in 1953 with a degree in mathematics, and he then served in the U.S. Air Force as a special weapons officer in the Strategic Air Command. He joined IBM in 1957, and was engaged in developing programs for the SAGE Air Defense System. Between 1960 and 1962 he served as project engineer in charge of the diagnostic engineering effort for the IBM STRETCH system. He joined the Data Processing Division in 1962, assuming responsibility for defining system diagnostics for special configurations and applications. Two years later, as a member of the System Development Division, Mr. Dent organized the diagnostic effort for IBM's compatibility features (Emulators) on the System/360, models 40, 50, and 65. His present responsibilities include unit diagnostics and system diagnostics for the System/360, models 65, 67 (Time Sharing) 75, and IBM's image processing products and custom systems diagnostics. He is a member of the Association for Computing Machinery.

Approaching nuclear power (page 105)

W. A. Chittenden, a partner in Sargent & Lundy, Engineers, Chicago, Ill., and head of the Nuclear Division, received the bachelor of science degree from the University of Illinois in 1950 and, two years later, joined Sargent & Lundy.

First engaged in the mechanical-engineering design of a number of steam-electric generating projects, since the formation of the Nuclear Engineering Division in 1955 he has been associated with most of the firm's activities relating to the application of nuclear power. He served as a project engineer for various engineering design assignments of the firm, and he also directed work associated with a number of economic and feasibility studies relating to the development and application of various reactor concepts for several national laboratories, reactor manufacturers, and utility companies.

Mr. Chittenden is a member of the American Society of Mechanical Engineers, the American Nuclear Society, and the Western Society of Engineers. He is a registered professional engineer in several states, served in various offices of the Chicago Section of ANS, and is currently serving as vice chairman of the U.S.A. Standards Institute Committee on Nuclear Piping.

Albert Nathan joined Sargent & Lundy, Engineers, Chicago, Ill., in 1950 and, since that time, has been involved in a number of conventional fossil-fueled, and nuclear power plant projects. His work since 1959 has been concentrated in electrical auxiliary systems and control and instrumentation systems for nuclear power plants. He has written and presented papers on various aspects of nuclear plant design and development to client conferences and engineering society meetings.

Mr. Nathan has, in addition, worked on numerous feasibility and cost evaluation studies in connection with system expansion investigations for various utilities. His nuclear power plant experience includes assignments on the La Crosse Boiling Water Reactor Power Plant and the Southwestern Experimental Fast Oxide Reactor Facility (SEFOR). Presently, he is assigned to the Fort St. Vrain Nuclear Power Plant Project, a high-temperature gas-cooled reactor plant being built by General Atomic for the Public Service Company of Colorado.

Mr. Nathan received a degree in electrical engineering from the Illinois Institute of Technology, and he is also a graduate of the Signal Officers Career Course of the U.S. Army Signal School. He has taught communications principles and applications at the State of Illinois Officer Candidate School.

Commercial satellite communications experience

Early Bird and successive communications satellites are paving the way for many forms of instant, worldwide communications. They also will play an important role in the success of the Apollo project

E. J. Martin, W. S. McKee

Communications Satellite Corporation

The operational Intelsat I (Early Bird) and Intelsat II satellites and associated earth stations that comprise the present commercial satellite communications system provide a capacity of close to 720 voice circuits between major earth stations. All but a limited portion of the inhabited globe is now within the line of sight of a commercial satellite relay. In addition, it is expected that by 1968, with the operation of Intelsat III, a truly global international satellite communications system will exist.

Early Bird, the first commercial communications satellite, was boosted into orbit on April 6, 1965, by a thrust-augmented Delta vehicle, heralding a unique venture in international cooperation in space. Recently a second generation of satellites, advanced versions of Early Bird, have been deployed over the Pacific and Atlantic Oceans. These new satellites, known as Intelsat II, will provide satellite communications from the United States to the west as far as Bangkok, Thailand, and to the east as far as the Indian Ocean. In addition, third-generation high-capacity satellites (Intelsat III) are under development for deployment in 1968, by which time a truly global international satellite communications system will have become a reality.

The experience that has been obtained to date with the Early Bird satellite and the second-generation Intelsat II already positioned over the Atlantic and Pacific Oceans are summarized in this article. The major elements of the communications system are depicted in Fig. 1; the system is composed of the Early Bird and Pacific and Atlantic Intelsat II, along with their associated earth stations.

Overall system configuration

Early Bird is positioned in a nearly stationary orbit over the Atlantic to provide communication between a U.S. earth station at Andover, Maine, and four European earth stations. The Andover station, originally built by AT&T for Telstar, was modified for Early Bird and leased to Comsat for transatlantic service. The station was subsequently purchased by Comsat and will be one of six U.S. earth stations owned and operated by a consortium of U.S. communications carriers. A Canadian station at Mill Village, Nova Scotia, has also served as the North American terminus about one day per week since late 1966. It has, in addition, carried Early Bird traffic during the Intelsat II launches to release the Andover station for launch support operations.

Three European earth stations at Goonhilly Downs (England), Pleumeur Boudou (France), and Raisting (Germany), all of the same performance class as Andover, serve alternately in the roles of operating station and standby. They are connected by microwave links and submarine cables, which permit all European traffic to be carried by any one of the stations. A fourth smaller station at Fucino, Italy, also participates, acting as a

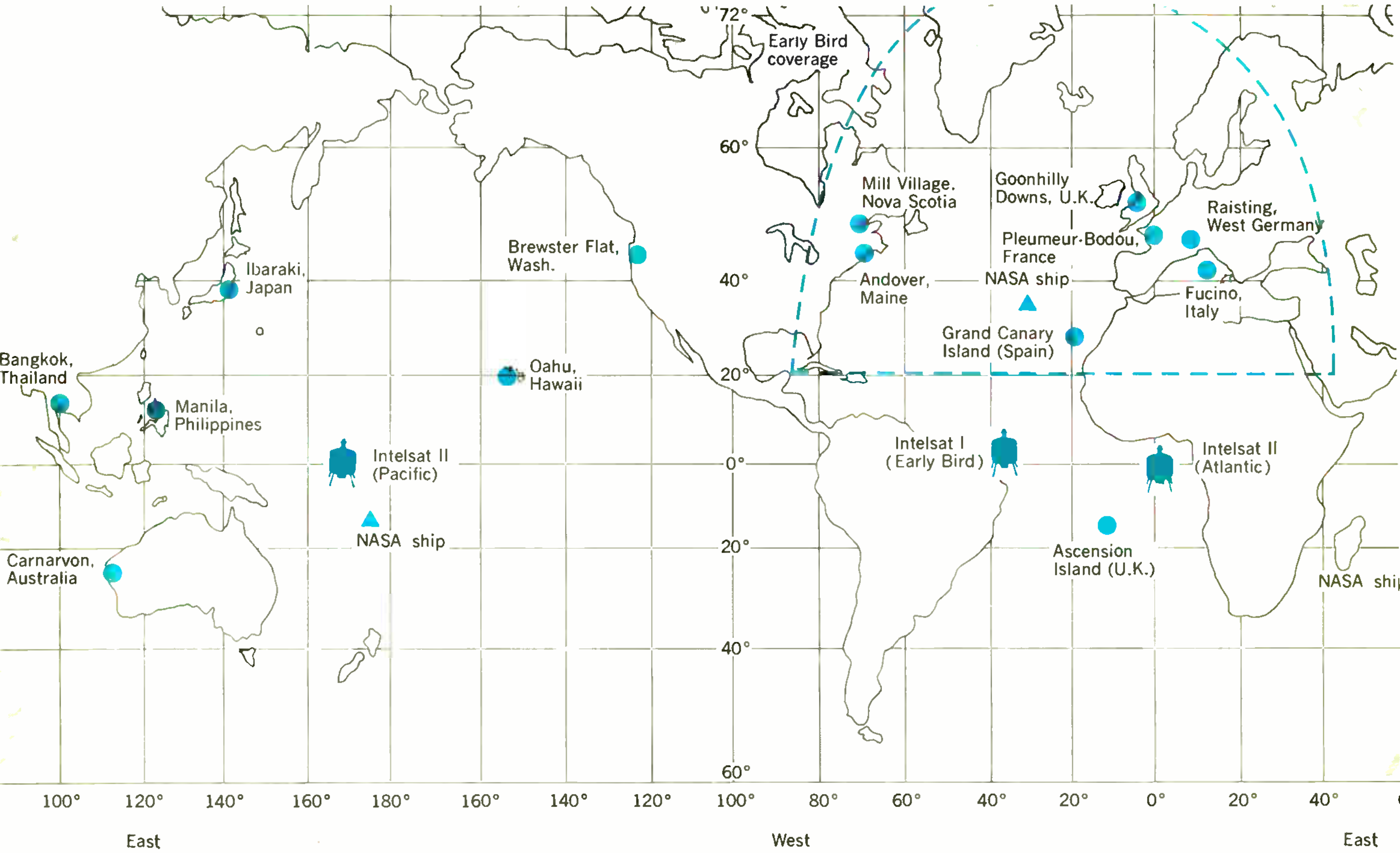


FIGURE 1. Intelsat I and II communications systems, including satellites and associated earth stations.

European terminus for weekend traffic. A larger facility under construction at Fucino should begin full-time operation this summer with the Atlantic Intelsat II satellite.

The first Intelsat II (F-1), launched on October 26, 1966, was intended for a "stationary" synchronous orbit position over the Pacific. Because this satellite failed to achieve stationary orbit, a second Intelsat II (F-2) was launched on January 11, 1967, and was successfully positioned in a stationary orbit to provide multiple-access communications in the Pacific area.

Voice and teletypewriter communications will be provided between one of the U.S. gateway stations in Washington and Hawaii and two NASA Apollo tracking stations: one in Australia and one on board a ship in the Pacific. The balance of the satellite capacity will be used to support communications to stations in Japan, Thailand, and the Philippines.

The third Intelsat II (F-3) was successfully launched on March 22, 1967, to provide multiple-access communications between a second earth station at Andover and NASA Apollo support stations on Ascension Island, the Canary Islands, and two ships located in the Atlantic and Indian Oceans.

The station-keeping constraints imposed by the system elements for the three satellites make an interesting comparison. Early Bird is maintained between longitude limits of 25°W and 40°W, sufficient to ensure higher-than-minimum antenna elevation angles for the participating earth stations. As a result, station-keeping maneuvers for Early Bird are necessary only at eight- to nine-month intervals. Stations communicating with the Pacific Intelsat II are at much greater distances from the subsatellite point, the two furthermost being Brewster Flat, Wash., and Bangkok, Thailand. The small mutual region of earth station coverage, coupled with the satellite's orbital inclination of 2.0°, requires that the satellite be maintained between longitude limits of 174°E and 176°E to ensure that antenna elevation angles for these two stations do not become less than 5°. Station-keeping maneuvers for this satellite will be required approximately every two months. Allowable longitude drift of the Atlantic Intelsat II will be less than that allowed for Early Bird, because coverage will be required between Andover and a ship in the Indian Ocean. The orbital inclination achieved with the Atlantic Intelsat II was 1.5°. Consideration of the orbital inclination and the mutual coverage area from the Andover and Indian Ocean ship stations results in allowable longitude limits for station keeping of 1.5°W to 11.5°W. Station-keeping maneuvers will be performed at correspondingly shorter intervals than is required for Early Bird.

Comsat, as system manager, maintains a communications and operations center at its headquarters in Washington, D.C., which serves as a focal point for directing system operations. Maneuvers associated with launch and station keeping are directed from this center with actual commands sent from the Andover station for

I. Design parameters for Intelsat I and II

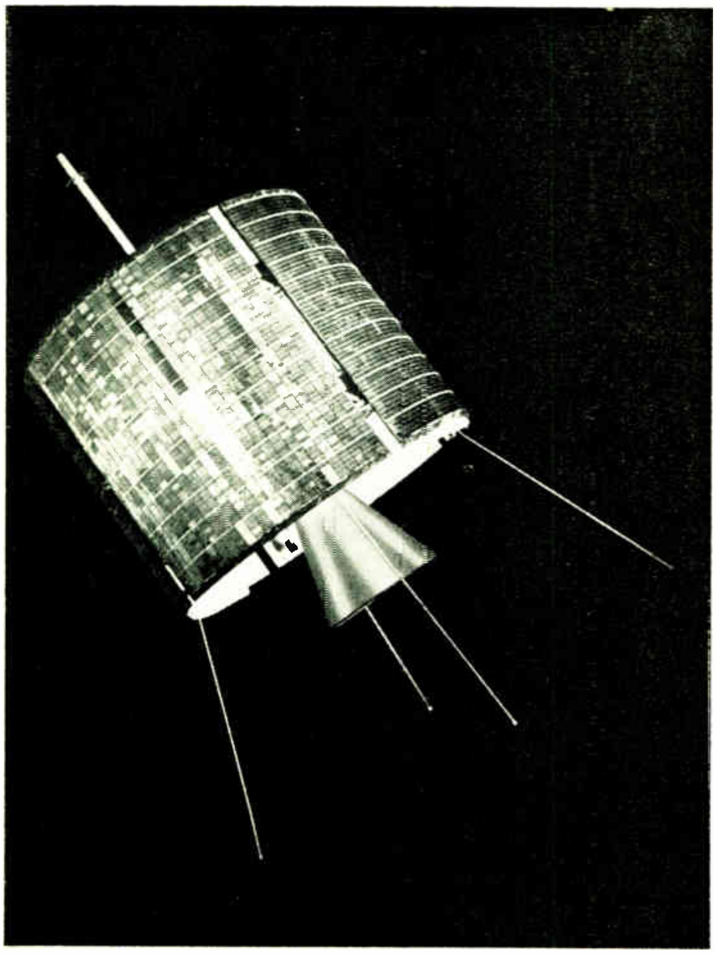
Parameter	Intelsat I (Early Bird)	Intelsat II
Size		
Solar array diameter	72 cm	142 cm
Solar array length	59 cm	67 cm
Weight		
At launch	68 kg	175 kg
Postapogee motor fire	38.5 kg	87 kg
Communications		
Repeater	Two independent frequency translation repeaters, each with 25-MHz bandwidth	Two redundant linear repeaters, each 126-MHz bandwidth
Power amplifier	Two 6-watt TWTs, one operating and one spare	Four 6-watt TWTs; one, two, three, or four operating at one time
Antenna	Receive, omni; transmit, collinear slot array; 11° beam centered at +7° aspect angle	Receive, omni; transmit, multiple-element biconical horn; 12° beam centered at equator
Frequencies		
Up-link	6301 ± 13 MHz, 6390 ± 13 MHz	6345 ± 63 MHz
Down-link	4081 ± 13 MHz, 4161 ± 13 MHz	4120 ± 63 MHz
Telemetry		
Transmitters	Eitner of two microwave beacons or two 1.8-watt 136-MHz VHF transmitters	Same
Data	Two reaction controls for system pressures, bus voltage, two battery voltages, temperature, spin rate, sun angle	Same
Command	One decoder for each receiver, 12 commands	One decoder for each receiver, 15 commands
Electric power		
Solar array	Predicted solar array power capability of 33 watts at 23.5° sun angle after three years in orbit	Predicted solar array capability of 75 watts at 23.5° sun angle after five years in orbit
Batteries	Two nickel-cadmium batteries with total of 1.5-ampere-hour capacity; no eclipse operation	Two nickel-cadmium batteries with total of 9-ampere-hour capacity; eclipse operation capability
Control system		
Configuration	Two independent H_2O_2 fueled systems, one radial and one axial jet each system, two fuel storage tanks	Same except four fuel storage tanks
Capability	Total velocity increment of 183 m/s, predicted station-keeping lifetime of three years	Same velocity increment, designed for station-keeping lifetime of five years
Apogee motor	Velocity increment of 1405 m/s	Velocity increment of 1790 m/s

Atlantic satellites and the Paumalu, Hawaii, station for the Pacific satellites.

Satellite design parameters

Syncom technology provided the basis for the Intelsat I (Early Bird) design. The Intelsat II design evolved from that of Early Bird, although it is approximately twice as large to allow use of a wide-band repeater with much higher output power and multiple access capability. Principal design parameters are summarized in Table I. The Intelsat I and II (Figs. 2 and 3) are spin-stabilized cylinders in configuration with an integral solid-propellant motor for injection from an elliptical transfer orbit into a nearly synchronous orbit. Both satellites use a body-mounted silicon n-on-p solar cell array for electric power in sunlight and two nickel-cadmium battery packs for operating power during launch and eclipse periods, as well as power for apogee motor squib-firing. The solar array also supplies the electric energy required to charge the batteries. In sunlight, the Intelsat I and II solar arrays supply approximately 40 and 90 watts respectively.

The Intelsat I batteries cannot support operation of communications circuits during the two 6-week periods each year in which solar eclipses may be encountered daily by synchronous satellites near the equatorial plane. During the eclipse season, the Early Bird repeater must be turned off at the beginning and end of each daily eclipse; the longest daily eclipse is approximately 70 minutes. The Intelsat II batteries are designed to support communications through the daily eclipse periods for true 24-hour coverage year round. Both satellites contain two completely redundant reaction control systems for orientation and station-keeping maneuvers. They are fueled with H_2O_2 under pressure and provide thrust vectors along and normal to the spin axis. The primary difference in the control systems is that the Intelsat II contains approximately double the fuel storage capacity of Intelsat I in order to compensate for the additional spacecraft mass.



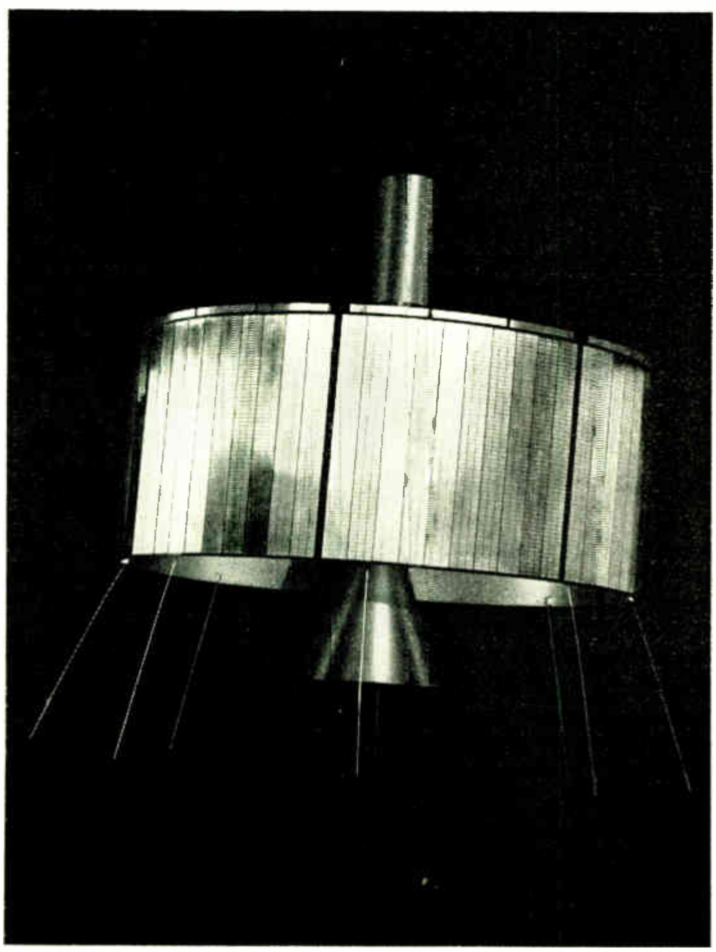


FIGURE 2 (left). Intelsat I (Early Bird) spacecraft.

FIGURE 3 (lower left). Intelsat II spacecraft.

The major difference between the two satellites is in the communications system. Both satellites contain repeaters that translate 6-GHz receive frequencies to 4-GHz transmit frequencies, but the repeaters differ greatly in operation and capability. Intelsat I uses two repeaters performing the functions of IF amplification and frequency translation, the bandwidth of each being 25 MHz. The repeaters drive one of two redundant traveling-wave-tube (TWT) power amplifiers. The microwave antenna consists of multiple-element collinear arrays. The receive portion has an earth-coverage omnipattern and the transmit portion has an 11° beam width at the half-power points with the beam center squinted 7° into the Northern Hemisphere to maximize satellite ERP in the direction of Europe and North America.

The Intelsat II communications system uses two redundant linear repeaters of the RF translation type, each having a bandwidth of 126 MHz. The repeater drives any combination of four output TWTs, although power limitations will restrict operation to a maximum of three. The microwave antenna is a multiple-element biconical horn; as in Intelsat I, the receive portion has an earth coverage omnipattern. The transmit portion has a 12° beam centered at the equator to handle traffic simultaneously in both the Northern and Southern Hemispheres. Both satellites provide a capacity of approximately 240 two-way voice circuits between major earth stations, but Early Bird is limited to use by two earth stations within the Northern Hemisphere coverage zone. Use of the Intelsat I type of squinted-beam antenna on the Intelsat II would double its communications capacity; however, its communications coverage area would be limited to that of Early Bird.

Similar telemetry subsystems complete the major spacecraft elements of both satellite designs. Both use redundant encoders that modulate VHF and microwave beacons. The VHF beacons are used primarily as a tracking aid during the launch operation and are not normally required after the spacecraft has been positioned in synchronous orbit. Telemetry is usually obtained by monitoring the microwave beacons.

Launch and station keeping

Launch operation and station-keeping maneuvers were conducted in generally the same manner for both Intelsat I and II. A brief description of the Intelsat I Early Bird launch and orbit history illustrates the mission objectives of both satellites.

Early Bird was thrust into a transfer orbit with an 11-hour period and 36 520-km apogee on April 6, 1965 (see Fig. 4). An attitude reorientation was performed at second apogee and a velocity increment was added with the reaction control system at fourth apogee. This permitted the subsequent achievement of a nearly stationary orbit over the Atlantic with the impulse from the apogee motor which was fired at sixth apogee. A final attitude reorientation

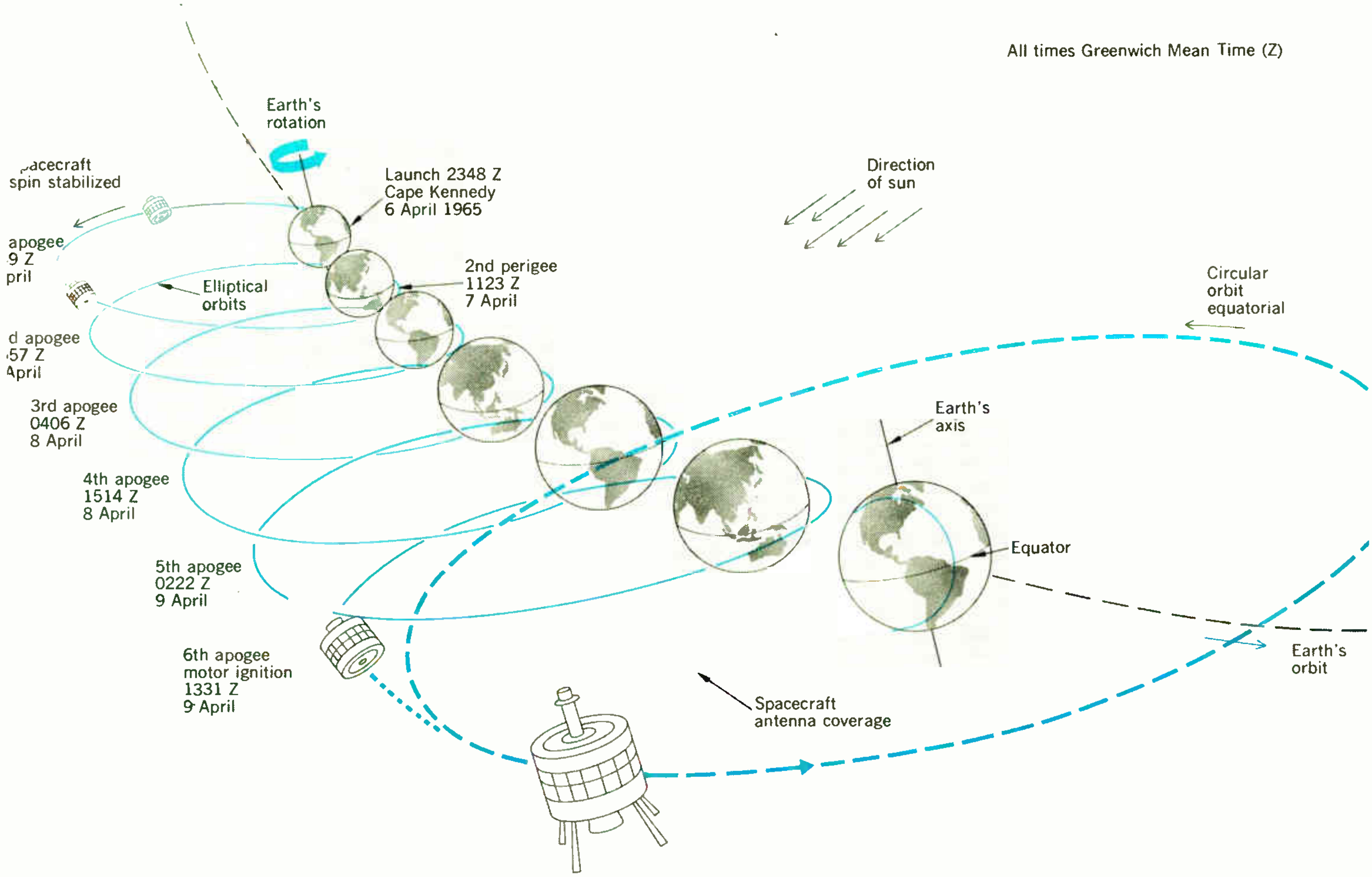


FIGURE 4. Orbit sequence of Intelsat I (Early Bird) satellite.

was performed and transatlantic operations begun approximately three days after launch. A few days later, velocity corrections circularized the orbit; at this point the satellite was drifting westward at a $\frac{1}{4}$°-per-day rate. On April 23, a 0.9-m/s velocity correction resulted in an eastward drift rate of 0.05° per day. Calculations, subsequently confirmed, indicated that earth triaxiality forces would reduce that drift to zero near 28°W longitude and the satellite would then begin drifting in the opposite direction. Only two station-keeping maneuvers have been required since that time to maintain Early Bird in a region for communication between North America and European earth stations.

The first Intelsat II intended for Pacific service was launched on October 26, 1966. The apogee motor underperformed, burning for approximately one quarter of the planned 16 seconds. Therefore, the satellite was left in a modified transfer orbit with an inclination of 17° and a period of 12 hours, well outside the range of any possible correction to stationary orbit by the on-board reaction control system. Further calculations and tests showed that during transfer orbit the apogee motor had been subjected to environmental temperatures lower than it could withstand. This is believed to have resulted in a contraction of internal insulation material, which allowed hot gases to reach the apogee motor case. An apparent separation of the nozzle from the motor case caused a reduction in chamber pressures, and propellant combustion ceased as a result of the pressure drop. The orbit was modified slightly through use of the on-board reaction control system to provide part-time communications coverage in the Pacific area until the second Intelsat II satellite could be launched for its intended full-time Pacific service.

The apogee motor problem was remedied for the second Intelsat II satellite and it was launched on January 11, 1967. The remedy was through three approaches: (1) The launch window was chosen such that the sun would always illuminate the apogee motor during transfer orbit. (2) A thermal shield was installed over the previously exposed apogee motor case. (3) A small electric heater was installed at the apogee motor nozzle boss to be turned on if telemetry indicated that the nozzle temperature was approaching predetermined low levels. Apogee motor fire and subsequent maneuvers performed with the reaction control system were normal. The satellite is now positioned over the Pacific in a stationary orbit.

Perturbations of the sun, moon, and unsymmetrical earth combine to tilt the orbit plane of any satellite slowly with respect to the earth's equatorial plane. The initial rate of inclination for an equatorial synchronous orbit has been found to be approximately 0.9° per year. For this reason the F-2 attitude was intentionally adjusted so that the apogee motor thrust not only would circularize the orbit but also would result in an orbital inclination biased away from equatorial. The sun–moon–earth perturbations will decrease the 2.0° inclination obtained due to apogee motor thrust, through zero degrees and back to a magnitude of 2.0° in approximately four years. This means that the daily latitude change will not exceed ±2.0° for the first four years of operation, assuring 24-hour coverage for the long communications link between Brewster Flat and Bangkok during that period.

The third Intelsat II satellite was launched on March 22, 1967, for positioning over the Atlantic. Inclination of the F-3 orbit was biased in the same manner to provide maximum time near the equatorial plane; the actual inclination achieved was 1.5°.

Spacecraft performance

Although the performance of Early Bird's subsystems has been generally as anticipated, some in-orbit difficulties have been experienced. The pressure buildup of peroxide system no. 1 has leveled off at the designed relief-valve crack pressure, but system no. 2 has lost pressure during each of the three eclipse seasons so far encountered. Between eclipse seasons, when average spacecraft temperature is higher, the pressure in system no. 2 rebuilds to pre-eclipse levels.

Bus voltage has been lower than originally expected. Laboratory tests showed that low bus voltage was probably the cause of a deeply discharged state of the batteries discovered during an attempt on June 11, 1965, to operate both VHF transmitters with the microwave repeater. Since August 1965, VHF transmission has been limited to short periods for special purposes only and bus and battery voltages have maintained acceptable levels. Periodic degassing of the spare TWT had been planned by occasional operation in place of the primary tube. Two attempts to operate the spare TWT in mid-October 1965 resulted in loss of carrier after a short operating time. A third attempt produced normal operation. No explanation is available for the temporary failure encountered. The primary tube has accumulated more than 17 000 hours of use, with no signs of performance degradation. It should be noted that all difficulties experienced have been in redundant or noncritical systems. Furthermore, several hundred commands have been issued to Early Bird and there have been no difficulties attributable to spacecraft operation.

The highly inclined (17°) orbit of Intelsat II F-1 has resulted in its being subjected to a more hostile environment than it was designed to withstand. Sun angles have been much lower than planned and temperature has ranged from unusually high during noneclipse seasons to unusually low during eclipse seasons. In addition, the spacecraft is subjected to two eclipses per day during the six-week eclipse season instead of the one per day experienced by stationary satellites. These conditions result in low solar array efficiency and restricted operation; at present, only one of the four TWTs can be operated continuously. During the period of part-time Pacific service, a second tube was turned on for eight- to ten-hour communication periods and turned off again at the end of each period to prevent battery discharge to levels that would not sustain normal operation.

The high temperatures have also resulted in a very high hydrogen peroxide decomposition rate and attendant higher pressure buildup rates. Control system lifetime has been severely reduced due to the rapid rate of fuel decomposition. Orbit and environment have reduced the usefulness and predicted life of F-1 to the point that its role in the global system does not extend beyond that of the partial communications service it provided in the Pacific prior to the F-2 launch.

The Intelsat II F-2 has been positioned on station over the Pacific and is performing as predicted. Results of in-orbit communications tests correlated very well with those of prelaunch tests. All systems are functioning normally and the satellite is in commercial operation with three TWTs operating on a 24-hour-per-day basis.

The Intelsat II F-3 is presently in circular orbit over the Atlantic. The on-board reaction control system was used to impart a small longitudinal drift to move the satellite to a stationary operating position. During the drift period, tests of the control system, power system, and telemetry and command systems, as well as the communications system, indicated satisfactory operation of all satellite functions. The satellite went into commercial operation on April 7, 1967, upon completion of system performance tests (station to station). Orbital parameters are listed in Table II for all four of the Intelsat I and II satellites now in orbit.

Communications history

The establishment of reliable commercial satellite communications service requires more than the successful

II. Comparison of orbital parameters

	Intelsat I (F-1)	Intelsat II (F-1)	Intelsat II (F-2)	Intelsat II (F-3)
Coverage area	Atlantic	Atl./Pac.	Pacific	Atlantic
Orbit date	3/7/67	3/24/67	4/2/67	4/6/67
Subsatellite longitude	29.2°W	...	174.1°E	17.0°W
Longitude drift per day	0.014°W	...	0.018°E	0.747°E
Inclination of orbit plane	1.86°	17.06°	1.97°	1.37°
Declination	−89.8°	−78.1°	−87.4°	−88.7°
Apogee altitude, km	35 800	37 070	35 790	35 780
Perigee altitude, km	35 770	3 280	35 780	35 670
Period of orbit, minutes	1436.23	717.90	1436.10	1433.19

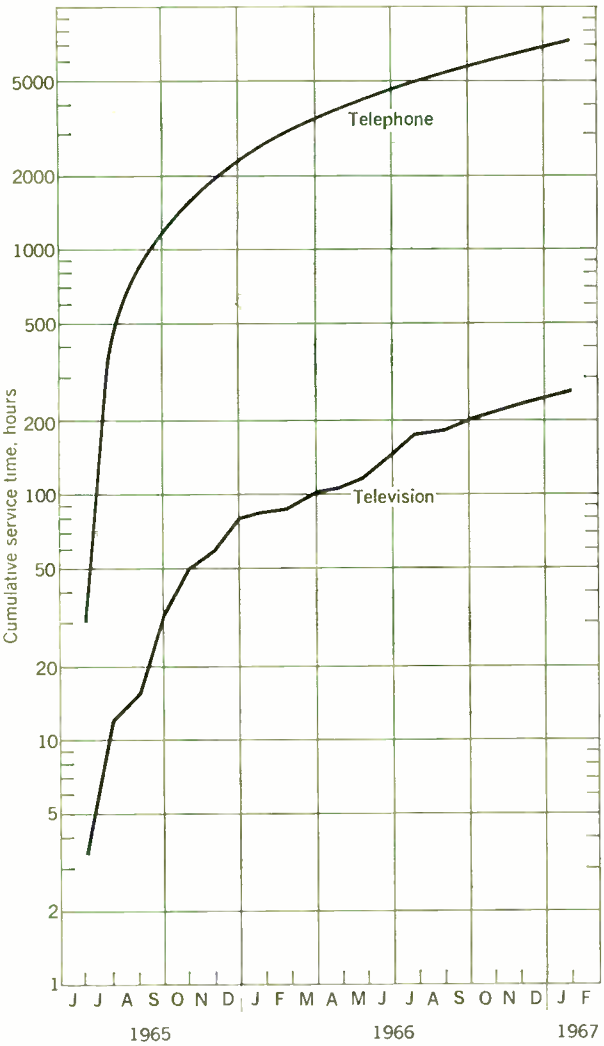


FIGURE 5. Cumulative service time for Early Bird.

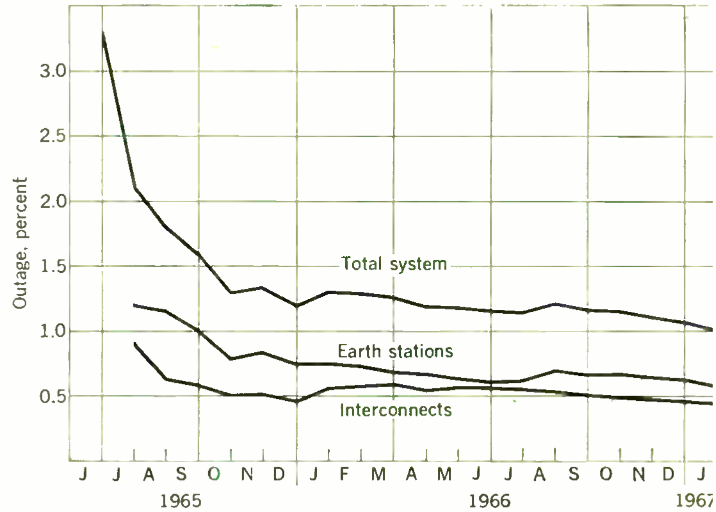


FIGURE 6. Cumulative outage performance for Early Bird.

placement of functioning satellites at their operating stations in space. Under Comsat tariff arrangements, the first commercial operations were initiated in June 1965 for 12-hour-per-day transatlantic service using the Intelsat I F-1 (Early Bird). This system provides a capacity of 240 voice circuits or one television channel using equivalent 25.8-meter-diameter antenna earth stations. Despite the fact that these facilities have been used well below capacity, this early experience can be described as an unqualified commercial success. A variety of experiments have been performed in connection with the Early Bird system, including such noteworthy firsts as computer-to-computer exchange and live television coverage of Gemini recoveries at sea.

Originally intended as an 18-month operational experiment to establish the suitability of synchronous orbit relays for commercial voice traffic, the Early Bird system attained its second anniversary of commercial service on April 6, 1967. Figure 5 summarizes commercial usage for telephone and television traffic through January of this year. The relatively small percentage of television usage may be attributed, in part, to the fact that voice service must be interrupted in order for a television signal to be transmitted. Limitations on simultaneous voice and television transmission are expected to be less severe with Intelsat II.

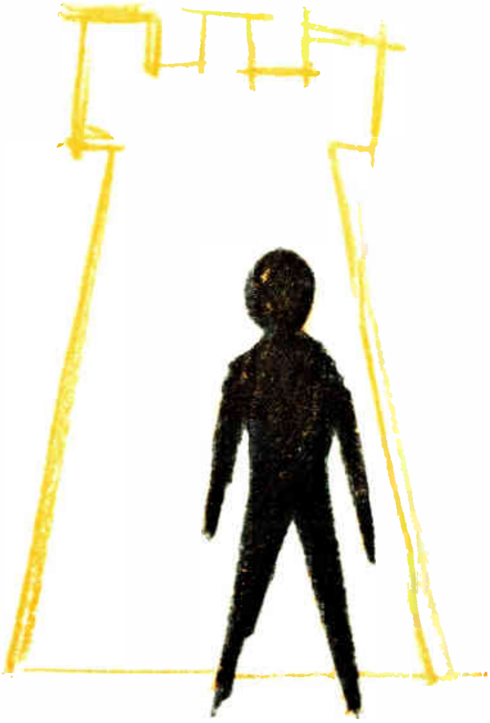
As has been noted, the satellite contribution to out-of-service time has been zero in the Early Bird system. All loss of service has been attributable to the seven major earth segment elements—the five earth stations and the U.S. and European interconnects. The reliability history of this service is shown through early 1967 in Fig. 6 in terms of cumulative percent outage time. The effect of operational experience on improving overall service reliability is particularly evident during the early months of operation.

The second phase of the global system establishment has begun by virtue of the recent initiation of commercial services with the Intelsat II series of spacecraft. Although the Intelsat II F-1 failed to achieve stationary orbit, the successful functioning of the communications subsystem permitted an 8-hour-per-day voice service to be established between the earth stations in Hawaii and Brewster Flat. The highly elliptical, inclined orbit reduced the commercial utility of the F-1 satellite to negligible proportions, but the experience with its functioning subsystems proved to be invaluable in preparing for operations with its successors.

Commercial multiple-access service with the Intelsat II F-2 was initiated on January 27, 1967, when the first full-time transpacific links in the evolving global network were established. Although it is too early to report on significant communications operational experience, spacecraft performance to date indicates that the first truly 24-hour-per-day satellite communications capability is now in existence.

The Intelsat II F-3, launched on March 22, 1967, had successfully completed in-orbit tests and station-to-station tests on April 7, 1967, in preparation for commercial operation. As of that point in time, only a limited portion of the inhabited globe is not within the line of sight of one of the commercial satellite relays.

Revised and updated version of a paper by Mr. Martin presented at the Northeast Research and Engineering Meeting (NEREM), Boston, Mass., November 2–4, 1966.



The human side of engineering

It has become alarmingly apparent that the engineer must abandon the ivory tower of purely technological considerations and concern himself with the interaction between sociological and scientific ideas

Edwin M. Turner

Wright-Patterson Air Force Base

The engineer, trained to evaluate hard facts objectively, is becoming increasingly involved in interpersonal and intersocial situations. As he advances toward managerial responsibilities, he must communicate with and administer personnel. Additionally, society is demanding that the engineer–scientist pause and evaluate the effects of the scientific disciplines upon the world at large. Frequently, the engineer belatedly discovers that he possesses neither the formal training nor the practical insight necessary to approach these problems effectively.

The very newness of the engineering sciences is illustrated by the fact that over 90 percent of all the scientists and engineers who ever lived are still working today. This youthful technology creates its own particular variety of growing pains; for in this modern age the engineer must concern himself with human relations as well as with computers, equations, and experiments.

The necessity of dealing with sociological problems is frequently a source of concern and frustration to the engineer, who is trained to concentrate on and work with hard, precise facts. Engineering technology, however, is the bedrock of our modern civilization; without it, progress would stifle. But with every major technological advance, society must adjust anew; advances in communication and transportation alone have changed the whole complexion of civilization in the past 50 years. Modern society without the benefits of chemistry, medicine, electricity, mechanics, and related products would be inconceivable.

Profile of an engineer

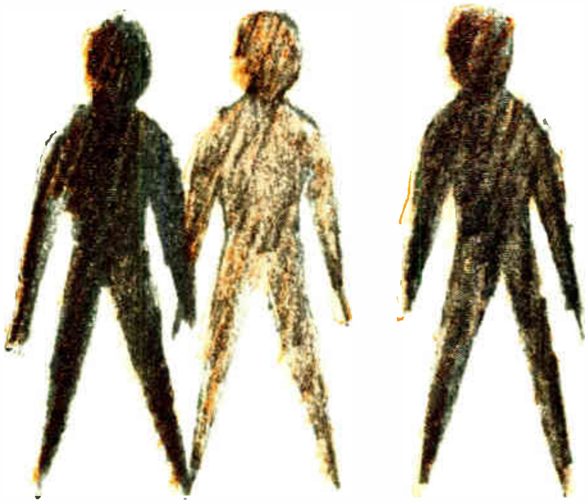
The individual engineer, however, is usually not sensitive to the psychological, economic, social, and political aspects of the environment that engineers have, in actuality, created. Often, these aspects assume importance equal to or greater than technological considerations, with a potentially profound effect on civilization.

The average engineer is a well-disciplined self-starter and hard worker, needing a minimum of supervision. He is trained to apply pertinent logic to the solution of technical problems; he is inherently objective, having little patience with subjective considerations. He must possess independent judgment, initiative, and technical insight. Above all, he must be able and willing to supplement his education continuously with new knowledge.

It would be no exaggeration to say that one half of our national budget is spent either directly or indirectly on some phase of science or engineering. This fact alone leads many individuals into the profession for dollars rather than dedication. The major goal of all engineering effort is to satisfy the needs of our advancing society and its people. Research is performed, managed, paid for, and accepted or rejected by the people. Finally, the engineer is trained, stimulated, and biased by the population of his social system; therefore, the criteria for success or failure are measured by the prejudices he absorbs from his society.

Training of the engineer

Dr. George Hawkins, dean of the Purdue University Engineering School, recently observed that engineering is a lifelong, self-imposed discipline, and that no one can hope to succeed in it without constant study and adaptation. Dean Hawkins characterized the "four phases" of the profession as follows: "During the first five years of one's career, he complains about the lack of practical training to do his job; during the next ten years, he is unhappy about a deficiency in his theoretical background;



during the next 15 years, he laments about his deficiencies in management training; finally, when he is ready for retirement, he is unhappy about his lack of training in the humanities and in the arts."

Of all professionals, the engineer in particular must beware of obsolescence. As it is, problems arise when the successful engineer is promoted to the managerial level, for he finds himself suddenly dealing with human relations as well as technology. He is often unprepared to delegate responsibility and work loads evenly, institute and enforce discipline fairly, negotiate with labor unions, and to make intuitive decisions based on vague facts.

Managerial administration requires the engineer to develop insight into the nebulous psychology of creative personnel. It is the conviction of this writer that the following factors are essential to creative individuals: (1) a high degree of motivation; (2) fearlessness of criticism; (3) skepticism; (4) ability to circumvent problems; and (5) confidence in the intuitive process.

Unfortunately, many situations tend to stifle the creative individual, and the administrative engineer must surmount such roadblocks. Creativity sometimes involves undirected effort; and many private organizations are unwilling to finance projects that are not immediately productive. Government laboratories, on the other hand, are concerned with requirements, evaluations, and administration, leaving little time for creative endeavor. Finally, the individual is likely to find little reward for creativity, even in the most progressive organizations. One is likely to evoke harsh criticism, even from his own colleagues, for advancing a "far out" idea.

Responsibilities of the engineer

The engineer of today has definite obligations to the world at large; and the world has many problems requiring his dedication. An adequate food supply must be provided for an ever-growing population, destined to increase from its present 3 billion to 50 billion by the year 2500. The world's supply of fresh water and air must be protected from pollution. Economic and cultural benefits of the affluent society must be extended to the remaining two thirds of the world. Finally, a more humane and less destructive alternative to war must be found to resolve man's disputes. It appears, unfortunately, that society is advancing technologically while regressing sociologically; this trend could lead ultimately to disaster.

It has been traditional for the engineer to concern himself solely with the technical aspects of his job, leaving its moral aspects to the politicians, businessmen, and sociologists. The impact of his profession has become so profound, however, that in this writer's opinion such recourse is no longer either wise or valid. Engineering has completely changed the way of life for one third of the world's population and has made the remaining two thirds restless—impatient to receive its benefits, but often without the training or resources to use those benefits. A visit to any one of the more backward countries will demonstrate the revolution of the centuries in transportation, communication, education, and in the general standard of living. The more astute observer is also aware of the simultaneous revolution in economics, warfare, and politics.

The engineer must consequently be aware of his obligation to work not only with the sociologist, economist, and industrialist, but also with the theologian, psychologist, physician, and politician. Indeed, scientists and engineers have a great role to play in shaping the future. With the goal of a more stable and substantial civilization in mind, they must enlarge their outlook to include the human side of engineering.

As of the present moment, the slowest process in computer technology is data preparation and the input operation. Among the proposed solutions to this problem are direct-reading handprinting input devices that boast of at least 95 percent accuracy. Character recognition is accomplished by means of linear decision functions, which are designed using an adaptive procedure. This technique allows flexibility in the character set, and the system may be tailored to a variety of users. The only drawback to widespread use is the present lack of economically feasible applications.

Handprinting input

J. G. Simek, C. J. Tunis
International Business Machines Corporation

There is currently an increasing emphasis on the development and improvement of input/output devices for digital computers. Improvements to conventional devices (card readers, punches, and printers) are being made, but there is much interest in completely novel means of communication. Audio-response and graphical display equipment are coming into increasing use as system outputs; the light pen and the Rand Tablet[1] are two relatively new forms of computer system input.

The reasons for this emphasis are at least twofold. One is the prosaic goal of simply reducing the cost and time spent in communicating with the machine. The other is the more profound goal of allowing intimate and immediate man-to-machine interaction—"man-machine symbiosis."

Optical character readers have been used for some time because of the speed and economy with which they can perform tasks formerly done by keypunch equipment. New character-recognition machines under development will be capable of handling not only typewritten but handprinted information as well. Because of this trend and the trend toward on-line communication with time-shared machines, on-line recognition of handprinted characters has been the subject of much recent research.[2] Not only might such on-line communication with a machine be convenient in itself, but the nature of the character-recognition problem may be simplified. For instance, advantage could be taken of the information available as the character is being generated or traced out. This use of such "sequential" information has already been investigated.[3-5] A second feature of on-line recognition is that "errors" or misidentifications can be corrected immediately (assuming there is some feedback in the system), and therefore the error rate of a machine may be merely annoying rather than costly.

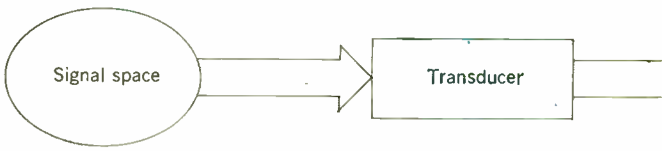
This article is part of a study concerned with the design and evaluation of on-line handprinting input terminals as part of a computer system. Possible applications for such a system would be in department stores, banks, stock markets (as clerk terminals), and teaching machines, as well as telephone, timekeeping, and work-scheduling recording systems. A necessary requirement for the system, and a novel contribution of this present work, is the development of an extremely low-cost handprinting transducer. Functionally, it is similar to the Rand Tablet, but conventional paper and pencil may be used. A somewhat similar approach, with different objectives, has been explored elsewhere.[6]

Although this particular work emphasizes handprinting input, the transducer is clearly applicable to allowing the input of more general information such as mathematical symbols, drawings, or graphs. These applications are not discussed here, although it is to be admitted that the greatest potential lies in these graphical fields where the conventional keyboard is too limited.

This study applies one particular pattern-recognition procedure to the classification of the handprinted characters as sensed by the transducer. There are a variety of basic approaches that can be taken in any investigation of a recognition procedure. Teitelman[3] discusses two recognition research methods that are at opposite extremes of the spectrum. One of these methods (Cyclops-1) is a sophisticated investigation of the character-recognition problem with typical research goals and a maximum of generality[7]; its aim is as much to learn more about the problem as it is to solve it. The other (the Stylator approach of Dimond[8]) is a specific attack on the problem of recognizing some limited set of printed characters, wherein the user can be expected to operate under a specific set of constraints. Dimond's main aim is to "solve" the problem.

The approach taken here is more toward generality. A minimum amount of constraint is put on the user; the flexibility of the character set is considered important; and the goal is to accomplish the task, indicating not only its performance, but also the inevitable compromises and alternatives involved.

FIGURE 1. Generalized pattern-recognition system.



device for computer systems

Practical handprinting transducers show promise of increased speed and flexibility in computer input techniques. The use of alphanumeric characters represents only a beginning as research progresses toward line input of graphs, charts, and equations

General operation

The output of the printing transducer is transmitted into an IBM 1620 data-processing system. The 1620 is used in this case to allow some flexibility in manipulation of the data representing the handprinted characters—different data transformation and recognition procedures can be simulated. Of course, the actual system that is visualized is one in which a number of handprinting terminals are tied into a central data-processing system. A multiplexing unit would control the various terminals, and might also embody the recognition logic. The central processor would thus be fed the recognized, printed messages from the multiplex unit. The terminal would also have facilities to receive responses from the central processor; these might be a voice answer-back facility or a strip printer. Thus, each transaction could be verified, and corrected if necessary.

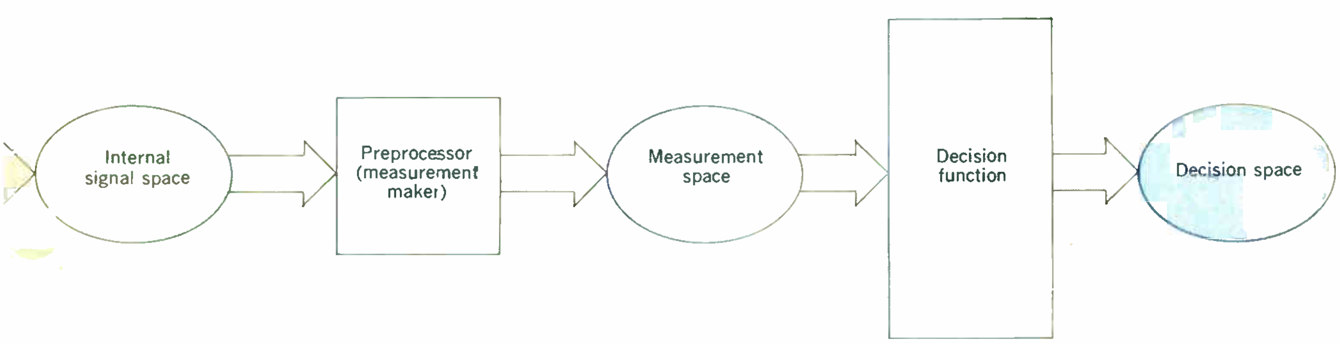
In one mode, the output of the transducer can be used to create a raster image of the printed character in the digital storage of the 1620. This raster pattern contains a relatively large amount of information, requiring the transmission of many bits between the transducer and the computer. This would be undesirable in any real-time terminal system, since much line time would be consumed. Using the IBM 1620, various transformations were applied to the raster image to simulate different measurement spaces. A measurement space requiring fewer bits than the original raster description of the pattern and less dependence on registration was chosen. Appropriate electronic hardware was built so that these measurements could be obtained directly from the input transducer.

The decision procedure uses adaptively derived linear boundaries. This form of decision function has been described previously.[9,10] The use here of the adaptive system allows the recognition to be trained for a variety of individual printers and for a variety of printing fonts.

The measurement space

In general, pattern-recognition systems consist of three main functional sections (see Fig. 1); the transducer, which obtains a representation of the input pattern; the preprocessor, which measures or senses features of the detected pattern; and the decision function, which assigns the pattern to one of the possible classes.

A recent paper in the character-recognition field[3] has stated that "research in pattern recognition may be characterized as a search for invariants." Indeed, much of the ingenuity of character-recognition machine design has been devoted to this search. The search is complicated by the fact that not only is it necessary to find invariants



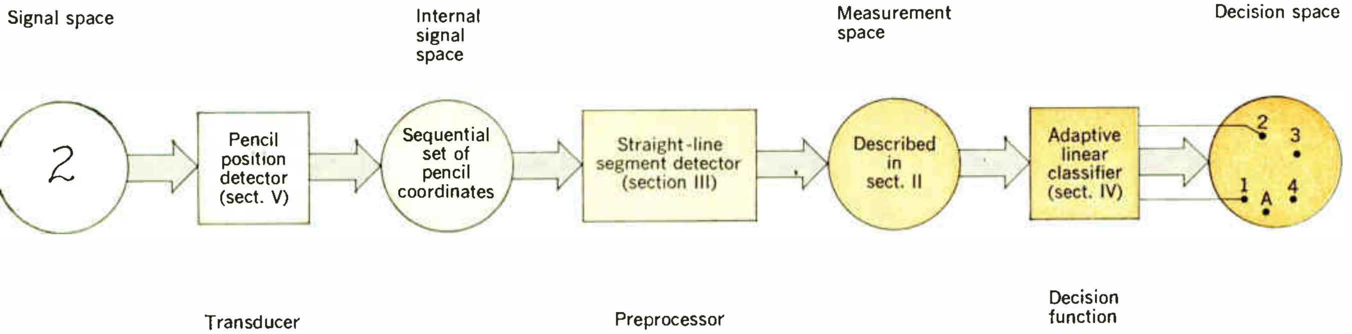


FIGURE 2. The handprinting recognition system.

(measurements) to obtain good recognition performance, but it is desirable to use a minimal number of invariants to reduce the amount of hardware required.

One common approach employed in character-recognition machines is to scan the printed character with either mechanical image dissectors or electronic beam scanners and recreate the image of the character in raster form in a binary register in the machine. This raster pattern can have the decision function applied directly to it in the form of a set of Boolean logic statements. However, even for one particular character of one particular font, the character image in the digital register will have many variations. These will be due not only to the quantization of analog signals into binary states, but also to the "print quality" of the character on the page as well as the orientation and registration of the character on the document. The problem of registration is sometimes handled by using a digital shift register to store the image of the character, and by applying logic as the character is shifted to determine some nominal position. The effect of other character variations is sometimes overcome by applying a set of measurements to the pattern in the shift register that are relatively insensitive to missing or added bits and slight variations in orientation. The measurements are typically a set of Boolean logic statements on the bits of the register. These logic statements are generally determining features of the character shape that are invariant; that is, the Boolean logic statements obtain a set of measurements on the character. Once a particular measurement set has been designed, there are a variety of statistical decision methods that can be applied to them to perform the final classification.

The handprinting transducer provides information sufficient to produce a raster image of the printed character (see Fig. 2). It is also capable of providing sequential information, since the character is sensed as it is actually being printed. This information (the order in which the *X*, *Y* intersections occur) is used to distinguish between a horizontal line drawn from left to right and one drawn from right to left, providing two distinct measurements, or segments. However, the sequential order in which the segments occur is not retained for the following three reasons:

1. The sequential approach has already been investigated and reported in the literature.[2,4,5]

2. We did not wish to prescribe the sequence of strokes with which particular characters are printed or to provide sufficient storage to retain all the variations in sequence that occur.

3. It was felt that the adaptive decision procedure, applied to a set of measurements, was an appropriate one for this application. The type of adaptive system that is employed here is not easily applied to sequential information.

The measurement space used describes the character in terms of straight-line segments, a natural way to describe many of the characters of the Roman alphabet. Eight types of segment are allowed, corresponding to the eight directions of the compass. Note that a horizontal line, drawn from left to right, is considered a different segment from a horizontal line drawn from right to left. This is the extent to which sequential information is used here. The printing area is divided into nine zones, or regions, and the region in which a particular segment begins determines where in the measurement space its occurrence is recorded. The number of times the printing stylus is lifted off the transducer during the creation of a particular character is also recorded (and is referred to as the "lift-off" measurement), as is the number of times a particular segment occurs in a particular zone. Figure 3 is a representation of the measurement space (entries are binary).

The columns of the matrix in Fig. 3 correspond to the eight slopes and the lift-off measurement. The zones (rows) correspond to the location in the printing area (a region assigned to each possible character position) where a measurement or segment has occurred. A particular measurement is recorded by entering a binary 1 in the appropriate position. The printing area is partitioned into nine distinct zones represented by coded combinations of the symbols A, B, C, and D. Figure 4 shows the relationship of the zones to the print area. If a measurement occurs in area AB, it is recorded in both the rows assigned to zone A and zone B.

Returning now to the matrix shown in Fig. 3, the rows are grouped into five sections of four rows each. The first four correspond to zones A, B, C, D. The fifth section is used to record the number of times that a particular measurement has occurred during the character formation.

Each section has four rows, thus allowing for four individual occurrences of each measurement in that section. As measurements occur, a bit is recorded in the appropriate zone or combination of zones. The first occurrence causes an entry in the topmost row; successive occurrences of the same measurement in the same zone will result in additional bit entries.

If the raster image of the character had been used,

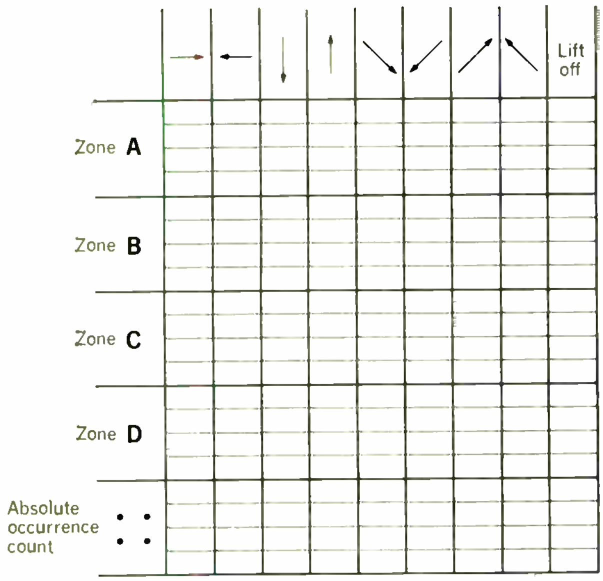


FIGURE 3. Static format of measurement space.

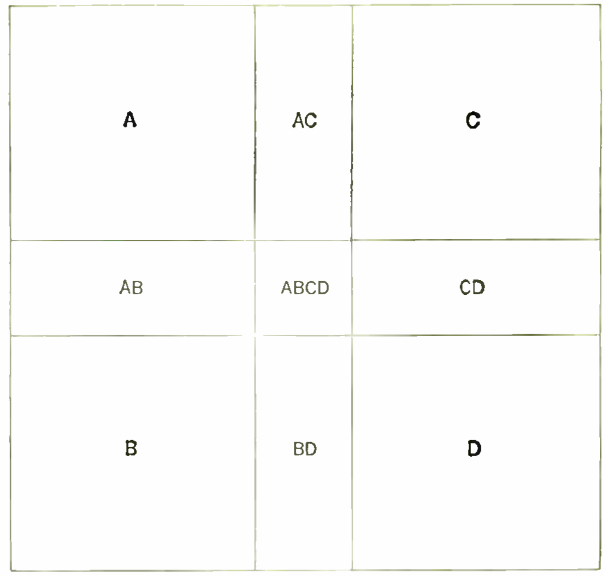


FIGURE 4. Zone location in a print area.

256 bits (as will be made clear later) would have been required to represent it. The raster pattern would be subject to character registration and orientation variations. The transformed (hopefully invariant) measurement space requires only 180 bits; experiments have been done with a measurement space of 90 bits and are reported in a later section. Further reduction of the number of bits in the measurement space is probably feasible without too much degradation in performance.

The sequence of stylus (pencil) location points as sensed by the transducer is transformed by electronic circuitry (on-the-fly) into the set of measurements just described. Figure 5 is a flow diagram of the logical operations. The circuitry tests the sequence of x, y location points for incremental changes. When a predominance of these changes is established in one of the eight possible directions, a valid measurement has occurred. An indication of the segment is transmitted to the computer; subsequent segments are recorded only if they differ from the segment immediately preceding them. For printed alphanumeric characters, an average of seven such measurements is obtained per character; for the character "8" a maximum of 14 such measurements occur. Note that we are breaking up curved sections of a character into a sequence of straight-line segments.

When the stylus is lifted from the transducer for some prescribed amount of time, the lift-off indication is transmitted to the computer. The specific time interval is both a function of the particular character and the human printer. Attempts to use a longer period of stylus lift-off as an indicator of the "end of character" were not successful. In many cases, the time between separate strokes of a character was longer than the time between characters. A separate panel push button was used to denote end of character.

Decision procedure

The decision problem can be formulated as follows: The input to the decision function is a description of the unknown pattern in terms of a set of measurements $x_1x_2 \ldots x_p$ (which we shall refer to as the measurement X). In this example the x's are binary, and testify to the presence or absence of particular segments, such as a horizontal line drawn from left to right (see Fig. 3).

Given a measurement X, we must decide which character, out of a set of possible characters, was printed. The theoretical solution to this problem in terms of statistical decision theory has been formulated by Wald.[12] A decision in favor of a particular character class C_i is made by choosing the maximum of the set of conditional probabilities

$$\Pr(c \in C_i|X) \qquad i = 1, 2, \ldots, N \tag{1}$$

or, in other words, the probabilities that, given some measurement X, the input signal (character) c is a member of the class C_i. We must compute these probabilities for each class i, where i runs from 1 to N, with N being the number of different character classes we recognize.

In general, we do not know the probabilities as given in (1). Bayes' rule, however, allows us to rewrite (1) in the following way:

$$\Pr(c \in C_i|X) = \frac{\Pr(C_i)\Pr(X|c \in C_i)}{\Pr(X)} \tag{2}$$

where

$\Pr(C_i)$ is the occurrence probability of characters of the class C_i. In most cases this is either known or the character classes are equally probable. In further development of (2) we assume the classes are equally probable.

$\Pr(X)$ is the occurrence probability of a given measurement X. This term is constant in each expression as i runs from 1 to N. It can be omitted in the comparison process.

$\Pr(X|c \in C_i)$ is the probability that an input character that is a member of a class i will result in a measurement X.

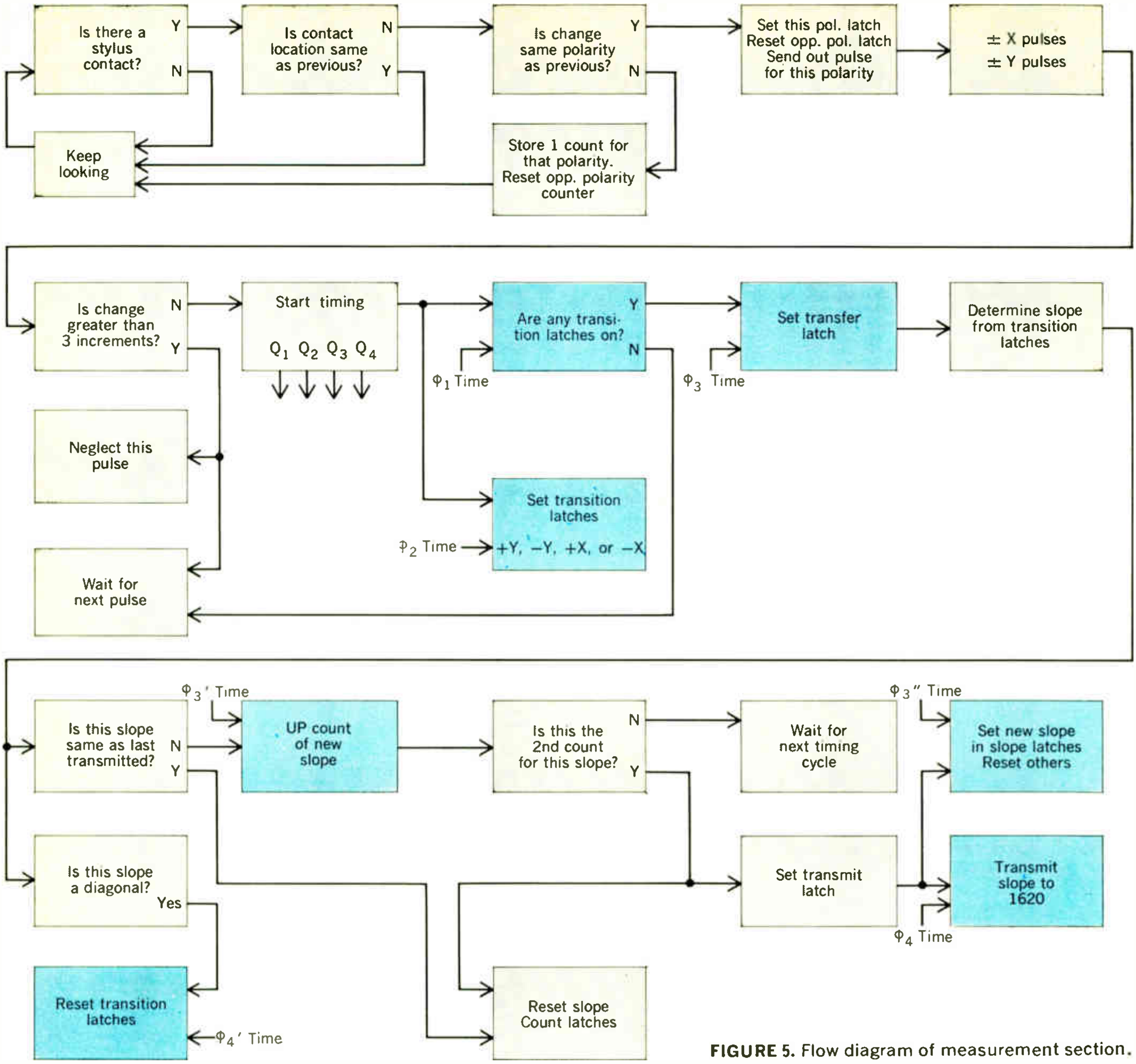


FIGURE 5. Flow diagram of measurement section.

Thus we can translate the problem of determining the maximum conditional probability of (1) to that of determining the maximum of

$$\Pr(X|c \in C_i) \qquad i = 1, 2, \ldots, N \qquad (3)$$

Under the assumption that the components of the measurement X, $(x_1, x_2, \ldots, x_p)$ are independent, *each* expression of (3) can be written as

$$\Pr(X) = \prod_{j \in v} p_{ij} \prod_{j \in \bar{v}} q_{ij} \qquad (4)$$

where

v is the set of j's for which $x_j = 1$ for $j = 1, 2, \ldots, P$.
$\bar{v}$ is the complement set to v
p_{ij} is the probability that measurement $x_j = 1$ for class i
q_{ij} is the probability that measurement $x_j = 0$ for class i and $q_{ij} = (1 - p_{ij})$

The foregoing product computation can be altered to an addition computation if ln (p_{ij}) and ln (q_{ij}) are used; this operation can be done with a familiar network device (Fig. 6). A single summing device of this type would be used for each class i and the maximum sum thereby detected. The class corresponding to the maximum sum is chosen as the identification of the input pattern c.

We must still determine the quantities p_{ij}, q_{ij} for each class of character i, but this operation is relatively straightforward. One possibility is to make statistical observations on the data to gain the information needed, and then implement the required network. Alternatively, the required p_{ij} and q_{ij} (or w_{ij} in the network) may be estimated by the use of an adaptive procedure that adjusts the weights, w_{ij}, during a "training" period when a (hopefully) statistically representative set of measurements X

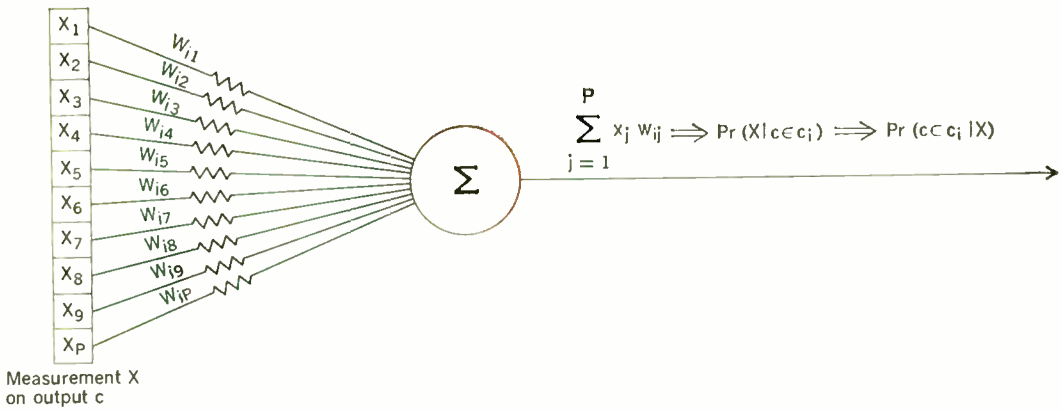
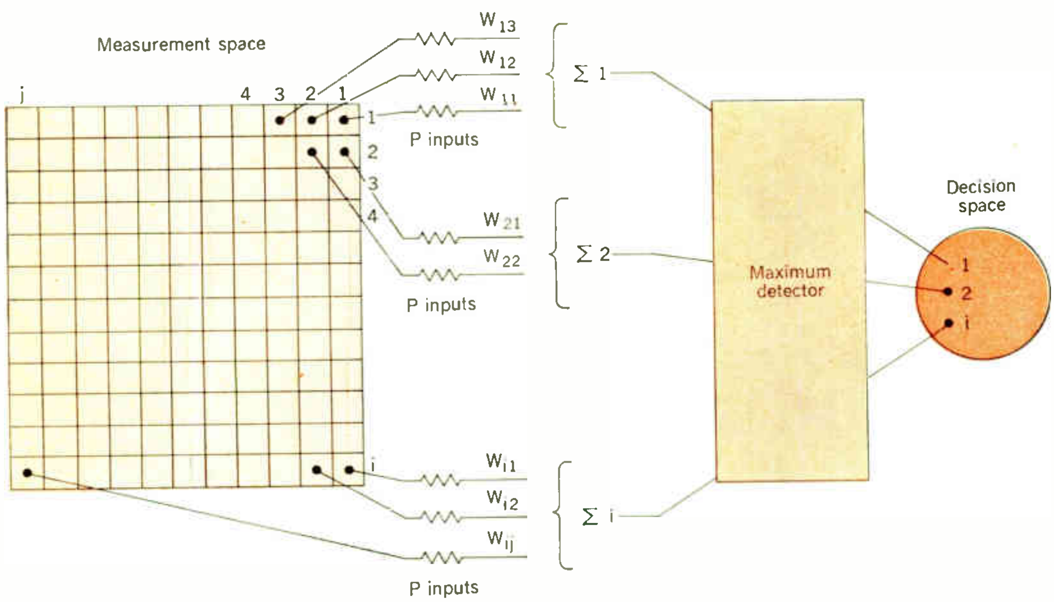


FIGURE 6. Linear summing device.

FIGURE 7. Linear decision function.



corresponding to known classes C_i is presented to the network. This adaptive procedure is described fully elsewhere. Roughly, the procedure used here is as follows:

A sample set of measurements corresponding to a known sequence of printed characters is obtained. The measurements are iteratively presented to the network of linear functionals in Fig. 7 (simulated in the IBM 1620). A well-known algorithm is used to adapt the set of weights w_{ij} so as to set up linear decision boundaries[9-11] that provide a minimum number of classification errors on the design sample. Thus, a suitable decision function can be designed for a variety of individual printers or printing fonts, provided that a sample set of measurements is obtained. The performance of such decision functions is given later on in the section on experimental results.

The handprinting transducer

The transducer consists of two separated layers of orthogonal conducting lines; when subjected to the external pressure of a writing instrument, they are forced together to produce ohmic contact (Fig. 8). The bottom set of conducting lines is created by photoetching a copper-clad glass-epoxy substrate. The top set of lines may be created by either copper photoetching or conductive-ink printing on some flexible surface member (in this case Mylar film was used). Both the top and bottom sets of lines are plated with gold. This arrangement of the two orthogonal sets of lines, at a density of 12.6 per centimeter, produces a structure similar to a multiple-point crossbar switch with a density of 159 contact points per square centimeter (Fig. 9). The two surfaces must be separated under the "no load" condition.

The actual writing is done on ordinary paper (with any desired printed format) placed on top of the array. The pressure of normal printing forces the paper and the top membrane down onto the bottom set of conducting lines, producing a contact point between the two orthogonal sets of lines in the location directly under the printing stylus. As this stylus moves from point to point, successive momentary contacts are made, and a raster-type image of the character, as it is traced out, is signaled to the electronic system. The stylus can be any natural implement, such as a ballpoint pen or a lead pencil.

It is quite feasible for the paper overlay to include multiple carbon copies. The transducer thus permits relatively inexpensive construction, freedom from mechanical constraints on the printing implement, and hard copy output.

The two sets of grid lines originally were separated by maintaining the top Mylar sheet under tension above the lower substrate with a separation of about 0.076 cm. The elasticity of the top sheet causes this separation to be maintained when there is no stylus pressure. However, after significant usage, a deformation of the copper lines causes the flexible upper sheet to develop wrinkles that can give rise to spurious contacts (contacts in areas not directly associated with the position of the stylus). Furthermore, it was necessary to ensure that the only

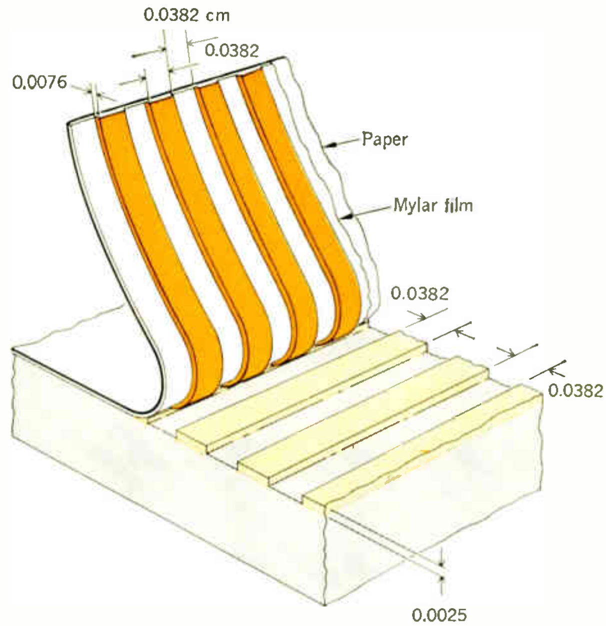
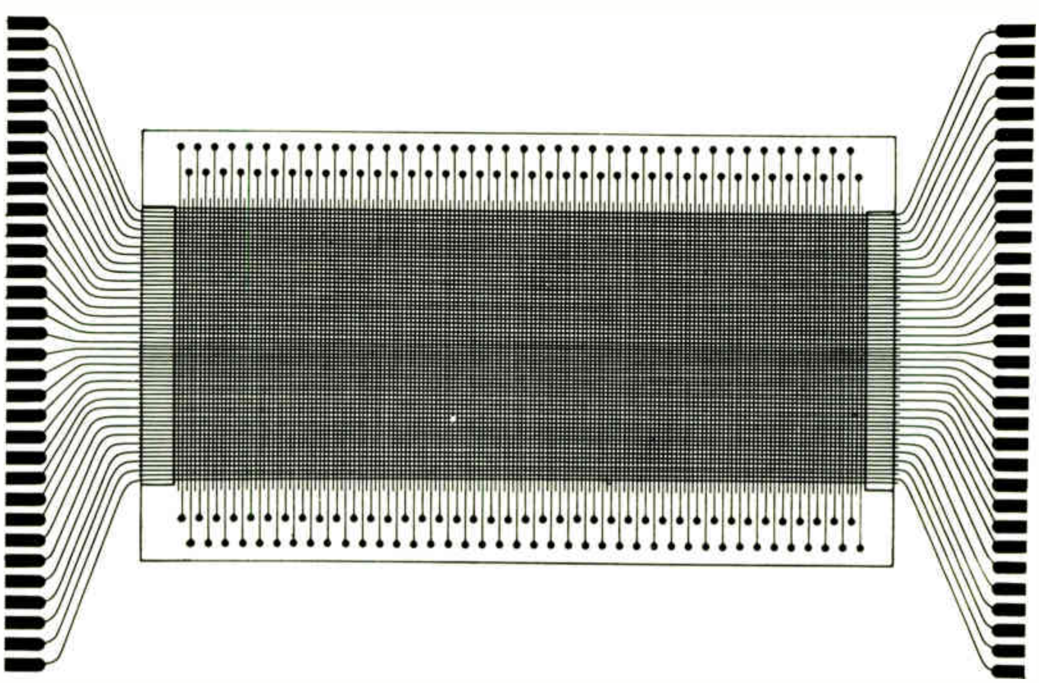
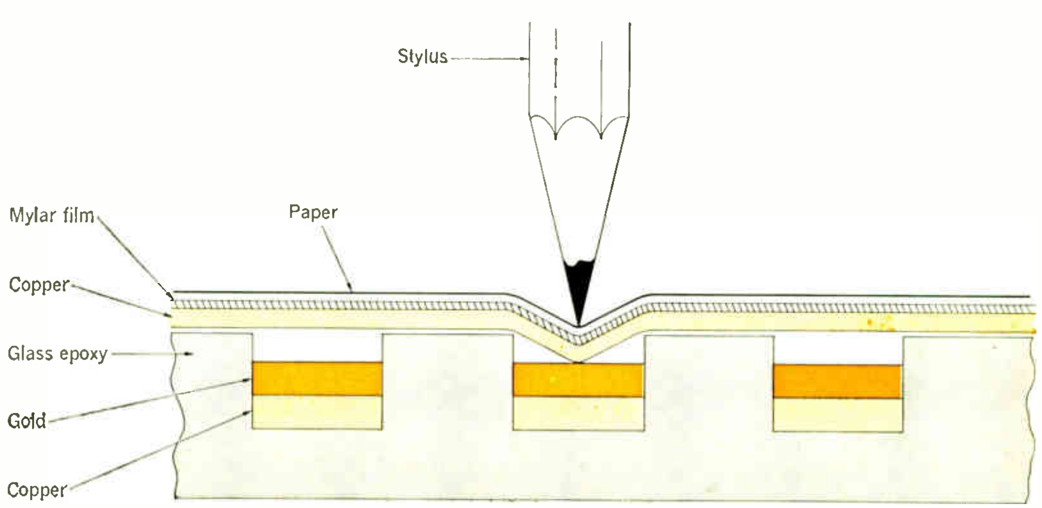


FIGURE 8 (right). Cutaway view of handprinting transducer.

FIGURE 9 (below). Top view of contact surface grid.

FIGURE 10 (bottom). Depressed base substrate.





force applied to the top membrane was that due to the stylus, for any extraneous pressure (such as that caused by the printer's hand) would also cause contacts.

By fabricating the base as shown in Fig. 10, the problems referred to earlier are eliminated. The copper-clad epoxy base is first photoetched to give the pattern of lines used in the grid; then, successive layers of gold and copper are plated on the lines. The metal lines are then "flushed" even with the surface of the epoxy by the application of heat and pressure to the base. The final step involves etching away the top plated layer of copper, using the gold as a resist, thus leaving the gold lines some distance below the epoxy surface. This dimension is directly controlled by the thickness of the top copper plating. By having the bottom set of grid lines below the surface of the epoxy, an accurate and reliable separation exists. An alternative would be to insert an insulator, with openings at the desired contact points, between the two layers. However, at high contact densities, registration and dimensional accuracy of the openings make this very difficult.

The dimension of the gold line depression is somewhat critical. If it is too deep, erratic contacts may occur, and there will be a "bumpy" feel as the printing stylus is used. If the depression is too shallow, spurious contacts may be made as a result of the nonuniformity of the gold-plating or hand pressure applied during printing. A dimension of approximately 0.0013 cm (for the line width of 0.038 cm) seems to be a compromise that gives adequate separation, a smooth printing surface, reliable contacts with nominal stylus pressure, and no contact when a broad (hand-pressure) force is directed to the top surface. With this system of separation, the top membrane can rest directly on the base without producing contact.

However, after significant use (that is, many weeks of operation), the copper lines on the top membrane tend to take a permanent set and "match" the depressions in the base, causing spurious contacts. (The copper lines, not the membrane, become deformed.)

An alternate approach is to print the lines on the Mylar sheet with a conductive ink. The ink will not "take a set" under prolonged use, and any deformation would result solely from the membrane. The physical stability of Mylar film is far better than that of copper; therefore, this combination produces a more stable top membrane. If necessary, the top sheet could be replaced periodically under heavy usage.

The specific transducer used in this experimental investigation has an overall printing surface of 5.1 by 12.7 cm. This is partitioned into forty 1.27-cm squares with a matrix of 16 × 16 grid lines per square. In each direction, every 16 lines are commoned, making a total of 40 equivalent 16 × 16 grids on the printing surface. In order to insure printing registration within this 1.27-cm-square matrix, a white Mylar overlay, with 1.27-cm-square grids marked off, is placed over the top membrane. These grid lines show through the printing paper. Sample characters are then printed on the sheet so as to fill as much of the square grid as possible without carrying over into the adjacent grid spaces. The character size is somewhat large, mainly because of the density of the conducting lines used in this case. Higher densities of grid lines (about 20 to 40 lines per centimeter) would allow smaller printing.

As has been stated previously, a character is sensed as a number of point contacts produced during printing. Each of these points can be described by an x, y coordinate system. The number of contact points produced during printing will vary according to the size and shape of the character. The sequence of x, y coordinates are the input signals to the "measurement circuitry" described previously.

The system in use

The transducer is attached to the printer's console that contains the electronics for implementing the measurement space as well as the circuitry interface needed for transmission to the 1620. The console display panel contains indicator lights used in monitoring coordinate contact location, output of the segment transformation circuit, and computer control. Push buttons for transmitting "end of character" and "end of transmission" are also located on the panel.

In operation, the user places a blank overlay sheet of paper or any compatible paper form on the transducer. This paper becomes a hard-copy record of the printed characters and also provides a writing blank sheet and guides the printer for correct character placement. After each character entry, the operator must depress the "end of character" button. When the desired message has been inserted, the "end of transmission" button must be depressed to free the computer of the remote connection.

There is, of course, both a "training" phase and a "recognition" phase. During the training phase, samples of an individual's printing are entered into the IBM 1620 via the transducer. This training sample or known sequence of characters (or analysis sample, as it is sometimes called) is used to determine the weights w_{ij} for the linear decision function. The weights for a particular individual are stored, in this experimental system, in the computer memory. Recognition weights for a variety of writers may be stored, and the approximate ones used, if the writer identifies himself to the system before he begins to print his message. This scheme allows the recognition system to be tailored for each writer and to a variety of printing styles and fonts.

Experimental results

Several sets of experiments were conducted to obtain an indication of the performance of the on-line hand-printing transducer and recognition system. The experimental results are based on fairly large samples (approximately 300 alphabets, where an "alphabet" is one example each of the characters to be treated) obtained from two individuals. The samples were obtained from the individuals at the rate of approximately ten alphabets per day. Only two individuals are represented thus far because of the time required to generate the data. It is felt that the recognition results obtained here are indicative of those that would be obtained on the printing of any individual, since the two printers were not trained and used their normal printing style. One example of the printed character structure treated in the investigation is shown in Fig. 11.

Four different experiments were conducted; the objectives of each are summarized as follows:

1. The first experiment was conducted to determine the recognition rate of the system on the printing of a given individual when the system has adapted to that individual. To do this, the total sample of printing is divided into

two portions, one being used for learning (the analysis sample) and the other to be used for recognition (the test sample). The performance on the test sample is dependent on the size of the analysis sample. In effect, the analysis sample must statistically represent the test sample.

2. A further set of experiments was conducted to determine the effect of alphabet size on the recognition rate of the system. Three distinct alphabet sizes were used: a 10-character numeric alphabet, a 15-character numeric and arithmetic symbol alphabet, and a 26-character Roman alphabet. The nominal versions of the printed characters are shown in Fig. 11.

3. Experiments were performed to determine the applicability of weights obtained from the analysis sample of one printer to a test sample prepared by a different printer. In this case, the two printers were asked to use the same font—that is, to shape their characters the same way. Presumably, if good recognition performance could be obtained under these conditions, then a "universal" set of weights could be used to recognize a variety of individual printers.

4. An additional set of experiments was conducted using variations of the original measurement space. These were to determine whether a reduction of bits in the measurement space could be made without a significant degradation in recognition performance.

FIGURE 11. Nominal alphabet samples.

FIGURE 12. Confusion matrix based on an experimental test sample consisting of 100 alphabets.

	Recognized as														
Pattern	1	2	3	4	5	6	7	8	9	0	−	+	x	/	=
1	97											3			
2		100													
3			99		1										
4				99			1								
5					100										
6						100									
7							100								
8								100							
9				1					99						
0										100					
−											100				
+	1			1								97	1		
x													100		
/														100	
=															100

The results of the experiments are as follows:

1. The recognition performance on the test sample of one individual, given a set of weights adapted on the analysis sample of that individual, showed that both individuals obtained a recognition performance in excess of 99.4 percent. If the handprinting system included some form of feedback to the printer, so that the occasional errors could be retransmitted, then this is quite a tolerable performance level. Only five characters per thousand would have to be retransmitted. Figure 12 shows a confusion matrix based on the experimental results of one individual. This kind of matrix is convenient to illustrate particular confusion pairs in the alphabet. The number of alphabets in the analysis sample was also recorded. A relatively large sample was required in this case because of the many variations of the way in which the characters were printed and sensed. For a significantly lesser number of alphabets in the analysis sample (approximately 30 alphabets) the recognition rate is in the vicinity of 96 percent.

2. Table I shows the variation in recognition performance for different-size alphabets. As might be expected, an increase in alphabet size results in a decrease in recognition performance. However, with the largest alphabet size used, the recognition rate is still in excess of 92 percent, a level which should be acceptable in many applications if some method of feedback, allowing retransmission or correction, is included.

3. The performance level achieved by one individual, using the weights generated by the analysis sample of a different individual (but both using the same font), is in the vicinity of 95 percent. This is, of course, significantly lower than when the individual uses his own set of weights. It thus appears that there are minor variations in forming the characters which would necessitate the use of an analysis sample from any given individual to obtain the maximum performance level on that individual. Table II summarizes the experimental data. It might be parenthetically noted here that when the analysis samples of two individuals are combined, the recognition performance achieved by either individual is as good as or better than the performance obtained by using his own set of weights. These data are also shown in Table II.

4. The original measurement space (Fig. 3) represents any given line as being in one of eight directions. Thus, there is a difference between a horizontal line drawn from left to right and one drawn from right to left. This information is dependent not only on the individual printer but also on the sequence with which a particular character is drawn. The measurement space was reduced by eliminating information pertaining to the direction in which a particular line was drawn, and there were thus only four directions. Table III shows the recognition result using this truncated measurement space. It appears that the direction in which a line is drawn is an information-carrying entity.

The original measurement space also contains information referring to the location of the beginning of each line segment. It is apparent from a cursory study of the numeric characters that this information is vital to the separation of such characters as the 2, the 5, the 0, and the 6. No experiments deleting these measurements were conducted.

In another study, samples of handwriting were fed into the system, employing intentionally "sloppy" char-

I. Recognition of different alphabet sizes

Alphabet Size	Analysis Sample Size	Test Sample Size	Recognition Rate
10	230	100	99.5%
15	230	100	99.4%
26	87	50	92.0%

II. Mixed printer recognition rates for 15-character alphabet

Test Sample Printer	Analysis Sample Printer	Analysis Sample Size	Test Sample Size	Recognition Rate
1	2	230	100	97%
2	1	230	120	92%
1	1 and 2	200	200	95%
2	1 and 2	200	120	98%

III. Recognition using different measure spaces

Measure Space	Analysis Sample Size	Test Sample Size	Recognition Rate
With direction	230	100	99.4%
Without direction	230	100	97.0%

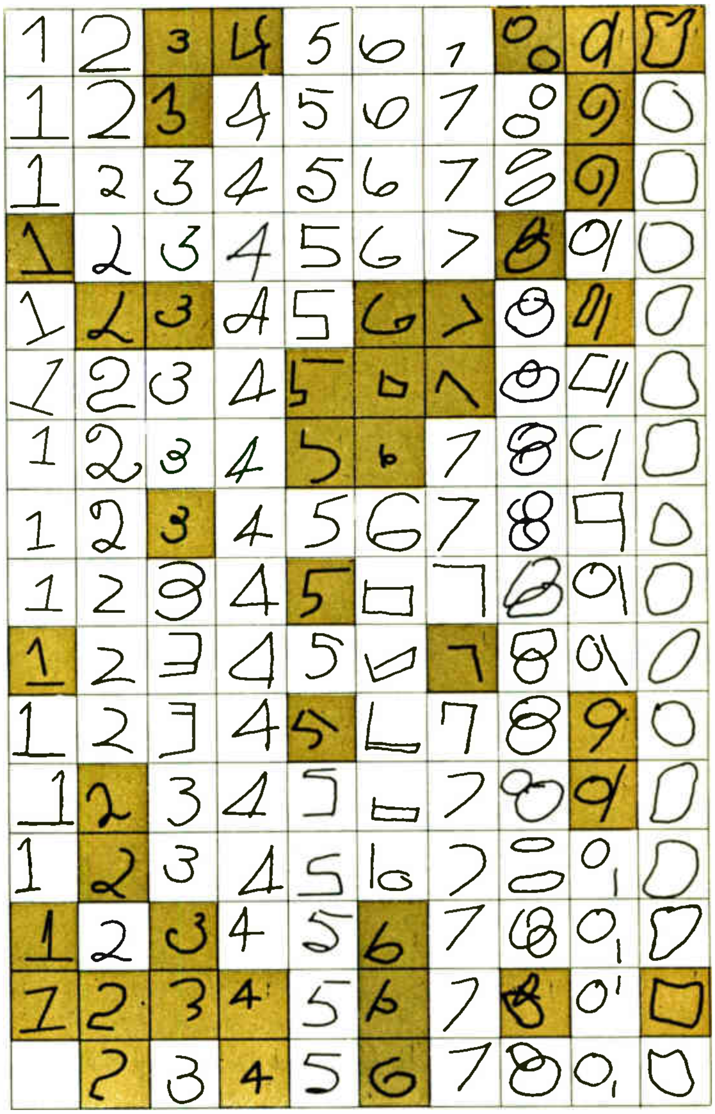


FIGURE 13. Samples of character extremes. Shaded areas were not recognized.

acters to test recognition capability (Fig. 13). The set of weights used was based on a sample of 200 alphabets of printer 1. The shaded characters were not identified correctly. Many of the characters identified incorrectly are noticeably smaller; they were undoubtedly not recognized because of system resolution.

Conclusions

The handprinting input device and recognition system described has been shown to be technically feasible. Its use awaits its economic justification in particular applications.

The measurement space is similar to that used by other workers. It has been shown to allow the reduction in constraints imposed on the human printer, and to be a suitable input to a linear decision function. Adaptive linear decision functions perform well in this application and are particularly convenient, allowing the tailoring of the system to a particular character set and a particular human printer.

The identification rates obtained by such a system are approximately 95 percent correct, depending on the alphabet size. Such a rate should be quite usable in many applications, particularly in an on-line system with feedback.

The work of C. P. Eller in the development and fabrication of the transducer and of L. J. LaBalbo in the implementation of the electronic hardware is gratefully acknowledged. C. E. Kiessling and Mrs. E. R. Ide are responsible for the simulation of the measurement transformations and decision functions.

REFERENCES

1. Davis, M. R., and Ellis, T. O., "A man-machine graphical communication device," *Proc. 1964 Fall Joint Computer Conf.*, pp. 325–331.

2. Groner, G. F., "Real-time recognition of handprinted text," *Proc. 1966 Fall Joint Computer Conf.*, pp. 591–601.

3. Teitelman, W., "Real-time recognition of hand-drawn characters," *Proc. 1964 Fall Joint Computer Conf.*, pp. 559–575.

4. Brown, R. M., "On-line computer recognition of handprinted characters," *IEEE Trans. Electronic Computers (Short Notes)*, vol. EC-13, pp. 750–752, Dec. 1964.

5. Mermelstein, P., and Eden, M., "A system for automatic recognition of handwritten words," *Proc. 1964 Fall Joint Computer Conf.*

6. "Elastic diaphragm switch technology," *Electron. News*, p. 43, Oct. 11, 1965.

7. Marill, T., *et al.*, "Cyclops-1: a second-generation recognition system," *Proc. 1963 Fall Joint Computer Conf.*, pp. 27–33.

8. Dimond, T. L., "Devices for reading handwritten characters," *Proc. Eastern Joint Computer Conf.*, pp. 232–237, 1957.

9. King, J. H., and Tunis, C. J., "Some experiments in spoken word recognition," *IBM J. Res. Develop.*, Jan. 1966.

10. Griffin, J. S., King, J. H., and Tunis, C. J., "A pattern identification system using linear decision functions," *IBM Systems J.*, vol. 2, pp. 248–267, 1963.

11. Nilsson, N., *Learning Machines.* New York: McGraw-Hill, 1965.

12. Wald, A., *Statistical Decision Function.* New York: Wiley, 1950.

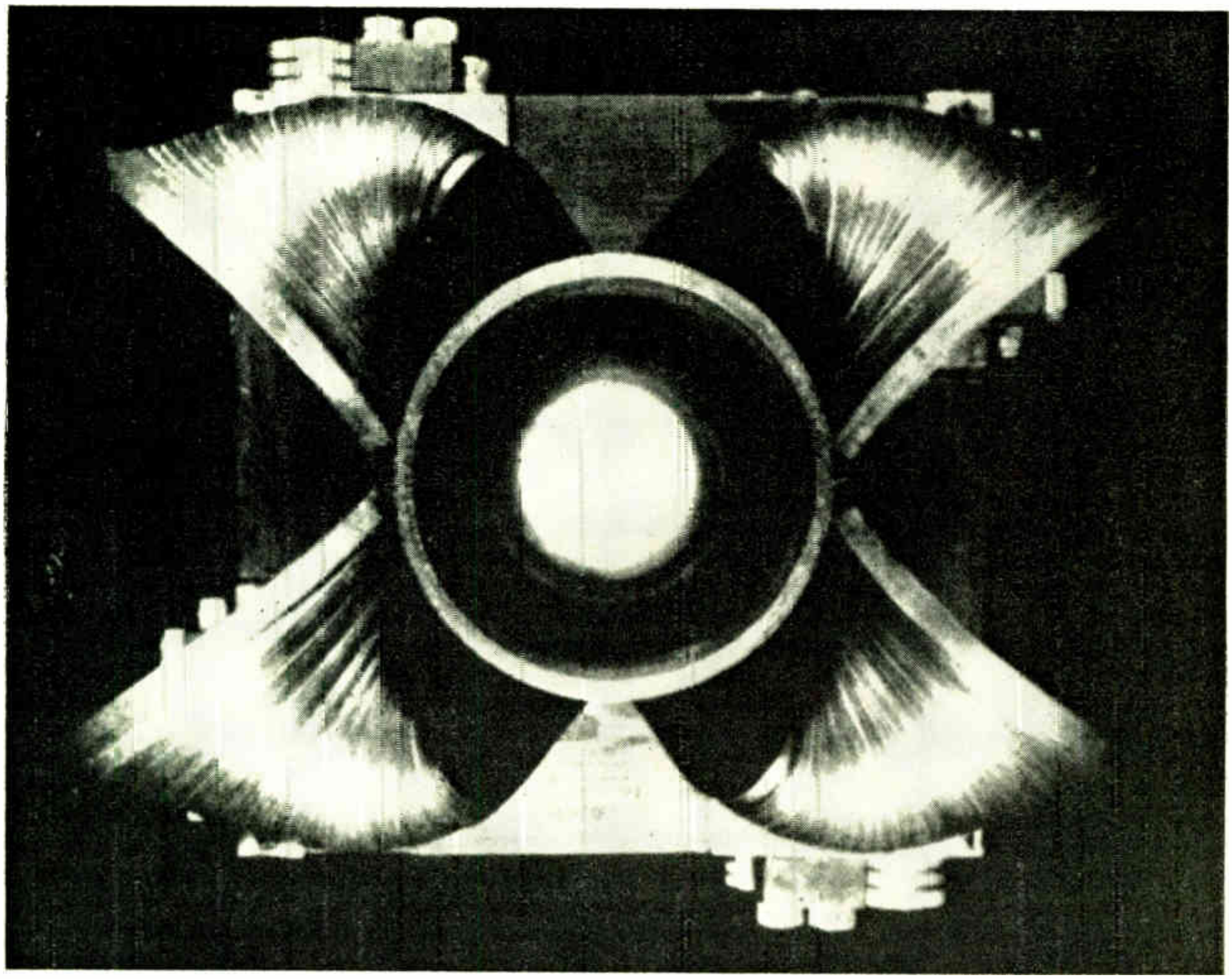


FIGURE 1. End view of a superconducting quadrupole magnet used for focusing of high-energy particles at the Brookhaven National Laboratory. The diameter of the bore is slightly larger than 2.5 cm.

In the broad sense the term "quantum electronics" relates to the motion of electrons as governed by quantum mechanical laws, but used more specifically it refers to electronic processes involving transitions between discrete energy levels. The latter category includes the laser, and it is to this development, which is responsible for most of the new horizons in the field, that this article is primarily devoted. Two specific examples are described: a tunable light oscillator and an intense-light-pulse generator. It is pointed out that, despite their promise, the ultimate impact of the laser devices on industrial technology cannot be predicted.

This article is not an attempt to predict the future of quantum electronics, but rather to sketch some new vistas that have been opened by recent developments. One might ask: "What are these scientific developments good for?" In answer to that question, I would like to remind you of Faraday's reply to the British prime minister who wondered what purpose the discovery of electromagnetic induction could have. Faraday did not predict the rise of an electric power industry; what he said, in effect, was: "My brain child has a function similar to that of a newborn British subject: you may tax it later." The history of science has shown that it is much safer to say that some technological applications will follow scientific advances than to deny future applications. The economic scale and impact of the applications are, however, very difficult to predict and I shall not venture into a realm that is beyond my competence.

Defining our terms

Presumably the term "quantum electronics" relates to the motion of electrons, as governed by quantum mechanical laws. Although the term is relatively new, the field defined in this manner is four decades old, starting in the late 1920s, or perhaps earlier if we include the photoelectric effect and photocells as part of quantum electronics. Looking back to past horizons to get our bearings, we find that the quantum electronic effect with the largest technological impact is probably the existence of several types of carriers of electricity in solids—electrons and holes with different effective masses. This fact, which was

New horizons in quantum electronics

The development of the laser has opened new horizons in quantum electronics, but although the future promises even greater possibilities for its use, experience teaches us that we cannot safely predict the economic scale or impact of such applications

N. Bloembergen *Harvard University*

first clearly stated by Peierls and Wilson in the late twenties, is at the basis of semiconductor electronics and transistors.

The tunneling of an electron through a potential barrier is another typical quantum mechanical effect; the Esaki and Zener diodes are successful applications. The field-emission microscope, with which individual molecules adsorbed on a metal point may be made visible, is also based on this effect but has not had a large-scale technological impact. These examples should remind us that there is no one-to-one correspondence between exciting discoveries in electron physics and important developments in electronic technology, although there is, of course, a large amount of correlation, and the economic scale of applications is difficult to predict.

The phenomenon of superconductivity is another manifestation of the laws of quantum mechanics relating to the motion of electrons. Although the effect has been known since 1908, only in the last 20 years has it been understood, and it has become technologically interesting only in the past five years or so. Type II superconductors allow the construction of superconducting magnets with field strengths of about 150 kilogauss and zero power consumption. These magnets do, however, consume liquid helium. Nevertheless, because of advances in cryogenic technology, liquid helium has become acceptable in technological operations, and so superconducting magnets may be the answer to magnetohydrodynamic power generation and to containment of plasmas for controlled fusion. I shall not venture into prophecies about these fields, which still harbor many uncertainties, but superconducting magnets are definitely useful in the laboratory. Figure 1 shows a quadrupole superconducting magnet for focusing high-energy particles; its compact size is a distinct advantage. Superconducting phenomena, such as tunneling of electrons between superconducting junctions, have also opened new horizons beyond which there may lie new applications. Superconducting films may be used in switching elements and junctions may be developed as sensitive detectors in the most inaccessible region of the far infrared. Very recently the effect was used for a new precision determination of the fundamental constant h/e. This may be said to be a metrological application.

The term "quantum electronics" is often used in a narrower sense also. It then refers more specifically to electronic processes, which involve transitions between discrete energy levels—as opposed to the continuum of energy levels, which are involved in semiconductors, superconductors, and plasmas. Discrete energy level devices are numerous and many were well established before the term came into vogue, so perhaps "discrete electronics" would be more accurate. In this field familiar horizons of the recent past are formed by the skyline of applications of magnetic resonance and microwave spectroscopy; the gyrator, including isolators and circulators, magnetoacoustic delay lines, and microwave masers, are well-established landmarks. The cesium-beam atomic clock has been adopted as the new standard of time and man finally has decoupled the measurement of time from the motion of the earth around the sun; he has turned from an astronomical precession to the precession of electrons inside the atom. And the cesium clock may well be replaced as a frequency standard by the atomic hydrogen maser if this instrument becomes more universally available. It is capable of a short-time stability of $1:10^{13}$ and a long-term stability and resettability of $1:10^{12}$. The hyperfine splitting of the atomic ground state of hydrogen is known to 11 significant decimal places.

The three-level solid-state maser is still the ultimate in low-noise microwave reception. This device, conceived in 1956, quickly became operational in the ground stations of satellite communication systems, in a few radiotelescopes, and in some radar installations. Its technological use, however, is not widespread and in most applications it is not competitive with simple cooled parametric devices, which were developed shortly afterwards. Here we have another example of how hard it is to predict technological application. Although the device was soon perfected to the point that very little further development is now necessary, alternative solutions for low-noise reception had advantages of economy and simplicity, and thus the technological impact of the solid-state maser was severely limited.

The pumping principle employed in the maser to create a medium with an electromagnetic gain, as shown in Fig. 2, has endured in many forms of lasers. And lasers have, of course, opened up most of the new horizons in quantum electronics to which the remainder of this article will be devoted.

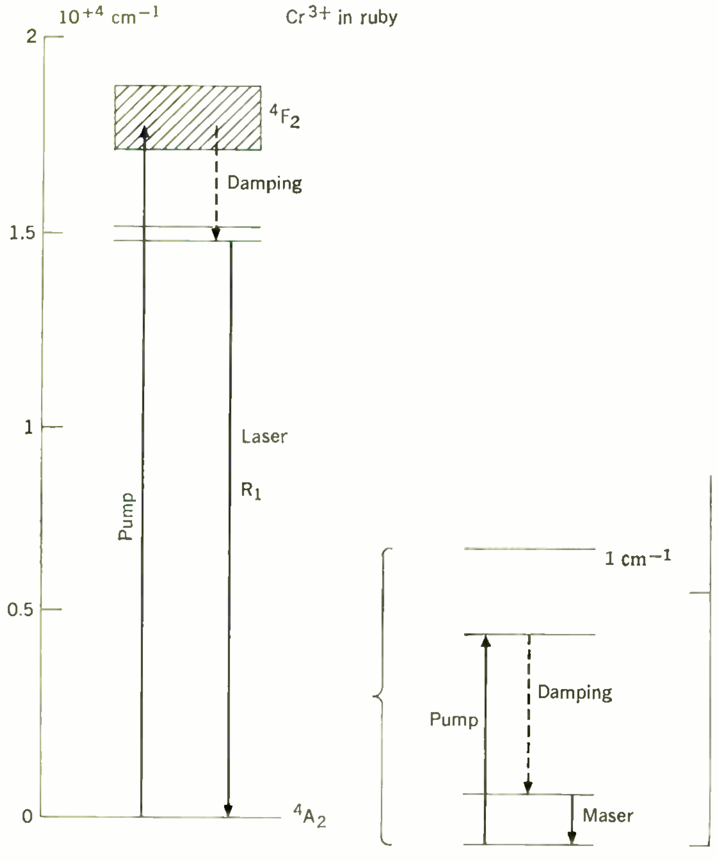
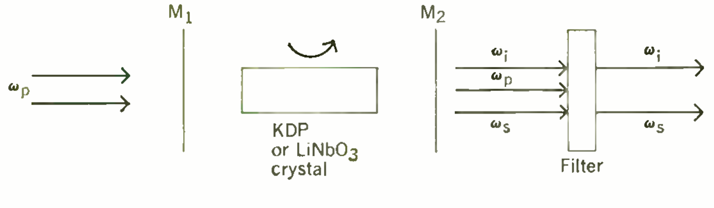


FIGURE 2. Energy levels of the chromium ion in ruby. The four levels of the spin quartet ground state $^4A_2$ are shown on an enlarged scale at right. The pumping and damping mechanism essential for maser operation was first proposed and demonstrated at microwave frequencies. The same principle is used in the ruby laser. The different energy levels at light frequencies are shown at left.

FIGURE 3. Schematic of optical parametric down-converter. Mirrors M_1 and M_2 transmit the pump laser beam at ω_p. Depending on the orientation and temperature of the crystal, frequencies ω_i and ω_s are generated, such that $\omega_i + \omega_s = \omega_p$.



The new devices

Very succinctly, but not too inaccurately, one might say that the quantum electronics of lasers consists of doing at light frequencies what is already done at radio and microwave frequencies. Although the difference in time and spatial scale (the frequency 10 000 times higher, the wavelength 10 000 times smaller than for microwaves) makes the physical embodiment and design of such items as coherent tunable light oscillators, light amplifiers, light modulators and demodulators, harmonic generators, light flip-flop and logical circuits, parametric converters, and light waveguides radically different from their radio and microwave counterparts, the underlying principles of Maxwell's theory and quantum mechanics are the same at all frequencies. I have chosen two recent examples to convey an idea of what lies ahead.

My first example is the tunable light oscillator, which is shown schematically in Fig. 3. An intense green light beam, obtained by harmonic doubling of a neodymium-glass laser beam in an oriented crystal of lithium niobate, is sent into another crystal, which has reflective coatings for infrared light. In this second crystal the green quanta split up into pairs of smaller quanta. The exact frequencies of these smaller quanta depend sensitively on the geometry and optical properties of the crystal. Data obtained for signal and idler by variation of the temperature of a niobate crystal are shown in Fig. 4. A tunable coherent-light oscillator results. To date, operation between 7800 angstroms and two micrometers has been achieved. With a crystal of potassium dihydrogen phosphate (KDP) and a pump beam in the ultraviolet, tunable laser-like beams in the visible region of the spectrum have been obtained. If such an oscillator were perfected and could be obtained commercially as easily as a radio signal generator, for example, the field of optical spectroscopy would be revolutionized. One would dial the wavelength of the laser-like beam.

The second example is an intense light pulse generator with pulse durations as short as a few picoseconds. The

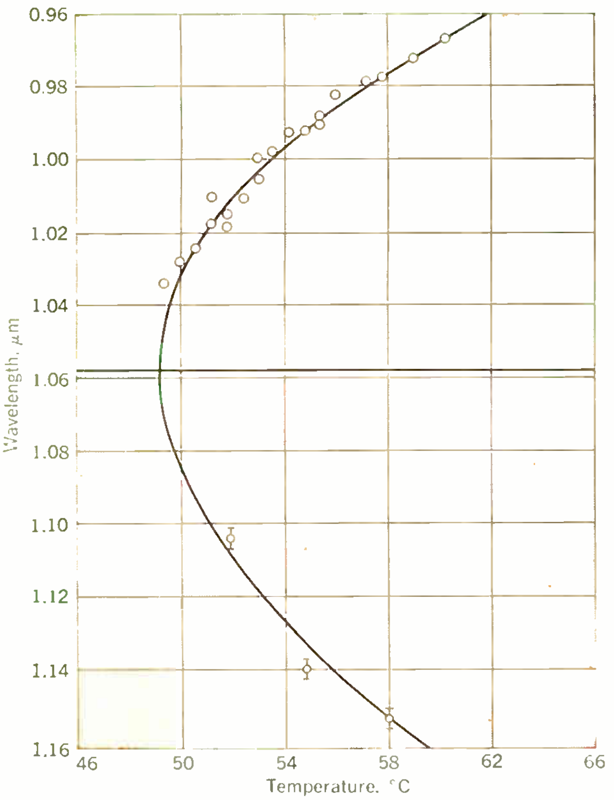


FIGURE 4. Wavelengths of parametrically down-converted light as a function of the temperature of an $LiNbO_3$ crystal.

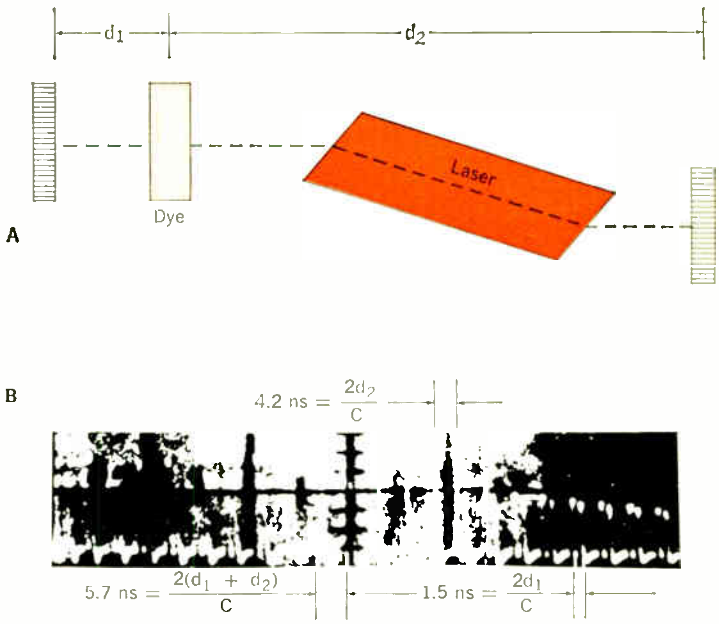


FIGURE 5. The operation of a Q-switched mode-locked neodymium-glass laser. (A) Experimental arrangement. (B) Oscillogram of the output at a sweep speed of 2 × 10^{-8} seconds per division.

principle of operation is based on mode locking of a *Q*-switched laser. A saturable dye solution is placed in the laser cavity, as shown in Fig. 5. After one round trip through the laser rod the light is so amplified that it bleaches the filter. The population in the excited state of the dye becomes equal to the population in the ground state. After the intense amplified light pulse has passed through, the dye returns rapidly to its normal absorbing state. The process repeats after another round-trip time of light in the cavity. The duration of the extremely short pulse is measured by the scheme shown in Fig. 6. The light pulse is split into two pulses of orthogonal polarization, which are recombined in a piezoelectric crystal of such orientation that second-harmonic light is formed only when both polarizations are present. The second-harmonic generator of light acts as an extremely fast coincidence counter. If the light path in one arm is increased by a few millimeters, no second harmonic is generated. The pulse duration can then be measured, and is found to be shorter than 10^{-11} second. This clearly opens new vistas for time measurement and ultrafast switching. The power flux density in such pulses can be staggering, reaching 10^{12} watts/cm², corresponding to light field amplitudes of 10^{8} volts/cm. If such fields persisted for longer times, which still means only 10^{-9} or 10^{-8} second, the material would break down mechanically and electrically. Somewhat less intense pulses of about 10^{-8}-second duration from powerful solid-state or pulsed gas lasers are used for ranging. Light radars with extremely high Doppler resolution and fast response at short-distance ranging appear promising, but again the competition from extremely sophisticated microwave radar techniques is hard to beat.

Continuous-wave power levels of a few kilowatts have been attained with CO_2 lasers at a 10.6-micrometer wavelength. An infrared beam of this type of only 100 watts can, when focused, easily cut through a standard two- by four-inch board or vaporize the most refractory materials. Although it is unlikely that a carpenter would use a laser in his daily work, the cutting of emery paper in a factory that at present rapidly wears out its cutting tools is a possible application. Laser beams have also been used for drilling holes in diamond dyes for wire drawing and for precision tooling operations, but electron-beam machining is a powerful competitor. Medical applications, such as retina welding and cutting of tissues with a laser beam, also are being tested.

When hot CO_2 gas is suddenly expanded, the excited vibrational state has a long enough lifetime that population inversion results with respect to rapidly depleted lower-lying rotational levels. This is the principle of gas dynamic lasers, which promise CW oscillation levels considerably above the kilowatt level.

Lasers are definitely well established as laboratory tools; they are in wide use for aligning and testing optical instruments, for demonstration and teaching, and for precision metrology. A laser slaved to a reference crystal, kept permanently in ultrahigh vacuum at liquid-helium temperature, may well provide the length standard of the future. The laser will transform Raman spectroscopy from a time-consuming tool of limited usefulness to an important analytical technique; for example, the hour-long exposures of Raman spectra on photographic plates are eliminated. Raman spectroscopy with gas-laser beams should have widespread application in analytical chemistry and solid-state physics.

It also appears fairly certain that the far-infrared region will be investigated more intensively and will be con-

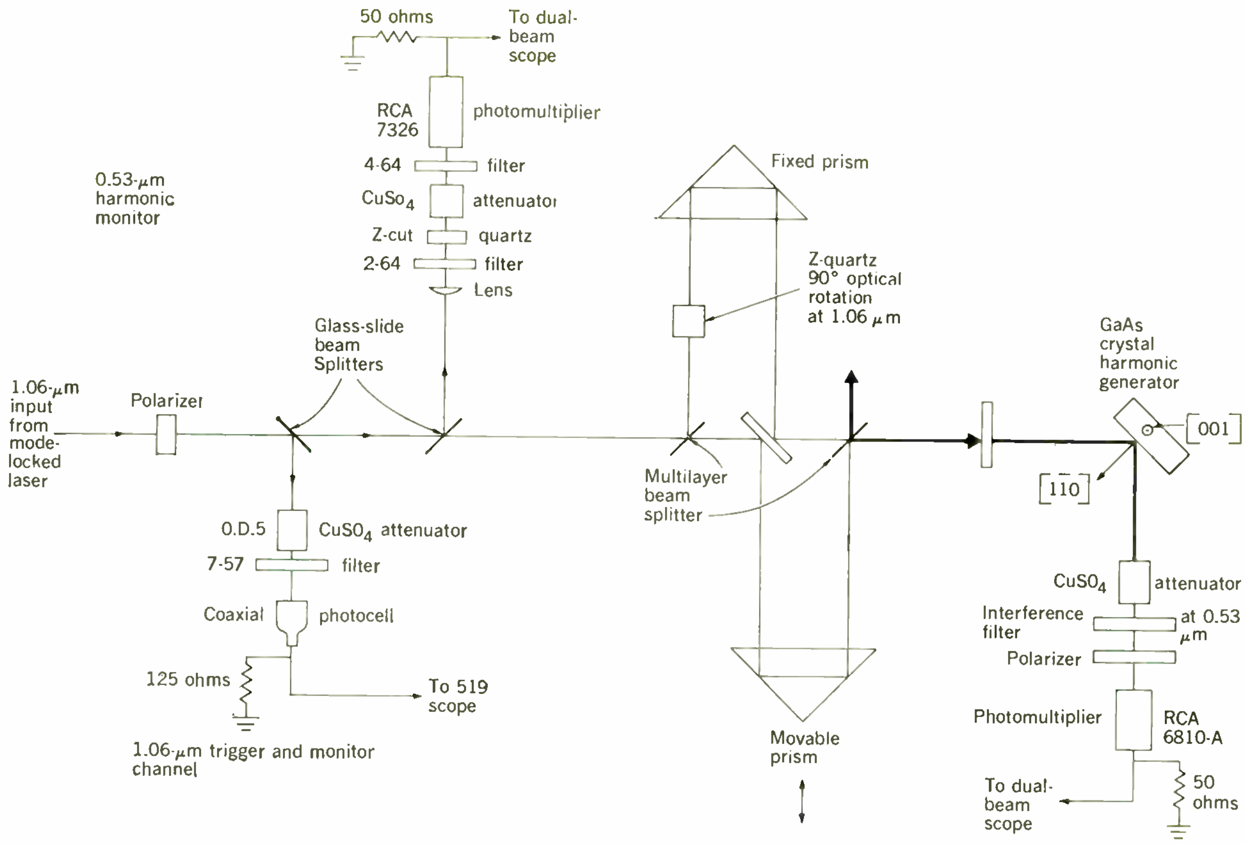


FIGURE 6. Diagram to measure the duration of light pulses shorter than 10^{-11} second.

quered by new techniques. The gap between the millimeter-wave region and wavelengths shorter than 100 micrometers is rapidly disappearing. Gas-laser sources and difference-frequency generation techniques are useful both as sources and as detectors.

Holography is another field for which the laser has opened many possibilities. Perhaps it will find useful application in pattern recognition and in storage of three-dimensional information as a Fourier transform. A bad spot in a photographic image will not spoil all bits of information completely; the Fourier transform of such a plate will still give a good image. It is too early, however, to tell how much impact holography will have on our society. Perhaps it will only be used as a tempting three-dimensional display for advertising purposes. Three-dimensional displays of airfield approaches in the cockpit of a jet liner with the correct viewing angle from the position of the aircraft would be a more interesting application. The ultimate dream of making visible three-dimensional X-ray pictures of crystals and molecules seems remote. How can one control the dimension to within a quarter wavelength for X ray during the photographic processing?

What of the future?

These are some of the new horizons that have come into view because of recent research, but they are horizons not yet attained. Even if the devices that seem promising are fully developed, the size of their impact on industrial technology cannot be prejudged, as past experience in other fields has shown.

Is the enormous increase in bandwidth offered by light as a carrier frequency in communications needed? For transmission in space the acquisition and aiming of the light beams pose formidable problems. In the atmosphere, rain, smog, fog, haze, snow, etc., make a light a poor competitor of microwaves. Can a system of enclosed tubes with controlled atmosphere and light repeater stations be built on a technologically sound and economically competitive basis?

The application of light in the computer area is also the subject of speculation. Superficially, it appears attractive to have fast switching, high storage density, direct visual display. Such developments would depend heavily on the availability of cheap, small, high-quality semiconductor lasers. If these were available, the entire organization of computers using them would probably be different. Could such a system compete with an existing and rapidly developing computer technology that thrives without lasers? The burden of proof is on the laser.

One should not expect the near future to bring sudden dramatic technological change. Rather, a gradual widening of the perspective and an increasing number of varied applications on a smaller scale in many different fields of endeavor appear a more likely course for the historical development. Even if full development is realized, the strength of the technological impact behind our new horizons is difficult to estimate. However, experience tells us that it would be remarkable if so many new scientific possibilities did not lead to some technological change.

Essentially full text of a paper presented at the Symposium on New Horizons in Science and Technology at the 1967 IEEE International Convention, New York, N.Y., March 20–23.

Cryogenic random-access memories

Cryoelectric memory systems for computers promise improved capacity, reliability, and speed. The memories are comprised of strip lines that display low characteristic impedance, high propagation velocity; and modest peripheral electronics requirements

A. R. Sass, W. C. Stewart, L. S. Cosentino

RCA Laboratories

Three-wire cryoelectric memory cells and the hybrid AB system organization that utilizes coincident-current selection are examined from the standpoint of batch fabrication requirements, redundancy, electrical parameters, tolerances, and noise immunity. These advances, demonstrated with experimental subsystems, are described in relation to previous work, and are shown to place cryoelectrics as a strong contender for achieving systems with capacities of more than 10^8 bits.

For the past decade, computer engineers have grappled with the problem of applying the phenomena of superconductivity in order to achieve a highly efficient central processor. The concept of an element displaying zero or full electrical resistance, depending on the magnetic field of an applied control current, was exciting from the standpoint of achieving a perfect switch. Combining this switch with shunt paths having zero electrical resistance gave rise to memory with an absence of half-select and delta noise, and logic with zero standby power and negligible state-change power. Moreover, the diamagnetic nature of superconductors and their adaptability to thin-film forms resulted in lossless, low-impedance lines for both devices and device interconnections.

The only threat to the supremacy of this new "cryoelectric" technology was the need for a low-temperature environment. It seemed that the technical problems associated with this situation were not as pressing as the economic: The capital investment in a refrigeration system for the processing unit could conceivably price cryoelectrics out of the market. It was apparent that the answer to this situation would be to introduce a cryoelectric system that would be (1) larger than any existing system, so that the cost per element would be competitive, and (2) able to perform functions unrealizable with any other technology, so as to justify the total system cost. The cryoelectric technology seemed suited to a batch fabrication process by which arrays of these thin-film superconductive devices could be formed at very high densities. The method was basically one of thin-film deposition and pattern forming using stencils or photoetch processes.

Probably one of the first overall decisions to be made was exactly how much of the central processor should be cryoelectric. If both logic and memory were to be cryoelectric, there was the associated question of whether they should be fully integrated with respect to physical location so as to produce a highly efficient processor; at least one laboratory actively pursued this approach. Another approach, inherently more cautious, was the development of only a random-access memory, admittedly the heart of the processor, which can be made by fabricating many repetitions of a basic element and thus is ideally suited to a batch fabrication process.

Why did cryoelectrics, which displays the foregoing general electrical characteristics so amenable to computer systems, and which was explored using the foregoing approaches, fail to deliver a working system in the projected time schedules initially envisioned? Based on the answers to this question, a new approach has yielded a workable cryoelectric system whose feasibility has been experimentally demonstrated.

Although the basic phenomena of superconductivity in relation to cryoelectric computer systems were never in question, it became apparent with the passage of time that the details associated with the phenomena posed rather serious problems. This point is illustrated best by the most severe problem that faced not only cryoelectric workers who were developing the fully integrated processor but also those at the other end of the spectrum who were developing the random-access memory: the lack of sufficient control of the electrical thresholds of the devices under consideration. This situation was rather serious when coupled with the fact that a batch fabrication process was required to produce a system whose capacity would justify refrigeration cost.

An example of this problem can be seen in the random-access memory area. Discussion of this area is pertinent since we are dealing here with solutions of the problem for cryoelectric memories. The projected "break-even" capacity, including refrigeration cost, for a cryoelectric memory is approximately 10^7 bits. A memory whose capacity is in excess of 10^7 bits, and whose word lengths are of the order of 10^2 bits, requires some type of coincident-current selection. However, such a selection scheme, when coupled with peripheral electronics requirements, places limitations on the variations of memory cell performance throughout the memory. Coupled with the batch fabrication process by which cells are made and "wired" in large quantities simultaneously, the problem of performance variation becomes substantial. Here is a situation in which elaborate designs, conceived with only partial information, were rendered unworkable in the final analysis.

In the past two years basic advances have been made that have changed the situation sufficiently that cryoelectric technology must again be considered seriously by computer designers. Advances were made by first re-emphasizing that the advantages of the superconductive phenomena and the process technology are best suited to the achievement of a random-access memory rather than an integrated processor.

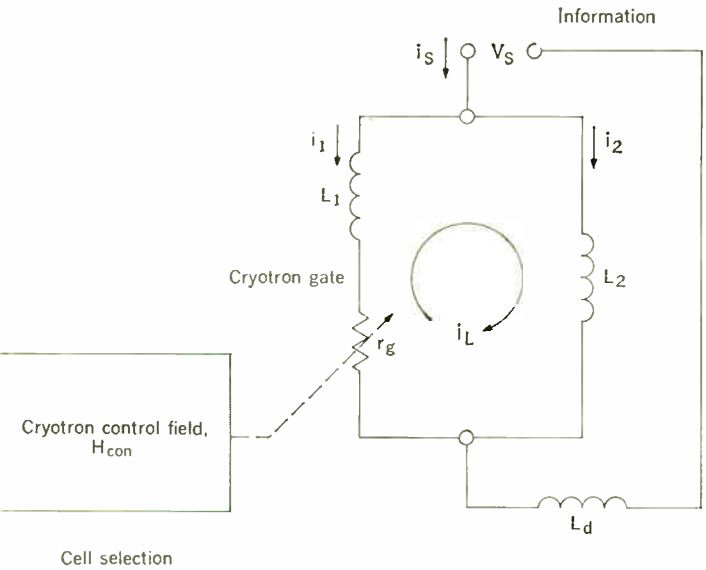
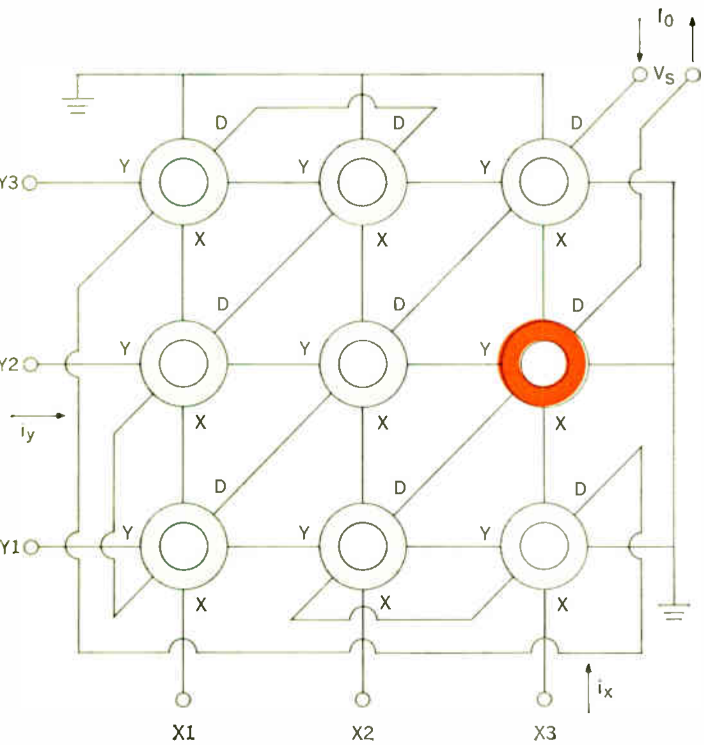


FIGURE 1. Schematic representation of a basic memory cell.

FIGURE 2. Schematic representation of a coincident-current CFC cell array. Cell X3, Y2 is shown selected by coincident currents.



Three-wire CFC memory cell

Basic cell. The basic cryotron memory cell consists of a persistent-current loop containing a cryotron gate along with sets of leads, one of which provides current to the loop and the others of which produce a variable magnetic control field for switching the gate between the superconductive and normal states. This is shown schematically in the equivalent circuit of Fig. 1. For the purpose of this discussion, the gate resistance r_g is zero when the control field H_{con} is less than some threshold value H_t, and is nonzero when $H_{con} > H_t$. The change of approximately 3 percent in L_1 produced by switching the gate normal considered by Meyers[1] is neglected in this model. The principle involved in obtaining memory lies in the well-known fact that the total flux linking a loop (or, more precisely, London's fluxoid) remains constant as long as the entire loop is superconducting. This fact is implicit in the network equations for the equivalent circuit of Fig. 1. Although three distinct modes of information storage have been proposed by Sass *et al.*,[2] the convention used here is that binary "one" or "zero" corresponds to presence or absence, respectively, of i_L.

Readout is performed destructively by applying a control-field pulse $H_{con} > H_t$. The presence or absence of a sense signal at the information terminals corresponding to the decaying of i_L signifies that a "1" or a "0" was stored in the cell. The information is rewritten by applying a pulse of amplitude I_0 or zero, corresponding to a "1" or "0," to the information terminals in coincidence with the control field. If the control pulse ends before the information pulse of I_0, a loop current of $I_0L_2/(L_1 + L_2)$ remains circulating in the loop. Note that a transient voltage (write noise) appears across the information terminals whenever a pulse of i_s is applied.

One of the major advantages of this basic cell stems from the higher degree of structure it contains in comparison with two-wire devices proposed previously. With the existence of well-defined current paths and switching fields obtained in the shielded crossed-film cryotron (CFC), the equivalent circuit of Fig. 1 provides a highly accurate representation of the memory cell.

The peak sense signal at the cell itself is of the order of 15 μV, with a pulse width of about 100 ns at its base.

Considerable passive step-up, of about 15:1, is obtainable in matching the low-information line impedance to that of the external circuitry. The bandwidth associated with this step-up is compatible with the system timing to be discussed.

Array operation. The CFC cell is applicable to a three-wire coincident-current organization[2] with the following features:

1. Memory is in the form of a persistent current in a superconductive loop.
2. The selection of one cell in an array is accomplished by the coincidence of currents in an *x* line and a *y* line that "intersect" at the physical location of the cell. The threshold amplitude for performing the selection is ideally independent of the stored current level.
3. The *x*- and *y*-selection currents do not contribute to the stored loop current; the information current is applied independently.
4. Slight variations in film thickness do not affect selection current levels.

These characteristics are obtained by connecting in series the "information" terminals of all the cells that comprise the same digit position in the memory words. The control fields for the cryotrons are formed by superposing two insulated control lines—one for *x* current and one for *y* current—of the same width but at right angles to the cryotron gate. These lines are arranged electrically in a two-dimensional matrix so that any one cell along the digit line is selected by energizing a unique pair of *x* and *y* drive lines through a decoding network. Such an arrangement is depicted schematically for one digit position of a nine-word (3 × 3) array in Fig. 2.

The operating range of such an array can be shown by means of a family of cryotron characteristics, as in Fig. 3. Here, the sum of the control currents $i_x + i_y = j$ is plotted on the abscissa as a function of the gate current i_g (common to all cells) on the ordinate. For any value of i_g, which is now the digit current, the sum of the threshold control currents for all cryotrons in the array lie between J_{upper} and J_{lower}. This band of values includes the effects of possible material inhomogeneities, the widths of the field transitions in individual gates (the range over which partial switching occurs), and line-width variations in the cryotron gates and control lines themselves; that is, the curves represent experimentally measurable quantities. It is apparent that disturb-free coincident-current operation is obtained for *x*- and *y*-selection currents lying between the values $I_{min} = J_{upper}/2$ and $I_{max} = J_{lower}$.

I_{max} and I_{min} are plotted against i_g in Fig. 4 with the shaded area depicting the operating region. Note that partial switching and half-select noise is nonexistent in the operating range defined here. Choosing $i_g = I_0$ in Fig. 4 provides an operating point for which the required cell-selection current is essentially independent of the stored information current, thus satisfying criterion 2 mentioned previously. Criteria 1 and 3 are satisfied inherently in the conception of the basic cell and the use of the noninductively coupled crossed-film control lines. The films are of sufficient thickness that the critical field is that of the bulk material.

The separation of the cell-selection process and the information-storage process to independently controllable excitation sources provides an additional degree of freedom in choosing operating conditions over two-wire cryoelectric memories. If all the cryotrons in the array were identical with perfectly sharp transitions, the shaded area of Fig. 3 would shrink to zero width and give $I_{max} = 2I_{min}$. The selection currents could be set at the mid-range value $(I_{max} + I_{min})/2$ and be allowed to deviate $\pm 33\frac{1}{3}$ percent, the theoretical maximum tolerance.

Bridge and loop cell. An early CFC cell that meets the three-wire requirements previously cited is the bridge cell proposed by Ahrons[3] and further investigated in

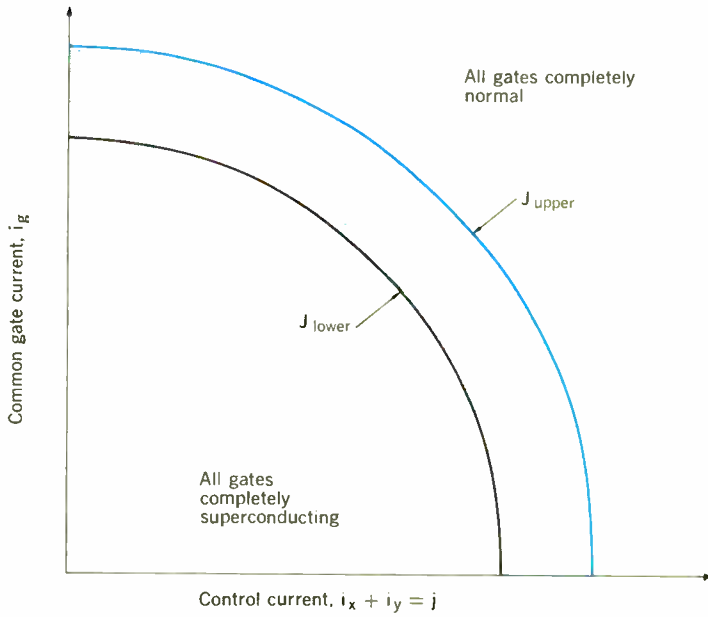
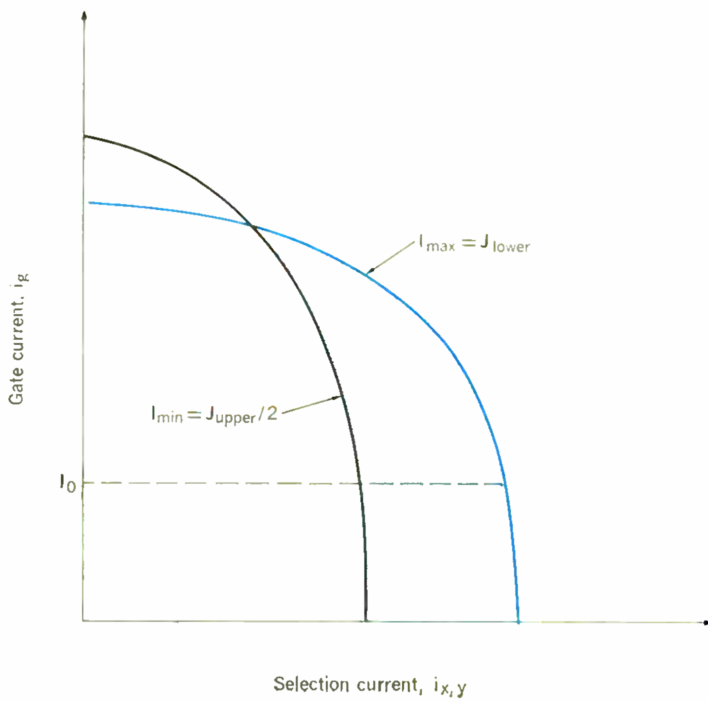


FIGURE 3. Schematic diagram of the characteristic curves of all CFCs in an array.

FIGURE 4. Diagram showing maximum and minimum permissible array selection currents as a function of common gate current.



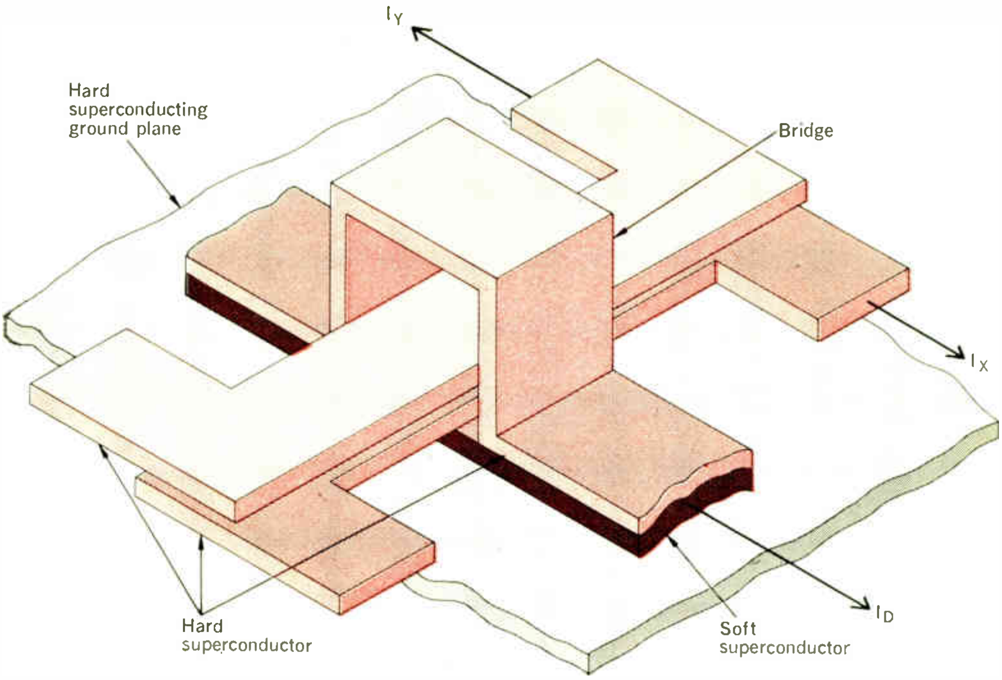
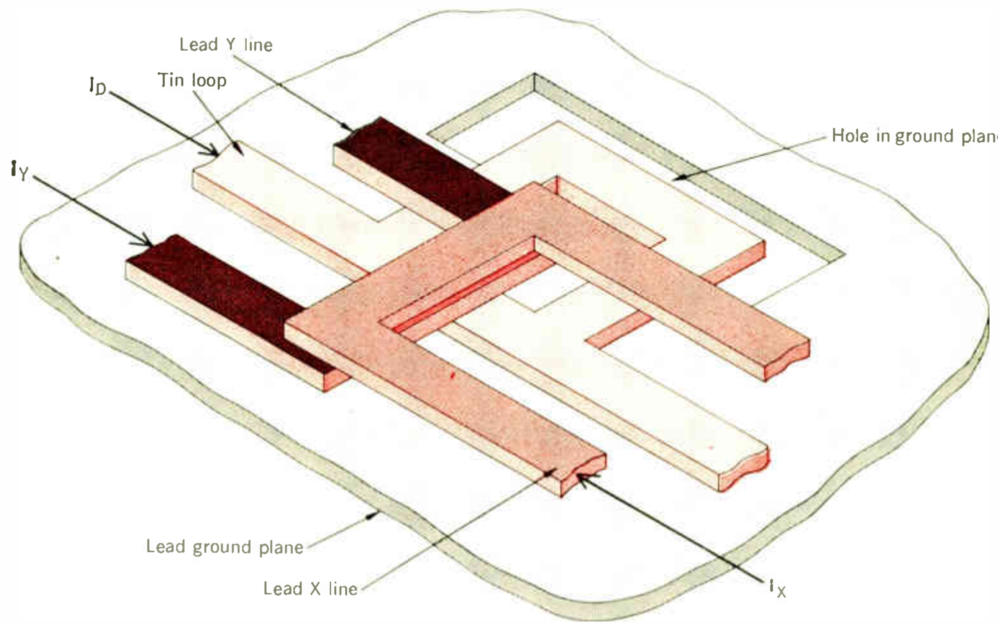


FIGURE 5. The bridge cell.

FIGURE 6. The loop cell.



bit-organized arrays by Sass *et al.*[2] This cell, shown in Fig. 5, has a storage loop which is perpendicular to the plane of the substrate. The upper branch of the loop is made of lead and the lower branch of tin. A superconducting ground plane exists beneath the cell. Note that for the bridge cell, the equivalent circuit of Fig. 1 applies with $L_1 = 0$, and all the digit current is stored.

Since the sense signal amplitude is directly related to the inductance of the upper branch, this inductance should be as large as possible. This means that the bridge should be as high as is practicable. It is also most important that the upper lead line make superconducting contact with the lower tin line, otherwise stored current will decay to zero. Both of these factors create problems if a process is to be used involving evaporation of metal layers with subsequent photoetching of desired patterns.

To accommodate this technology, a three-wire cell that is more easily fabricated was designed and developed by Gange.[4] This memory element, called the loop cell, is shown in Fig. 6. The schematic circuit of Fig. 1 is directly applicable to this cell. It is desirable to have $L_2 \gg L_1$ so that most of the digit current will be stored. It can be seen that a large value for L_2 is achieved by the use of a hole in the ground plane beneath the high-inductance

I. Experimental results obtained from three-wire memory cells

Cell Type	Line Width, μm	Number of Cells on Substrate	Packing Density, cells/cm²	$I_{x,y}$ mA	I_D, mA	Peak Sense Signal, μV
Bridge	127	8 × 8 = 64	6.9	400–600	100	200
Loop	63.5 (drives) 127 (digit)	8 × 8 = 64	13.2	200–300	50	500
Loop	50.9	8 × 8 = 64	1010	170–260	20	120
Loop	50.9	16 × 64 × 6 = 6144	1010	170–260*	20*	120*
Loop	25.4	8 × 8 = 64	2020	100–150	10	40

* Data pertain to a random sampling of 384 cells. Further data on a stack of such planes are available.[6]

branch of the loop. As a result, the insulating layers can be deposited in sheets rather than spots. The entire storage loop is now etched from a single tin layer so that no difficulties arise with superconducting contacts. Also, the number of layers needed per sample is reduced. The modification of the three-wire cell for fabrication purposes does not appreciably affect operating tolerances or performance. Good results have been obtained with the loop cell as well as the bridge cell.

Hundreds of samples with three-wire memory cells have been made and thoroughly tested. The results have been in accord with the models and equivalent circuits. Tolerances were excellent and typically averaged ±20 percent for planes with large numbers of cells. As cell size has been shrunk and cell packing density has been increased, drive currents and sense signal amplitude have scaled down as expected, while tolerances have remained high.

Table I lists typical experimental results obtained on many different specimens. $I_{x,y}$ is the magnitude of the current that can be applied to any x and y line without deteriorating the performance of any cell on the plane. The values for peak sense signal voltage pertain to the output of a 1:10 step-up transformer located at the sample.

Reproducibility was also excellent from sample to sample and was illustrated by the fact that a number of individual planes could be stacked with only a small reduction in overall tolerances.

Statistical studies indicate that, with present fabrication uniformity, large commercial-size arrays will have operating tolerances of about ±12 percent.

Random-access memory system

Hybrid AB system. In order to tailor a memory system specifically to the unique properties of the three-wire cryoelectric cell, the hybrid AB system was devised by Gange.[4,5] The name is derived from the fact that the organization is word bit in character rather than being purely bit-organized or word-organized. Another important feature of this design involves the use of room-temperature decoders rather than decoders deposited directly on the substrate. Because of the special properties of a cryoelectric memory, to be discussed later, only two such decoders are required for the complete system. In addition, many other advantages are gained if the decoder is not included on the substrate. The main benefits are that fabrication and evaluation of memory planes are greatly simplified, thereby increasing yield; standard commercially available decoders can be used at room temperature; many more memory cells can be put on one substrate; cycle times are no longer limited by cryotron decoding trees; and, finally, perfect planes are no longer absolutely necessary. Concerning the last point, extra elements can be included on the plane and only the good portions actually used. The number of input wires and interconnections needed without a cryotron decoder is greatly increased but still falls within the capability of present technology.

Recently, a laboratory model of a 14 000-bit hybrid AB system, employing loop cells as storage elements, was experimentally demonstrated by Gange, Nagle, and Scheible.[5]

Preliminary checkout of a 262 144-bit memory plane having a capacity of 4096 words, 64 bits per word, using the AB three-wire organization is under way. The size of the unit is approximately 10 × 14 cm.

The hybrid AB system employs three-wire cryoelectric memory cells as basic elements. Thus a write operation requires current coincidence of an A drive line, a B drive line, and a D digit line, whereas a read operation requires only AB current coincidence.

The A lines provide the word dimension and thread from plane to plane through a memory stack. Let A_T be the total number of such lines. The B lines are segmented into d parts and only B_T/d lines appear on each plane where B_T is the total number of B lines in the memory. The B lines on each plane thread through a number D_p of digit lines on a plane, which is equal to the number of digits in a word. Each D line includes a number of digit strips equal to B_T/d, which are serially connected together. Each digit strip is orthogonal to the A lines and contains a number of cells equal to A_T. Each B line threads through one strip of a D line, then turns and threads through one strip of the next D line, and so on through the plane. Thus the number of words on a plane W_p will be equal to $A_T \times B_T/d$. The total number of words in the system will be $W_T = P \times A_T \times B_T/d$, where P is the total number of planes. Similarly, the number of bits on a plane N_p will be equal to the number of bits on a strip A_T, multiplied by the number of D lines on a plane d. Thus $N_P = A_T \times B_T$. The total number of bits in the memory will then be $N_T = P \times A_T \times B_T$.

Since each D line on a plane is associated with one digit of all words on that plane, corresponding digits on different planes can be serially connected. Write current is then applied serially to all D lines in one group, but the output sense signal will come from only one member of the group. The number of D lines that can be so grouped will be limited by loading effects and speed considerations.

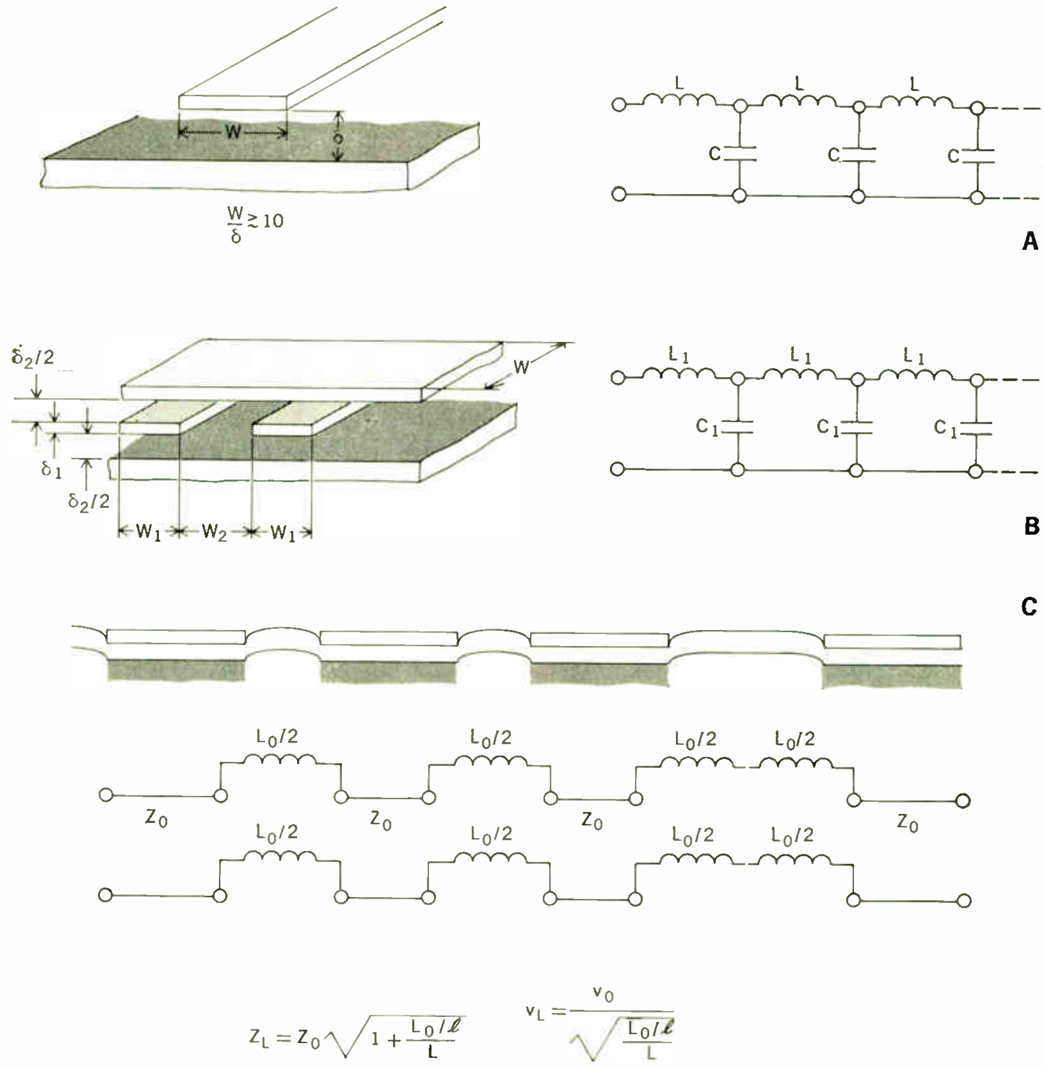


FIGURE 7. (A) Superconducting strip lines. (B) Superconducting strip line crossing over other superconducting strip lines. (C) Interconnected superconducting strip lines; selection and digit lines are serially connected from one substrate to the next. L_0 is inductance of interplane connections.

Extra A, B, and D lines can be included on a substrate and the good portions selected after a preliminary check-out. Thus, redundancy is an important feature of the hybrid AB system and 100 percent defect-free planes are no longer required.

If we assume an A-line decoder of a levels, then $A_T = 2^a$. Similarly, $B_T = 2^b$ with a B-line decoder of b levels. Then the capacity or total number of bits in the memory is $P \times 2^{a+b}$, with a total number of words $W_T = P \times 2^{a+b}/d$. The number of wires into the system is minimized when $P = d$; for this case, $a = b = m$ and $A_T = B_T$. Then $N_T = P \times 2^{2m}$ and $W_T = 2^{2m}$. For any memory, the capacity and number of digits in a word will then determine the system structure. For a 10^8-bit memory with a 100-digit word and $P = d$, the numbers of wires and interconnections have been estimated to be about 5000 and 250 000 respectively.[4]

General electrical characteristics. In order to examine the electrical characteristics of a superconducting memory system, some rather general relationships regarding inductance, capacitance, characteristic impedance, and phase velocity should be considered. The results can be applied here to the AB system; it will be seen that they also pertain to other organizational layouts since they are rather basic in nature.

Since the digit lines and the selection lines are thin films over a superconducting ground plane, it is useful to consider first the electrical characteristics of the basic structure of a superconducting strip line over a ground plane; see Fig. 7(A) and (B). An insulation thickness δ of 10 000 Å will be assumed. (London field penetration in the superconducting films will be neglected.)

Inductance/square $= L_s \approx 10^{-12}$ henry $\qquad L_s \propto \delta$
Capacitance (2-mil-wide line)/square
$= C_s \approx 1.5 \times 10^{-13}$ farad $\qquad C_s \propto 1/\delta$
Resistance/square $= 0$
Characteristic impedance $= Z_0 \approx 3$ ohms $\qquad Z_0 \propto \delta$
Phase (group) velocity $= v_0 \approx c/2.5$

where the relative dielectric constant of the insulator (SiO) is taken to be six.

There are several impressive features of the preceding results as applied to the hybrid AB system. An analysis of the electrical characteristics of the system is based on equivalent circuits such as those in Fig. 7. The results of the study are as follows:

1. Since the transmission lines are lossless and the dielectric is fairly lossless, pulse degradation is negligible. (The case in which lines on many substrates are interconnected does not alter this conclusion.)

2. Characteristic impedances are extremely low (of the order of 10 to 20 ohms when loading effects of the interconnections are included), and thus the requirements on peripheral driving sources are modest. It is interesting to note that for a 200-mA drive current the back EMF across the line in the memory having the highest characteristic impedance is only 4 volts. These results are of paramount importance, since they are independent of system size and demonstrate the feasibility of the AB system.

3. Reasonably high propagation velocities are the only limit on system timing, because of the cell characteristics and cases 1 and 2.

Actually, this last comment is the most powerful of all—that is, the only electrical characteristic limit to system size is the desired cycle time, not half-select noise, delta noise, or excessive requirements for peripheral electronics. Thus, from Fig. 8, the worst-case cycle time of the memory is

$$T_T = 2T_p + 3T_s + 2T_c$$

where T_p is the longest propagation time along a selection line, T_s is the longest propagation time along a digit line, and T_c is the cell time constant.

To reiterate, the cryoelectric memory is made up of strip lines, which, though interconnected from plane to plane, display low characteristic impedance and high propagation velocity, and require modest peripheral electronics. Therefore, propagation velocity is the only real limit to memory cycle time. Typical cycle time for the 10^8-bit AB system mentioned earlier is approximately 1 μs.[5]

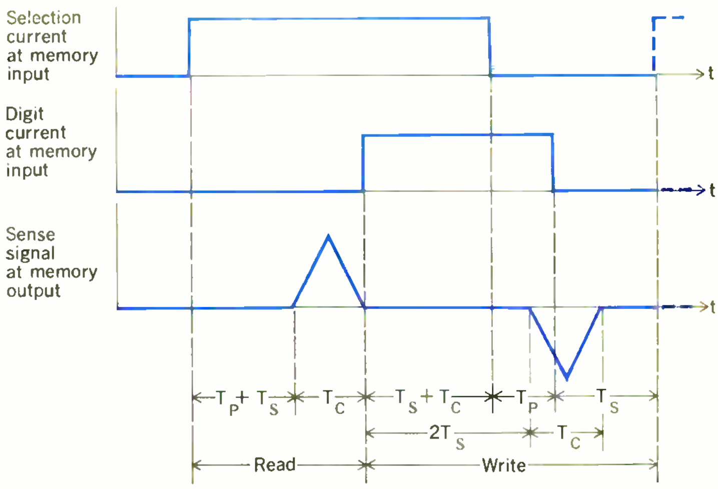


FIGURE 8. Worst-case cycle time.

Noise considerations

Deterministic noise. One of the features of the three-wire cryoelectric memory cell is that the sense signal is picked up from the terminals on the array into which digit current is fed, as shown schematically in Fig. 9(A). The transient produced by a pulse of digit current is considerably larger than the sense signal and, in general, will overload the sense amplifier. Since it is necessary for the amplifier to recover before the next read operation occurs, the recovery time may constitute a significant additional portion of the read–write cycle time.

One solution to this problem is obtained at the expense of adding a separate sense line to the memory array. It is laid out in a way such that coupling of adjacent cells (or groups of cells) on the digit line is of opposing polarity, as shown in Fig. 9(B). The resulting sense signal inversion for half the cells is of no consequence, and the write noise is bucked out on a cell-by-cell basis. An experimental test of this scheme on a small 64-cell array showed that the write noise was reduced below the limit of resolution of the test equipment, which was a factor of 10^3. An alternative scheme, which does not involve the addition of another line to the array, provides for write noise cancellation in the coupling transformers at the terminals of the digit line.[2] The digit line is divided into an even number of groups containing equal numbers of cells, and the transformer secondaries are connected in series opposition as shown in Fig. 9(C). Well-matched transformer cores are required in this case. Other passive cancellation techniques of a similar nature are, of course, possible. The point to be made here is that the existence of write noise in the basic three-wire system presents no fundamental limitation to the operation of the system.

When cell-selection currents are applied to the A and B drive lines at the beginning of each read–write cycle, the possibility arises that stray coupling may produce a read noise transient in the sensing circuit, from which the sense signal must be discriminated. In-mode coupling is effectively eliminated by suitable array layout and holder design. Capacitances between metallic layers in the arrays are not negligible, however; and there is inductance in the lines from the stack to the peripheral electronics at room temperature. It is instructive to consider the order of magnitude of balancing required to eliminate the common-mode component. Consider, for example, that the selection lines each carry 200 mA, the sense signal to be detected is 100 μV at the 50-ohm input of a single-ended sense amplifier, and that a signal-to-read noise ratio of 4 to 1 is desired. The stray current into the hot side of the amplifier must then be 118 dB below the drive current. Novel solutions to this problem are being considered; circuitry that appears to satisfy these requirements has been conceived and tested on a moderate scale.

Random noise. Since cryoelectric memories of the type described here give sense signals smaller than those obtained in conventional core memories, consideration has been given to the effect of random noise in the system. The question to be answered is at what frequency random noise spikes occur with sufficient amplitude to be mistaken for a sense signal. The results of an analysis by Blatt[6] of this type of problem relating to magnetic film memories can be easily extended to the situation of interest here.

The major source of random noise in the sense signal path is generated in the first stage of the sense amplifier; the remaining stages, up to the threshold detecting element, are considered to constitute a linear low-pass

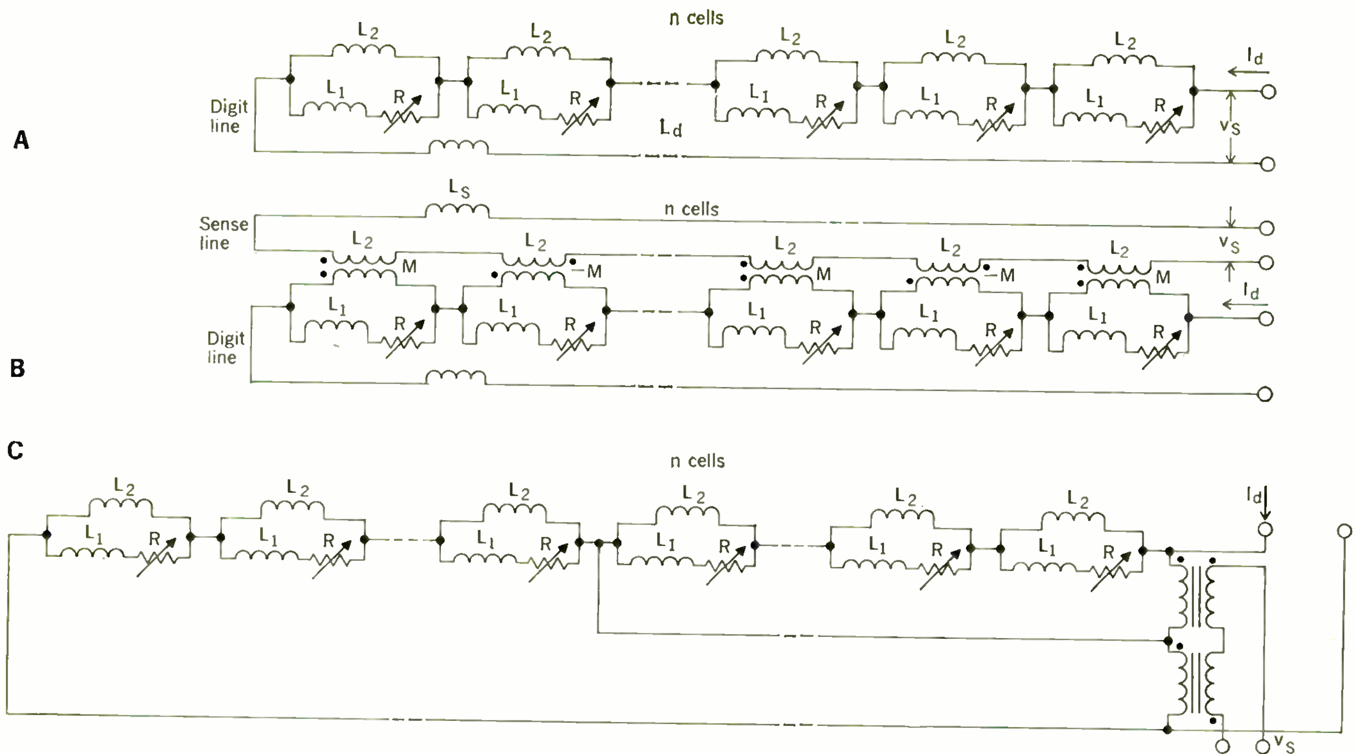


FIGURE 9. (A) Schematic diagram of n cells along a digit line. Cell-selection lines are not shown. (B) Write noise cancellation using separate digit line. (C) Write noise cancellation using center-tapped digit line.

filter, and the noise contribution from the normal cryotron gate in the cell at 3.5°K is negligible. The signal-to-noise ratio at the input to the threshold detector determines the mean time between errors (MTBE) as given by the following expression:

$$\text{MTBE} = \frac{1}{fm}\left[\frac{2}{1 - \text{erf}\left(\frac{\alpha}{\sqrt{2}}\right)}\right]$$

where f is the number of readouts per second, m is the number of digits interrogated in parallel, α^2 is the signal-to-noise ratio, and MTBE is in seconds. Limiting the amplifier frequency response to the reciprocal of the sense signal duration gives $\alpha \approx V_m/\sigma_0$, where V_m is the peak sense signal amplitude and σ_0 is the rms noise voltage of the amplifier. For an MTBE of 10 000 hours (over a year) or better, V_m/σ_0 should be equal to or greater than 8 in a memory in which 100 bits are read in parallel at a 1-MHz rate. If we assume a 100-μV sense signal at the input to a sense amplifier with an upper cutoff at 5 MHz, the rms noise voltage should be 12 μV or less. This figure is well within the capability of existing circuitry.

Conclusion

The design philosophy for memory cells that are insensitive to materials variations has been examined. It has been shown that this philosophy can be applied to a cell and memory organization suited to a batch fabrication process. The device advantages of superconductivity, which have been known from the outset of cryoelectric research, can now be applied to the realization of a large-capacity cryoelectric memory. Although such a memory has not yet been produced, the important cornerstone concepts have been experimentally demonstrated on subsystems. There remain many production and systems problems that must be solved in order for the final goal to be achieved; however, these problems are not of a fundamental nature.

Thus the cryoelectric random-access memory is a very real contender for large-capacity memory applications and deserves the attention of computer development engineers.

This article is a revised version of a paper that was presented at the 1967 International Magnetics Conference, Washington, D.C. April 5–7. The original paper will appear in the September issue of the IEEE Transactions on Magnetics.

The authors are grateful to L. L. Burns and J. J. Carrona for their constant encouragement throughout the course of the work described here, and to R. A. Gange and H. Scheible and the RCA Cryoelectric Devices Laboratory for communicating to them the results of their research prior to publication.

REFERENCES

1. Meyers, N. H., "Switching speed and dissipation in fast, thin-film cryotron circuits," *Proc. IRE*, vol. 50, pp. 2452–2464, Dec. 1962.

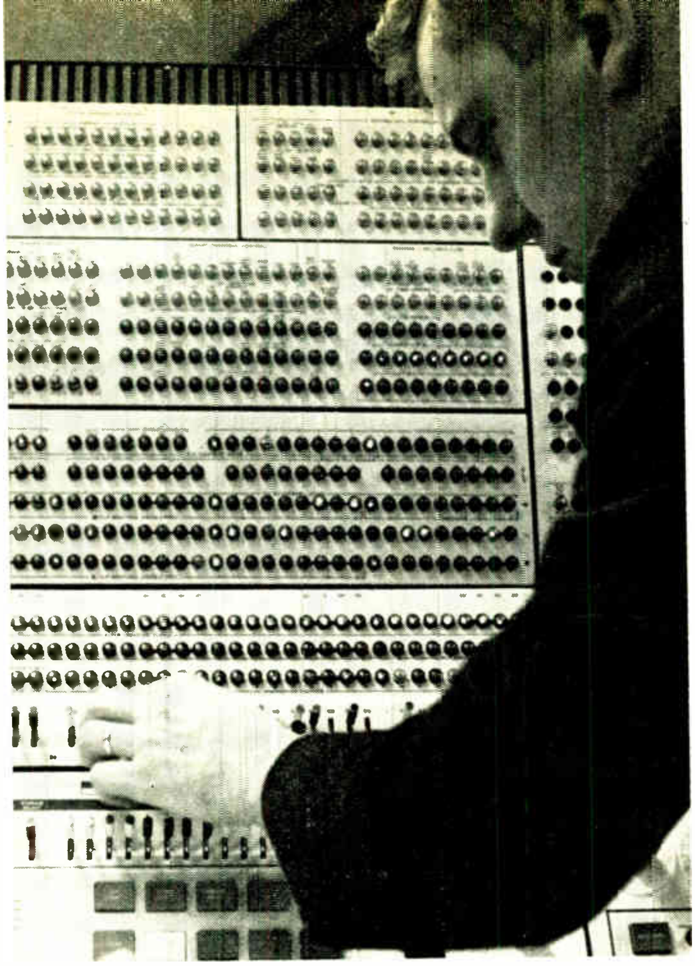
2. Sass, A. R., Nagle, E. M., and Burns, L. L., "Three-wire cryoelectric memory systems," *IEEE Trans. Magnetics*, vol. MAG-2, pp. 398–402, Sept. 1966.

3. Ahrons, R. W., "The bridge cell—a new superconductive memory cell for random-access word-organized memories," *Proc. 1965 Intermag Conf.*, pp. 9-1-1–9-1-5, Apr. 1965.

4. Carrona, J. J., Gange, R. A., Scheible, H. G., and Space, E. V., "Manufacturing methods for cryoelectric memories," *Tech. Rept. AFML-TR-66-106*, Apr. 1966.

5. Gange, R. A., "Taking cryoelectric memories out of cold storage," *Electronics*, vol. 40, pp. 111–120, Apr. 17, 1967.

6. Blatt, H., "Random noise considerations in the design of magnetic film sense amplifiers," U.S. Gov. Rept. AD605323, Aug. 1964.



Maintenance console of a typical large data-processing system.

Diagnostic engineering

The growing size and complexity of computer systems has created the need for a particular type of engineer—the specialist capable of putting into practice the various techniques for detecting and isolating system errors

John Dent *International Business Machines Corporation*

One of the formidable problems addressed by diagnostic engineering is to develop a means for verifying the design of a sophisticated system having over 100 000 interconnected logical circuits. After the system is designed, built, and tested, there is still the problem of maintaining it in good working order. The optimum strategy for testing, detecting, and isolating malfunctions must be found. Perhaps the most important problem is how to design a system in such a way that it can be tested automatically, so that errors can be detected and isolated in a reasonably optimum manner.

The phenomenal growth of the data-processing industry in the past decade has created many new career opportunities (and demands) in the various computer technologies. One of these demanding professional specialties is diagnostic engineering, sometimes referred to as diagnostic programming. Diagnostic engineering includes the disciplines of programming, logic design, and system design, as well as aspects of reliability, maintainability, and human factors. This article discusses some of the fundamental concepts of diagnostic engineering as it applies to all data-processing systems. It does not cover in detail some of the additional requirements unique to time-shared and real-time applications.

Manually testing extremely complex systems is so time consuming that it is, for the most part, out of the question. Computer programs—which are sequences of coded instructions (commands), in the computer storage, that control the operation of the system—provide a means for automatic testing. These programs are the primary tools of diagnostic engineering for the detection and isolation of system errors.

A diagnostic program is an automated procedure, defined in the form of computer instructions, for testing and/or analyzing results. Individual tests are arranged in a sequence, and each test result determines which point in the sequence will be executed next. Logical analysis is accomplished automatically in the same way. A predetermined "decision tree" is executed via coded instructions.

The term "error" is used in this article to refer to any deviation from the expected operation or results, including design errors, fabrication errors, component malfunctions, and, in some cases, operator or program errors. It would be very convenient to categorize these errors, define each category, and develop separate detection and isolation procedures. Unfortunately, the distinction among the various types of errors is part of the diagnostic problem. Where does one start when a complex system does not operate in the expected manner? This question is the beginning of a diagnostic procedure that encompasses all categories of possible errors.

The life cycle of a data-processing system begins with an exhaustive series of tests. Each piece in a complex system must be tested and, in addition, the various pieces must be tested for proper operation with each other; finally, the system must be tested in a manner simulating the typical operational use of the system. Thus various levels of testing are implied. The two predominant levels are (1) unit tests, which involve the testing of single units operating with the central processor, and (2) system tests, which involve the testing of several or all units operating concurrently in a typical or simulated operating environment.

Generally, unit diagnostics are written for each input/output (I/O) unit and the various logical subdivisions of the central processor. Typically, a unit diagnostic begins its test under a controlled set of conditions (that is, one I/O device at a time) in order to isolate errors more effectively. The testing sequence normally follows a logical structure. On the other hand, a system test program departs from the controlled environment of a unit diagnostic. It is designed to provide a comprehensive system test similar to an operational environment.

There are various levels of unit and system tests; sometimes the distinction between the two is arbitrary. In addition to levels of testing, a variety of test environments must be considered, such as engineering, manufacturing, shipping, installation, and field maintenance. The unique requirements of each of these environments must be considered in planning a diagnostic strategy for a data-processing system.

System design

Diagnostic engineering begins in the initial phases of system design. A maintenance strategy is defined and the system is designed to include features necessary to meet the requirements of this strategy. Special features, known as "diagnostic handles," are needed for testing the system automatically, and for providing adequate error isolation. The diagnostic program must be able to control the system logic in slightly different ways from the operational programs. Examples of the things a diagnostic engineer looks for in the design stage are:

1. Special circuits to inject errors for testing error-detection circuitry under program control.
2. Automatic feedback loops for programmed checking and analysis of output devices.
3. Program-controlled switches to reduce the need for manual interventions.
4. Programmed sensing of the status of key control circuits and internal registers not normally available under program control.

The diagnostic engineer plays a multiple role as part of the design team. His knowledge of logic design and his programming background allow him to obtain a good grasp of the total system design. He can readily correlate the impact of a particular logic design in one area on the overall functioning of the system. As the design begins to take shape, a detailed review of this logic is performed to develop the diagnostic plan. This review often uncovers some design errors even before fabrication. Throughout the design cycle, frequent and close communication is required with the design engineers.

Engineering test

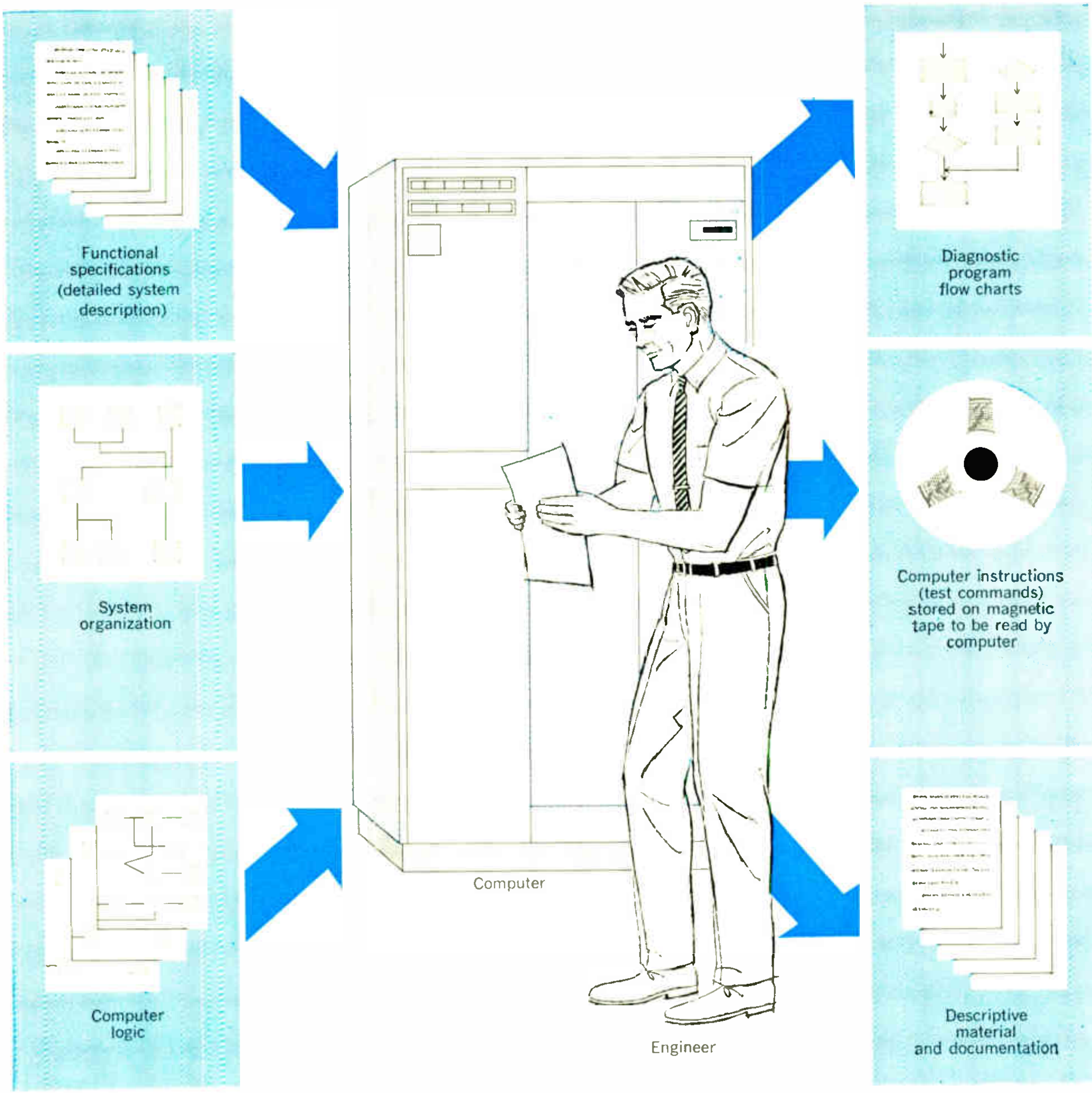
The first system to be fabricated (the engineering model) has unique test requirements. Many errors will exist both in the system and in the test programs. Simple test programs are used to "bring up" the system, followed by a complete set of functional tests to verify that the system performs all the functions in the precise manner described in a set of functional specifications or the user's manual. For example, each computer instruction (command) will be executed with various options to insure that the proper result occurs. To prepare these functional tests, it is important that the *functional* specifications (what the machine should do) be referenced and not the *actual* logic design (what the machine does); otherwise, the end result is a test of the machine as it is designed and not necessarily as it was supposed to be designed.

Functional testing alone, as exhaustive as it may seem, is not sufficient to test a large data-processing system. The internal logic of the system must be referenced to ensure that all the logical paths are tested. Then finally there is that illusive type of error that does not occur according to a logical pattern. Even though all the logic blocks (circuits) and all the paths (interconnecting wires) have been tested, some unique set of timing conditions or some unique pattern may cause a failure.

It is impossible to test for all possible data patterns and timing conditions. A simple 32-bit register has 2^{32} possible combinations, not to mention gating-in, gating-out, and control circuits. The number of possible combinations for a system with thousands of logical circuits approaches infinity. Even if the test programs were available there would not be enough time in the life of the system to run all of them.

Although 100 percent detection is not possible, it can be approached. Worst-case test patterns are developed using knowledge gained from past experience and the knowledge of the design engineers.

The difference between functional tests and circuit level tests can be illustrated with an example from IBM's System/360. The central processors of models 30, 40, 50, 65, and 75 are all program compatible; that is, there is a basic instruction set that operates in the same way in each computer. The description of the instruction set is identical for all these models and the same program can be used to operate all of them. However, because of differences in internal speed, core storage capacity, features, etc., the internal logic design of each central processor is quite different, and thus two types of tests are needed. There is a set of functional tests that can be used to test all models for functional compatibility. These programs were written by referencing the System/360 principles of



THE DIAGNOSTIC ENGINEER uses a computer to assist in the preparation of a diagnostic program. At the left are some of the sources of information; at the right are the principal parts of a diagnostic program.

operation (the programmer's manual). On the other hand, there are diagnostics, unique for a given System/360 central processor, that test the unique logic and use the internal logic design as a reference point.

Production and maintenance requirements

When a data-processing system progresses from the system design and engineering test phases into the production phase, there are additional diagnostic requirements. The diagnostic engineers who wrote the programs and assisted in debugging the first models must now make these programs available for others to use. Good documentation of the diagnostic programs is essential so that other test engineers will know when and how to use the program and how to recognize and interpret the error indications.

Although the design errors have been removed from the system, the diagnostic tests must be effective in handling multiple fabrication errors and component failures in the manufacturing test environment. Testing begins with a small subsystem, and progressively encompasses more equipment until the entire system is operating properly. Then, the final shipping test should be the most stringent test possible. Since this last test provides a "go/no-go" decision, speed and thoroughness take priority over the degree of isolation.

When the system reaches its destination, an installation test similar to the final shipping test is required to ensure that nothing detrimental happened to the equipment in transit. If the system is actually a subsystem of a larger system, another level of test may be required. After a successful installation, the requirement is for good field maintenance for the life of the system. Ideally, single errors are detected, isolated, and repaired one at a time. A diagnostic program, written strictly for the maintenance environment, can take advantage of the single-error assumption in making a diagnosis. Since we do not live in an ideal world, however, the field maintenance diagnostic must be capable of coping with multiple failures.

Usually, one program or set of programs is written to satisfy each of the various test environments. Functional specifications and design logic are referenced. Multiple options are designed into the program, which the test engineer can select or delete to fit each particular error situation. In some cases, a diagnosis is attempted; if the diagnosis is incorrect, or none is offered, the error symptoms are printed or displayed to assist the test engineer in making a diagnosis.

Diagnostic techniques

Diagnostic programs generally employ a mixture of testing and isolating techniques. The mix will depend upon the equipment being tested and the purpose of the test. The basic techniques are as follows:

The building-block approach starts with the simplest test involving the smallest amount of circuitry possible. Each additional test includes a small additional amount of circuitry. Failure to pass a given test at any point in the test sequence indicates that the probable cause of error is the additional circuitry being tested. This technique is particularly effective when there are multiple machine malfunctions. Since each error is repaired as it is detected, before the test sequence is continued, multiple failures should not confuse the diagnosis. In addition, it is not necessary to rely on a defective computer to analyze test results. Although this technique is disadvantageous in that it is time consuming, the time required is insignificant for the central processor because of the speed at which instructions are executed.

The go/no-go approach usually starts with a quick test of a large amount of circuitry. The strategy is to "zero in" on an error. The system program is designed to isolate errors to a small area of the system (that is, a specific unit or logic function). A more specialized diagnosis is then used for that specific area in order to narrow down the trouble further. This technique provides the fastest test under successful conditions, and is often used to provide rapid checkout of a large system. When the diagnosis is correct, the isolation time is optimized. On the other hand, an incorrect diagnosis can lead the test engineer up a time-consuming blind alley. Multiple errors are difficult to handle with this approach.

The multiple-clue approach uses a series of individual tests, saving the success or failure indications of each. Each test in the series will take a slightly different path through the maze of logic. At the end of the series, all the test results are analyzed and a diagnosis is made. For example, if paths 1, 3, 7, and 10 failed, any circuit common to these paths would be suspect. This technique is particulary useful for circuitry (logic) that cannot be adequately diagnosed by any one test; that is, the smallest amount of circuitry used by any single test is still too large.

Error-environment recording depends on special circuitry for detecting errors. Records are kept of the error and the machine's "environment"—that is, the status of indicators, contents of registers, etc. Isolation of the error is then attempted by analyzing this recorded error environment. This technique, unlike the previous ones, is effective for diagnosing the intermittent error, particularly if the malfunction occurs in normal operation but cannot be recreated in a diagnostic test. The success of this "one-shot" diagnosis depends on the resolution of the error indicators, the exact environment saved, and the speed—relative to the speed of the computer—with which errors are detected and the environment captured for later analysis.

Preventive maintenance techniques belong in a separate category because they predict and repair abnormal conditions before they result in system errors. A failure to pass a test indicates a marginal condition but not necessarily a failing condition. The required adjustment can be made before it results in a loss of computer time for the customer.

It is important to point out that a certain amount of basic control and logical circuitry must be functioning before the computer itself can be used effectively to control the sequencing of tests and to analyze test results. This operating hard core involves a substantial amount of circuitry and demands special attention in developing the diagnostic strategy. First it must be defined, usually in the form of a set of basic computer instructions or commands, and then these instructions and associated computer circuitry must be tested without using other instructions or depending upon the computer logic for analysis. Either manual techniques or additional circuitry, or both, are used for this purpose. This additional circuitry, when used, now becomes the "hard" hard core—if you will pardon the expression. It is obvious that a hard-core test is the first step of a building-block approach.

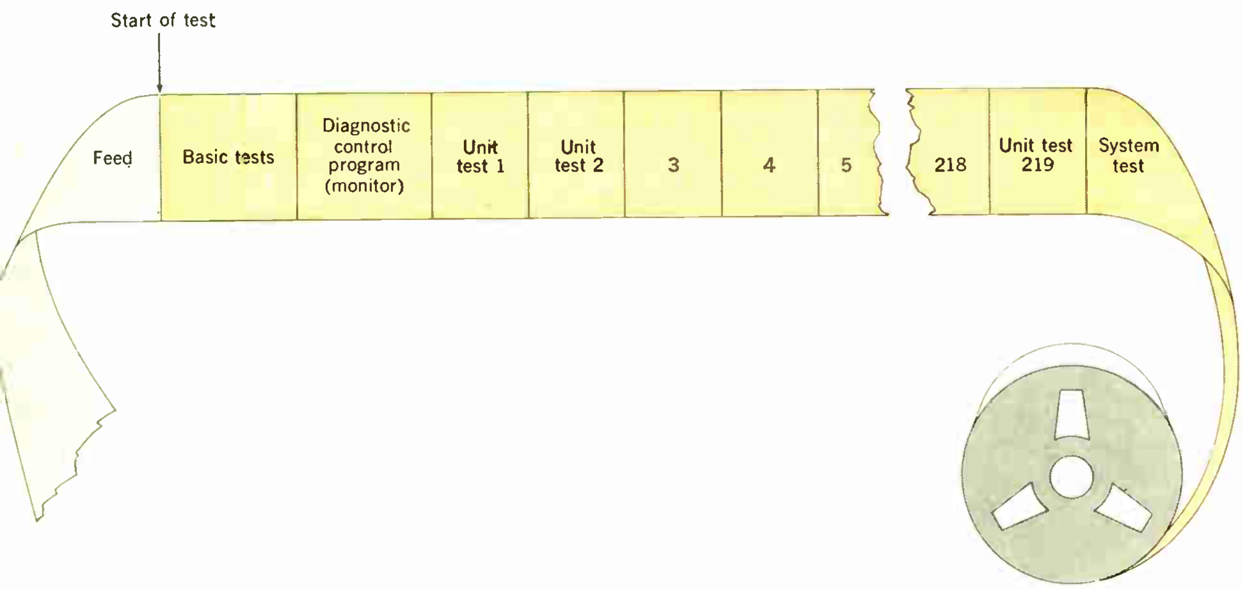
Human factors

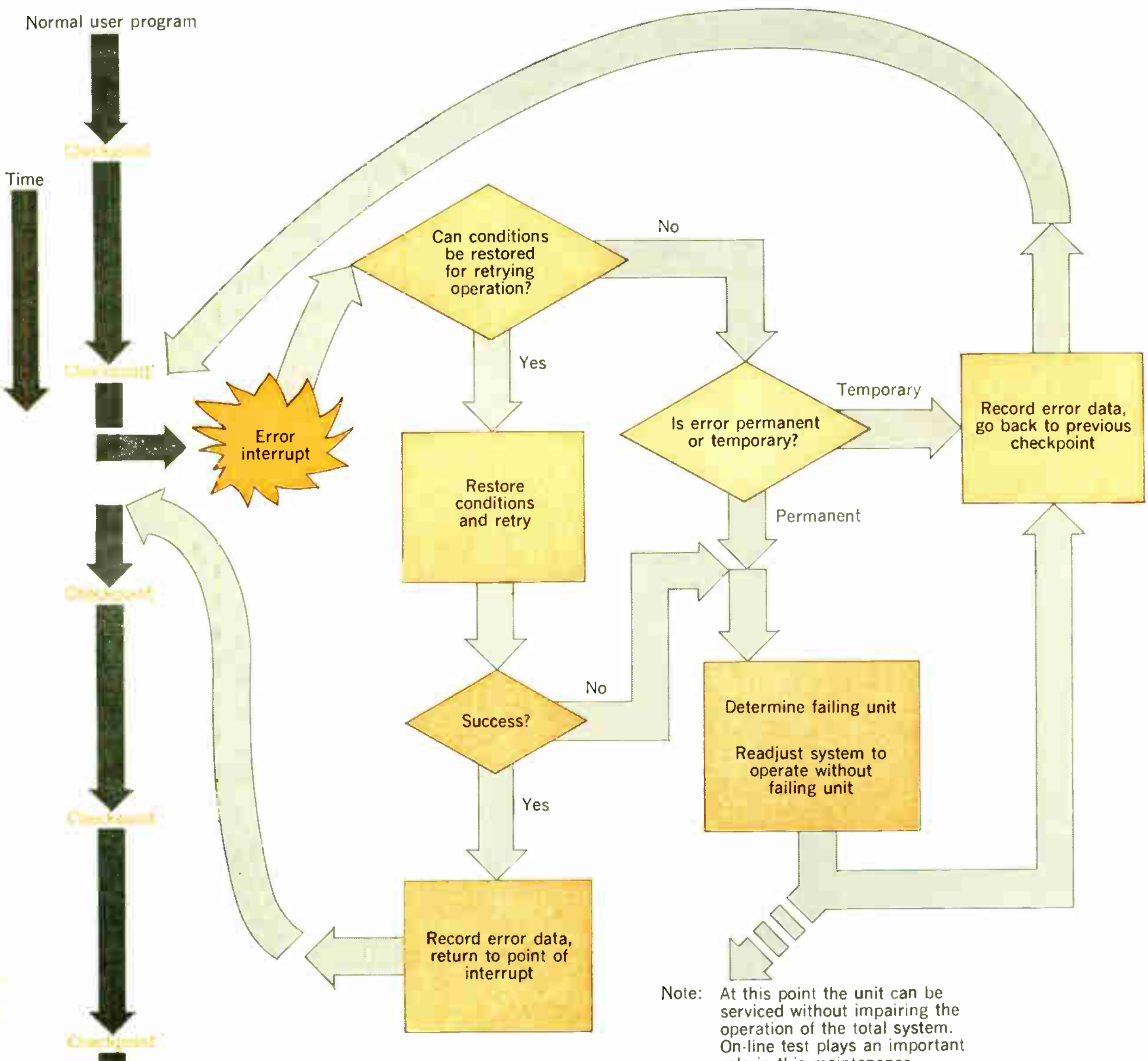
If we view the diagnostic problem from the point of view of the test engineer, we can immediately appreciate the enormous human-factor problems. The test engineer is faced with a sophisticated system that, for some unknown reason, is not operating properly (or not operating at all). The design of this same system required a team of specialists in circuit design, logic design, system design, mechanical engineering, programming, etc. The system was manufactured by experts under controlled factory conditions. The customer engineer's job is to find the cause and take corrective action.

In order to assist him, the computer manufacturer has supplied a multitude of aids: a complete set of manuals describing the theory of operation of the system and each

A TYPICAL DIAGNOSTIC TAPE for a data-processing system includes many programs. The system must pass basic tests before diagnostic program is run. Each unit must be tested before a more sophisticated system test is used.

FLOW CHART, illustrating the concept of total error management of a data-processing system.


Start of test
Feed
Basic tests
Diagnostic control program (monitor)
Unit test 1
Unit test 2
3
4
5
218
Unit test 219
System test


Normal user program
Time
Error interrupt
Can conditions be restored for retrying operation?
No
Yes
Is error permanent or temporary?
Temporary
Permanent
Record error data, go back to previous checkpoint
Restore conditions and retry
Success?
No
Yes
Determine failing unit
Readjust system to operate without failing unit
Record error data, return to point of interrupt
Note: At this point the unit can be serviced without impairing the operation of the total system. On-line test plays an important role in this maintenance.

of its subassemblies; a complete set of logic diagrams, showing each circuit and all point-to-point wiring; manuals, flow charts, and listing for the operating programs; numerous diagnostic programs with their write-ups, flow charts, and listings; maintenance aids, including numerous lights, buttons, and switches; an oscilloscope, tool kit, spare parts, etc.

The test engineer requires many aids because he is faced with a variety of environments and situations. He may want to verify rapidly that the system is functioning properly (go/no-go test). If he has no idea where the trouble is, he may want to begin with the very first test program on the diagnostic tape and automatically sequence through all tests. If he has a good idea where the trouble is, he may want to select a specific program on the test tape. He may want to loop on a given program, or one test routine for scoping purposes. He may want to delete error printouts, stop on the first error, and skip certain tests. In short, the test engineer requires a flexible system of automatic tests. However, one of the problems of diagnostic engineers is to make this wealth of materials more manageable and useful. Conventions must be established and an interface provided through which the test engineer can communicate with this system. This interface must also provide a medium for the individual diagnostic programs to communicate meaningful test results.

The diagnostic monitor or control program

As a partial solution to the human interface problem between the test engineer and the system of automated diagnostic tests, the diagnostic monitor or control program has evolved. It is to the maintenance system as the operational control program is to the operating system (user's application program). Some of the functions performed by the diagnostic monitor are:

1. To control the sequence of tests automatically, by selecting the test from the diagnostic file, initiating and terminating the test, and controlling the mode of operation.
2. To accept and respond to input from the test engineer by defining the system to be tested and the diagnostic mode of operation, selecting a specific test, terminating a test in progress, and altering the test sequence.
3. To communicate test results to the test engineer, including success or failure indications, failure data, failure diagnosis, and other test information.

The numerous diagnostic tests required for a sophisticated system are frequently written by different people, at different times, and even at different locations. It is essential to establish a diagnostic architecture or framework in order to get some consistency in the human interface. The diagnostic monitor or control program provides this framework.

The trend toward larger and more complex data-processing systems continues to tax the imagination of diagnostic engineers. Technological improvements have been continually providing us with faster and more reliable components. The failure rate per function is therefore decreasing. The same technology is also providing smaller components, thus giving higher logical density; consequently, the failure rate per cubic foot tends to increase. Coupled with this trend is the continual trend toward more sophisticated systems, with large equipment complexes, and sophisticated programming packages. Because of the sheer quantity of components, such systems must be designed to tolerate malfunctions. Techniques have been developed to detect errors automatically, and to take corrective action in milliseconds, faster than any human being could possibly react. Techniques have been (and continue to be) developed to detect a faulty unit, to adjust the system to operate without that unit, and to allow the unit to be serviced.

These system-error management and serviceability techniques are accomplished through a combination of special equipment design (logical circuits) and special programming techniques. Possibilities considered in the design include:

1. Error-detection circuitry using some form of redundancy for the quick detection of errors.
2. Save mechanisms (circuitry) to preserve the system environment and error data for later program analysis.
3. An interrupt system (circuitry) designed as an integral part of the system to stop the operation in process and turn the control over to a diagnostic routine (program).
4. Diagnostic routines (program) to analyze the error data and environment information to determine the corrective course of action (for example, retry the operation in the case of an intermittent error and continue with the system operation at the point of error and, in the case of a solid error, readjust the system to continue operating without the faulty unit).
5. On-line test routines that are part of the operating system (program). These routines allow testing and servicing of portions of the system while the rest of the system is engaged in useful work.

Future requirements

The state of the art of circuit technology, system organization, and application development has been advancing at a rapid pace. New challenges are continually facing diagnostic engineering. Some of the more pressing diagnostic needs are for

1. Total error management.
2. Techniques that will approach 100 percent detection in the most efficient manner.
3. Methods for the economical production of diagnostic routines that automatically isolate to the replaceable component.
4. Diagnostic techniques for distinguishing between equipment errors and program errors.
5. Drastic reduction of human-factor problems. The quantities of sophisticated equipment, programs, and reference materials facing one test engineer are getting unmanageable.

The earlier data-processing systems, which used vacuum tubes, relied almost entirely on programs to detect and isolate malfunctions. Solid-state circuits, with continual reductions in size and cost, made it practical to include redundant circuits for the purpose of error detection. Today's systems have error-detection circuitry that will detect a majority of malfunctions. Isolation, for the most part, is still performed by programs, which are getting increasingly complex. As the cost per logical circuit decreases, more of the isolation functions will probably be shifted from programs to logical circuitry.

Approaching nuclear power

Practical guidelines are offered for integrating a nuclear power plant into an existing utility system. Steps for obtaining licenses, choosing the plant site, and training personnel are considered

W. A. Chittenden, A. Nathan *Sargent & Lundy*

Because of economic, system expansion, and air pollution considerations, the committed capacity of nuclear installations in the United States has risen from 3500 MW to approximately 16 000 MW in the past two years. The factors that should be considered in the integration of nuclear energy into a utility system are examined in this article. These include site location and size, scheduling, authorization, public relations, and personnel training. For a typical plant, it is estimated that some 90 man-years are required for personnel training.

The application of nuclear energy for generating electricity has recently assumed a significant and influential role in the United States. In the past two years, the committed capacity of nuclear installations has risen from 3500 MW to a value approximating 16 000 MW (see Table I). Interest by the utility industry in this energy source is increasing because of economic considerations, expansion needs, and the problems of air pollution from coal-fired plants. The recent past has witnessed a narrowing of capital cost differences that have existed between conventional coal-fired installations and nuclear units, as well as a reduction in nuclear fuel cycle costs.

Site criteria

As with all condensing power plants, the question of an adequate source of cooling water is of vital concern when

I. Recent nuclear plant contracts

Utility	Manufacturer	Unit Size, MW	Operating Date
Commonwealth Edison	General Electric	800	1969
Conn. Power & Light, Western Mass. Elec. & Hartford Elec. Light	General Electric	600	1969
Rochester Gas & Elec.	Westinghouse	450	1969
Consolidated Edison	Westinghouse	870	1969
Boston Edison	General Electric	600	1970
Consumers Power	Combustion Engrg.	770	1970
Florida Power & Light	Westinghouse	670	1970
Commonwealth Edison	General Electric	800	1970
Wisconsin Elec. Power	Westinghouse	450	1970
Carolina Power & Light	Westinghouse	700	1970
Commonwealth Edison	General Electric	800	1970
Northern States Power	General Electric	550	1970
Public Service of Colo.	General Atomic	330	1971
Florida Power & Light	Westinghouse	670	1971
TVA	General Electric	1100	1970
Duke	Babcock & Wilcox	820	1971
TVA	General Electric	1100	1971
Duke	Babcock & Wilcox	820	1972

the available sites within a given utility system are being considered. Although all of the conventional means of supplying this requirement—such as rivers, lakes, and cooling towers—are applicable to either a conventional or a nuclear plant, a marked difference in the total requirements for plants of the same capability may exist. Differences between light-water-reactor and conventional power plants are noted in Fig. 1; the curves are based on a net plant heat rate of 11.2×10^6 J/kWh for light-water-reactor plants and of 9.4×10^6 J/kWh for conventional installations. This significant difference in plant heat rates and its consequent effect on cooling-water requirements

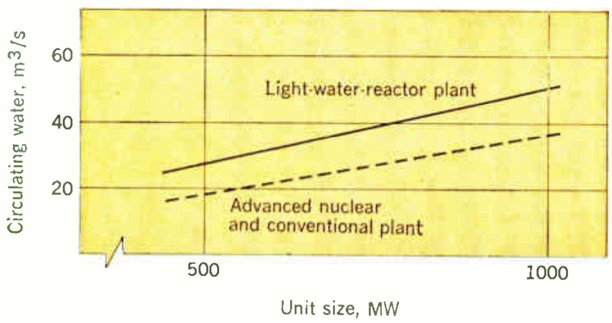


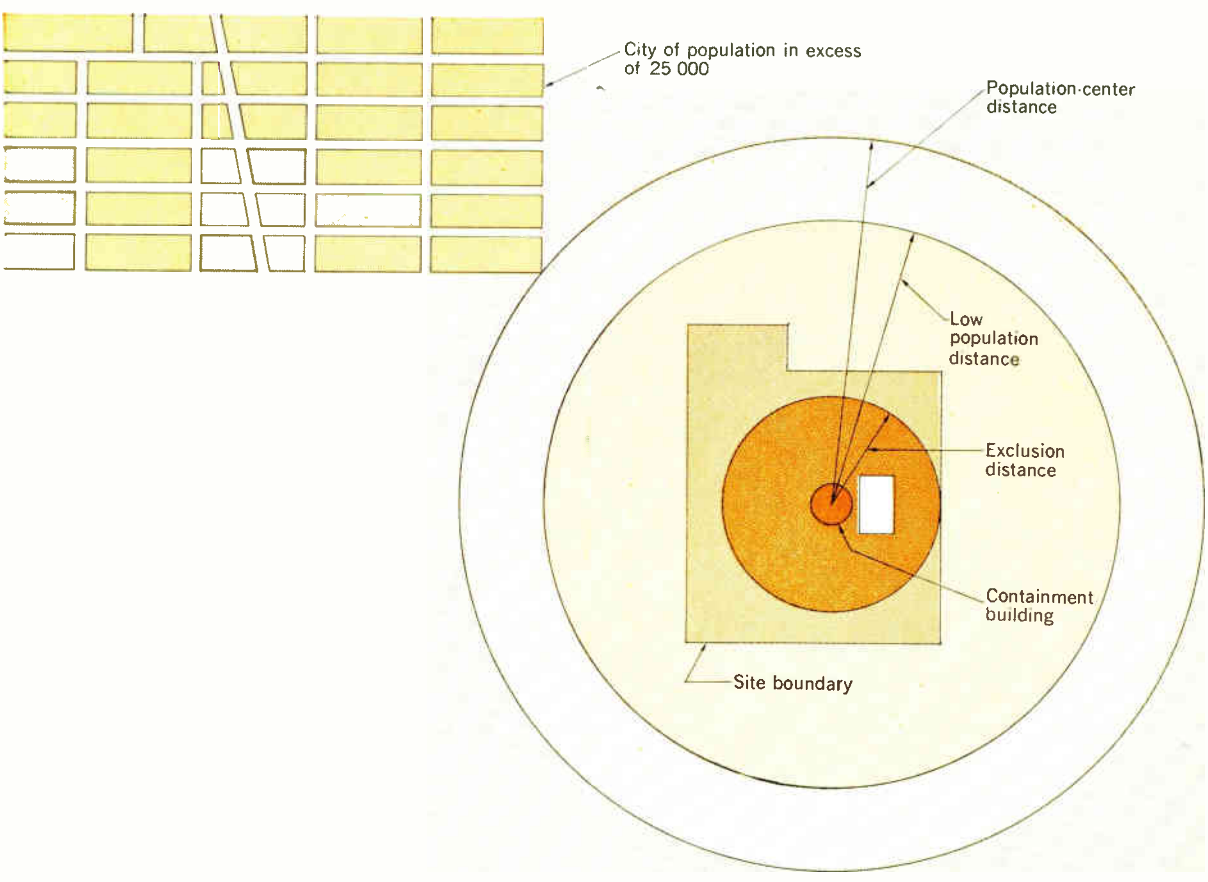
FIGURE 1. Typical plant circulating-water requirements.

should be allowed for in estimates of water flow and heat-sink capability of a particular site.

Site location and size. When the radiological aspects are brought into focus, the nuclear plant generally requires a substantial increase in land area compared with a conventional coal-fired plant. A by-product of nuclear fission is the formation of radioactive material, which is normally confined within the reactor core or system; the resulting radioactivity presents a negligible hazard. There are, however, two hazards associated with core meltdown and the attendant fission product release into the containment structure: direct radiation and the ingestion of radioactive material. Protection against direct radiation may be accomplished by providing sufficient distance between the source and receptor or by providing an adequate intervening shield structure. The release of gaseous and volatile fission products leads to ingestion and inhalation hazards, which are generally the limiting hazards against which the plant design must protect the general public.

The suitability of a proposed site may be evaluated by considering its hydrology, meteorology, and seismology in conjunction with the population distribution and land use adjacent to the site. The basic criteria for evaluating a postulated reactor core meltdown hazard are defined as well as the upper limits for the amount of fission product release and leakage from the containment and the allowable maximum for radiation exposure to the public. With the application of conservative meteorological conditions and demonstrable containment leakage rates, values for

FIGURE 2. Hypothetical nuclear plant site and environs.



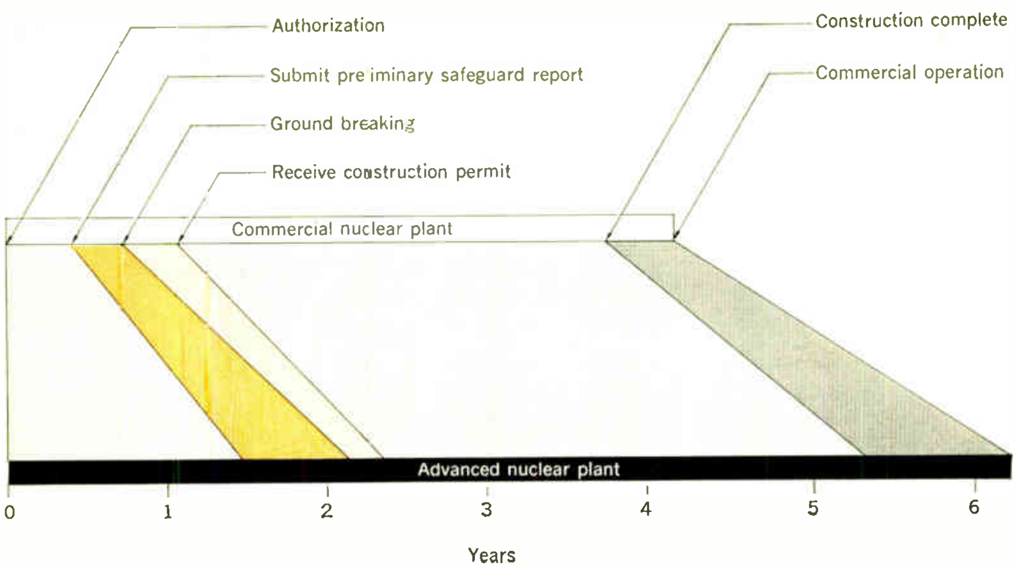


FIGURE 3. Representative project schedule.

plant exclusion area, low population zone, and population center distance may be derived. In this context, the exclusion area is of such size that a person on its boundary for an elapsed time of two hours immediately following fission product release would receive a whole body radiation dose of less than 25 rems or a total dose of less than 300 rems to the thyroid from iodine. Likewise, the low population zone is of such size that a person on its outer boundary would receive equivalent doses when exposed to the radioactive cloud resulting from fission product release; a population center distance 1.33 times the distance from the reactor to the outer boundary of the low population zones is assumed (see Fig. 2). Engineered safeguards built into recent plant designs have permitted significant reduction in distance requirements as compared with those of less sophisticated designs.

Schedules

A four-year schedule from plant authorization to commercial operation has been typical for recently announced large-scale nuclear plant projects (see Fig. 3). The first hurdle to overcome after plant authorization is the preliminary safeguards report, which is normally submitted four months after authorization. This report includes a preliminary plant design and the necessary criteria for a detailed plant design to insure that the unit can be operated "without undue hazard" to the general public. Ground breaking takes about seven months following project authorization, and the construction permit usually follows by four months.

Plant design criteria of light-water-reactor systems are based on sufficient experience that the details may be developed during the early months following project authorization and be submitted to the U.S. Atomic Energy Commission with a request for a construction permit. Regulations require that field work on the permanent structure for the plant may not begin until a construction permit has been issued by the AEC. Such a permit is based on a finding by the AEC that the proposed plant may be operated without creating an undue hazard to the general public. Therefore, the only field work that may begin before receipt of the construction permit is that of site preparation and excavation. Construction usually takes some 44 months to complete after authorization; commercial operation starts after four months of testing and shakedown operation.

An overall six-year schedule is suggested as representative of the first large-scale application of an advanced reactor concept. As much as 1.5 years may be set aside in the schedule to develop the preliminary safeguards report and the construction permit application. Further, it is reasonable to expect that the time interval for issuance of the construction permit may be longer than for a light-water commercial reactor. Completion of construction is estimated at a little over five years after project authorization with the remaining year of the six-year schedule allocated for initial plant operation and checkout prior to commercial operation.

The sequence of procedures required under the Atomic Energy Act is illustrated in Fig. 4. The preliminary safeguards report is submitted to the Division of Licensing and Regulation (DLR) and the Advisory Committee on Reactor Safeguards (ACRS) for their independent review and analysis. Copies of the report are also placed in the Public Document Room for the interested public. During the course of the review, questions that may arise must be answered by the applicant; typically, there may be several meetings between the applicant and the respective review boards to resolve any questions in connection with the plant design or criteria. Once it has concluded that the proposed nuclear plant could be operated at the site without creating an undue public hazard, the AEC appoints an Atomic Safety and Licensing Board to review the application and to conduct a public hearing. The hearing affords the general public an opportunity to present any arguments in support of or in opposition to such construction. Upon conclusion of the public hearing, the Atomic Safety and Licensing Board reaches a decision and, if favorable, directs that a construction permit be issued. Following this, the commissioners of the AEC may set the ruling aside or allow it to stand. Once the construction permit is received, field work may commence immediately.

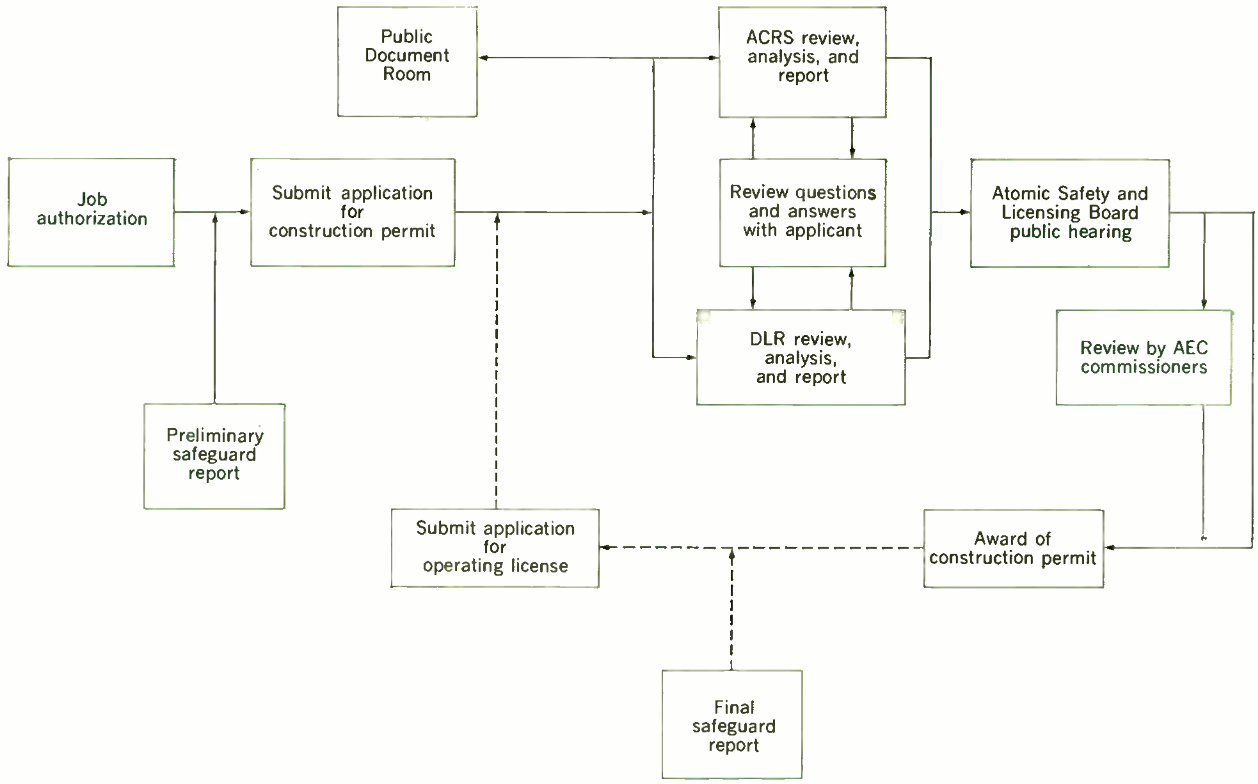
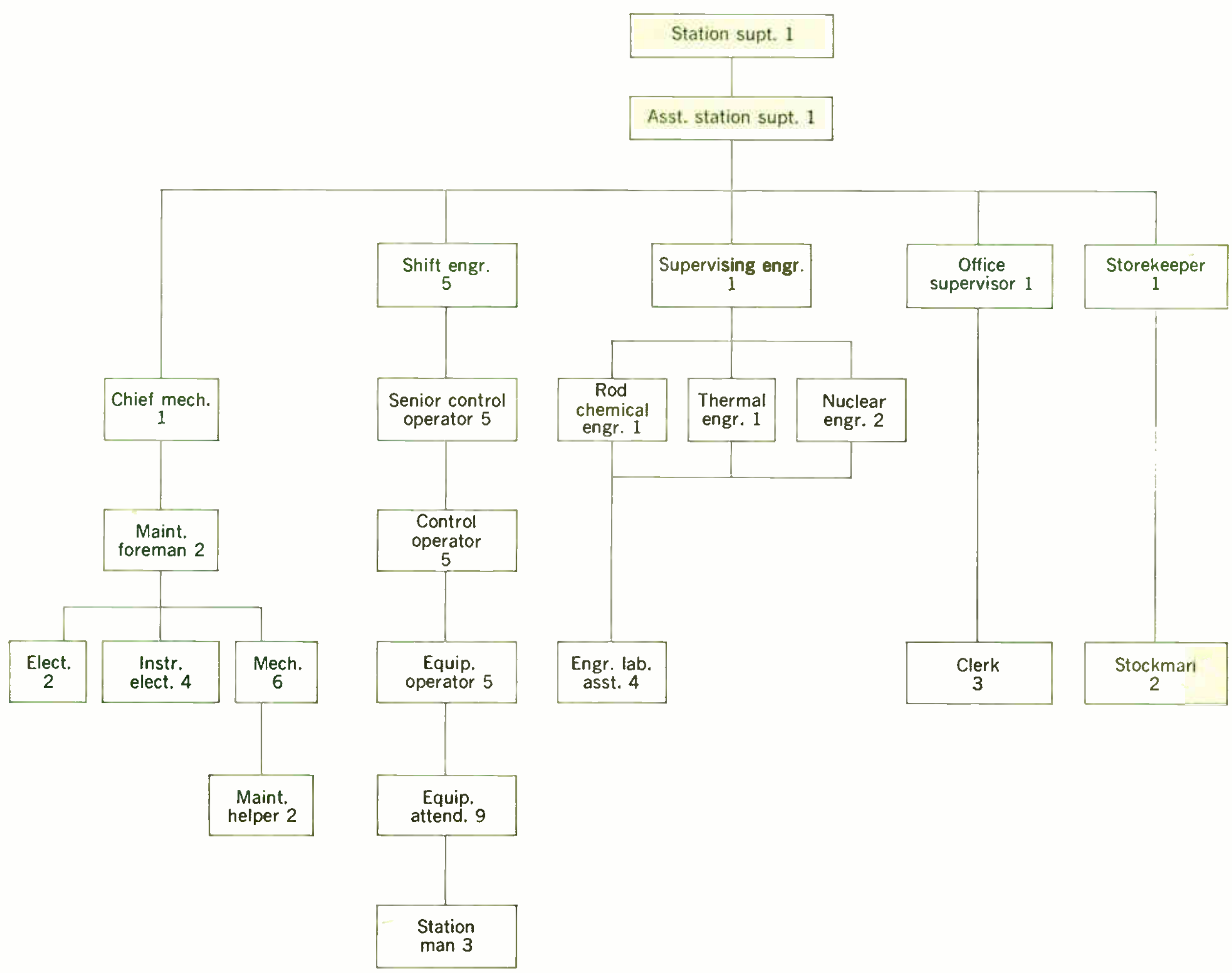


FIGURE 4. Outline of regulatory procedures.

FIGURE 5. A typical station organization chart.



Several parts of the foregoing sequence have established time limits. The AEC must issue a 30-day notice to the public for holding a public hearing; a ten-day period must elapse between the directive to issue a construction permit and its effective date. In addition, a 45-day limit is placed on the review period during which the commissioners of the AEC may set a finding aside. The remaining factors are dependent upon the nature of the plant design and its criteria, the separate independent review boards, and any developments arising from the public hearing. A similar procedure is followed upon substantial completion of plant construction to obtain an operating license for the plant.

Even after the plant is built and an operating license obtained, the Federal licensing and regulatory aspects are not at an end. The AEC Division of Compliance periodically inspects the plant to insure that the unit is being operated in accordance with the limitations of its operating license. Procedures for modifying the operating license for increased power output or other changes follow well-established lines that parallel, to some degree, the original licensing procedures.

Because the licensing of nuclear plants is a function of the Federal Government, the various states have had, to date, a limited influence on the design and construction of nuclear plants. In the future this situation is likely to change somewhat from the standpoint of the effect of operation of such plants on the ecology of the plant environs. At present, it would seem that state requirements will be similar to those associated with conventional plants with respect to safety codes; however, radiation-exposure control of operating personnel will receive particular attention.

Public relations. The most successful public relations programs have been extensive and responsive to all levels of public interest. These programs have included a speakers' bureau, specific arrangements for tours and programs for various levels of local schools, and provisions for accommodating the public at the plant site. If a plant is not under active consideration, a long-range approach to the matter of educating the public for acceptance of nuclear energy may be incorporated in the normal utility public relations programs. Certainly, the utility itself is in an excellent position to gauge public opinion in its service area; on the basis of these judgments, custom public relations programs may be developed to cope with the problems in any specific area.

Personnel and training

An important factor to consider early in the development of a nuclear plant organization is the creation of a special plant staff concerned with nuclear matters only. The nuclear plant organization should reflect the standard practices of the utility's conventional plants wherever possible, particularly with respect to the use of manpower, the location of the plant within the system, union agreements, maintenance philosophy, and the projected plant loading schedules. The functional responsibilities and the number of people needed to operate and maintain a nuclear plant may be based on an expansion of the organizational assumptions developed for the system's conventional units. These factors include the projected number of other nuclear plants within the system, source of major overhaul and maintenance personnel, and the extent of technical staff responsibilities for continuous plant operation and of company policy with regard to backup personnel for vacation and illness.

Experience has suggested that the vast majority of personnel requirements can be met from within the utility's present organization, with a limited number of personnel being recruited from the nuclear industry. A representative organization for a 500-MW plant is illustrated in Fig. 5. This chart covers the functional responsibilities associated with station management, maintenance, operation, and technical support and plant services; a total staff of 67 persons is indicated.

Training programs. The training requirements for plant personnel may be broken down into separate categories for management, technical personnel and senior control operators, foremen and control operators, and all other plant personnel. Specific classroom and on-the-job training and special programs for each category of personnel are required.

The current practice for an initial nuclear installation is to have the reactor manufacturer conduct a series of courses for senior personnel. These personnel may then become an integral part of the teaching staff for the remaining people at the plant site.

Nearly all of the plant personnel will be actively participating in the development of operating procedures and in the conduct of preoperational tests prior to plant start-up. These practices develop the requisite familiarity with and knowledge of plant design and operating characteristics as well as the nuclear theory necessary for the licensing of reactor operators by the AEC. Typically, the assistant station superintendent, the supervising engineer, and the control operators may each be licensed to operate the plant. For the representative plant organization shown in Fig. 5, approximately 90 man-years are required for personnel training.

Prior to operation, senior personnel from each of the plant divisions may be in training programs of upwards of two years, covering basic nuclear technology, radiation protection, work training in nuclear engineering and respective specialty areas, and operator on-the-job training. Fewer senior personnel will be involved for shorter periods of time in less rigorous training programs which emphasize those topics most important for the individual job assignment. The training program covers all personnel; for example, clerks, storekeepers, and janitors receive a course in radiation protection.

Conclusion

The significant number of recent commitments for large-scale nuclear units throughout the United States attests to the acceptance of this new energy source by the electric utility industry. Implicit in this widespread acceptance is an economic position that suggests that an increasing number of utilities will be looking to nuclear energy to serve their expanding system requirements. Further, the problems of air pollution that are associated with coal-fired plants have contributed to the need for this activity.

The factors discussed in this article are the prime considerations entering into the decision-making process. That the overall capabilities for developing this energy concept exist in utility organizations is demonstrated by the experience of those who have made nuclear commitments. The utilization of nuclear power by a large number of utilities seems just a matter of time.

IEEE publications

scanning the issues
advance abstracts
translated journals
special publications

Scanning the issues

A Single Page. Integrated electronics s now officially defined by the IEEE is "that portion of electronic art and echnology in which the interdependence of material, device, circuit, and system-design considerations is especially sig-nificant; more specifically, that portion of the art dealing with integrated cir-cuits." This definition, and seven others —element, integrated circuit, mono-ithic integrated circuit, multichip in-egrated circuit, film integrated circuit, nybrid integrated circuit, and substrate —constitute a single but significant page n the current issue of the IEEE JOUR-NAL OF SOLID-STATE CIRCUITS. The ignificance of these definitions is that lespite the almost universal interest among IEEE members in integrated electronics, there has been a prolonged absence of an adequate set of terms and definitions pertinent to this important ield. This single page then constitutes omething of a long overdue achieve-nent.

A suggestion of the kinds of difficulties hat obstructed progress on this issue is ontained in the brief introduction by ditor James D. Meindl. The IEEE ef-ort to select and define appropriate erminology for integrated electronics, Meindl says, can be traced to co-ordinated AIEE IRE attempts in 1960. Although the need was recognized learly, progress toward a Standard was mpeded by the number of diverse view-points that influenced these early at-empts. As a consequence of reorganized fforts following the formation of the EEE and the appearance of a measure of onsensus within the electronics com-nunity, tangible progress toward a tandard occurred in late 1964. A first raft was produced within the Solid-tate Standards Committee of the Group n Electron Devices in early 1965. After nodification, it was submitted to the EEE Standards Committee and ap-roved in September 1966.

And meanwhile, as the years passed, other groups and organizations were drawing up their own definitions; thus, as Meindl says, the long absence of a national consensus has made it difficult to offer a unified U.S. position.

Considering the years of labor that have gone into the birth of this single page, it is almost diabolical to take the Terms Task Group to task for split-ting their infinitives, but editors can never rest easy when confronted even by such minor imperfections. It shouldn't, however, take more than another year to bring about complete perfection. ("Definitions of Terms for Integrated Electronics," *IEEE Journal of Solid-State Circuits*, March 1967.)

Electron Devices. A number of the re-cent issues of the IEEE TRANSACTIONS ON ELECTRON DEVICES have included ar-ticles of more than ordinary interest. In the May issue, for instance, "Device Modelling" by John J. Sparkes clears up some points raised by earlier authors on the problems of modeling electronic devices. In particular, Sparkes questions the wisdom of trying to achieve a gen-eral theory of modeling. That is, models of electronic devices have to express the physical structure of the device, represent the significant conduction processes of a device, and lead to use-ful equivalent circuits for circuit anal-ysis. Sparkes suggests that a deliberate separation of these different func-tions might be preferable to attempting to lump them together.

Since modeling is such a basic in-gredient in all types of engineering, it does no harm to repeat Sparkes' statements on its aims:

1. We want a one-to-one correspond-ence between the model parameters of the device and the physical processes so that we can calculate one from the other and consequently can interpret performance requirements of the device in terms of physical structure and prop-erties.

2. We want our circuit model simply to help us design circuits using the device. (Since, incidentally, this is the ultimate objective of all device modeling, therefore, it should per-haps be regarded as the primary ideal.)

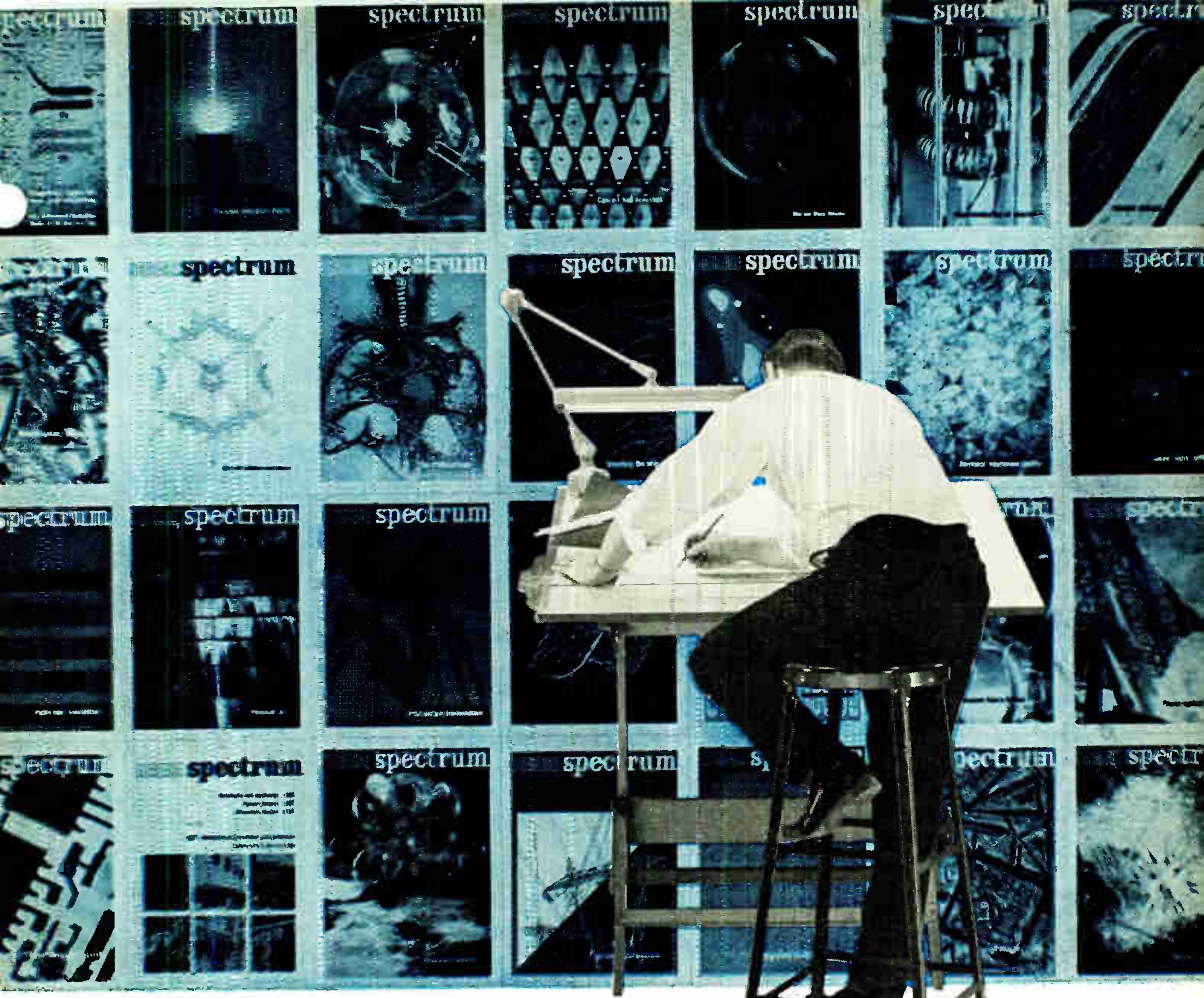
3. We want understanding and insight partly for educational purposes and partly so that we can visualize the effects of change of environment, cir-cuit performance, etc. It is here that the general-purpose model is needed, provided its generality does not con-fuse its purpose!

In the article, Sparkes discusses in a clear and straightforward style each of these aims, with particular reference to the active region of oper-ation of the junction transistor, since, he says, it highlights many of the points in question.

Of somewhat more specialized interest is an article in the April issue, "The Accuracy of Numerical Solutions for Electron Gun Design," by Vladimir Hamza and Gordon S. Kino. These authors have successfully attacked an important facet of digital computer programming for analysis of Pierce-type electron guns, and have thereby added a novel concept to what is now a well-established technique.

For those who do not know, Hamza and Kino point out that during the last five years there has been a consider-able effort to design electron guns by using high-speed automatic digital com-puters. In the past, space-charge-flow problems were solved primarily through analog techniques such as the electro-lytic tank, resistance network, or rub-ber membrane. Some of these analog techniques require complicated and ex-pensive equipment, and all have limited accuracy. Numerical analysis, with high-speed automatic digital computer, is eminently suited to solving problems of this type.

By various means discussed in the article, Hamza and Kino find it possible to increase the accuracy of numerical solutions by an order of magnitude or more. The editor of these TRANSAC-



the only magazine serving the Whole e/e engineer

Electrical/electronics engineers rely on but one magazine to keep fully abreast of every important development in their fast-moving field. Sure, there are many "new product" or "news of the industry" magazines, a handful of "how to do its" and a wide selection of "specials" that cater to small vertical audiences within the industry. But do you know of one magazine besides *Spectrum* that encompasses the entire electrical/electronics universe?

Spectrum is designed to be broad. One of its basic policies is to educate its readers away from the pitfall of over-specialization. Its articles, solicited from the top men in the industry, are well illustrated, carefully edited—often rewritten to make them appealing and easy to digest. As a result, they're eagerly sought by *Spectrum's* 145,771* engineer/executive audience. This interest extends to the advertising. These are the men who can act on your ads—decision makers who are involved in every type of electrical/electronic purchase. You can reach this audience for about one cent per page per prospect.** It's the biggest buy in the industry—at the lowest rate per thousand.

IEEE spectrum

* BPA audited ** 12 time rate

Proceedings of the IEEE

The acknowledged leader for highly technical, in-depth articles on specific electrical/electronics subjects. 40,000 of the world's most advanced electrical/electronics engineers pay $3.00 a year for it in addition to their IEEE dues. They design for the future, specify products in the prototype stage and rely on *Proceedings* to keep them well in advance of the times.

the institute of electrical and electronics engineers inc.

ADVERTISING OFFICES: New York 72 West 45th Street, 212 MU 2-6606
Boston 617 MI 3-2714 • Fairfield, Conn. 203 ED 5-3938 • Chicago 312 FI 6-8242
Minneapolis 612 FE 6-5357 • Dallas 214 WH 3-4266 • Los Angeles 213 HU 3-8530
San Francisco 415 EX 2-6794 • Paris Central 2716 • Tokyo 291-0205

Technical correspondence

Evolution or revolution?

Is it desirable that we make our patent laws "uniform with those of the rest of the world?"

Whatever our patent shortcomings may be, the United States system has been one that promotes publication, has advanced technology in the United States at what appears to me to be a greater rate than that of England or France, and has given me, as an independent inventor, my licenses, and the public some degree of satisfaction.

With reference to "Spectral lines" in the April issue of SPECTRUM, my reactions to the three "consequential" changes are, first, that patenting to the first to file would be most unjust. The present U.S. system provides for the patent to be granted to the first to invent. Although the "first to file" concept might lead to fewer interferences, it would retrogress to a system we abandoned nearly a century ago. Second, "preliminary" application would revive the old "caveat," which was abandoned. Third, computer programs are somewhat like "blank forms" and blank graph sheets; they are not currently patentable nor recognized by copyright law. Trademarks would not seem to offer protection. Patents would be of doubtful value since enforcement would be almost impossible. Yet, when one *invents* a new, useful, time-saving programming method, I'd favor his being permitted to reap some reward. Perhaps a special class comparable to plant and design patents could be evolved.

Personally, I'd rather see our present system evolve than suffer a revolution.

Paul W. Klipsch
Hope, Ark.

Solid-state opportunities

I read with interest the article entitled "Limitations in Solid-State Technology" by E. G. Fubini and M. G. Smith (see SPECTRUM, pp. 55–59, May 1967). I would like to emphasize one important point not discussed in the article and which seems to be overlooked under the present technological developments.

If we examine the past five or six years we see that what we have done, with the exception of some isolated cases, such as surface studies, is to decrease the size of elements in a given area for the sake of improving the gain–bandwidth product of circuits. Nobody can deny the advances in the fabrication technology and its effectiveness in realizing improved devices. If we extend our memories beyond six years, we find that one of the advantages of a solid-state device was discovered to be its functional property. For example, a Shockley diode, connected properly to a battery and a resistor, could perform a bistable flip-flop action as well as a stable multivibrator action with frequencies in the range of a fraction of a cycle.[1] Similarly, the tunnel diode later proved to be a useful device both in analog and in switching areas.[1]

Thus, the negative-resistance properties of the above devices proved to be very useful, provided they were being utilized properly. Unfortunately, not enough time was devoted to the proper utilization of such devices. Other useful semiconductor elements, such as delay elements, unijunction transistors, etc., were neglected for the sake of improving the fabrication technology. For example, delay elements can be used very effectively in building parallel to serial converters in digital systems and, most important of all, they can be used as summing elements in additive amplifiers to improve the gain–bandwidth product of existing video amplifiers far beyond their present level.[2] Another important concept is the use of silicon dioxide in connection with active elements to realize memory devices.[2]

The above examples are only a few possibilities among an infinite number of opportunities. With the improved fabrication technology, it is possible to achieve certain functions in a much simpler way and to realize functions that were not previously possible with existing technological concepts.

Vasil Uzunoglu
South Hampton, Pa.

REFERENCES

1. Uzunoglu, V., "Tunnel diode flip-flops and pnpn diodes," *IRE Trans. Electronic Computers*, vol. EC-11, Dec. 1962.
2. Uzunoglu, V., "Some future aspects of microelectronics," presented at WESCON, Aug. 23–26, 1966.

International units

The advantages to IEEE members of using siemens (symbol S) instead of mho as the unit of conductance are: (1) We would at last have a truly internationally accepted term and symbol, which would facilitate the exchange of scientific information on a worldwide basis. (2) The symbol S is much more convenient[1] for use in the laboratory and for the reader, printer, or typist than ℧, Ω^{-1}, or A/V.

I would ask your correspondents (see SPECTRUM, pp. 154–157, Apr. 1967) to remember the difficulties caused by the lack of international agreement when trying to teach a group of students from different continents. One of the most disheartening and frustrating experiences is to find that one's laboratory equipment and associated handbooks, including that from the United States, have terms and symbols that do not follow international standards.

It may be of interest to note that correspondence published elsewhere[1] did not provoke any reaction against the adoption of the term siemens and its symbol S.

G. May
Surrey, England

1. May, G., "Use of S.I. units (Letter)," *The Radio and Electronic Engineer*, vol. 32, p. 232, Oct. 1966.

When units are selected, the total scope of convenience or annoyance is often difficult to anticipate. My typist liked cps much better than Hz because it did not involve use of the shift key on her typewriter. Shifting up and down takes longer than typing one character, so cps types much faster than Hz. My students have adapted, but the vocalized units sound like: "Hurts, Kurt's, Myrt's and Gert's."

I am rather afraid of what they would do with siemens.

Edward J. Gauss
College, Alaska

Mr. May's comments are well taken, but although the name siemens has been recommended for the unit of conductance by the International Electrotechnical Commission, there is some question whether it will become a "truly internationally accepted term." In April 1967 the Consultative Committee on Units of the International Committee of Weights and Measures (ICWM) met in Paris and discussed new short names for derived units in the International System. Both siemens and pascal (for newton per square meter) were considered, but the Consultative Committee decided not to recommend any more short names. Unless the ICWM overrules its own committee, these names will not be added to the International System in the near future.—Bruce B. Barrow.

The problem of standardization of units and symbols is one of expressing the most information in the clearest possible manner, using the least number of words and/or symbols. Any language that does this, be it in words, symbols, or abbreviations, must be a natural tool of the user for him to derive the greatest benefit from it.

In order to serve the purposes of transmitting information clearly, rapidly, and with ease, the language must be a habitual part of one's approach to a job. If ease of expression is suppressed for standardization, the very real danger of a lack of interchange occurs, with the high probability of disastrous consequences. Thus, the language must be developed for the user—not the habits of the user for the language—if optimal communication is to result.

If for some *logical* reason the majority of users agree to a modification of this language, be it for standardization, clarification, or just for arbitrary change, then, to the extent that the modification satisfies the majority will, such a change should be made. My sampling of the majority is limited to my immediate associates and thus not truly representative. However, this group, with diverse educational and geographical backgrounds, disagreed—to the last man—with *all* of the changes discussed in the previous articles and with some changes not mentioned in these articles. When closely questioned, the unanimous consensus was that there is no logical basis for the change of the names of units as proposed.

The new system of units is not natural. It requires a translation from the stated unit to one of familiarity and, in doing so, serves no useful purpose to the user, for whom the language should be a servant, not vice versa.

It is agreed that there is some logical justification in the change to "hertz" for the unit of frequency. There is some chance that a misunderstanding due to the possible suppression of the $\sec^{-1}$ in cycles/second would occur. However, to my knowledge, this has never occurred and the use of cycles/second is a very useful tool for explaining to the uninitiated what frequency is, using graphical means.

To add insult to injury, a colleague

telephoned an editor of a representative trade journal on this question. This editor stated that, on a national basis, the editors of similar magazines had met and chosen their own standard set of notation, different from that recommended by the IEEE, and, unless severe criticism resulted, such notation was fixed.

We are thus faced with three sets of notation: (1) our natural, historically developed one; (2) the IEEE recommended notation, with its slight variations on the international standard; and (3) the publication standard, which differs slightly from (2). Why should we change "db" to "dB," "h" to "H," "g" (gauss) to "G," "pf" to "pF," "mc" to "MHz," etc., when for years these have been perfectly acceptable and usable units, engendering no confusion and admitting to ease of understanding and ready application?

If we must change something for the sake of consistency, let it be the rules (for example, Section 3.2 of IEEE Standards Publication No. 260) so that the language described and guided by them is truly that of the user, not an artificial imitation (for instance, see Webster's New Collegiate Dictionary, "The Use of Capitals," 1959, p. 1155).

D. E. Rogers
Shrewsbury, N.J.

I have read the letter on "International Units" by R. W. Beatty in the April 1967 issue of SPECTRUM and would like to make a few comments on it.

Mr. Beatty states that "the adoption of the International System of Units as a whole was probably a good thing" but he wants to make a few *exceptions*, such as the "hertz" and the "dB." He thinks that these are very unpopular. This is a matter of opinion, since I think they are very popular. However, this looks to me more like a case of: "I am all for it, *but* don't touch my pets." If Mr. Beatty would take the time he would find that "hertz," for instance, has been in use in many countries outside the United States for many decades. It is nothing new, it is just new to a few people here.

In this connection, I would like to commend the IEEE for taking this forward-looking stand on the use of proper units and having the courage to support a change long overdue. I think this is progress. I would like to see more editors and publishers adopt a firm policy and insist on using the International System of Units. If there is no mandate to adhere to the rules, then there will always be a few who want to have their own pet units and we will never get on the same "language" with other nations. The *Journal of Applied Optics* has set a very good example. They had the courage to state in their Information For Contributors: "The metric system of units *will* be used in all papers published in Applied Optics." I think this is a very necessary decision to ensure the uniform use of the International System of Units.

This is also to let you know that I like IEEE SPECTRUM very much because it contains articles of general interest to everybody in the IEEE and which are not too highly specialized. I consider IEEE SPECTRUM an important and necessary link between all the specialized Groups of the IEEE.

M. Bodner
Burbank, Calif.

Error of omission

It has been called to my attention that a minor error appeared in my article "Interaction Between Light and Sound" (see SPECTRUM, pp. 42–54, May 1967).

The equations for I_1/I_0 in the last paragraph of the section on "Efficiency" on page 48 are given as if efficiency did not depend on light wavelength. Actually, the numerical factor in these equations is proportional to the inverse square of the light wavelength; the figures given hold for 6328 Å, the wavelength most commonly employed in laser experiments.

This is not mentioned anywhere.

Robert Adler
Chicago, Ill.