RCA Review

December 1969 Volume 30 No. 4

PRINTED IN U.S.A.

RCA REVIEW, published quarterly in March, June, September, and December by RCA Research and Engineering, RCA Corporation, Princeton, New Jersey 08540. Entered as second class matter July 3, 1950 under the Act of March 3, 1879. Second-class postage paid at Princeton, New Jersey, and at additional mailing offices. Subscription price in the United States and Canada: one year $4.00, two years $7.00, three years $9.00; in other countries, one year $4.40, two years $7.80, three years $10.20. Single copies up to five years old $2.00. For copies more than five years old, contact Walter J. Johnson, Inc., 111 Fifth Ave., New York, N. Y. 10003.

RCA REVIEW

a technical journal
Published quarterly by
RCA LABORATORIES
in cooperation with all subsidiaries and divisions of
RCA CORPORATION

VOLUME 30 DECEMBER 1969 NUMBER 4

CONTENTS

RCA REVIEW

BOARD OF EDITORS

A PHOTODETECTOR ARRAY FOR HOLOGRAPHIC OPTICAL MEMORIES

By

J. M. Assour and R. D. Lohman

RCA Laboratories
Princeton, New Jersey

Summary—A photodetector array specifically designed for use in a holographic read-only memory has been fabricated and tested. The array is operated in a nonsynchronous word-organized mode. Each element of the array consists of a PIN diode and two p-channel MOS transistors. There are 256 such elements in the array arranged as 16 words of 16 bits each. The array combines the attributes of high sensitivity, fast operation, and a large discrimination between selected and unselected elements. In addition, it is fabricated by a sequence of relatively simple processing steps.

Introduction

A TYPICAL CONFIGURATION for a holographic read-only optical memory is shown in Figure 1. A laser beam is deflected digitally to one of x^2 positions. (x being the number of positions in each of two orthoganal directions) At each position, the beam strikes a hologram in which is stored a page of information consisting of y^2 bits. The real image of each page appears at a given point in space regardless of which hologram is being struck by the beam. Thus, a single photodetecting array of y^2 elements is required to sense the output of the hologram and convert the pattern of light intensities to electronic signals.

The performance of the memory[1-6] is largely dependent on the char-

[1] V. A. Vitols, "Hologram Memory for Storing Digital Data," *IBM Tech. Disclosure Bull.*, Vol. 8, p. 1581 (1966).

[2] F. M. Smits and L. E. Gallaher, "Design Considerations for a Semipermament Optical Memory," *Bell Syst. Tech. Jour.*, July 1967.

[3] L. K. Anderson, S. Brojdo, J. T. LaMacchia, L. H. Lin, "A High-Capacity, Semipermanent Optical Memory," IEEE Conf. Laser Engineering and Applications, June 1967, *IEEE Jour. Quantum Electronics*, June 1967.

[4] J. L. Reynolds, T. Y. Tao, and R. S. Schools, "An Experimental Optical Memory System," *NEREM RECORD*, p. 78, Nov. 1968.

[5] D. H. R. Vilkomerson, R. S. Mezrich, and D. I. Bostwick, "Holographic Read-Only Memories Accessed by Light-Emitting Diodes," *AFIPS Conf. Proc.*, Volume 33, p. 1197 (1968).

[6] R. L. Gamblin, "High Capacity Read-Only Digital Information Storage Using Holography," Symoposium on Microelectronics, March 3—7, 1969, Toulouse, France.

acteristics of the photodetector array. Since the total capacity of the memory is x^2y^2, the larger the detector array the smaller the number of deflection positions can be for a given total number of bits. The sensitivity of the detecting element is also important, since the total laser power is divided by y^2, in addition to experiencing large optical attenuation in the deflector and hologram itself. Finally, the speed with which the array responds to low-level light signals and the ease with which these signals can be separated from addressing noise determine in part the overall memory access time.

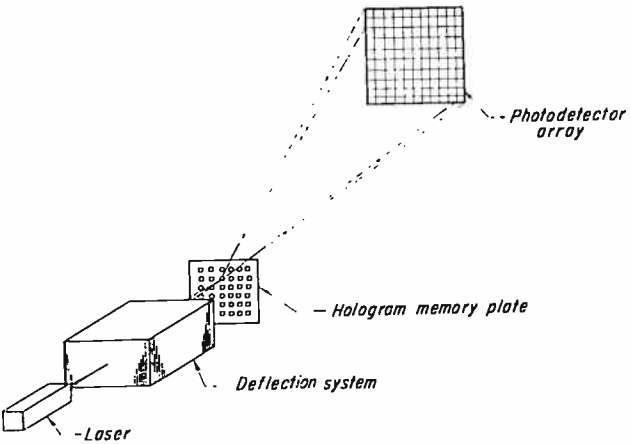


Fig. 1—Typical structure of a holographic read-only optical memory.

Configuration and Principal of Operation of the Detector Array

For application in a holographic memory, it is usually desirable that the detecting array be word organized.[7] Thus, once the light pattern has struck the array, addressing consists of energizing one row (word line) and sensing simultaneously all the columns (digit lines). Since operation of the array is nonsynchronous, no attempt was made to use the detecting device in a charge-storage mode.

The array was designed with silicon p+n−n+ photodiodes as sensors and metal-oxide-semiconductor (MOS) field-effect transistors as switching elements. Silicon p+n−n+ photodiodes are capable of very fast response to pulses of light, typically a few nanoseconds.

[7] R. S. Mezrich, D. H. R. Vilkomerson, and J. M. Assour, "A High-Speed Word-Organized, Photodetecting Array," *Proc. IEEE International Solid-State Circuits Conf.*, p. 58, Feb. 1969.

Their spectral response extends from the ultraviolet region (4000 Å) to the near-infrared region (11,000 Å) and thus matches closely the spectral output of the available light sources, including both solid-state and gas lasers. They are highly sensitive to optical signals with a quantum efficiency of nearly unity. MOS devices are capable of high-speed switching operation when operated as current switches. Their extremely small size (6×10^{-6} cm^2 of chip surface), high input impedance (10^{14} ohms), bilateral current flow, and simple process requirements make them ideal devices for addressing the PIN photodiodes.

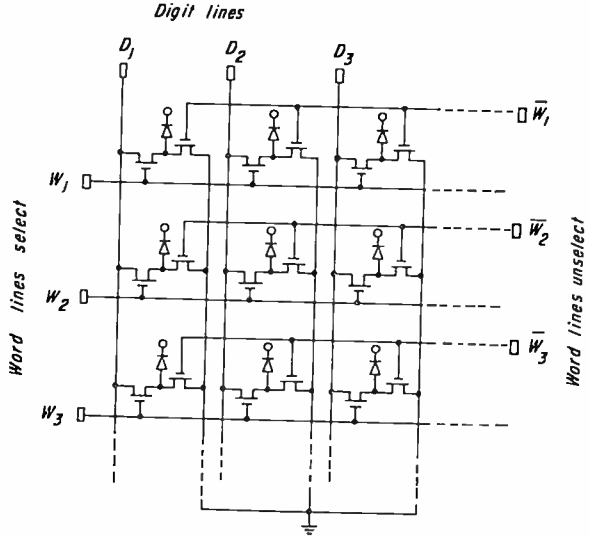


Fig. 2—Circuit configuration of photosensor array. The common cathode of all photodiodes is biased positively.

The configuration of the array is shown in Figure 2 and its cross section in Figure 3. The individual photocell is composed of a photodiode and a pair of p-type MOS transistors. The p+n−n+ photodiodes have a common cathode. Their anodes are connected to the source electrodes of the MOS devices. The gate electrodes of the two MOS devices represent a pair of complementary address lines. The drain electrodes of the MOS devices represent a pair of complementary sense lines. The array is organized with 256 photocells arranged in a 16 words × 16 bits-per-word configuration. As shown in Figure 2, each word-select lines connects to 16 MOS gates, while the correspond-

ing word-select line connects to the 16 adjacent gates. Each digit line connects to the 16 MOS drains, while the digit line connects to the 16 adjacent drains. To operate the array single ended, the second set of digit lines is grounded, as in Figure 2. In the cross section of the photosensor (Figure 3), light is incident on the p+-regions of the photodiodes, which represent the photosensitive surface of the array.

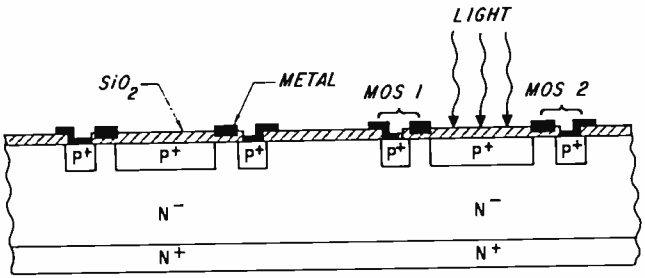


Fig. 3—Cross-section of the MOS-P+N−N+ photosensor array.

In operation, the photodiodes are biased in the reverse direction by applying a constant positive bias to the N+ region. With no radiation incident on the device, the current flowing through the diodes is due mainly to thermally generated electron–hole pairs throughout the semiconductor layers. This dark current is small, of the order of 10^{-7} A/cm^2. Under illumination, the absorbed radiation in the photodiode generates a photocurrent that is ultimately channeled to the appropriate digit line for read-out. Two complementary address lines are used to select a word. With reference to Figure 2, assume that all of the $\overline{W}$ lines are low and all of the W lines are high. All the photocurrent is thus shunted to ground through the ON MOS's driven by the $\overline{W}$ low level. The MOS's connected to the digit lines are all OFF, since their gates are high. When a word is selected, the $\overline{W}$ line for that word is made high and the W line low. The photocurrent for each bit in the word now flows through the digit line, generating a signal for the sense amplifier. For sense-amplifier input impedances less than a few thousand ohms, the current is switched with no appreciable voltage change at the photodiode anode. The discrimination between a selected bit and an unselected bit is approximately R_{OFF}/R_{ON}, where the R's are for an MOS either OFF or ON, and the discrimination between a selected bit and the sum of all the unselected bits is approximately R_{OFF}/NR_{ON}, where N is the number of words. Since the addressing signals are complementary, transient addressing noise caused by capacitive feedthrough from the word lines to the digit lines tends to cancel.

The array can be driven by a single line if the complementary line is fixed biased or grounded, and a differential output can be obtained by ungrounding the complementary digit line.

Fabrication of Photodetector Arrays

The realization of this array has been greatly enhanced by the fact that the p+n−n+ photodiodes and the p-type MOS devices are compatible devices. Both require only one controlled p-type diffusion for the active junctions, followed by a clean oxide layer, as shown in Figure 3. An extremely simple process has been used to fabricate the array. A detailed photomicrograph of one photocell is shown in

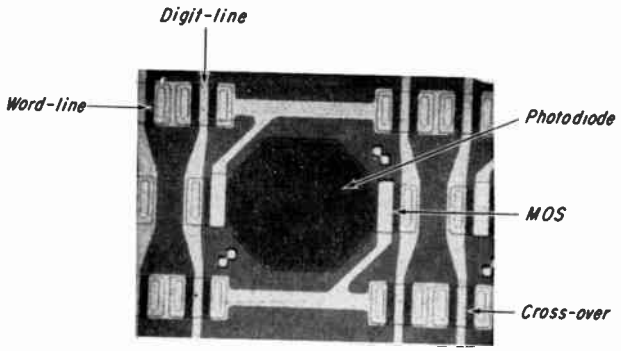


Fig. 4—Detailed photomicrograph of a photodector cell (85 ×).

Figure 4. The photosensitive area of the diode is equal to 3.9×10^{-4} cm^2 with a center-to-center spacing between adjacent photodiodes of 3.9×10^{-2} cm. These dimensions were chosen in the experimental array to match a fly's-eye lens system used in making the hologram storage plates. The MOS devices have a channel area of 12.5×70 microns and a drain area of 37×70 microns, before diffusion. Each photocell contains four diffused crossover regions for the continuity of word and digit lines. Thus, the complete array is composed of 256 photodiodes, 512 MOS devices, and 1024 crossover regions, all addressed by 65 electrical leads. A photomicrograph of a packaged array is shown in Figure 5. Its size is 0.76×0.76 cm.

The processing steps employed during the fabrication of the arrays are given below in their proper sequence. The starting wafers are n-type silicon with a resistivity of 50 ohm-cm and (111) orientation.

The wafers are mechanically lapped and polished on both sides to a thickness of 75 to 85 microns. A layer of thermally grown silicon dioxide is used to mask the silicon surfaces against impurity diffusion. Standard photomasking techniques are used to etch windows in the oxide layer for the diffusion of the p-regions. Heavily doped oxide layers deposited on silicon heated to 325°C are used as the diffusion sources. Silane gas mixed with diborane is used to first deposit p-type doped layers. This step is followed by mixing silane with phosphine to deposit the n-type doped oxide layer. Both impurities are then simultaneously driven in at 1200° C to a depth of 4 microns. The doped oxide layers are etched away and the silicon wafer is passivated

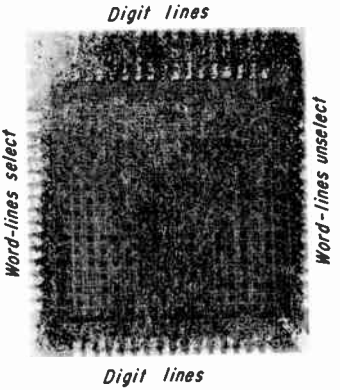


Fig. 5—Photomicrograph of a packaged 16 × 16 photodetector array (approximately 4 ×).

with a clean oxide layer. Its thickness is adjusted for maximum transmission of incident light at a wavelength of 5000 Å, which corresponds to the spectral output of argon lasers. This oxide layer also serves as the gate insulator for the MOS devices. Finally, aluminum is evaporated to provide the necessary electrical contacts and interconnections. Because of the relatively large size of the photosensor chip, only four arrays have been fabricated on each silicon wafer. The yield per silicon wafer and the uniformity of characteristics of the devices in the array have been found quite satisfactory. These results are illustrative of the many advantages[7] in using doped oxides as diffusion sources.

Characteristics of Photodetector Arrays

The performance of the photosensor array has been measured in terms of dc measurements, pulsed operations, and electrical character-

istics of individual devices. P-channel MOS transistors and P+N−N+ photodiode test vehicles were fabricated within and outside the array with dimensions equal to those devices in the array. These test devices are conveniently used to control the applied processing steps, because their electrical characteristics accurately reflect those of the integrated devices which are not readily accessible for testing.

Typical drain characteristics of the p-type MOS transistors are shown in Figure 6. The threshold voltage measured for a grounded

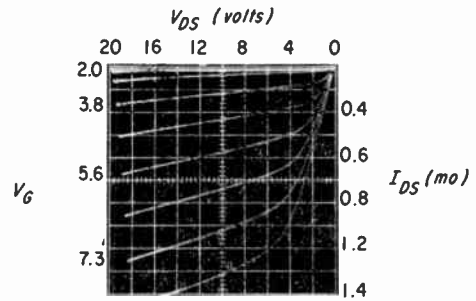


Fig. 6—Typical drain characteristics of the p-type MOS transistors.

substrate is $V_T = -2$ volts. Under dynamic operation of the array, however, a positive potential is applied to the substrate in order to reverse bias the photodiodes. The applied body-to-drain voltage, which is +5 volts, causes an increase of one volt to the initial threshold voltage of the MOS device. Because of this effect, it was found imperative to develop a controlled process for the reduction of the MOS threshold voltage. The transconductance of the MOS device measured at $V_G = -8$ V, $V_{DS} = 10$ V, and $I_{DS} = 1.8$ mA is $g_m = 500$ μmho, and is in agreement with that theoretically calculated. The dynamic resistance is $r_d = 12$ kilohms. Dark currents between the active junctions, insulators, and the silicon substrate do not exceed 10^{-10} ampere at an applied bias of −10 volts. Similarly, dark currents measured for the P+N−N+ photodiodes are less than 10^{-9} ampere. The quantum efficiency of the photodiodes was measured with three different light sources; argon laser ($\lambda = 4880$ Å), He-Ne laser (6238 Å), and GaAs laser (8480 Å). The quantum efficiency at $\lambda = 4880$ Å is $\eta = 0.55$; it increased slightly at higher wavelength. The response time of the photodiodes to optical signals was measured under a GaAs laser. Typically, the photocurrent response consisted of two parts. The fast part of the response, which represented more than 70% of the total

current output, was less than 5 nanoseconds, i.e., the resolution limit of a Tektronix 585 oscilloscope. The remaining slow part of the photocurrent varied up to 400 nanoseconds and was dependent on the reverse biasing voltage. It must be noted that since the n− region is approximately 70 microns thick, the total photocurrent is composed of both drift and diffusion currents, thus giving rise to the above transient response. The diffusion currents can be considerably reduced by decreasing the thickness of the n− region and increasing the substrate resistivity.

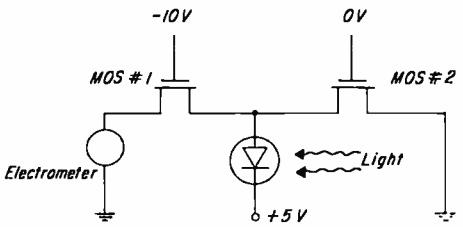
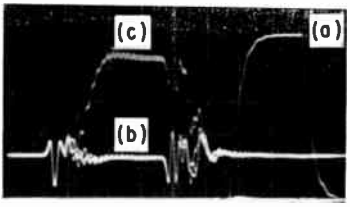


Fig. 7—Circuit diagram of a photocell under illumination.

The dc characteristics of the photodetector array and the uniformity of response of the photocells to optical signals is conveniently described in terms of the diagram in Figure 7. A test jig designed with manual switches was used to apply the appropriate voltage to the gates of the MOS devices. A beam of visible light was focused with a microscope to illuminate the area of each photodiode. First, with no light incident on the array, MOS #1 is turned on by applying −10 volts to the gate electrode, while MOS #2 is turned off by applying 0 volts to its gate electrodes, as shown in Figure 7. The detected dark current is 1 nanoampere, which corresponds to the leakage current of the photodiode. When both MOS devices are turned on, the value of the dark current decreases by one half, as expected for a pair of MOS transistors with matched characteristics. With light incident on the photodiode and MOS #1 turned on while MOS #2 turned off, the photocurrent measured is three orders of magnitude higher than the dark current for an incident light level of 6×10^{-1} W/cm^2. If, however, MOS #1 is turned off and MOS #2 is turned on, the value of the current decreases to the nanoampere level observed under dark conditions. Thus, a high discrimination ratio between selected and unselected bits is observed as predicted.

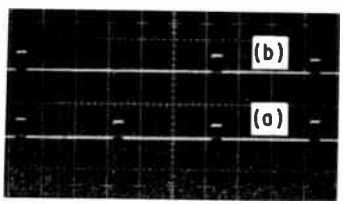
To determine the performance of the array under dynamic condi-



(a) Word-line input pulse (2 V/cm).
(b) Digit-line output response with no light (10 mV/cm).
(c) Digit-line output response with light (10 mV/cm, time scale 0.1 μsec/cm).

Fig. 8—Photoresponse of a photodetector cell.

tions, a driver was designed to deliver a pair of complementary pulses, each 10 volts in magnitude, to the gate electrodes of the MOS devices. The substrate was biased at +5 volts. Visible light illuminated the entire surface of the array. The output at each digit line was measured across a load resistor with an oscilloscope. The detailed response of a photocell is shown in Figure 8. The word-line input pulse (shown offset in time in Figure 8(a)) has a rise time of about 80 nanoseconds. In Figure 8b, with no light incident on the array, the output consists only of noise typical of stray capacitance coupling. In Figure 8c, the output response measured at each selected digit line also has a rise time of 80 nanoseconds, which is limited by the driving circuit. Noise due to stray capacitance is also seen superimposed on the output pulse. A result illustrating the high discrimination between bits achieved in the present array is shown in Figure 9. Response (a) is the output from four words detected at one digit line with the entire surface of the array illuminated. Response (b) is the output of one digit line when the entire array is illuminated except one photocell.



(a) Output of one digit-line when the entire surface of the array is illuminated.
(b) Output of one digit line when the entire array is illuminated except one photodetector cell (scale 5 mV/cm, 20 μsec/div).

Fig. 9—Photoresponse of the photodetector array.

Conclusion

The performance of a novel word-organized photodetector array has been shown to satisfy the basic requirements for application in holographic read-only optical memories. The experimental array has a capacity of 256 bits arranged in an x-y configuration. The array consists of 256 p+n−n+ silicon photodiodes, 512 p-type MOS transistors, and 1024 crossover regions. The photodetector array offers a high discrimination ratio between selected and unselected bits, different modes of operation, and simplicity of fabrication. It is believed that with current state-of-the-art semiconductor technology, larger capacity arrays of this type should be readily realizable.

Acknowledgment

We wish to acknowledge the valuable assistance of D. Tarangioli and J. Murr during our studies.

A REFLEX ELECTRO-OPTIC LIGHT VALVE TELEVISION DISPLAY

By

D. H. Pritchard

RCA Laboratories
Princeton, N.J.

Summary—A high-brightness television display device that utilizes the electro-optic effect in crystalline material has been developed and evaluated. An element composed of electro-optic crystalline material mounted in a vacuum chamber is scanned by an electron beam, and a charge pattern is produced on its surface by intensity modulation of the electron beam. Polarized light from an external source is passed through the electro-optic material in a reflex optical system. The intensity of the light is controlled by the charge pattern and results in a television image projected on a screen. Using an "off-the-air" television signal, this display device has demonstrated a resolution limit of 250 lines per inch on the electro-optic element, a decay time of 1/30 second, and a large area contrast ratio of at least 50:1. Picture brightness on a one-square-foot unity-gain screen of between 50 and 100 foot lamberts has been demonstrated using tungsten light sources ranging between 50 and 250 watts of input power and a variety of light-collection systems. This brightness was achieved with a unique reflex optical system whose light transmission efficiency was 32%. Similar performance has been obtained with electro-optic crystal elements of potassium dihydrogren phosphate (KDP) and potassium dideuterium phosphate (KD_2P). Out-gassing and sealed-off tube life problems have not been investigated with these materials.

Introduction

In LIGHT VALVE displays, where an independent light source is used, the brightness is limited by such constraints as optics, cost, and size; this is in contrast to cathodoluminescent displays, where brightness is controlled by energy density limitations. In light valves, the light image is controlled by a medium that modifies light from an independent source in accordance with a spatially distributed electrical charge image. Electron-beam modulation and deflection provides a convenient technique for writing the image charge pattern on the control medium.

In practice, this type of light valve is composed of a small cathode-

ray tube containing an electron gun and an electro-optic crystal element as a target. An external light source together with polarizing optics is employed to project television images on a viewing screen. The basic light-modulation phenomenon is the Pockel's effect employing a first-order, linear, electro-optic material operated in the longitudinal mode. The electron beam is scanned over one surface of the crystal target depositing a charge pattern corresponding to the video information contained in the intensity modulation of the beam. The intensity of the light traversing the crystal and polarizing optics is controlled in accordance with the charge pattern. Imaging the crystal plane on a screen produces a television picture.

Off-the-air television pictures observed using this system have a large-area contrast of at least 50:1, speed of 1/30 second, and limiting resolution of 250 television lines/inch on a one inch by one inch electro-optic element. The electron-beam current required at 15-20 kilovolts averages 10 microamperes. Operating picture-brightness values of between 50 and 100 foot-lamberts on a one-square-foot unity-gain screen have been produced using tungsten filament light sources ranging between 50 and 250 watts of input power and a variety of light-collection systems.

Some of the system limitations lie in the areas of subjective picture sharpness, uniformity, vacuum operation, life, and material-handling problems.

Electro-Optic Materials

The specific electro-optic materials investigated were single crystals of potassium dihydrogen phosphate (KDP) and potassium dideuterium phosphate (KD_2P). These crystals are of the $\bar{4}2m$ crystal class and exhibit a linear electro-optic effect when an electric field is applied.[1,2] The electric field when applied in the direction of the Z axis of the crystal, causes an induced birefringence, or a change in the refractive indices, that is proportional to the magnitude of the field. If light is passed through the crystal in the Z direction, the induced birefringence causes a differential retardation between the ordinary and extraordinary rays and is called the longitudinal mode of operation. When plane polarized light is passed through the crystal, the induced birefringence produces elliptically polarized light. A "crossed" analyzer

[1] B. H. Billings, "The Electro-Optic Effect: Uni-axial Crystals of the Dihydrogen Phosphate Type I, Theoretical," *Jour. Opt. Soc. Am.*, Vol. 39, No. 10, p. 797, Oct. 1949.

[2] R. D. Carpenter, "The Electro-optic Effect in Uni-axial Crystals of the Dyhydrogen Type: III. Measurements of Coefficients," *Jour. Opt. Soc. Am.*, Vol. 40, No. 4, p. 225, April 1950.

added to the system results in the selection of the orthogonal direction of polarization. Thus, the intensity of the light passing through the system can be modulated in proportion to the applied electric field (see Figure 1a).

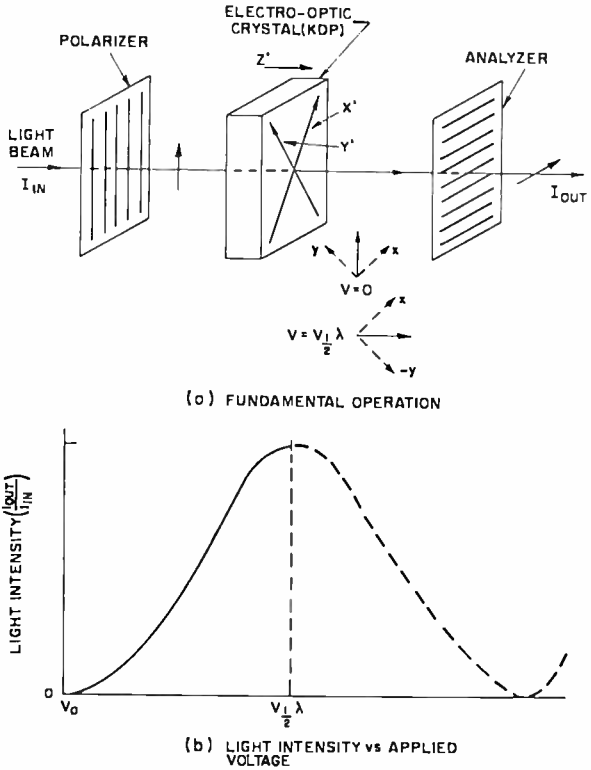


Fig. 1—(Top) Fundamental operation and (bottom) relative light intensity versus applied voltage.

The light intensity varies as a sine-squared function of the applied voltage (Figure 1b) in accordance with the following expression:[3, 4]

$$I_{out} = I_{in} \sin^2 (\pi\Gamma)$$

where I_{out} = output light intensity

I_n = input light intensity

[3] *Electro-Optic Light Modulators*, Baird-Atomic, Inc., March 1965.

[4] J. F. Nye, *Physical Properties of Crystals*, Oxford University Press. London, 1st Ed. 1966.

Γ = retardation = $(n_o^3 r_{63} V_z)/\lambda$,

n_o = ordinary index of refraction

λ = wavelength of light in air

V_z = applied voltage in Z direction

r_{63} = electro-optic coefficient.

The relative light transmission is given by

$$T = \sin^2\left(\frac{\pi}{2}\frac{V}{V_{\lambda/2}}\right),$$

where T = relative light transmission

V = instantaneous applied voltage in the Z direction

$V_{\lambda/2}$ = half-wave retardation voltage.

The light transmission characteristic without surface reflection losses in the visible range is about 96% for both KDP and KD_2P (see Figure 2) for a 1-centimeter-thick crystal.[5] Crystal thicknesses that are required to achieve the best resolution and brightness performance (about 10 mils) have negligible absorption losses. Reflection losses may be minimized by index matching the crystal to the supporting substrate and use of an anti-reflection coating on the outer surface of the substrate. Thus, by eliminating the far infrared (below 1 micron) from the light source with suitable filters, heating of the crystal by intense illumination in the visible from sources up to 500 watts is not a serious problem.

The discharge time constant is determined by the combination of the materials dielectric constant and effective bulk resistivity.

Experimental Light Valve

The electro-optic light valve display device is shown in Figure 3. The single crystal material, 1×1 inch square by 0.010 inch thick is operated with the Z axis perpendicular to the flat faces of the crystal. Plane polarized light is directed through the crystal essentially parallel to the direction of the Z axis, reflected from a segmented aluminum mirror, analyzed, and imaged on a viewing screen with suitable projection optics. The electron-beam addressing requires that a trans-

[5] *Optical Crystals*, Isomet Corp., Bulletin #1101, 1965.

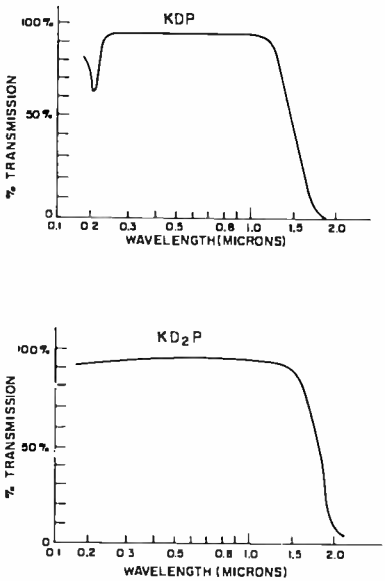


Fig. 2—Transmission characteristic for KDP and KD_2P.

parent conductive coating be placed on one face of the crystal and be maintained at a fixed potential. This potential serves as a reference for the electric field to be produced across the crystal and also as the accelerating potential for the electron beam. The electron beam is directed on the opposite face of the crystal and is scanned over the

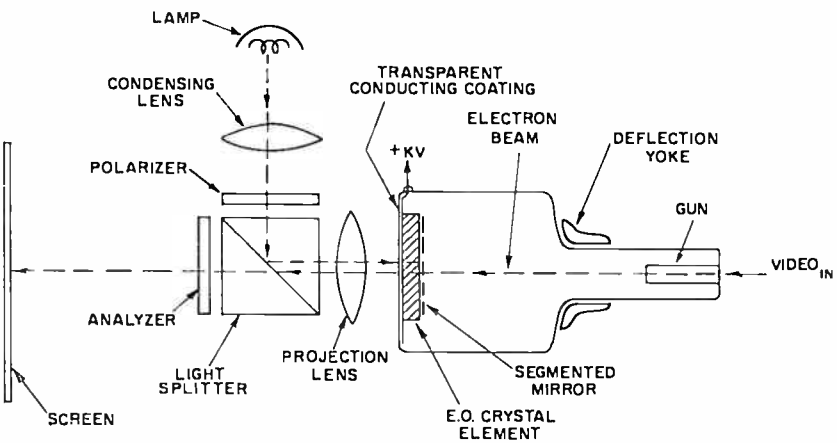


Fig. 3—Reflex on-axis system.

surface of the crystal in a conventional television raster format. The reference potential is maintained at a value higher than that represented by the "second crossover" point, or the potential at which the secondary emission ratio becomes unity. Thus, the electron beam produces a charge on the beam side of the crystal that is negative with respect to the reference potential. The amount of deposited charge is determined by the intensity of the electron beam, which, in turn, is controlled by the video signal applied to the control grid. During picture signal "black", no beam current flows, no charge is produced, and no light is allowed to pass through the analyzer system to the viewing screen. During picture "white", beam current produces a local charge in accordance with a particular picture element. A corresponding birefringence in the crystal is induced, effectively rotating the plane of polarization, and light passes through the system to the screen. The light is thereby modulated in accordance with the video information and a television display appears on the screen.[6-9]

Reflex, On-Axis, Configuration

The electro-optic display device may take any one of several configurations that are classified with regard to the optical system arrangement, the electron beam direction, and whether the light beam transverses the crystal one or more times. The "reflex" optical configuration causes the light to pass through the electro-optic element twice and may be used with either "on-axis" or "off-axis" electron beam alignment. The reflex system may be preferred over single-pass methods due to the two-to-one reduction in voltage requirements for 100% modulation. The lower voltage reduces breakdown problems and increases the sensitivity by requiring less electron-beam current. When the electron beam is operated off-axis the beam-spot shape is distorted and the raster suffers keystone distortion. On-axis beam operation removes these problems, but suffers from a loss of light efficiency due to reflection losses in the mirror placed on the electro-optic element. Figure 3 shows the reflex optical system with the on-axis electron-beam operation that was chosen for these tests. Since a varying charge pattern is formed on the crystal surface by the electron beam, the

[6] V. K. Zworykin and G. A. Morton, *Television*, John Wiley & Sons, New York, N. Y., 1st Ed., 1940.

[7] H. R. Luxemberg and Rudolph L. Kuehn, *Display Systems Engineering*, McGraw-Hill Book Co., 1st Ed., New York, N. Y., 1968.

[8] J. S. Donal, Jr., and D. B. Longmuir, "A Type of Light Valve for Television Reproduction," *Proc. IRE*, Vol. 31, p. 208, May 1943.

[9] E. J. Calucci, B. C. Gangler, and A. C. Clough, "Solid State Light Valve Study," AD602666 Tech. Report.

mirror cannot be electrically continuous. Both dielectric and segmented metal mirrors were considered for this purpose. The dielectric type consisted of continuous evaporated layers of zinc sulfide and chiolite forming a dichroic mirror. Typically, eleven layers were used.

The segmented metal mirror consisted of a thin film (1000 Å) of aluminum evaporated through a mesh (see Figure 4). The dimensions of the mesh used provided 50% coverage and had a pitch of 250 lines/inch (greater than the limiting system resolution).

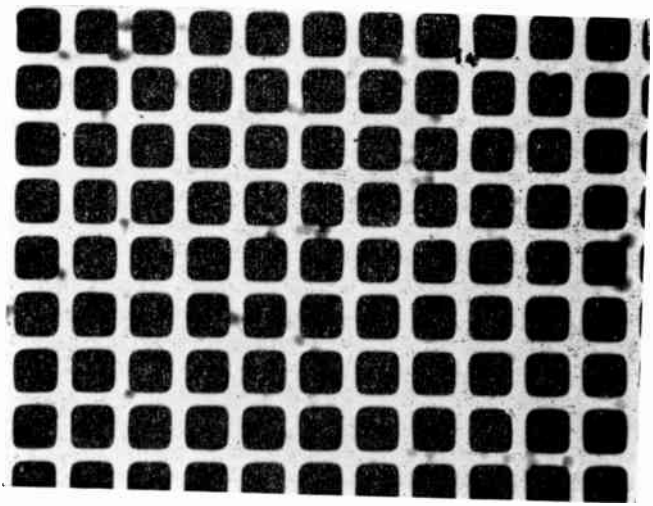


Fig. 4—Photograph of aluminum segmented mirror.

Optical System

The optical system was optimized to provide a transmission efficiency of 32%. The optical system, when a 100% coverage mirror was substituted for the electro-optic element, provided a contrast ratio of about 250:1 as measured over the standard luminosity curve range.

In most light-projection systems, the cone-angle spread is not critical for determining contrast. Therefore the system is optimized for the highest collection efficiency for a given aperture size. In this system, however, the cone angle is important in determining contrast because of the type of electro-optic material used in the tests. When the cone angle spread results in light rays that are not parallel to the crystal Z axis, the off-axis component (orthogonal to the crystal Z axis) is not modulated and is manifested in a reduction of available contrast range. Therefore, the collection efficiency is determined by the maximum allowable cone angle consistent with the desired contrast range. A direct trade-off may be made between system brightness and system contrast

for a given source size and for a given aperture by trading light collection versus cone-angle spread.

An efficient polarizer–analyzer combination can be achieved by the use of a dichroic mirror set between two prisms. Such a configuration is shown in Figure 5. Nonpolarized light enters the system normal to one surface of the prism and passes to the mirror at the Brewster

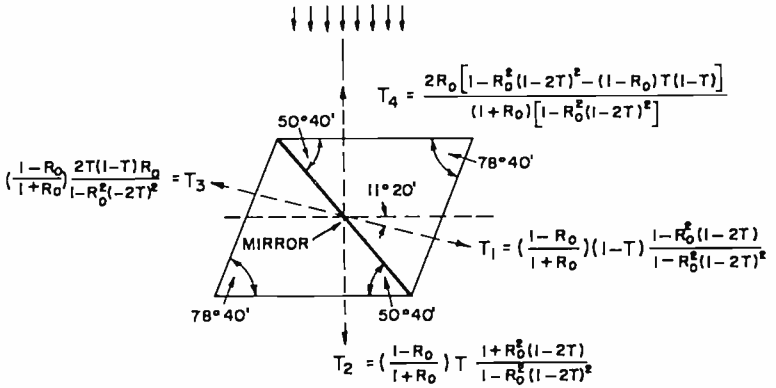


Fig. 5—Dichroic-mirror–prism configuration.

angle.[10, 11] Reflection from the six layers of the dichroic mirror reinforces one direction of polarization and cancels the orthogonal direction of polarization. Thus, essentially 50% of the incident light is reflected as linearly polarized light in one direction and the other 50% passes straight through the system and is polarized at 90° from the reflected rays. For polarized light returning through the system from the electro-optic element, the dichroic mirror acts as a crossed analyzer unit. For parallel light at one wavelength, the efficiency is very high and approaches the theoretical limit of 50%. In the actual case, for the cone angle involved (7-9°) and for operation over the "visible" spectrum, an effective transmission efficiency of 32% is obtained for a 100:1 contrast ratio. Some color nonuniformity results over the image field due to the nonparallelism of the light and the broad spectrum of the source. Improved uniformity and increased contrast ratio results from the addition of pre-polarizer and post-analyzer units. The configuration of the composite unit is shown in Figure 6. A simple and inexpensive method of construction was used for the mirror-

[10] Born & Wolf, *Principles of Optics*, Pergamon Press, New York, N. Y., 1959.

[11] F. A. Jenkins and A. E. White, *Fundamentals of Optics*, 2nd Edition, McGraw-Hill Book Co., New York, N. Y., 1950.

prism units. A box was built with the walls made of anti-reflection-coated glass where light rays must enter or leave the system. The prism boxes were then filled with mineral oil (Nujol) and the dichroic mirrors, previously prepared by evaporation of zinc sulfide and chiolite on a glass slide, were placed in the oil at the proper angle to form the

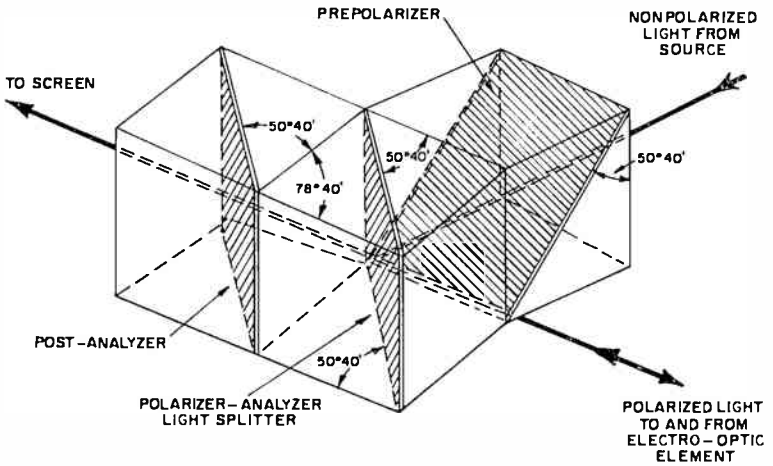


Fig. 6—Dichroic-mirror–prism optical system.

prism. The optical system performance using this dichroic mirror and prism arrangement is given below.

Figure 7 is a diagram of the optical system. The light emerging from the condensing lens is considered as 100% and the percent transmission for each element is indicated. The values on the light path in Figure 7 indicate the percentage of light remaining at each point in the system.

Table I summarizes the performance in tabular form. The overall transmission efficiency is about 32% and is in close agreement with the range of theoretical figures.

Screen brightness values of 40-50 foot lamberts were measured with this system, and operating television pictures from an off-the-air signal source have been demonstrated using a 250-watt light source. Operating picture contract was measured at about 50:1 with 10-mil-thick electro-optic elements. The contrast ratio of the optical system exclusive of the electro-optic element was measured as 250:1. The operating contrast is therefore controlled primarily by the thickness of the electro-optic element and the 7-9 degree cone angle employed. Figure 8 is a curve of system contrast versus electro-optic element thickness determined experimentally.

Table I—System Summary

	% Transmission Individual Ideal Elements	% Light Remaining at Each Stage		Ft. Lamberts Measured 100% Mirror	Ft. Lamberts Measured (Comp. System) 50% Mirror
		Theoretical (100% Mirror)	Measured (100% Mirror)		
1. Condensing Lens Output	100%	100%	100%	31,300	31,300 (1½-inch dia.)
2. Orthogonal Comp. Loss	50	50	50	—	—
3. Pre-Polarizer Light Splitter	90	45	45	—	—
4. Proj. Lens (1st pass)	90	40.5	40.5	—	—
5. Electro-Optic Element & Mirror	100	40.5	40.5	12,500	6,250 (1½-inch dia.)
6. Projection Lens (2nd pass)	90	36.4	36.4	—	—
7. Light Splitter Post Analyzer	90	32.8	28-32.5	—	—
8. Screen	Unity Gain	32.8	28-32.5	107-125	40-66.25 (1 sq. ft.)

Notes:

1. Theoretical efficiency: 29%-32% with white light, anti-reflection coated surfaces, 42%-48% with monochromatic light at 5300 Å.
2. Measured efficiency: 28%-32.5% (white light).
3. Theoretical screen brightness = 123 foot lamberts (100% mirror)
4. Measured screen brightness = 107-125 foot lamberts (100% mirror)
5. Operating screen brightness with 50%-coverage aluminum segmented mirror and electro-optic element = 40-66.25 foot lamberts.

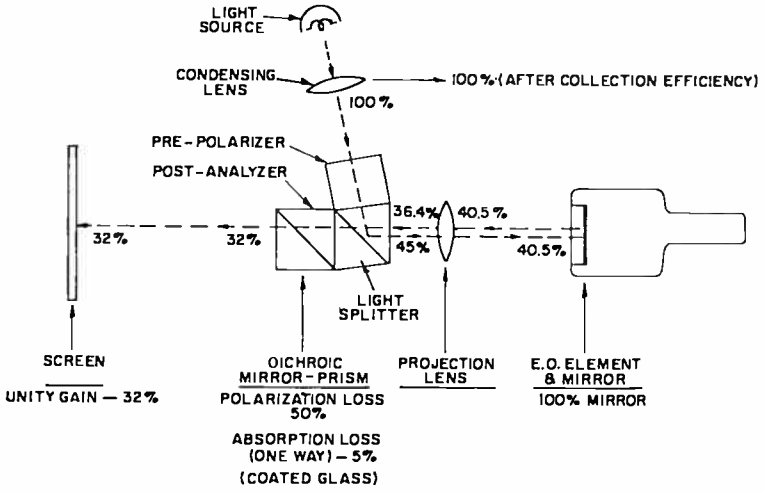


Fig. 7—Present optical system.

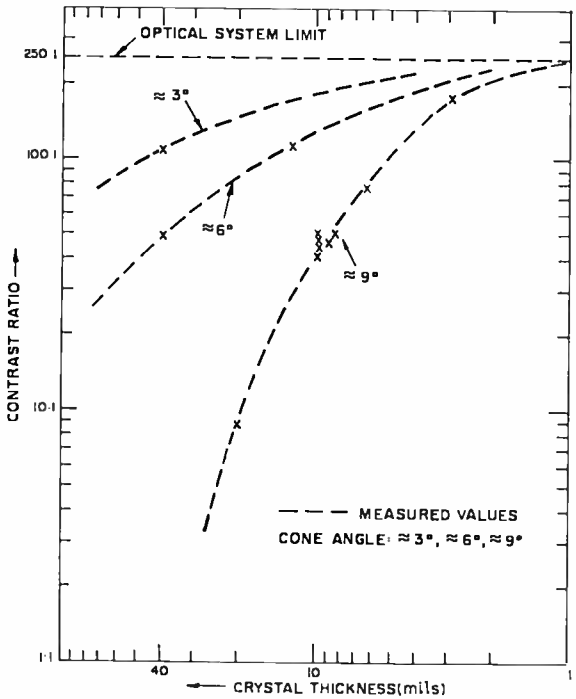


Fig. 8—Contrast ratio versus crystal thickness.

The theoretical analysis of the optical system contrast ratio indicated that the addition of the pre-polarizer would be more effective than the analyzer. Measurements made to confirm this analysis are summarized in Table II.

Table II

	Contrast Ratio		Brightness	
	100% Mirror	Complete System	100% Mirror	Complete System
Light Splitter Only	22.9:1	22.6:1	125 Ft. L.	46.25 Ft. L.
Pre-polarizer Added	100:1	39.6:1	115 Ft. L.	42.5 Ft. L.
Post-analyzer	225:1	47.4:1	107 Ft. L.	40 Ft. L.

The spectral response of the dichroic mirrors is not flat over the visible light range. The use of the light splitter and pre-polarizer alone resulted in some color shading (during picture black). Adding the post-analyzer eliminated the color shading and adds some degree of contrast improvement without appreciable loss. Therefore, the complete unit consists of a single container employing a dichoic-mirror pre-polarizer, dichroic-mirror light splitter, and dichroic-mirror post-analyzer as shown in Figure 6.

Performance Characteristics of Experimental System

A vacuum demountable with the associated electronic equipment was assembled in order to determine the operating performance characteristics of this display device (see Figure 9). The demountable is so designed that both the electro-optic target sample and the electron gun can be readily removed and replaced. A conventional television magnetic deflection system is used and produces a raster scan over the target area. The electron gun is a 5TP4 type using electrostatic focus and is positioned with respect to the target to have a 6-inch throw distance. The electronic circuits feeding the video signals to the grid of the electron gun can be adjusted to maintain flat amplitude-versus-frequency response to a bandwidth of 5 MHz. The signal sources consisted of test-signal generators, flying-spot scanner with photographic test pattern and scenic slides, and live off-the-air television program material.

A phosphor-coated plate is interposed between the target and the electron gun so that the beam-spot size and position can be observed and measured. This plate, mounted on a hinge arrangement, can be

removed from the beam path when the electro-optic target is being tested.

The light source and associated projection optics are mounted outside the vacuum chamber with the light striking the target through an optically flat window.

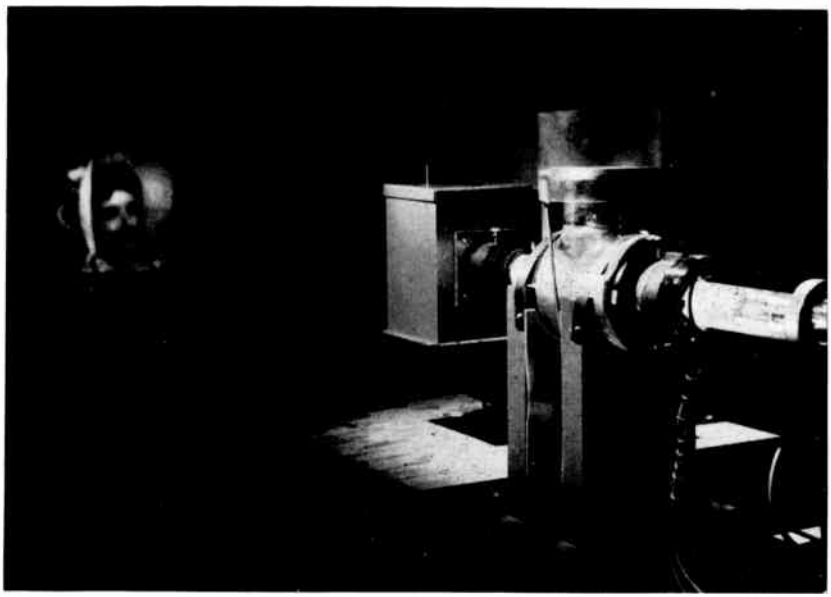


Fig. 9—Electro-optic display device in operation.

1. Resolution

Both the limiting resolution and the equivalent light amplitude-versus-frequency response characteristic are required in order to completely describe the resolution capability of a television display. In terms of frequency response, the limiting resolution occurs at a frequency, or repetition rate, represented by the spacing between adjacent picture points when the two light spots visually merge together and become indistinguishable as individual elements. This point may be readily determined by observation of standard converging-wedge television test patterns designed to indicate equivalent response either in terms of cutoff frequency or in terms of equivalent television lines of resolution.[6, 12]

[12] D. Fink, *Principles of Television Engineering*, McGraw-Hill, New York, N. Y., 1st Ed., 1940.

The equivalent light amplitude versus frequency roll-off characteristic is a direct function of the cross-sectional shape of the intensity of a single-picture element light-output spot. The cross section of the light-spot shape is usually Gaussian or at best a cosine-squared shape. Thus, the equivalent amplitude response rolls off gradually, approximating a 6-dB per octive characteristic. If a spot shape has even wider "skirts" the response will roll off at a faster rate and start decreasing at a lower equivalent frequency. A picture produced by a spot shape having "wide skirts" will subjectively appear soft even though it may have a limiting resolution equal to that of a picture produced by a spot shape having "narrow skirts".

The limiting resolution of the electro-optic light valve is determined primarily by the dimensions of the charge spot (beam diameter) and the charge spreading through the thickness of the crystal. The spacing of the segments of the mirror is made equal to or higher than the resolution capability of the crystal, so that the mirror is not the limiting factor. The beam spot size can be made to be equal to or smaller than the mirror segment dimensions. Thus, the determining factor in the resolution limit is the physical thickness and charge-field spreading in the crystal. Surface-charge spreading also affects the resolution limit and appears to be a constant factor with thickness. The practical limit on crystal thickness is reached at the point where dielectric breakdown of the crystal occurs. For the reflex configuration, the maximum value of voltage is one-half of the half-wave retardation voltage.

The mirror segments deposited upon the crystal surface are made of a metallic conducting material (aluminum). Therefore, when the beam spot lands on a mirror segment, the effective charge-spot shape and dimensions are that of the mirror segment. When the beam spot overlaps two mirror segments, the charge-spot shape is a combination of the electron-beam-spot shape and the mirror-segment dimensions. The resolution capability for a given thickness is somewhat less than that of the original beam diameter. Also, the light intensity spot "skirts" are wider than those of the Gaussian distribution of the electron beam, which results in an equivalent frequency-versus-amplitude characteristic roll-off at a faster rate than that of the beam alone. The beam diameter employed is approximately 3 mils and the beam limiting resolution, in a one-inch raster, lies between 300 and 325 lines/inch as measured by telescopic observation of the phosphor test plate. Thus, with a segmented mirror of 250 lines/inch, the electron beam itself is not a limiting factor in determining resolution capability of the crystal.

A theoretical analysis was made of the charge-field-pattern spread,

assuming the beam lands on a mirror segment. The potential distribution created by a charge on a conducting disk of radius r_o located in the $Z = d$ plane above the conducting ground plane at $Z = 0$ can be evaluated by adding the effect of an image charge of equal magnitude but opposite sign at $Z = -d$. The potential inside the crystal is

$$V_{(x_1, y_1, z)} = V_o \left[\tan^{-1} \frac{r_o}{\alpha_+} - \tan^{-1} \frac{r_o}{\alpha_-} \right], \tag{1}$$

where $\alpha_\pm^2$ are the positive roots of the equation

$$\frac{(Z \mp d)^2}{\alpha_\pm^2} + \frac{x^2 + y^2}{\alpha_\pm^2 + r_o^2} = 1. \tag{2}$$

The radial dependence of the voltage on the crystal surface is obtained from Equations (1) and (2). The variables $\alpha_\pm$ are found by substituting $Z = d$ into Equation 2,

$$V_{(r)} = V_o \left[\tan^{-1} \frac{r_o}{\alpha_+} - \tan^{-1} \frac{r_o}{\alpha_-} \right], \tag{3}$$

where now

$$\alpha_+^2 = r^2 - r_o^2$$

$$\alpha_-^2 = \frac{1}{2} (4d^2 + r^2 - r_o^2 + \sqrt{(4d^2 + r^2 - r_o^2)^2 + 16\, d^2\, r_o^2}). \tag{4}$$

The curves in Figure 10 show the radial dependence of the crystal voltage $V_{(r)}$ for different crystal thicknesses. The voltage at $r = r_o$ has been normalized to unity and the distances r and d are expressed in units of r_o. These graphs show the increased spreading of the potential with increasing crystal thickness.

The limiting resolution of the electro-optic light valve can now be determined for different choices of crystal thickness and mirror-element size. In a system using 2-3 mil mirror elements ($r_o = 1$-1.5 mils) the expected resolution would be 200-220 lines/inch, which is in fair agreement with experimental results (see Figure 11). The somewhat better experimental resolution may be due to the anisotropy of the crystal dielectric constant, since the large dielectric constant in

the crystal a-direction will help to shield the deposited charges in the radial direction.

An extensive series of tests were conducted to determine the limits and the relationship between resolution and crystal thickness.

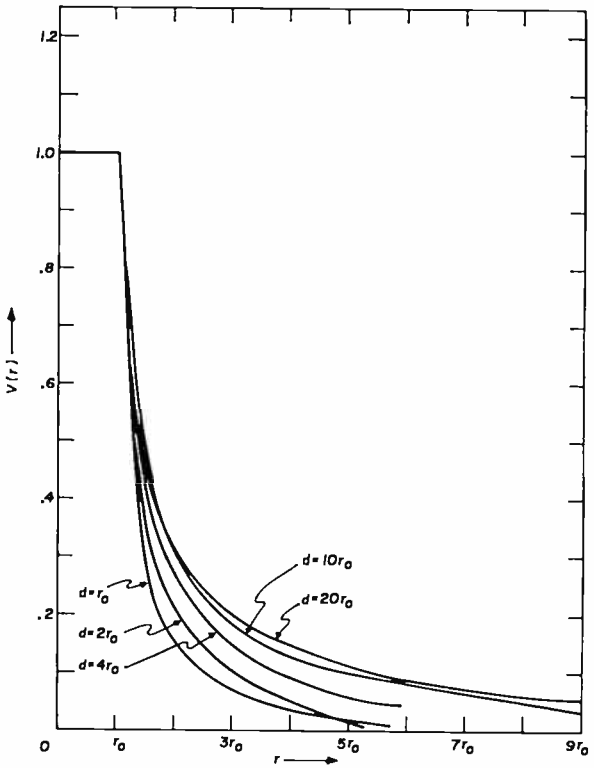


Fig. 10—Potential distribution on crystal surface.

The same crystal blank was used over a variety of thickness values ranging between 40 mils and 3 mils. The values of resolution obtained from these tests were confirmed later with a number of different crystals. The same type of tests were conducted with both KDP and KD_2P samples in order to measure and compare the relative performance of the two crystal types. Figure 12 is a plot of measured limiting resolution versus sample thickness along the Z axis. Figures 13 and 14 are actual photographs taken at the thickness values indicated and show the resolution and appearance of pictures produced by the electro-optic television light-valve system. The dimensions of the

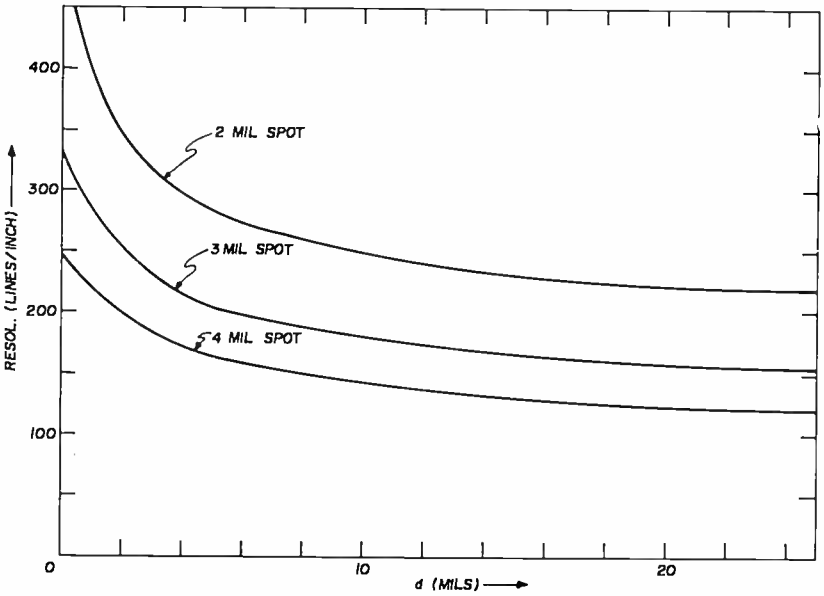


Fig. 11—Resolution based on 30% contribution from adjacent spots.

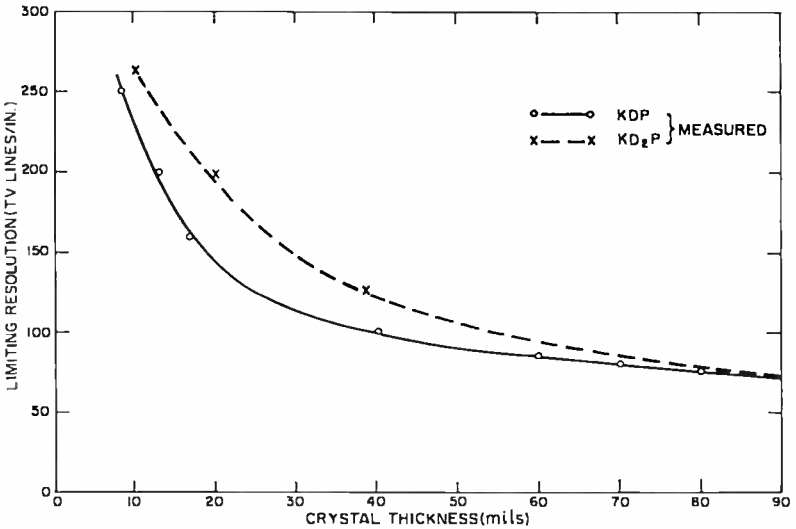


Fig. 12—Limiting resolution versus crystal thickness (measured).

crystal samples used were 1×1 inch, ground and polished to the desired thickness. The segmented mirror deposited on the beam side of the crystal was 250 lines/inch in all cases. The scenes as projected on the screen were photographed directly and consisted of locally generated test patterns, typical slide scenes, and live off-the-air standard television pictures. Figures 13 and 14 also indicate comparisons be-

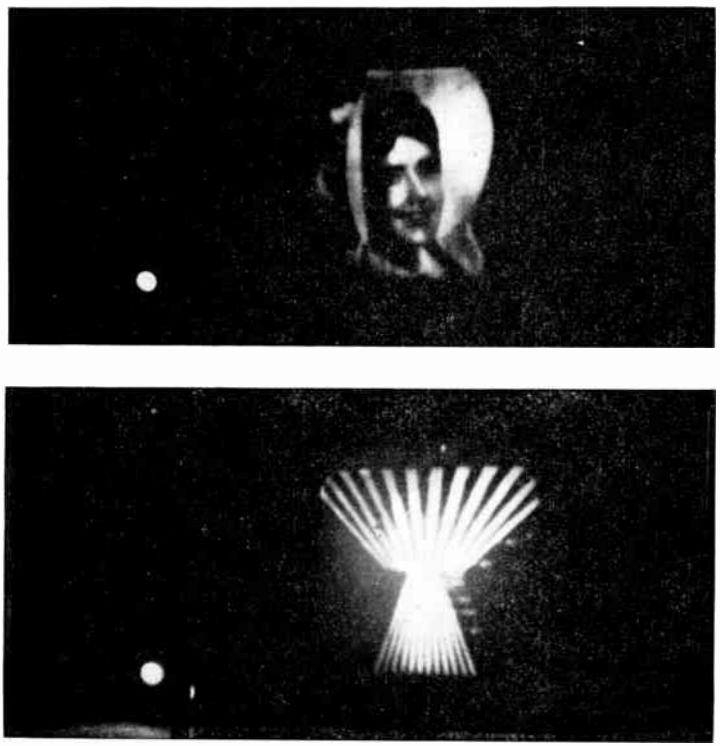


Fig. 13—Photograph indicating resolution (KDP, 8-10 mils).

tween KDP and KD_2P at equivalent thickness values.

As can be seen from the photographs, a limiting resolution of about 250 lines/inch is obtained at thickness values of between 8 and 10 mils for both KDP and KD_2P. A noticeable difference in equivalent frequency-versus-amplitude response is indicated when KDP and KD_2P samples of equal thickness are compared. KD_2P has improved low-frequency response as compared to KDP. This may be attributed to less field spreading, because KD_2P requires less voltage for 100% modulation than KDP (½ retardation voltage).

Thickness values of less than about 8 mils for both KDP and KD_2P

do not produce appreciable increases in limiting resolution and do not represent practical operating thickness values due to dielectric breakdown effects occurring at 100% modulation voltage values.

Effects of secondary-electron redistribution upon resolution and contrast were investigated. A fine mesh was supported ⅛ inch from the beam side of the crystal such that the beam passed through the mesh in order to strike the crystal. The crystal reference-plane potential and

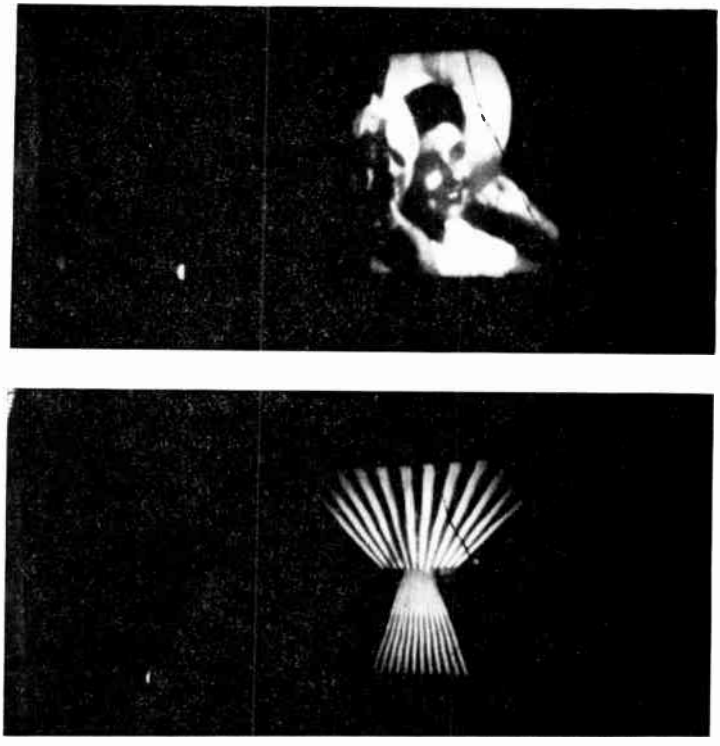


Fig. 14—Photograph indicating resolution (KD_2P, 8-10 mils).

the potential applied to the mesh were made adjustable with respect to each other. When the mesh potential was held more positive than the crystal reference potential, the secondary electrons were attracted to the mesh and prevented from redistributing on the crystal surface. No measurable improvement in resolution or contrast was found as a result of this mode of operation. An equivalent effect arises from the fact that the crystal is mounted in a metallic holder that is connected to the positive reference voltage. Thus, secondary electrons are collected by the crystal mounting. Redistribution effects were observed when the

collector mesh potential was purposely made equal to, or slightly more negative than, the charging potential of the crystal surface.

Standard television resolution may be achieved with this system utilizing KDP or KD_2P crystals whose dimensions are 8-10 mils thick and between 1.5 inches and 2.0 inches in the horizontal and vertical direction. Expanding raster tests on 1×1 inch crystals have demonstrated this to be a fact.

2. *Contrast*

The large-area contrast of the optical system, exclusive of the electro-optic element, is 250:1. Therefore, the operating contrast of the complete system is determined by the characteristics of the electro-optic element and lies in the 50:1 to 100:1 range. When light is passed through the crystal parallel to the Z axis and an electric field is impressed across the crystal in the same direction, the index ellipsoid[3, 4, 11] is distorted in a manner such that an effective rotation of linearly polarized light occurs. When the aperture cone angle spread (degree of collimation of the light) results in light rays that are not parallel to the crystal axis, the off-axis component in the orthogonal direction passes through the system and results in a reduction in available contrast range. Fortunately, the path length is short (in the order of 10 mils) and cone angles of between 5 and 10 degrees result in a contrast range of between 50:1 and 100:1. Figure 8 is a curve of large-area contrast measured over a range of crystal thickness values for fixed cone angles of 3 degrees, 6 degrees, and 9 degrees. Figure 15 is a curve of crystal thickness versus cone angle for a fixed contrast ratio of 100:1. In the actual test set-up the aperture cone angle was approximately 9 degrees. An angle of 7 degrees is required for a contrast ratio of 100:1 with a crystal thickness of 10 mils.[3]

Front-surface reflection from the various optical elements in the system do not represent a problem in this display. The reflections primarily maintain the same polarization as the incident light and are blocked from the screen by the action of the crossed analyzer portion of the optics. The reflections represent a loss in light efficiency, but this may be minimized by use of anti-reflection coatings.

Dynamic large-area contrast measurements were made by operating the crystal with a scanning electron beam adjusted to produce full $\lambda/2$ voltage (100% modulation). A scene involving a large black area and a large full white area was used and the relative brightness of the two areas was measured and compared with the Spectra Spot Brightness Meter. As long as the decay time of the crystal was at least equal to 1/30 second (total vertical scan time), the dynamic contrast range as

well as the highlight brightness agreed with the static contrast and highlight brightness within the limits of the measurement accuracy ($\approx 5\%$). The highlight brightness is reduced in relation to the ambient low-lights (black) as the decay time becomes less than the maximum allowed storage time of 1/30 second.

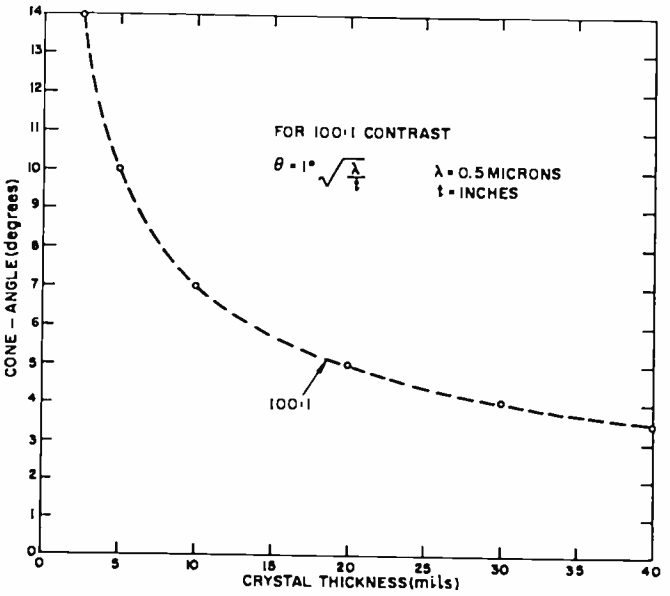


Fig. 15—Crystal thickness versus cone angle for 100:1 contrast.

3. *Transfer Characteristic*

The standard television system basically attempts to provide an overall linear transfer characteristic of light input to light output. Specific pickup and reproducing devices are usually nonlinear in a complementary fashion. The electron gun used in the electro-optic light valve display is operated in a conventional manner and provides a gamma exponent ranging between 1.8 and 2.7. In this system, however, an additional nonlinear element is introduced by the light-modulation characteristic of the electro-optic crystal element. This characteristic, relating electric charge across the crystal and light output, follows a sine-squared law (see Figure 1b).

Two modes of operation are possible to compensate for the nonlinear relationship of light output versus applied charge. The first is to operate over the nearly linear portion of the sine-squared curve at the expense of dynamic range and peak light output. Approximately 85%

modulation[1] may be achieved with this method. The second mode is to provide a relationship between the input video signal and the charge developed on the crystal that is sufficiently nonlinear to compensate for the nonlinearity of the sine-squared response. This may be accomplished by controlling the amplitude characteristic of the video amplifier feeding the control grid of the electron gun.

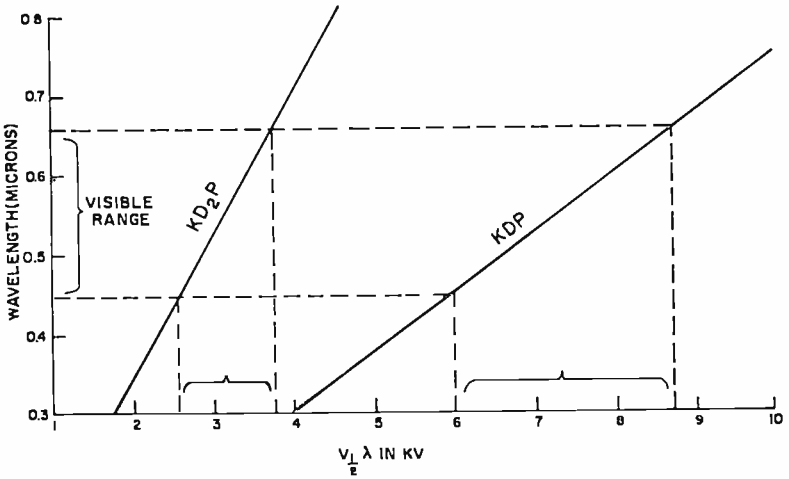


Fig. 16—$V_{\lambda/2}$ versus wavelength.

4. *Chromatic Dispersion*

The $\lambda/2$ retardation voltage for most electro-optic materials is not constant with wavelength of light. A curve of $\lambda/2$ voltage versus wavelength for KDP material is shown in Figure 16.[3,4] Thus, when white light is passed through the crystal the blue components reach 100% modulation at a lower applied voltage than the red components. A color shift occurs between low-light and highlight portions of a scene. If the applied voltage is increased beyond the first half-wave value, the color dispersion increases progressively.

With KDP and KD_2P the color dispersion is subjectively acceptable within the limits of the first half-wave voltage range. Although acceptable in the case of KDP and KD_2P, excessive dispersion in other materials could constitute a problem.

5. *Operating Speed*

The maximum discharge time constant for any one element of a standard television display is 1/30 second. When any point on the

crystal in the electro-optic display is charged to its half-wave voltage (100% modulation) the decay time must be no more than 1/30 second in order to prevent smearing of motion in the scene. A decay time of considerably less than 1/30 second results in a loss of storage time, indicated by a reduction in average brightness of the scene.

The decay-time mechanism in the electro-optic element involves a parallel combination of the crystal capacitance and resistivity on a picture-element by picture-element basis. Theoretically the decay time should be a linear function of the crystal thickness, since the capacitance and the absolute value of the bulk resistance vary linearly with thickness. In actual practice, it was found that the time constant decreased at a faster rate, i.e., in a nonlinear fashion, as the crystal was made thinner. Two factors may contribute to this effect. First, the surface resistivity does enter into the overall discharge-time situation, as evidenced by a change in average time constant when different mounting cements are used, the bulk resistivities of which vary over orders of magnitude. Second, local heating effects caused by the electron beam might be expected to reduce the effective bulk resistivity, since this parameter is known to be a sensitive function of temperature in crystal materials of this type. Bombardment-induced conductivity is probably negligible, since the electron-beam penetration depths are of the order of one or two microns as compared to the typical crystal thickness of 10 mils.

Measurements of operating time constant were made of the same crystal as the thickness was varied over the range of 40 mils to 3 mils for both KDP and KD_2P. A plot of the values obtained is shown in Figure 17. The average time constant for KDP reaches the desired value of 1/30 second at a thickness of between 8 and 10 mils, while KD_2P has a considerably longer decay time and decreases at a slower rate as the crystal is made thinner. The average intrinsic bulk resistivity of KD_2P is somewhat higher than that of KDP. The value reported for KDP is 8×10^{11} ohm-centimeters.[1,2]

A special problem exists in the form of variations in bulk resistivity within the usable area of any one crystal. In a television-raster-scan mode of operation, the charge at any one point is repeated at a 30 cycle rate. If the discharge time constant at any point is longer than 1/30 second (high resistivity), the charge is replenished faster than it can leak off and higher brightness occurs than in an area whose discharge time is less than 1/30 second (low resistivity). This effect has been observed in a number of samples as shaded areas in the reproduced scene. The brighter areas have longer decay time (as much as 1-2 seconds). Examination by a polarizing microscope or by x-ray diffraction does not reveal any optical distortions in the nonuniform areas.

Figure 18 is a photograph obtained with a typical sample indicating this condition. Attempts to control the bulk resistivity during the crystal growth process have resulted in some degree of success in overcoming this problem.

Electron-beam current requirements are a function of discharge time constant (bulk resistivity and dielectric constant of the crystal material), half-wave voltage of the crystal material, raster scan area,

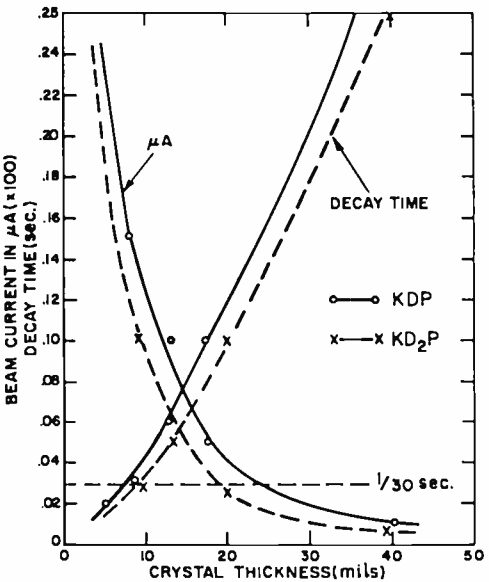


Fig. 17—Decay times versus crystal thickness (measured).

and whether the system is operated in a single-pass or reflex mode for the light beam as it traverses the crystal. The double-pass, or reflex, mode of operation results in a two-to-one reduction in beam current over the single-pass mode, due to the light passing through the crystal two times. In this case, the actual charge potential required for 100% modulation is one-half of the half-wave voltage of the crystal. For KDP this potential is about 3.6 kV (7.2 kV = $V_{\lambda/2}$); for KD_2P, it is about 1.8 kV (3.6 kV = $V_{\lambda/2}$). The measured current for an all-white scene is 10-15 microamperes for a KDP sample operated in the reflex mode and having dimensions of 1 × 1 inch by 0.01 inch thick and with a time constant of 1/30 second. A KD_2P sample required approximately half this current due to the lower half-wave voltage and the longer time constant for equal thickness.

The secondary-emission ratio enters into the beam-current requirements in that the electron beam must supply excess electrons in order to charge the crystal surface in a negative direction with respect to the positive ultor reference potential. The reference potential must be

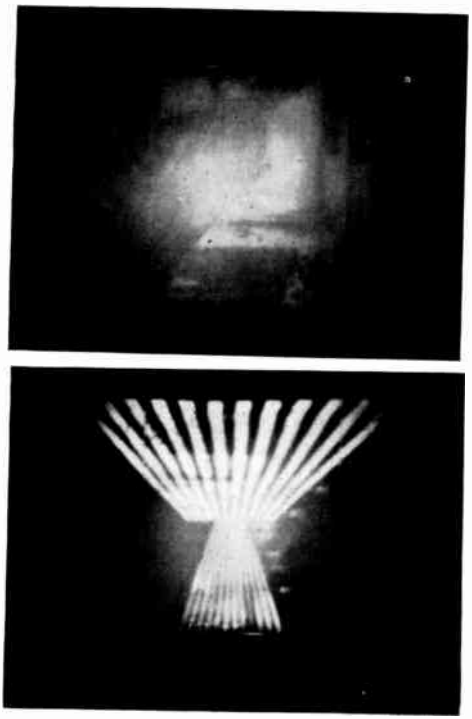


Fig. 18—Photograph indicating poor crystal uniformity (top: test pattern; bottom: blank radar scan).

sufficiently more positive than the "second crossover" point (unity secondary emission ratio) to allow the required half-wave voltage to be devoloped. The second crossover point was measured at 3.2 kV on a KDP crystal surface upon which an aluminum segmented mirror had been deposited (50% coverage and 1000 Å thick). Thus, the ultor potential, in a reflex system, must be at least 7 kV (3.2 kV plus 3.6 kV) to obtain full modulation. A 5-kV potential is adequate for KD_2P (3.2 kV plus 1.8 kV). Crystal samples have been operated to full modulation over an ultor potential range of from 5 to 20 kV. The lower voltages are desirable from the point of view of reducing beam-spot energy, while the higher voltage values are desired from a beam-spot-size consideration. This situation must be examined thoroughly if other materials whose secondary emission characteristics vary over a wide range are deposited over the crystal surface. An example might be in

the case of a continuous dielectric mirror employing alternate layers of magnesium fluoride and cryolite.

6. *Out-Gassing and Electron Beam Deterioration of the Crystal*

The vapor pressure for KDP and KD_2P is satisfactory for vacuum environment at room temperature. However, the out-gassing characteristics at elevated temperatures are unknown. The melting point of 252°C prevents standard bake-out procedures at 450°C from being employed. The life of the crystal in vacuum was not studied.

Long-time bombardment of the surface of the crystal with the electron beam may, or may not, cause changes in the optical or physical properties of the material. Short-term bombardment at velocities up to 25 kV and power levels of 0.25 watts caused no apparent damage. Suitable protective coatings, such as SiO_2, may be evaporated over the beam-side of the crystal surface if necessary.

Summary

An evaluation of the performance characteristics of an electro-optic light-valve display was carried out using commercial television standards as measurement criteria. Feasibility of the system was demonstrated using commercially available electro-optic crystalline materials. Workable techniques were developed to handle 1 × 1 × 0.01 inch crystal materials in an operating system. A high-efficiency optical system was developed for use with this display device.

Performance limits were determined and an operating system was constructed and demonstrated. Off-the-air television pictures were observed having a large-area contrast ratio of at least 50:1, a speed of 1/30 second, and a limiting resolution of 250 television lines/inch on the electro-optic element. Thus, a crystal having dimensions of between 1½ and 2 inches should provide standard television resolution. Operating picture brightness values of between 50 and 100 foot lamberts in a one-square-foot area, using a unity-gain screen, have been produced with tungsten filament light sources ranging between 50 and 250 watts of power input and a variety of light-collection systems.

Acknowledgment

The author would like to gratefully acknowledge the contributions in the areas of theoretical analysis, system planning, and materials handling techniques of J. J. Brandinger, E. Ramberg, J. Wilder, W. Gorkiewicz, J. van Raalte, A. C. Schroeder, and R. Hassell.

A TECHNICAL AND ECONOMIC APPRAISAL OF THE USE OF MICROWAVE ENERGY IN THE FREEZE-DRYING PROCESS

By

A. C. Grimm

RCA Electronic Components
Lancaster, Pa.

Summary—A number of investigators have demonstrated that the freeze-drying process can be greatly accelerated through the use of microwave energy to supply the heat of sublimation. This paper reviews the freeze-drying process in light of existing information to determine if the use of microwave energy in the process can be economically justified. It is concluded that, in general, equipment employing microwave energy is more costly and complex and the cost per pound of product processed is higher when microwave energy is used. Two basic limitations account for the unfavorable economics. First, accelerating the process does not reduce the amount of energy required to sublime the water from the product, and microwave energy is the most expensive of the various forms of energy considered for use in the process. Second, use of microwave energy requires that the processing chamber pressure be approximately an order of magnitude below that required when conventional radiant heat is used.

The use of microwave energy in the freeze-drying process will find its greatest acceptance in the processing of thick products or products whose quality is enhanced through the use of microwave energy in the later stages of the drying cycle. Here, a value trade-off between costs and improvement in product quality will be the determining factor. A typical plant for such products will probably be a continuous-flow system using a combination of radiant and microwave heating.

Introduction

THE OBJECT of this study is to appraise the potential for microwave energy in the food freeze-drying industry in light of published information and the results of independent investigations made at RCA during the past four years. The literature contains many excellent papers showing that the use of microwave energy to supply the heat of sublimation can accelerate the freeze-drying process. Rates of application of heat, vacuum conditions, type of foods, and

other parameters have been studied and reported; however, little has been published on the economics of using microwave energy in the process, probably due to the difficulty of obtaining definitive information on equipment costs. This study attempts to circumvent the problem of accurate costs by examining the various components that make up the freeze-drying system to determine how their ratings and characteristics are influenced by the use of microwave energy. The relative costs of the component for each system can then be compared. By

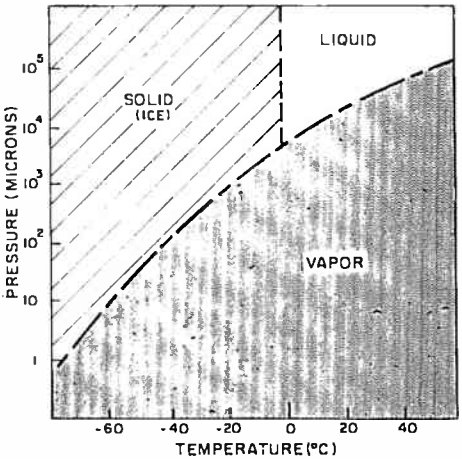


Fig. 1—Phase diagram for water at low pressures.

this technique, a relative cost comparison is made between microwave-accelerated freeze-drying and the conventional freeze-drying process both for new plants and for the expansion of existing processing plants.

Freeze-Drying Process

The freeze-drying process is dependent upon the phase-transition characteristics for water when pressures and temperatures are varied. As can be seen in Figure 1, which is a phase diagram for water at low pressures, ice will transform directly to vapor without liquifying at temperatures below 0.0098°C (32.01764°F) if the pressure is below 4567 microns. This process is called sublimation. To accomplish the transition from ice to vapor, heat must be added at the ice–vapor interface.

If a frozen material is placed in a chamber and carefully heated while the chamber is pumped so that both the temperature of the ice and pressure of the system are maintained in the sublimation region

of the phase diagram, water will be removed from the material without liquifying and, as a consequence, the material will be dried.

The advantage of this process for drying is that the structural integrity and concentration of solutes in the material are closely preserved; thus upon reconstitution by rehydration, the product more nearly resumes its natural appearance, texture, and taste than do similar products dried by other means.

Typical commercial equipment for freeze-drying consists of a vacuum chamber containing many horizontal shelves that are heated by pumping hot fluid through them. The product to be dried is placed in a thin layer on trays. The trays are suspended from a movable cart in cantilever fashion, so that when the cart is rolled into the chamber each tray is suspended between two heated shelves. The chamber is pumped through an interconnection of valves and plumbing by either a refrigerated condenser and vacuum pump or by a steam-ejection vacuum system.

The freeze-drying process consists of rapidly freezing the product, usually by rolling the movable cart with its trays into a refrigerated cold room. After the product is frozen, the cart is removed from the cold room and is inserted into the drying chamber, which is then pumped to the desired pressure. Heat is applied to the shelves where it is transferred to the product by conduction and radiation, causing the ice in the product to sublime. The water vapor thus created is condensed on the refrigerated condenser or pumped away by the steam-ejecting system. A typical process schedule is shown in Figure 2.

The amount of heat required to sublime each pound of ice is 1220 BTU or 357 watt-hours. When the drying process is started, heat reaches the surface ice layer of the product to be dried by direct radiation from the hot shelves and/or by conduction through the bottom of the tray. As the drying process progresses, the ice-phase boundary recedes leaving a layer of dried product surrounding the ice as shown in Figure 3. This dry layer usually has poor thermal conductivity. As a result, the rate at which the heat of sublimation can flow through this insulating layer decreases as the thickness of the dried layer builds up, unless the temperature at the surface of the product is greatly increased. Danger of scorching the product places a practical limit on the maximum permissible surface temperature, and thus a limit on the rate of heat flow and the product drying rate. The long time required to remove the last 10% of the moisture in thick products by radiant heat, see Figure 2, is a direct consequence of the above effect.

Recognition that microwave energy would couple throughout the volume of the frozen material has suggested to a number of in-

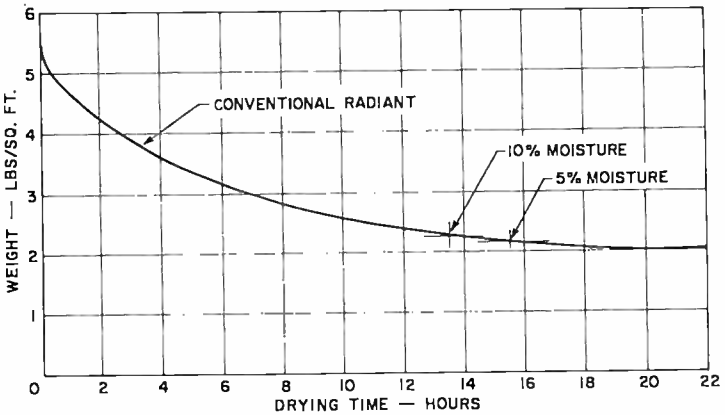


Fig. 2—Typical freeze-drying processing schedule using radiant heat (chopped beef, 1-inch thick).[1]

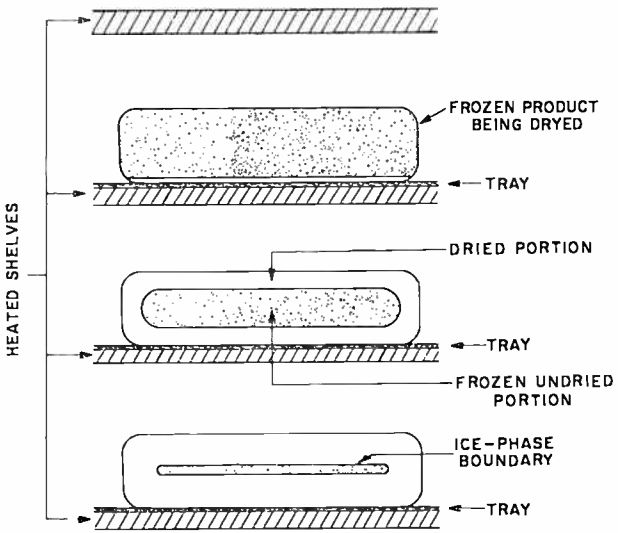


Fig. 3—Regression of ice-phase boundary during freeze drying.

[1] M. W. Hoover, A Markantonates, and W. N. Parker, "UHF Dielectric Heating in Experimental Acceleration of Freeze-Drying of Foods," and "Engineering Aspects of Using UHF Dielectric Heating to Accelerate the Freeze-Drying of Foods," *Jour. Inst. Food Tech.*, Vol. 20, No. 6, p. 103, 1966.

vestigators that it might be an ideal means for supplying the heat of sublimation at a rapid rate without danger of scorching the product. If this method could be used, the freeze-drying process could be greatly accelerated, particularly during the latter stages of drying when the ice layer is shielded from the surface heat source by a blanket of dried product.

Indeed, Hoover et al[1] showed conclusively in their experiments at the North Carolina State University Food Service Department that microwave energy could be used effectively to accelerate the drying rate of foodstuffs (see Figure 4). Their experiments also showed, however,

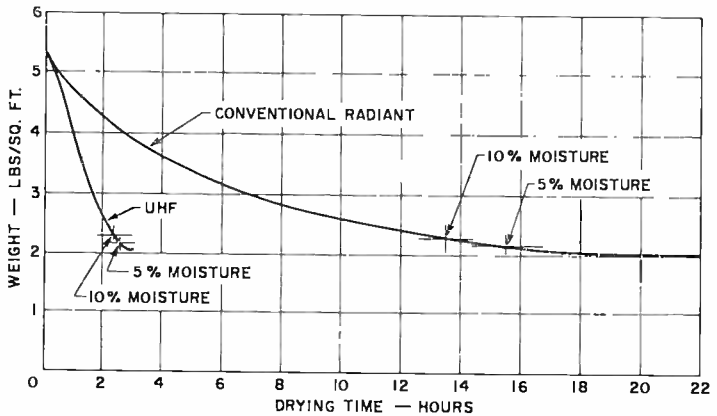


Fig. 4—Comparison of drying curves for UHF and radiant freeze drying of chopped beef 1-inch thick to moisture levels of 0, 5, and 10%.[1]

that unless the chamber pressure was below 50 microns, adequate rf power could not be coupled to the product to be dried without a glow discharge occurring in the chamber as a result of ionization of the residual gases. The glow discharge has several deleterious effects. It absorbs and wastes microwave energy, adversely affects the load on the rf generator, and produces a scorched flavor and taste in the food.[2] Figure 5 shows typical experimental data taken by Parker and Hoover at North Carolina for glow power versus chamber pressure. Similar

[2] D. A. Copson, *Microwave Heating*, p. 160, AVI Pub. Co., Westport, Conn., 1962.

power limitation problems caused by glow discharge were experienced by Parker in other experimental equipment,[3] and other examples have been reported in the literature by Copson,* Leatherman and Stutz,[4] and by Jackson et al.[5] Parker has concluded from his tests that a practical microwave system must operate in a pressure region below 50 microns. This was the same conclusion reached by Jackson et al[5] in 1957.

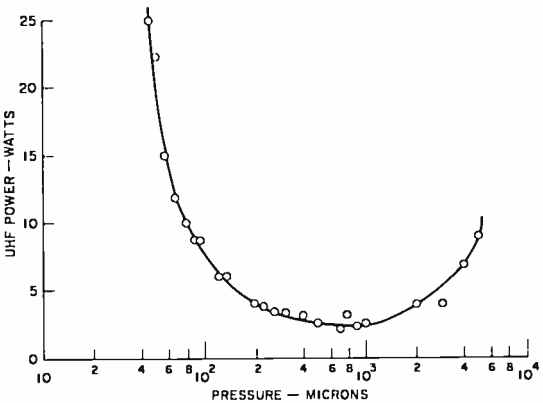


Fig. 5—Typical variation of glow power as a function of pressure for a beef patty 0.5-inch thick at a frequency of 915 MHz.[1]

Parker[6] also calculated the breakdown in air at reduced pressures for dc, 915 MHz, and 2450 MHz. These data, shown in Figure 6, verify that glow is most likely to occur in the region of 200 to 2000 microns—just the pressure region in which many commercial freeze-dry processing plants are designed to operate.

Burke and Decareau[7] state that normal chamber pressures for freeze-drying of foods range between 200 to 2000 microns. It has been

---

* See Reference (2), p. 159.

[3] Private communication.

[4] Leatherman and Stutz, *Freeze-Drying of Foods*, F. R. Fisher, ed., p. 92, NAS-NRC.

[5] S. Jackson, S. L. Rickter, and C. O. Chichester, "Freeze Drying of Fruit," *Food Tech.*, Vol. 11, p. 468, Sept. 1957.

[6] W. N. Parker, "Freeze-Drying," *Microwave Power Eng.*, E. Okress ed., Vol. 2, Academic Press, 1968.

[7] Burke and Decareau, "Recent Advances in Freeze-Drying of Food Products," *Advances in Food Research*, Vol. 13, Academic Press, 1965.

found beneficial to operate at the higher pressures in order to take advantage of the higher thermal conductivity of the dried solids at high pressure. For many fruits or other high-sugar-content foods, it is essential to lower the chamber pressures to 250 microns to prevent thawing and still obtain reasonable drying rates (sugar solutions have lower freezing points than pure water.) In a practical system, the pressure should not be significantly lower than that which permits good product quality, since heat transfer to ice interface may be reduced unnecessarily.

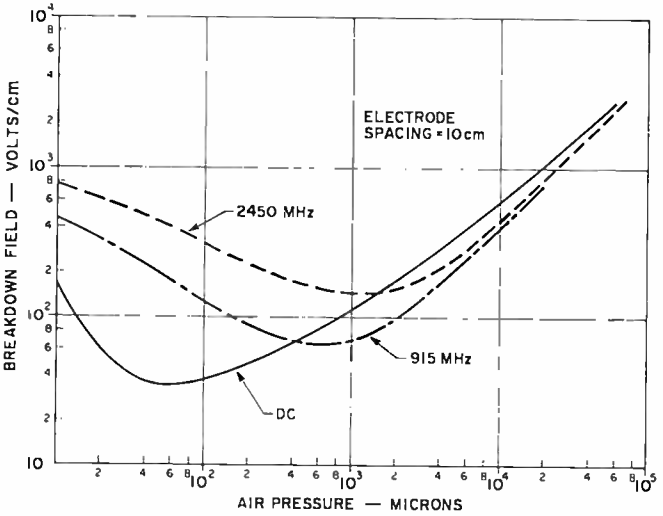


Fig. 6—Vacuum voltage-breakdown field strength as a function of dry air pressure for dc, 915 MHz, and 2450 MHz. The breakdown field is expressed in peak volts/cm.

Commercial equipment manufacturers recommend operation at 700-1000 microns for peas, green vegetables, mushrooms, diced meats and poultry; 300-600 microns for fruits, strawberries and mangoes to preserve color; and 100-200 microns for beverages such as instant coffee and orange juice. In general, it can be said that typical commercial equipment using radiant heat in the freeze-drying process can operate at a pressure level essentially an order of magnitude greater than the pressure level in a system using rf heat. The exception to this is equipment for processing beverages.

It is interesting to note that, although the literature describes extensive investigations directed toward a better understanding of the detailed mechanism of freeze-drying with the objective of accelerating

the process, no references have been found that apply the results of these investigations to determine whether it is possible to build a practical system that employs microwave energy and is cost competitive with conventional systems. Several studies of system economics (rf versus radiant heat) have been published, but these have inadequately considered the implication of glow discharge, i.e., the need for low operating pressure, on the equipment requirements. The following discussion attempts to fill this gap.

Economics of RF-Accelerated Freeze-Drying

Figure 7 is a block diagram showing the essential components that make up a typical batch-type freeze-drying plant.

Since the water rate, i.e., the amount of water removed from the product per hour, is greater during the early hours of processing than during the final hours of processing, it is common practice in large batch-processing plants to have two or more processing chambers operating on staggered processing cycles using valves to direct refrigerant to the condensers and the condensers to the various chambers as the water-load demand changes. With this mode of operation the installed refrigeration capacity is minimized, since the ratio of average-to-peak demand is lower.

If a plant is designed for continuous processing, the batch-processing chambers are replaced by a more complex chamber containing input and output air-lock systems and a product conveyor in the form of a moving belt or cart mover. Other basic components would remain the same except for minor changes in characteristics.

This study will concern itself only with systems using refrigerated condensers to remove the water vapor. Steam ejectors are ruled out, because they are not practical at the low operating pressure (50 microns) needed to prevent glow discharge when rf is used to supply the heat of sublimation. Steam ejectors are most practical where high vapor volumes must be handled at pressures of 1000 microns or above. Lower pressures may be achieved with steam ejectors, but the costs rise rapidly. Pressures less than 300 microns are not considered practical for steam-ejection systems.*

Two questions arise:

(1) When constructing a new facility, will the use of rf provide a plant of lower initial investment and/or lower operating costs?

(2) What is the most economical way to expand an existing plant.

* See Reference (2), pp. 166-170.

Will the use of rf achieve a lower cost for the facility expansion? What will be the relative operating costs with and without rf?

Several generalities are helpful in resolving these questions:

1. Regardless of how the heat of sublimation is supplied, 1220 BTU or its equivalent, 357 watt-hours of energy, must be supplied to sublime each pound of water removed from the product.

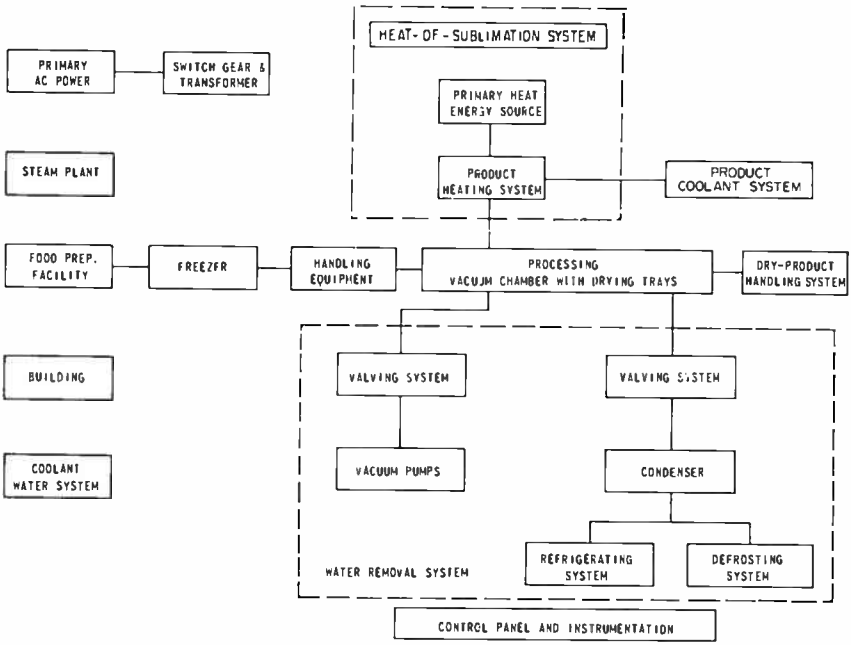


Fig. 7—Block diagram of freeze-dry processing plant.

2. One BTU from steam costs 1/3 to 1/5 the price of equivalent BTU from electricity and 1/10 to 1/20 that of an equivalent BTU from an rf source. The rf generator is assumed to be 50% efficient and tube cost is 2 cents/kWh generated. No capital depreciation or equipment maintenance costs are included in these numbers.
3. Refrigerated condensers will require refrigeration capacity at the rate of 1 ton of refrigeration for each 10 pounds of water removed per hour from the product being dried. The capacity in tons is independent of condenser temperature.
4. The water vapor pressure at the ice surface on the condenser

must be less than the vapor pressure in the product chamber to provide a positive pressure differential between condenser and product drying chamber, i.e., so that the vapor flow will be towards the condenser. This means the temperature of the ice surface on the condenser must be less than that of the ice-vapor interface in the product.

The approach to answering the questions concerning the economics of using rf energy versus conventional means for supplying the heat of sublimination will be to examine the effect that the change in operating pressure will have on the cost of the various components that make up the freeze-drying system. Examples will be based on comparing systems of equal production capacity operating with processing chamber pressures of 1000 and 50 microns.

Condenser

The condenser performs the functions of removing water vapor from the system and maintaing the vapor pressure of the system at the desired level. The first function is accomplished by condensing water vapor into ice on the surface of the condenser. The second function is accomplished by regulating the temperature at the surface of the ice on the condenser. Vapor-pressure-versus-temperature data for ice (see Figure 8) shows that to maintain a pressure of 1000 microns in the condenser chambers, the surface temperature of the ice must be less than −17.5°C. To achieve a pressure of 50 microns the ice surface temperature must be lowered to −46°C.

To maintain constant system pressure, the condenser must condense the water vapor into ice as fast as the vapor is created. To accomplish this, the condenser must remove essentially the same amount of heat that was added as heat of sublimation to each pound of product being dried. Very few additional BTU's must be removed to achieve a lower ice temperature once ice is formed on the condenser surface. For example, if 1220 BTU must be removed per pound of ice at −17.5°C, only 25 additional BTU's must be removed to reduce the temperature of the pound of ice to a temperature of −46°C. (Ice has a specific heat of 0.5.) In general, then, if two systems have equal productive capacity, the amount of water removed and the amount of heat removed by the condenser per hour from each system are essentially the same regardless of system pressure. This statement has two implications. (1) Low system pressure may be achieved without greatly increasing the refrigeration tonnage required (tonnage is basically a function of the pounds of water removed per hour). (2) System pressure has little effect in determining condenser surface area (con-

denser surface area is basically determined by the total pounds of water removed).

So far, only the surface temperature of the ice on the condenser has been considered. In practice, the temperature of the condenser at the

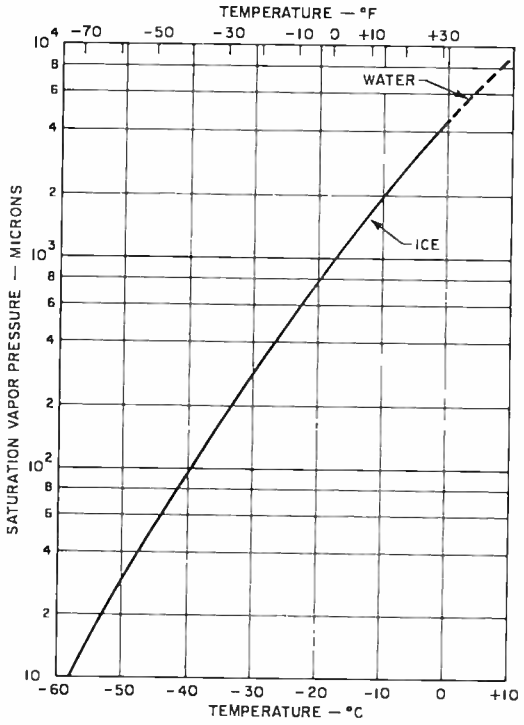


Fig. 8—Saturation vapor pressure as a function of temperature for pure water.

ice–metal interface must be lower than the ice surface temperature in order to maintain the heat flow towards the condenser when the ice layer builds up in thickness. Usually, the condenser is designed to operate about 6°C lower than the ice surface. Since the thermal conductivity of ice is fairly constant over the range of temperature of interest, the same 6°C temperature gradient across the ice layer is suitable for all practical system pressures. This, of course, assumes the condenser surface area per pound of water condensed is the same in each system.

It can be concluded from the above, that operation at lower system pressures will have very little effect on the design and cost of the condenser proper.

Plumbing and Valves

It is common practice to separate the condenser from the product-drying chamber and provide a valve to isolate the chamber from the condenser during periods of defrosting. It is important to determine the relative size of the valves and piping required to achieve equal flow rates for water, which is in the form of vapor, between the product-drying chamber and condenser for various system pressures.

In this case study, it is required that the systems being compared handle the same number of pounds of product per hour. Therefore, the same number of pounds of water must be transferred from the product to the condenser each hour. In other words, the mass transfer rate must be the same whether the pressure in the system is 1000 microns or 50 microns.

Commercial equipment is so designed that the vapor flows into the condenser at a rate such that the pumping speed is limited by the critical flow rate through the valve. Under this condition the condenser pressure is adjusted to be approximately one-half the drying-chamber pressure. Pressure differences greater than this do not increase the vapor flow rate.[5] The mass flow of the water vapor reaches a maximum when the pressure difference between condenser and drying chamber is equal to or less than 0.55 times the chamber pressure. This maximum value is given by the following expression, which is adapted from Guthrie's relation for flow at the critical pressure ratio,[5]

$$\frac{dw}{dt} \sim 24.5 \times 10^{-6} AP \text{ grams/sec at } 20^\circ\text{C}$$

where A is the aperture area in cm^2, P is the drying chamber pressure in microns of Hg, w is mass of gas in grams, and t is time in seconds.

Since it is desired to have the same mass-flow rate in a 50-micron system as in a 1000-micron system, the above relation shows that the valve area must be increased 20 times to compensate for the lower system pressure. A valve for a 50-micron system must, therefore, be 4.5 times the diameter of the valve used in the 1000-micron system. Of course, the size of any connecting pipes must be increased by the same factor to physically mate with the larger valves.

When the volume of vapor sublimed is reduced, as during the latter stages of drying, the vapor flows through the system as a viscous fluid. It can be shown that for equal mass-flow rates of viscous gases through

[5] Guthrie and Wakerling, *Vacuum Equipment and Techniques*, p. 16, McGraw-Hill Book Co., N.Y., 1949.

round pipes, the pipe diameters must be altered inversely as the square root of the entrance pressure at the pipes.[9] This relation is valid only if the ratio of exit and entrance pressures at both pipes are equal. As explained above, the normal mode of operation for freeze-drying systems is to operate the condenser at one-half the chamber pressure, regardless of system pressure. Under this condition, the ratio of exit-to-entrance pressure of all systems is equal; therefore, connecting pipes for systems operating at 50 microns must be 4.5 times the diameter of pipes for a 1000 micron system to achieve equal mass-transfer rates. This is the same factor computed by the critical-flow equation.

It can be concluded that the valve and pipe sizes must be very much larger for a system using rf than for a system using conventional heating, because a lower product-drying-chamber pressure is necessary when rf is used.

One obvious technique for circumventing the need for large pipes and valves is to place the condenser inside the product chamber. Indeed, this has been proposed, and one company, the FMC Corp., has produced units using integrated condensers. Burke and Decareau[7] sum up the pros and cons of integrated condensers as follows:

> "In general, where refrigerated condensers and mechanical vacuum systems are used, they are usually separate from the drying chambers. The arguments favoring segregation are: (1) if separate, the condenser can be isolated and thrawed during the end of the drying cycle, thus saving time for the turnaround; (2) if separate, the condenser can be more fully insulated; (3) direct heat losses from platens are reduced, thus minimizing refrigeration load. On the other hand, integrated drying and condenser chambers have the advantage of the following: (1) more direct path from food to condenser; (2) maximum utilization of space in cylindrical chambers. By far the majority of commercial freeze-dryers have separate condensers, if any. Some of the chief reasons advanced are:
>
> 1) Fear of ice melting and wetting food during unloading.
> 2) Desire to reduce refrigeration load.
> 3) Concern that the ice surface might be higher with internal condensers, with consequently higher pressure.
> 4) Cost of separate units believed negligible.
> 5) Pressure controls more easily adapted to a two-chamber system."

---

[9] VanAtta, *Vacuum Science and Engineering*, p. 24, McGraw-Hill Book Co., N.Y., 1965.

If the freeze-drying plant is built for continuous flow operation, multiple condensers are used. These are generally separated from the drying chamber by some sort of valving to permit periodic defrosting of the condensers. Such systems operating at low pressures must have very large valves if they are to operate at high product rates. If the condenser is built into the drying chamber in order to reduce the valving problem, some means must be provided to periodically scrape the ice from the condenser and to remove the ice shavings without disturbing the vacuum conditions within the drying chamber. This means that mechanical scrapers, shaved ice transfer mechanisms, and vacuum interlocks must be provided within the system.

Refrigeration

The tonnage of refrigeration required for the freeze-drying process equipment is directly related to the rate of removal of water in the ratio of 1 ton of refrigeration for each 10 pounds of water sublimed from the product or ice frozen on the condenser surface per hour. The tonnage required is essentially independent of the pressure in the system. However, the cost of each ton of refrigeration is dependent upon the condenser surface temperature and hence the system operating pressure. Figure 9 shows cost-estimating data for commercial refrigeration equipment. These data are for refrigeration equipment not installed. Low-temperature refrigerating machines are physically larger than high-temperature machines of equal tonnage. Consequently, more floor space is required per ton of refrigeration, and the cost difference between installed low- and high-temperature equipment will be somewhat greater than that depicted by Figure 9. Further, more horsepower is required to drive compressors for low-temperature systems, resulting in greater operating cost per pound of water removed from the product.

The magnitude of the cost difference for refrigeration can be obtained by considering the condenser temperatures for the systems being compared. To achieve a drying chamber pressure of 1000 microns, the condenser chamber pressure must be less than half this value, or below 500 microns. The condenser ice-surface temperature to maintain a vapor pressure of 500 microns from Figure 8 must be −25°C. Allowing 6°C additional for the temperature gradient across the ice layer, the refrigeration system must be designed for −31°C. For a 50-micron drying-chamber pressure, the condenser chamber vapor pressure must be less than 25 microns. The condenser temperature from Figure 8 for this vapor pressure is −52°C. Again, allowing 6°C for the gradient across the ice, the refrigeration system must be designed for −58°C.

Consulting Figure 9, the cost for −58°C (−72.4°F) refrigeration not installed is approximately 1.8 times that of −31°C (−23.8°F) refrigeration for equal tonnage capacity. No figures are available on installation costs.

It is evident from the above that operation at low system pressures increases both the capital cost of the refrigeration equipment and the amount of power required to drive the refrigeration equipment.

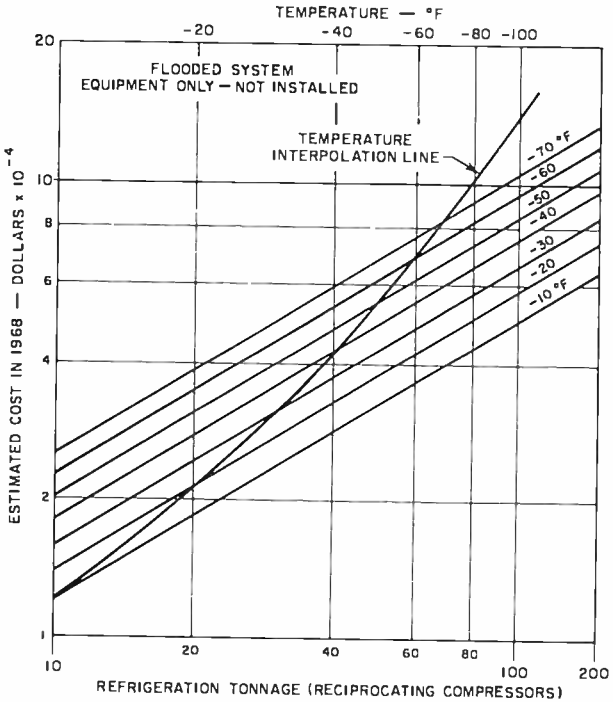


Fig. 9—Estimated refrigeration costs (1968) as a function of temperature.

Product-Drying Chamber

The size and number of product-drying chambers used in a plant is determined by the volume of product to be processed in a given period of time and the need to average the water-rate load to minimize the refrigeration requirements.

In conventional systems, shelf or tray area in the chambers is usually one square foot for each 2 to 3 pounds of wet product processed in a batch. Batching plants using radiant heat have a very high packing

density in the drying chamber. Ten to fifteen trays are closely stacked into a single chamber. Typical commercial chambers require only 150 square feet of floor space for 1,000 square feet of tray area.

If the process can be accelerated or altered to permit greater loading on the shelves, then the total shelf space required will theoretically be reduced in direct proportion to the decrease in process time and reduced in inverse proportion to the increase in product loading.

The use of rf to supply the heat of sublimation can affect both the processing time and shelf loading to decrease the required shelf space. Figure 4 shows the gain in process time for chopped beef achieved by Hoover and Parker. With rf, the process is accelerated about 7 times to the 5% moisture level and 9 times to achieve 0% moisture level.[10] It has also been demonstrated that the thickness of the layer of product can be increased when rf energy is used to supply the heat of sublimation. RF has the potential of reducing the product shelf area requirements by a large factor—possibly a 6 to 14 or more reduction can be achieved with some products.

Some of this reduction in shelf area stands to be lost by advances in techniques for supplying the heat of sublimation by conventional means. For example, Stokes has recently developed a system that moves the product through a number of deep channels in a long heated aluminum tray. Products such as diced meats, peas, and other products that can be cut into small pieces can be processed in 5.5 hours with this technique on a continuous product flow basis. The product is both moved along the channels and revolved or turned over continuously by mechanical vibration of the tray. This movement brings all sides of the product into contact with the heated walls of the channel and hence accelerates the process. With developments such as this, the advantage of rf as it affects shelf area will be greatest with equipment to process special products that are large in physical size, e.g., steaks, chops, large shellfish, etc.

While the use of rf can, theoretically, greatly reduce the shelf-space requirements and hence the number of drying chambers for certain products, the practical considerations of averaging the water-rate load and providing room within the chamber for rf applicators and larger vacuum piping will make it difficult to realize a reduction in number of chambers proportionate to the reduction in shelf area.

A chamber equipped to batch process materials using rf will undoubtedly be more expensive than a chamber using heated shelves. The problem of designing a chamber for batch processing that uniformly

[10] W. N. Parker, and H. F. Kazanowski, Private communication.

heats even one type of product with rf is a formidable one.[6] The design of a chamber to heat many different products uniformly during the entire drying cycle presents an even greater challenge to the rf engineer. It has been proposed that this problem be solved by limiting the use of rf to systems that operate on a continuous flow rather than a batch process. In such a continuous flow system, the rf-applicator design problem is simplified because the electrical properties of the product at a given stage of the processing are always the same, even though they may vary widely during the process cycle.

Heat of Sublimation

This whole study revolves around the technique for supplying the heat of sublimation to the product. Conventional systems use steam as the primary heat source. Heat in the steam is transferred by a heat exchanger to a fluid, such as Mobiltherm 600, which is pumped through the shelves in the processing chamber. The fluid system has the advantage of providing a means for cooling or even refrigerating the shelves of the drying chamber. This is accomplished by cooling the fluid prior to pumping it through the shelves. Cooling or refrigerating the shelves in the product chamber is advantageous for processing certain critical products, particularly biological and pharmaceutical items.

Electrical heating could be used to supply the heat of sublimation, but no commercial installation is known to use this system. It is reasoned that proportional controls for resistance heaters are too costly.

RF energy can, of course, be used to advantage to accelerate the process. However, the cost for the heat of sublimation per pound of product produced will be higher for a system using rf energy. The relative costs between steam, electricity, and rf are subject to much debate. If it is assumed the rf system has a 50% overall conversion efficiency, the cost for electric power for each BTU from rf will be twice that of a BTU obtained by electrical heating. To this figure must be added the replacement cost of rf power tube—estimated to be 2¢ per kWh generated. Steam has been reported to cost 1/3 to 1/5 the cost of electrical energy. The exact figure depends upon availability of steam and cost of primary electrical power. Without being specific about the cost of the different heat sources, it can be stated without much fear of contradiction that the order of ascending cost-per-pound of product produced by the three sources is steam, electric, and rf.

Miscellaneous System Components

The preceding sections have discussed those major components that

make up a freeze-dry plant that would be specifically affected by converting to or using rf energy. Cold rooms to prefreeze the product and dry-product handling equipment are not influenced greatly by the use of rf in the freeze-drying process. Controls and instrumentation would be affected in character but probably not greatly in cost. The building to house the plant would be affected as its size relates to floor space requirements. Product-handling facilities must be geared to whether the system is a batch or continuous-flow process, not whether or not rf is used in the process. Utilities requirements, i.e., steam, coolant water, and electric power, will vary with the type of system. The capital investment in utility facilities will bear a relation to the amount of the utility used to produce a pound of product. For example, transformers to supply power for rf energy must have twice the capacity of transformers when electric heating is used to supply an equal number of BTU. If steam is used, no transformer is needed, but a steam plant is required. Except where the study shows the use of rf will greatly affect the relative costs of these miscellaneous system components, it will be assumed that the slight differences in the capital invested in these facilities will not be a dominant factor in the comparison of the system economics.

Use of RF to Expand Existing Plants

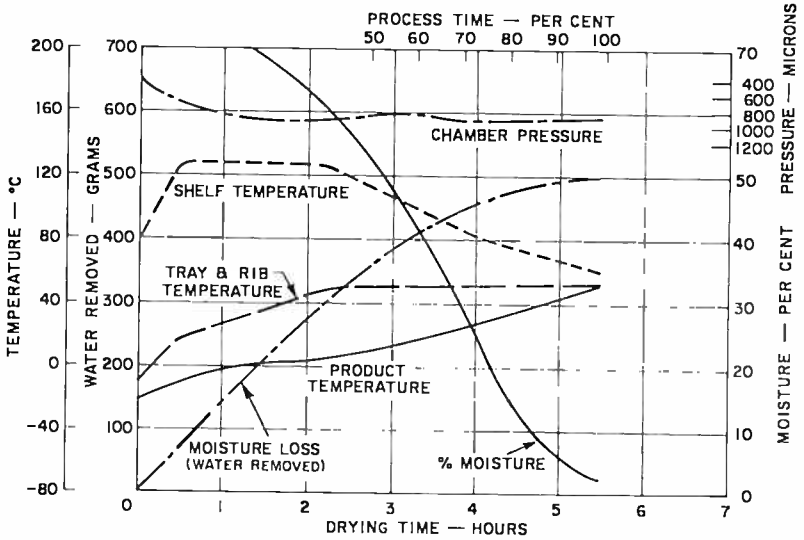
The economics for expanding the capacity of an existing freeze-drying plant through the use of rf to accelerate the final stages of drying will be explored by considering a specific problem. Would it be more economical to expand a batch-type plant capacity 25% by adding more of the same equipment, or by employing rf in the final stages of the drying process?

The process schedule of the plant to be considered is shown in Figure 10. Because of the difficulty in adapting rf to batch-processing chambers, rf processing in the system considered will be done on a continuous basis.

If it is assumed that all components of the existing plant are operating at capacity, 25% more capacity in chambers, condensers, refrigeration and floor space for these items must be added to achieve the 25% increase in capacity by conventional means.

When rf is used, the product has to be dried to some moisture level that will permit it to be transferred without damage from the batch-drying chambers to the continuous rf-drying chamber. The exact moisture level at which the product can be safely handled during transfer without product deterioration is not known, but it is speculated that a moisture level of 15 to 20% will be satisfactory.

If the plant capacity is to be increased 25% by adding rf in the final stages of processing, the existing drying chambers must handle 25% more wet product in the same period of time that they previously handled a capacity load. This means each batch must be removed from the drying chamber in 0.8 or less of the normal drying time. Reference to Figure 10 shows that in 0.8 of the normal process time the product



Representative test:
peas: pre-chilled trays 8¾ × 7½ inches
7 ribs: 1⅜ inches high, ⅞ inch spacing, load = 1.41 inches, sacrified, approx. 3 lb/ft².

Fig. 10—Processing schedule for freeze-dried meat (add ½ hour for load and unload per cycle).

will be dried to the 15% moisture level. The additional time required to load, unload and pump down 25% more batches will shorten the time that can be allocated to processing to slightly less than the 0.8 figure. Despite this loss, the time available for processing should permit the product to be processed to at least the 20% moisture level, and possibly slightly less. The moisture level criterion for safe transfer can therefore just be met with the existing chambers.

Using rf to expand the plant capacity by 25% requires that the total refrigeration tonnage must also be increased 25%; however, a portion of this increased tonnage must be more costly low-temperature refrigeration. If the product is predried to the 20% moisture level in

the existing chambers, it can be determined from Figure 10 that 465 g or 93% of the total water to be removed has been removed when the 20% level is reached. The remaining 7% must be removed in the rf-drying chamber. The refrigeration requirements for the predrying and rf-drying chambers must be allocated as follows:

Predrying-chamber refrigeration—

$1.25 \times .93 = 1.163$ times the original tonnage

RF-drying-chamber refrigeration—

$1.25 \times .07 = .087$ times the original tonnage

Of the added 25% refrigeration tonnage, 65% must be allocated to the predrying chamber and 35% to the rf-drying chamber. Of the 25% added refrigeration tonnage, 35% must be of the more costly variety.

The condenser capacity on the existing batch chambers must be increased to handle 16.3% more water load. This will mean adding another condenser unit and possibly increasing the size of the valves on existing condensers if the present valves limit the vapor-pumping speed.

Facilities must be provided to transfer the semi-dry product from the batch chambers to the continuous rf-drying chamber. The semi-dry product will be received as a large batch at the rf-drying chamber; thus, provision must be made to hold this partially dried product without damage until it can be loaded for processing through the rf-drying chamber. The transfer equipment will probably have to be refrigerated and of sufficient size to hold one full batch.

Added to these items must be an rf-drying chamber complete with its low temperature condenser, input and output air-vacuum ports, moving belt, rf applicator, and rf generators.

The above is summarized in Table I, which lists the basic components that must be added to the plant to increase its capacity 25% by each approach, i.e., by adding more facilities of the conventional type or by adding facilities to use rf to dry the product in the latter stages of the drying process. There can be no question after studying the table that the use of rf would complicate the system and increase the facilities cost.

The cost for additional primary power to supply the rf generators and larger motors on the low-temperature refrigeration unit will cause the operating cost per pound of product to rise when using rf. Additional labor costs involved in transferring the product from the batch-

Table I—Comparison of Conventional and RF Expansion when Increasing the Capacity of a Conventional System by 25%

	Conventional Expansion	*RF Expansion*
Refrigeration	25% more tonnage of same type in one unit.	25% more tonnage split in two units. Divided 65% of same type, 35% low temperature.
Chamber	One conventional batch chamber with 25% capacity.	One continuous belt chamber with 125% capacity of semi-dry product, air locks for input and output, plus rf-waveguide applicators.
Condenser	One with 25% additional capacity.	Two: one with 16.3% additional capacity to be added to existing batching chambers; one low-temperature unit with 8.7% capacity for the rf chamber.
Valves and Plumbing	One set for the added chamber.	Two sets: one for condenser added to existing chamber, plus one for rf chamber.
Heat of Sublimation	25% added steam capacity, plus pumps and heat exchanger.	16.3% added steam capacity with pumps and heat exchange to handle 16.3% more capacity, plus rf generators to supply 8.7% of the heat of sublimation.
Material Handling	25% more of same type.	Special facilities to handle 125% of product during transfer from batch driers to rf drier, plus storage facilities for semi-dry material to be fed to the rf chamber.
Instrumentation	25% more of same type.	Possible slight modification to existing facilities to accommodate the added condenser, plus complete new set of instruments for rf-drying chamber and associated equipment.
Floor Space	Approximately 25% more.	Unknown but greater than required to expand plant by conventional means.

drying chambers to the rf chamber will further increase the cost per pound of product processed. RF in the final stages of drying will minimize the versatility of the batch-processing plant unless both the air locks and rf applicators can be designed to handle an extremely wide variety of products.

It must be concluded that some desirable attributes, such as improved texture and taste, must be realized if the use of rf in the final stages of drying to increase plant capacity is to prove worth the cost. Based on economics alone, the best way to expand an existing plant is to add facilities similar to those already installed.

The question immediately arises: If the conditions were different from those assumed in the case study, could rf be economically used in conjunction with a conventional batch plant to "finish dry" the product? The answer to this question is probably negative for the following reasons.

(1) Variable operating costs per pound of product produced will always be more expensive when rf energy is used. The greater the total amount of water removed from the product by rf, the greater the cost for primary energy to process each pound of product; this greater cost is a fundamental limitation.

(2) Capital costs will also increase, because, as the moisture level at which the product is transferred from the batch to the continuous rf-drying chamber increases, the problem of storing the semi-dry product before placement into the rf-drying chamber without deterioration becomes somewhat more severe. The semi-dry product has to be transferred very rapidly to prevent thawing and then held under refrigeration until it can be fed into the rf-drying equipment. The cost for refrigeration per pound of water removed is greater at the low system pressure required for rf processing. Thus, the more water removed from the product by rf, the greater the cost of the refrigeration facility and the greater the cost to operate the facility.

(3) Additional labor must be provided to transfer the semi-dry product from the batch to the continuous-flow rf drier and for monitoring the output of the rf-drying unit. It is difficult to visualize a composite batch–continuous-flow drying plant that would have less labor content than a straight batch operation.

With operating costs, capital invested, and labor content all increasing with increased use of microwave energy in the composite batch–continuous-flow drying process, the use of this system does not, on the basis of cost analysis alone, appear attractive.

There are always exceptions to any general statement. The possible exceptions to the above statement might be a process with an extremely

long "tail" on the drying cycle, such as would occur with thick products or a process that requires, for product quality reasons, the use of pressures at or below 50 microns. Just what the market requirements are for products with these unique drying characteristics is unknown.

Use of RF in New Freeze-Drying Plants

In the previous section, we concluded that rf offered no cost advantage when expanding existing batch plant by 25% and that it is doubtful if rf can be economically justified when used at any level as an adjunct to existing batching facilities. In this section the feasibility of using rf in new plants will be explored.

New plants can be of two basic types, batch or continuous flow. Any comparison should be made between plants of similar types. Except in those instances where rf offers improved product quality that cannot be achieved by any other means, the use of rf must be justified on a purely economic basis.

If the two plants being compared (rf versus radiant heat) handle the same volume of product per hour and the product is handled by similar means (batch or continuous), then the cost of utilities, i.e., water, steam, electricity, etc., used per pound of dried product will be less favorable to the system using rf energy. This is a consequence of the higher cost of a BTU of rf energy compared to a BTU produced by steam or electricity and to the increased power required for the lower-temperature refrigeration needed for the rf system. For an rf system to be practical, the higher operating costs must be offset by a lower capital investment in the plant or else the use of rf must provide improved product quality.

The higher operating cost for the rf system can be minimized by using a composite system employing steam or electricity in the early stages of drying process to remove 90% or so of the product moisture and rf energy to remove the remainder. However, when combined heating is used, it must be accomplished without manually transferring the product from the conventionally heated product chamber to the rf processing chamber. Manual transfer will add labor costs and require costly equipment to hold the semi-dry product until it can be fed into the rf-drying equipment. Under these boundary conditions, can the use of rf reduce the capital investment sufficiently to offset the higher operating cost?

An rf system, because of its ability to process certain products in a shorter period of time, will require correspondingly less tray or belt area and, hence, fewer drying chambers and consequently less floor space. The refrigeration tonnage and condenser area required are un-

affected by the speed-up of the process as explained in the preceding sections. Cost of the refrigeration for the rf portion of the system will be greater because of the lower drying-chamber pressure required to prevent glow discharge. Condensers for use in rf portion of the system will be more expensive because of the need for larger valves and pipe sizes. Apparently, the only cost elements in a processing system using rf that can contribute to offset the higher operating costs are the savings in capital and carrying charges associated with the reduction in number of drying chambers required for equal plant capacity, together with any resultant saving in building floor space. Offsetting these savings are the higher cost of refrigeration equipment, cost of rf power generating equipment, and increased cost for a higher-speed vapor-removing system and whatever floor space is required for this equipment.

New Batch Plant Using RF in Final Stage of Drying Cycle

The drying tray or belt area required for a given plant capacity using any freeze-drying process is

$$\text{Tray Area (Ft.}^2\text{)} = \frac{\text{Plant Capacity (Lbs./Hr.)} \times \text{Process Time (Hrs.)}}{\text{Tray Loading (Lb./Ft.}^2\text{)}}$$

RF has the advantage of both decreasing the drying time and increasing the depth of product load on the tray. The magnitude of the improvement depends upon the type of product being processed. Items such as thick steaks, etc., can be processed much more rapidly with rf. Hoover et al have reported a speed-up of 6:1 for chopped beef 1 inch thick. Diced foodstuffs do not show nearly as dramatic an improvement However, these items can be processed with deeper bed loading using rf. The possible improvement in bed-loading depth that can be achieved with rf is not accurately known, but a 2:1 improvement seems feasible.

No single figure can be established for the saving in tray space that might be achieved with rf. Each group of products will have a different improvement factor. It could be as low as 2:1 and perhaps as high as 6 or 8:1, depending upon the products to be processed. In addition, some of the advantages of rf are lost because methods of transferring radiant heat to the product have undergone improvement over the past several years. Stokes* reports some of their early equipment is producing nearly twice its design capacity. The process

* Pensalt Chemicals Corp., Stokes Equipment Div., Phila., Pa.

speed-up has been achieved by improved tray design. Of course, to achieve this speed-up some additional refrigeration had to be added to handle the greater water load.

Present day batching plants using radiant heat have a very high packing density in the drying chamber. Ten to fifteen trays are stacked into a single chamber. Increasing the depth of bed loading reduces the number of trays that can be placed in a chamber of fixed volume. Space

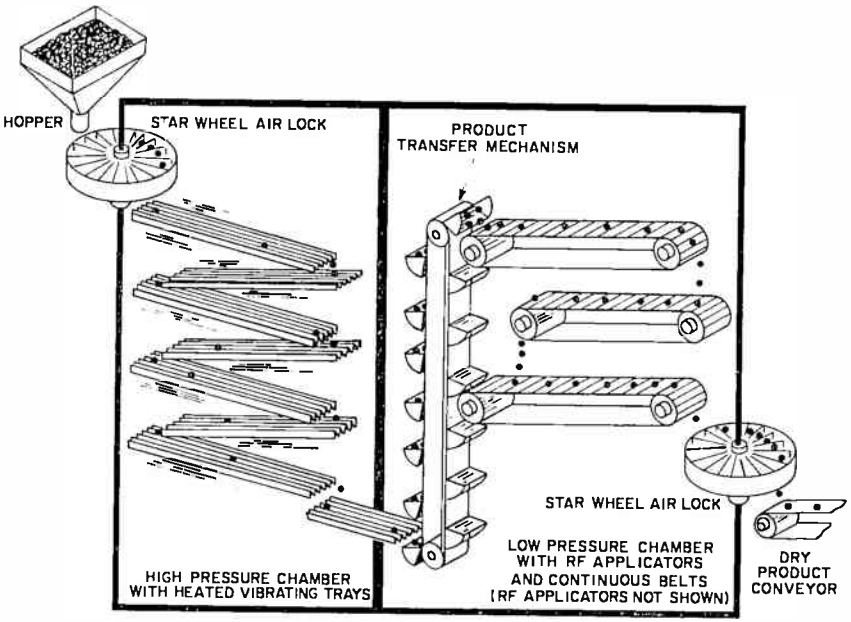


Fig. 11—Composite continuous freeze-drying chamber.

required in the chamber for the rf applicators will further decrease the tray packing density. These practical considerations tend to reduce the number of chambers that can be saved through the use of rf. The exact savings can only be determined by laying out batch-type chambers for both the conventional and composite systems and pricing the designs out.

A further practical problem exists in devising an rf applicator that will uniformly heat the product throughout the batch processing chamber since the product's rf characteristics change during the processing cycle. This problem becomes more complicated when processing several different products in the same equipment, or when using a combination of conventional and rf heating in the same chamber. The design of a practical applicator for batch processing has been considered too

formidable to tackle by those knowledgeable in rf techniques. As an alternative, they suggest concentrating on continuous-flow-processing systems.

Continuous-Flow-Process System

The high cost of rf energy does not justify its use throughout the whole of the drying process. Therefore, any practical continuous system would be a composite that would use radiant heat at high system pressure in the early stages to remove most of the water and rf at flow system pressure in the final stage to complete the drying process. Product transfer in a continuous system presents no problem, because the product can be automatically transferred within the drying chamber from the radiant-heat zone to the rf-heating zone provided the product flow rate in each zone is adjusted to be equal. A possible system is shown in Figure 11.

Again, the primary savings rf contributes to the overall system is a reduction in the amount of tray or belt space within the chamber. This, of course, affects the chamber size and, hence, plant floor space. Offsetting these savings are, again, the higher refrigeration costs for the rf portion of the chamber, applicator, and rf generator.

Conclusion

Although the use of rf energy to supply the heat of sublimation can accelerate the freeze-drying process, its use cannot in general be justified based on cost trade-offs alone. The use of rf in batch-processing plants requires complicated system designs. Continuous processing plants that make use of radiant heat in the early stages followed by rf to finish the drying process economize on chamber size and plant floor space. Whether the savings in floor space can offset the cost of the more complex processing chamber and the additional cost of low-temperature refrigeration equipment will have to be resolved by an equipment manufacturer after he has made a detailed layout of the drying chambers required for both an all-radiant-heat system and a composite system using both radiant and rf heat and then priced out both designs. The incentive for the equipment manufacturer to undertake this effort depends upon the potential market for this type of equipment.

Variable operating costs for the composite system will always be higher than for a system using only radiant heat. Thus, the market for the composite system will be limited to products whose quality (texture, taste, etc.) is enhanced by the use of rf processing and/or

to products whose drying cycles trail off to such an extent that the potential capital savings in plant floor space will offset the higher variable operating costs. The trade-off between over-all economics and product enhancement must be made on an individual product basis.

ACKNOWLEDGMENT

This study could not have been prepared without a great deal more effort were it not for the data and reference material accumulated by W. N. Parker during his past several years of work on the freeze-dry project. I am indebted to Mr. Parker for his cooperation and assistance in preparing this report.

DESIGNING HIGH-FREQUENCY DIFFUSED RESISTORS

By

R. E. Kleppinger

RCA Electronic Components
Somerville, N. J.

Summary—The frequency performance of a diffused resistor is developed by use of an RC transmission-line model and consideration of the effect of the metal-contact pads. This model is shown to agree with experimental data. Two different geometrical patterns (bar-bell and bar) are considered, and design equations incorporating both geometry and frequency are developed. The computer is used to generate design data on a sample problem. This sample problem shows that as high frequencies become important, the bar-bell resistor approaches the bar resistor as a limit. In addition, the importance of considering the size of the metal-contact pad is pointed out graphically.

Introduction

ALTHOUGH MUCH GENERAL information has been published on the design of diffused resistors, very little information can be found concerning the frequency performance of such resistors. One of the better models that can be used to describe a diffused resistor is the distributed *RC* transmission line.[1] However, results obtained with this model are not usable until the model is simplified to a lumped-constant *RC* configuration. In addition, the effects of the end pads, which play an important role in limting frequency response, must be considered. This paper describes a transmission-line model for diffused resistors and the use of this model when frequency response is of concern.

Electrical Model

For analysis, a diffused resistor can be considered to consist of three physical parts: (a) the input pad (C_{ip}), (b) the actual diffused resistor

[1] D. K. Lynn, C. S. Meyer, and D. J. Hamilton, *Analysis and Design of Integrated Circuits*, pp. 107-113, McGraw-Hill Book Co., N. Y., 1967.

(*RC* transmission line), and (c) the output pad (C_{op}). Figure 1 shows these parts and an equivalent circuit containing their electrical counterparts.

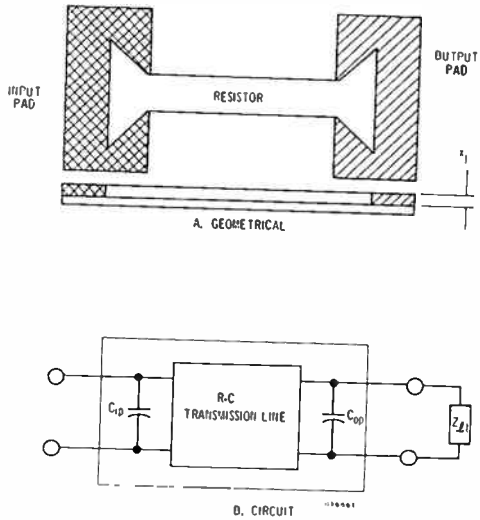


Fig. 1—Resistor models.

The steady-state solution of an *RC* transmission line is well known.[2] When such a line is terminated in a passive impedance Z_l, the input impedance Z_i is given by

$$Z_i = Z_c \frac{Z_{lt} \cosh \gamma l + Z_c \sinh \gamma l}{Z_c \cosh \gamma l + Z_{lt} \sinh \gamma l}, \tag{1}$$

where

$$Z_c = \sqrt{R/(j\omega C_T)}$$
$$\gamma = \sqrt{j\omega RC_T/l^2}$$

R, C_T = total resistance and capacitance of transmission line.

The normalized total resistor impedance is obtained by the parallel

[2] E. W. Kimbark, *Electrical Transmission of Power and Signals*, Chapter 6, John Wiley and Sons, N. Y., 1949.

combination of the input impedance of the transmission line and the input-pad capacitance:

$$\frac{Z}{R} = \frac{(Z_i/R)\,[1/(j\omega T_p)\,]}{Z_i/R + [1/(j\omega T_p)\,]}, \tag{2}$$

where $T_p = RC_{ip}$ and is the time constant of the resistor input-pad capacitor.

Because most diffused resistors have one end returned to ground through an evaporated-aluminum path, this type of termination is considered in this analysis. The impedance Z_{lt} of such a conductor is given by[3]

$$Z_{lt} = \frac{l\rho\sqrt{2}}{w\delta}\, e^{j\pi/4} \text{ ohms/meter}$$

where

l = length

ρ = resistivity in ohm-cm = 2.824×10^{-6} ohm-cm

w = width

$\delta = \sqrt{\rho/(\pi f \mu)}$ meters = $0.0845/\sqrt{f}$

μ = magnetic permeability = $4\pi \times 10^{-7}$ henry/meter

$|Z_{lt}| = 1.32 \times 10^{-9}\, l\sqrt{f}/\text{w}$ ohms

f = frequency in megahertz

It can be seen that Z_{lt} is insignificant compared to R up to very high frequencies (above 1 GHz) and that the transmission-line output can be considered a short circuit. The size of the output pad, therefore, is immaterial. The transmission-line input impedance given in Equation (1) can then be normalized as follows:

$$\frac{Z_i}{R} = \frac{\tanh\sqrt{j\omega T_D}}{\sqrt{j\omega T_D}}, \tag{3}$$

where $T_D = RC$ and is the resistor time constant.

Substitution of Equation (3) into Equation (2) yields the desired normalized input impedance:

[3] W. C. Johnson, *Transmission Lines and Networks*, p. 69, McGraw-Hill Book Co., N. Y., 1950.

$$\frac{Z}{R} = \frac{-j \tanh \sqrt{j\omega T_D}}{(k\omega T_D) \tanh \sqrt{j\omega T_D} - j\sqrt{j\omega T_D}}, \tag{4}$$

where

$$k = \frac{T_p}{T_D} = \frac{C_{ip}}{C_T} = \frac{\text{input-pad area}}{\text{resistor area}}. \tag{5}$$

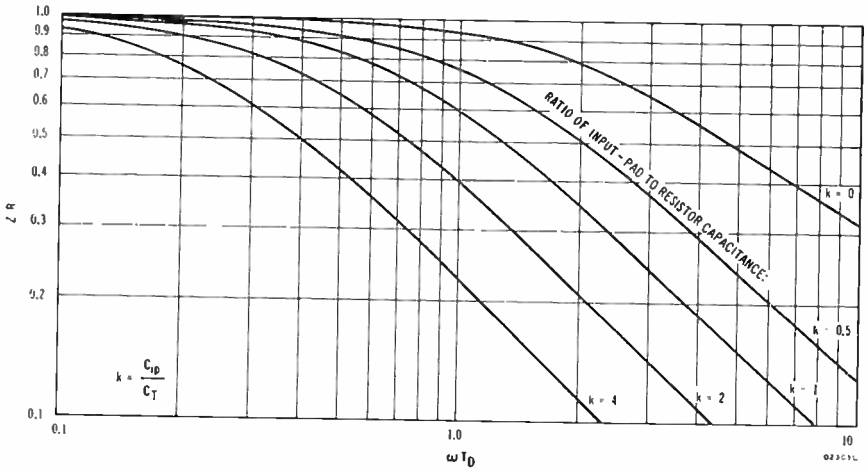


Fig. 2—Magnitude of resistor impedance as a function of normalized frequency.

The magnitude and phase of the normalized input impedance were computed for different values of k and ωT_{DS} as shown in Figures 2 and 3. The input-pad capacitance has a strong effect on the phase and magnitude. The curves of Figures 2 and 3 can be used to determine the minimum value of ωT_D for any value of k, if it is assumed that the magnitude remains within 10% of the dc value. Figure 3 also shows the phase angle, or amount of reactance for value of ωT_D used. When ω is set equal to the minimum circuit frequency, the maximum time constant of the device can be obtained.

As a check on this model, published frequency-response data[4] were compared to the curve obtained by use of the modified transmission-line model. Figure 4 shows the circuit, the measured data, and the cal-

[4] R. M. Warner and J. N. Fordemwalt, *Integrated Circuits*, Figs. 10-20. p. 262, McGraw-Hill Book Co., N. Y., 1965.

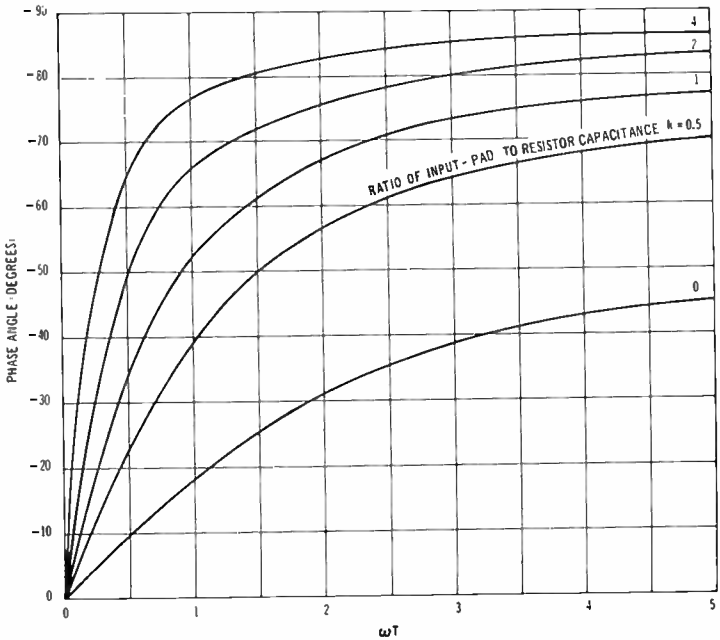


Fig. 3—Resistor phase angle as a function of normalized frequency.

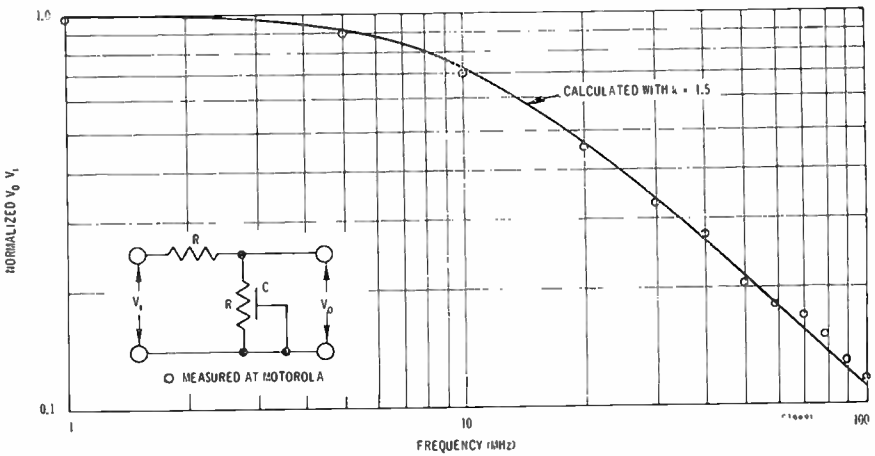


Fig. 4—Experimental verification of resistor model.

culated curve. Agreement between theory and experiment is very good for the assumption that the area of the input pad is 50% greater than the actual area of the resistor.

For resistors not terminated in a short circuit, the curves of Figures 2 and 3 can be recomputed provided the terminating impedance is known.

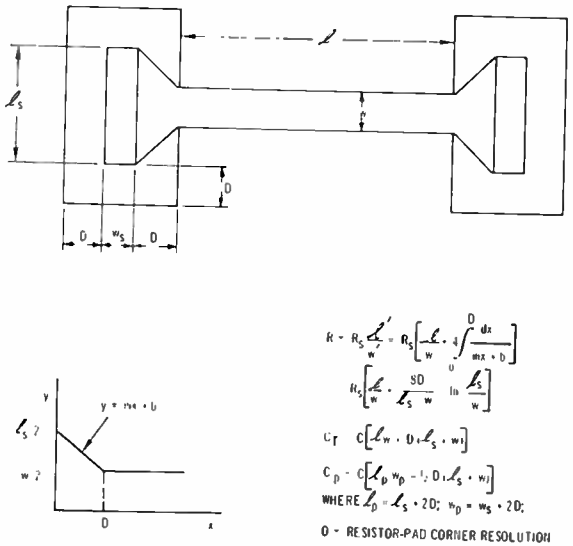


Fig. 5—Geometrical model for resistor.

Physical Model

Resistance and capacitance values can be determined from the device geometry, as shown in Figure 5. The terms in the equations of Figure 5 can be rearranged to yield the following normalized expressions:

$$\frac{R}{R_s} = (1 + CF_1)\,\frac{l}{w} \tag{6}$$

$$\frac{C_T}{C} = (1 + CF_2)\,\frac{l}{w} \tag{7}$$

$$\frac{C_{tp}}{C} = (1 - CF_3)\,l_p w_p \tag{8}$$

where

$$CF_1 = 8\left(\frac{1}{(l_s/w) - 1}\right)\frac{D}{l}\ln\frac{l_s}{w} = \frac{CF_4}{l} \tag{9}$$

$$CF_2 = \frac{D}{l}\left[\frac{l_s}{w} + 1\right] = \frac{CF_5}{l} \tag{10}$$

$$CF_3 = \frac{1}{2}\frac{D}{w_p}\frac{w}{l_p}\left(\frac{l_s}{w} + 1\right) \tag{11}$$

Equation (4) can then be defined in geometrical terms,

$$k = \frac{C_{ip}}{C_T} = \frac{(1 - CF_3)\, l_p\, w_p}{(1 + CF_2)\, lw}. \tag{12}$$

The value of C, capacitance per unit area, can be determined from published curves,[5,6] provided the following parameters are known,

Background concentration (C_b),

Diffused sheet resistivity (R_s),

Diffused junction depth (x_j),

Applied junction voltage (V), including contact potential.

Appendix I shows that this capacitance per unit area is constant for a wide range of base sheet resistances (R_s). This fact is an advantage in later calculations.

Design Equations

A set of useful resistor design equations can be obtained by combination of the appropriate equations from the electrical and physical models. In this way, the resistor design can be determined from the desired electrical characteristics. To solve these equations, several different approaches were tried, each imposing different constraints. The following procedure gives a unique solution that is consistent. The variable parameters in this method are

[5] J. C. Irvin, "Resistivity of Bulk Silicon and of Diffused Layers in Silicon," *Bell Syst. Tech. Jour.*, Vol. 41, p. 387, March 1962.

[6] H. Lawrence and R. M. Warner, Jr., "Diffused Junction Depletion Layer Calculations," *Bell Syst. Tech. Jour.*, p. 389, March 1960.

Circuit time constant (T_c),

Desired resistor value (R),

Capacitance/mil^2 (C) (calculated from the Lawrence–Warner curves[6] and left constant for $R_s = 50$ to 500 ohms/square),

Resistor sheet resistance (R_s),

Metal contact length (l_s),

Resistor-pad corner resolution (D),

Resistor width (w).

The parameters calculated are

Uncorrected resistor length (l),

Resistor delay time (T_D),

Ratio of input pad area to resistor area (k),

Input-impedance phase angle (Ph),

Metal contact widths (W_s),

Resistor area used (A),

The calculated parameters were obtained for different resistor widths over a predetermined range of widths. A brief outline of the mathematical steps involved is given below.

Length l is computed from Equation (6) as follows:

$$l = \frac{wR}{R_s} - CF_4, \tag{13}$$

where

$$CF_4 = \frac{8D \ln (l_s/w)}{l_s/w}.$$

Resistor delay time T_D is obtained by combining $T_D = RC_T$ and Equation (7) as follows:

$$T_D = RCw(l + CF_5), \tag{14}$$

where

$$CF_5 = D\frac{l_s}{w+1}.$$

The k-factor is then obtained from Equation (4) by imposing the constraint that the magnitude of the normalized impedance Z/R be 0.90 (i.e., the magnitude should be 90% of the dc value) and using the calculated resistor delay time. This step is performed best by an iterative process on the computer.

Metal contact width w_s is calculated from a rearranged Equation (12) and Equation (11),

$$w_s = \left(\frac{kw}{l_p}\right)(l + CF_5) + \frac{1}{2}\left(\frac{D}{l_p}\right)(l_s + w) - 2D. \tag{15}$$

The silicon pellet area A used can then be calculated,

$$A = 2l_p w_p + lw. \tag{16}$$

The resistor width is set in 0.25-mil increments from $w = 0.25$ mil to $w = l_p$. And widths that yield negative lengths l or metal widths w_s are eliminated automatically. Although negative lengths down to $l = -2D$ are practical, as shown in Figure 5, the correction factors derived in Equations (9) through (11) would no longer be correct because corner-pad resolution D is a function of length (see Figure 5 equations).

The bar-bell pattern of Figure 5 is changed to a bar pattern if we impose the constraint

$$l_s = w - 2D. \tag{17}$$

Sample Problem

A resistor from an integrated-circuit logic gate was chosen as a test problem. This resistor is part of the output circuit. Actual physical-electrical parameters are

$R = 500$ ohms (desired)

$C = 0.208$ pF/mil^2

$R_s = 200$ ohms/square

$w_s = 1.15$ mils

$l = 4.3$ mils

$w = 2.0$ mils

$l_s = 2.3$ mils

$D = 0.35$ mil

$T_c = 1$ nanosecond (desired frequency = 159 MHz)

As a first approximation, the resistor delay time can be calculated as

$$T_D = R_s Cl^2 = 0.77 \text{ nanosecond.}$$

A more accurate value, obtained by use of the equations developed in this paper, is 0.58 nanosecond. The calculated resistance value resulting from these equations is 520 ohms.

The effects of different sheet resistances R_s and resistor widths on resistor length, resistor delay time, ratio of input pad area to resistor area, metal-contact width, and silicon pellet area can be obtained by use of the design equations described previously. The curves in Figure 6 are identified by the order in which they were calculated (i.e., graph B was obtained from graph A, and so on). Each curve terminates when the value of length l or metal contact width w_s becomes zero, or when the value of width w equals pad length $(l_s + 2D)$. As expected, increasing sheet resistances provide shorter delay times because less resistor area is needed for the same resistance value and resistor width.

Similar data were calculated and plotted for different values of resistor-pad corner resolution D, as shown in Figure 7. In this case, increased pad-corner resolutions result in shorter delay times, because less resistor area is needed. The resistor area is decreased because a greater percentage of the resistor area is in the trapezoidal ends shown in Figure 5.

Figure 8 shows similar curves for different metal contact lengths l_s. As expected, shorter metal-contact lengths result in shorter delay times. The shorter delay time is the result of less resistor area; again, a greater percentage of the area is in the trapezoidal ends.

As an example of how to utilize these curves, it can be assumed that the resistor delay time is one-half the circuit delay time (i.e., $T_D = .0.5$ nanosecond). This constraint could have been used in the computer program so that only solutions with $T_D \leq 0.5$ nanosecond would have been printed. Instead, the constraint can be superimposed upon the curves to obtain the allowed solutions. The shaded areas in Figures 6 through 8 show graphically the allowed solutions for the parameter ranges considered. A sheet resistance value, resistor width, and resistor length that will be consistent with photoresist technology can be chosen from Figure 6. For example, if $R_s = 200$ ohms/square and $w = 1.5$ mils, then $l = 1.5$ mils, $T_D = 0.47$ nanosecond, and $w_s = 0.64$ mil. Figure 7 shows that l could be larger and T_D smaller, but $D = 0.35$ mil is a fairly reasonable choice. Figure 8 shows that delay time can be decreased to 0.34 nanosecond by use of $l_s = 2$ mils. From this last curve, the

resistor design can be obtained. The design values compare with the values used as follows:

Value	R_s	w	D	l_s	l	w_s	A	T_D	k	R
As used	200	4.3	0.35	2.3	4.3	1.15	19.7	0.58	0.47	520
Design	200	1.5	0.35	2.0	1.35	0.82	10.5	0.34	1.15	500

The effects of parameter tolerances can be obtained by use of a different sequence for calculating parameters. The same computer program can be used to determine the effect of changes in any computed parameter while all other parameters remain constant.

The effect of different sheet resistances on the parameters of the bar-shaped pattern is shown in Figure 9. Because the constraint on metal-contact length is different, the slopes of the curves of delay time as a function of width are different from those of the bar-bell curves in Figure 6. This difference is large enough for higher-frequency circuits to make the bar pattern more useful. Figure 8 shows that the bar-bell pattern approaches the bar pattern at higher frequencies because the metal-contact length is made smaller and smaller. These data are summarized as follows:

Shape	R_s	w	D	l	l_s	w_s	A	T_D	k
Bar-bell	200	1.75	0.35	1.95	2.3	0.5	10.6	0.5	0.65
Bar-bell	200	1.75	0.35	1.35	1.5	1.6	10.6	0.36	0.95
Bar	200	1.75	0.35	0.75	1.05	1.75	10.0	0.24	1.5

As in the case of the bar-bell pattern, increased pad-corner resolution yields shorter delay times with shorter lengths, as shown in Figure 10.

When the approach outlined above was used, the resistor design for either the bar-bell pattern or the bar pattern resulted in delay times considerably shorter than those actually used. Other shapes could be analyzed by this same method provided appropriate adjustments were made in the mathematics and in the computer program.

ACKNOWLEDGMENTS

The author thanks J. Rudolph for many discussions concerning the mathematical approach taken and the computer programming used, and Dr. J. Hilibrand for his encouragement and advice in developing this work.

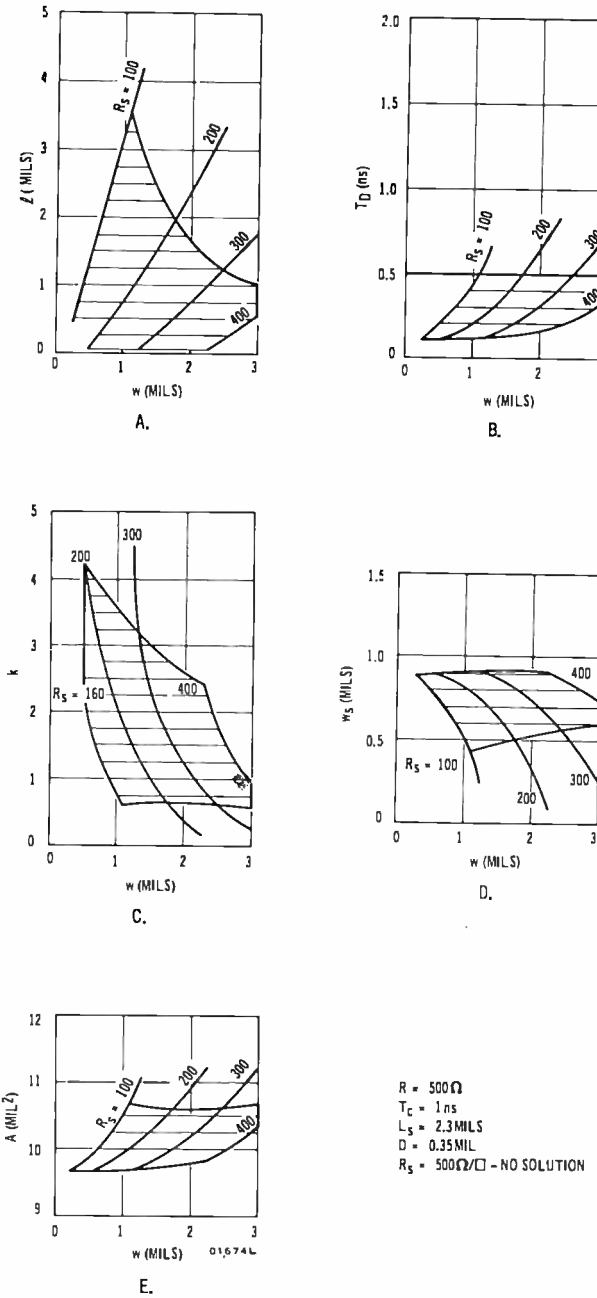


Fig. 6—Effect of varying sheet resistance in bar-bell resistor (shaded areas are allowed solutions).

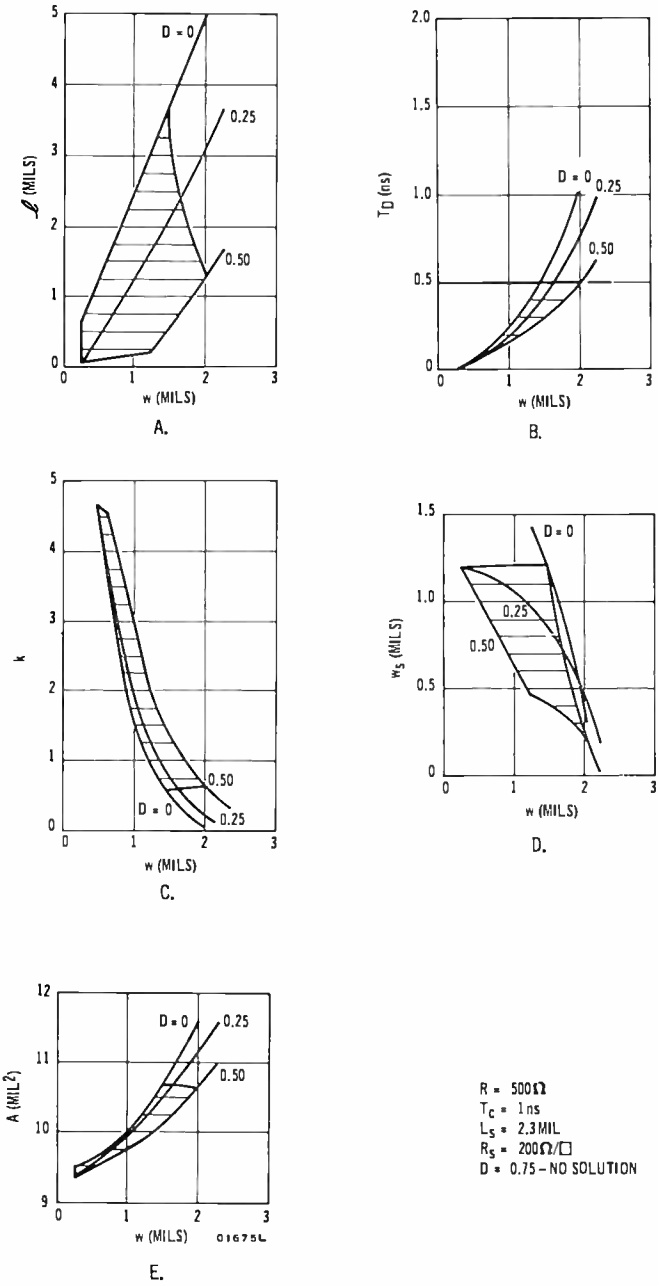


Fig. 7—Effect of varying pad-corner resolution in bar-bell resistor (shaded areas are allowed solutions).

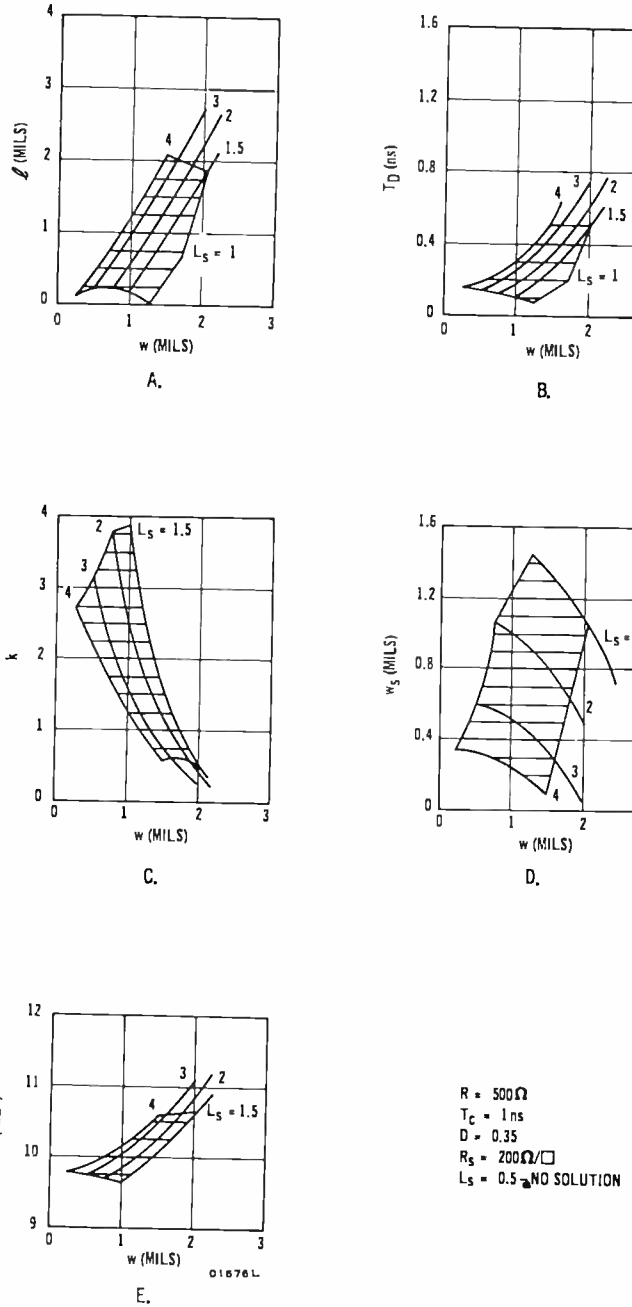
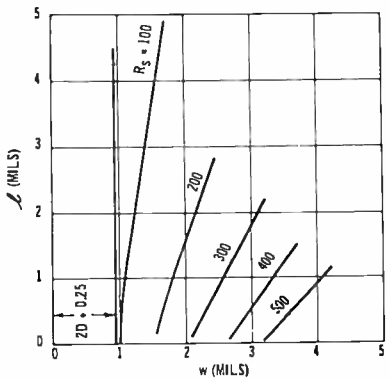
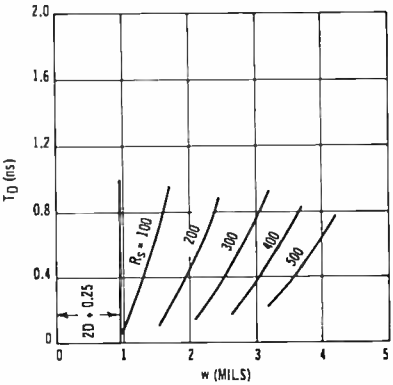


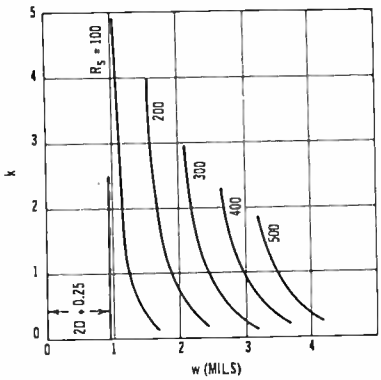
Fig. 8—Effect of varying metal-contact length in bar-bell resistor (shaded areas are allowed solutions).



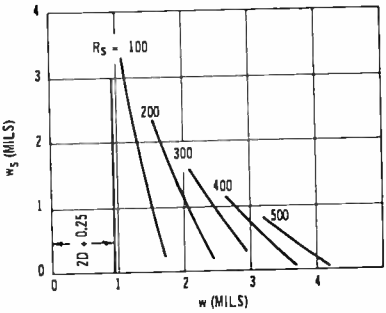
A.



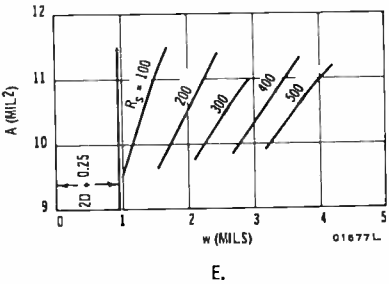
B.



C.



D.



E.

R = 500Ω
T_C = 1 ns
D = 0.35 MIL

Fig. 9—Effect of varying sheet resistances in bar resistor.

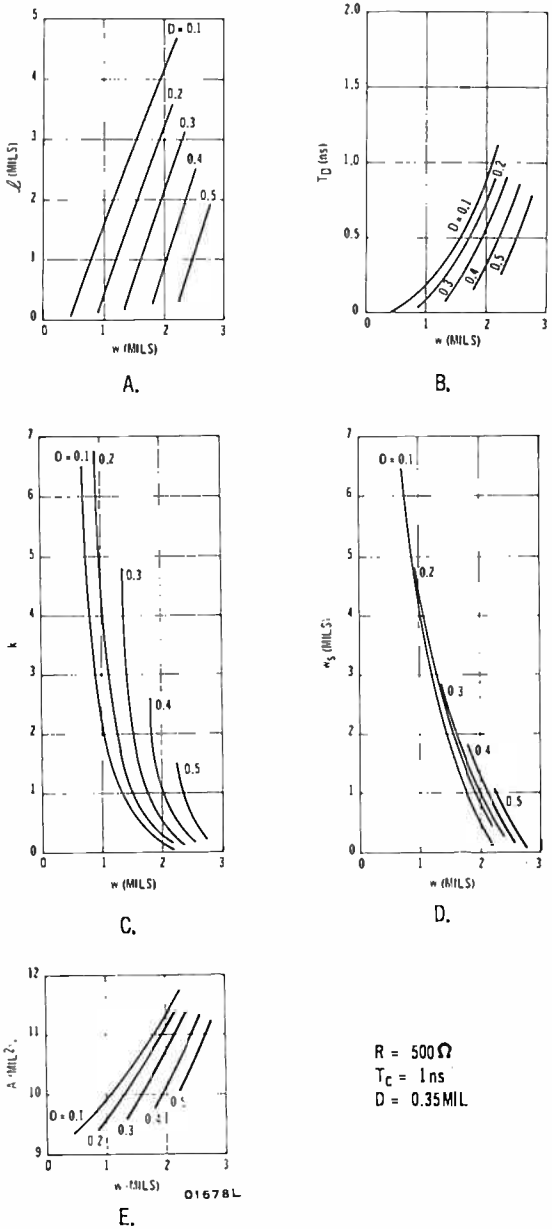


Fig. 10—Effect of varying pad-corner resolution in bar resistor.

Appendix I—Determination of Capacitance Per Unit Area of a Diffused Junction

The capacitance per unit area can be obtained by use of Irvin's curves[5] and the Lawrence–Werner curves,[6] provided the parameters C_b, R_s, x_j, and V are known. If it is assumed that $C_b = 10^{16}$ (0.6 ohm-cm, n-type) and $V = 1$ volt, the capacitance per unit area is

x_j (microns)	R_s (ohms/square)	C (pF/mil^2)
0.5	50	0.245
0.5	200	0.245
0.5	500	0.245
1.0	50	0.206
1.0	200	0.206*
1.0	500	0.206

These calculations indicate that the capacitance per unit area is independent of the diffused impurities.

* This value agrees surprisingly well with the average value of 0.208 pF/mil^2 obtained experimentally.

PULSED HEAT CONDUCTION IN A LAYERED SEMICONDUCTOR-METAL TRANSFERRED-ELECTRON OSCILLATOR GEOMETRY

By

Barry S. Perlman

RCA Electronic Components
Princeton, N.J.

Summary—A method of analyzing the temperature gradient within the active layer of a bulk-effect GaAs transferred-electron device during pulsed excitation is presented. The effects of material and source parameters are considered with respect to the thermal transient response and associated time constants for specific device geometries. The results have been presented in graphical form using normalized variables. The increase in average device temperature due to spreading-resistance effects has been calculated. The thermal time constants for devices using either GaAs or Si substrates have been computed using the active-layer length and relative capping-layer thickness as parameters.

When this device is characterized by an optimum product of carrier concentration and active-layer length, the heat density is inversely related to this length. The asymptotic steady-state temperature is linearly dependent on the n-region thickness, while the total dissipated power remains independent of the n-region thickness and is proportional to the cross-sectional area. The limiting thermal response is imposed by the active layer, where the time constants are related to the square of the different relative lengths. The asymptotic temperature has been evaluated from consideration of the inverse temperature dependence of the thermal conductivity of the semiconductor material.

Introduction

A CONVENIENT METHOD for analyzing the variation of temperature throughout composite multilayered media is the use of the Laplace transformation.[1] The heat-conducting properties of a complicated thermal geometry are closely analogous to an equivalent electrical noninductive lossy transmission line. The

[1]Carslaw, H. S. and J. C. Jaeger, *Conduction of Heat in Solids*, 2nd. Ed., Oxford University Press, New York, 1959.

thermal flux temperature distribution may be visualized as a current-density–voltage distribution in an equivalent electrical transmission system.

The problem of particular interest concerns the uniform generation of heat in a pulsed layer of active semiconducting material, i.e., the n-layer of a transferred-electron oscillator (TEO), and the resulting transient temperature distribution. Several analyses have provided significant insight into the thermal properties of similar materials in the steady state.[2-5] However, in many important high-power applications, where steady flow is virtually impossible, a pulsed mode of operation is used to maintain a safe heating limit for the device. During a typical pulse, a thermal transient is developed that results in many interesting, and usually undesirable, rf electrical and material phenomena (e.g., FM, jitter, redistribution of carrier concentration, thermally induced avalanche, etc.).[6,7] The ultimate rise and distribution of temperature within the active layer, which directly affects the electrical performance of the device, is related to the heat-sink thermal constants and spreading resistance, the source interpulse spacing, the layer thickness, and the input power density.

The primary aim of this analysis is to develop a quantitative picture of the transient temperature distribution in the layered geometry associated with a transferred-electron semiconductor device. Only one-dimensional heat transfer by conduction from the bulk active region representing a uniform source is considered, although the analytical methods used are readily applicable to other source conditions.

By establishing a close analogy between the thermal geometry and an equivalent electrical network, this problem may be more clearly interpreted and solved using several techniques of circuit analysis.*

---

[2] Knight, S., "Heat Flow in n^{++}-n-n^{+} Expitaxial GaAs Bulk Effect Devices," *Proc. IEEE (Letters)*, Vol. 55, p. 112, Jan. 1967.

[3] Swan, C. B., et. al, "Composite Avalanche Diode Structures for Increased Power Capability," *IEEE Trans. ED*, Vol. 14, p. 584, Sept. 1967.

[4] Clorfeine, A. S., "Power Dissipation Limits in Solid State Oscillator Diodes," *Microwave Jour.*, Vol. 11, p. 93, March 1968.

[5] Gooch, C. H., "The Thermal Properties of GaAs Laser Structures," *IEEE Jour. Quantum Elec.*, Vol. QE-4, p. 140, April 1968.

[6] Haitz, R. H., "Nonuniform Thermal Conductance of an Avalanche Microwave Oscillator," *IEEE Trans. ED*, Vol. 15, p. 350, June 1968.

[7] Brown, R. F., "Transient Temperature Distribution in Diode Lasers and the Time Duration of the Output Pulse at 300°K," *IEEE Jour. Quantum Elec.*, Vol. QE-4, p. 135, April 1968.

* The analogy to cascaded transmission lines used here in a one-dimensional analysis should be applicable to heat transfer in the active layer of this semiconductor device if the active layer is very thin compared with its lateral dimensions, the heat generation is uniform, and the contact layer does not have a larger surface area than the active area.

In addition to the use of the Laplace transform, we introduce a thermal input impedance, using matrix techniques, at the source–sink interface and utilize the principle of superposition to provide a solution for variable sink temperatures, V_A.

The heat diffusion through the sink geometry may be analyzed in terms of several thermal time constants, although a first-order approximation of a single constant is satisfactory. The limiting response is imposed by the relatively poor thermal characteristics associated with the active layer of GaAs material.

When this particular device is characterized by an optimum product of carrier concentration and active layer length, the resulting heat density is affected by the length used, with the asymptotic transient temperature is *linearly* related to this length. The equilibrium temperature associated with an infinite half space spreading resistance is independent of this length and linearly related to the device radius.

A possible complication in this analysis would be imposed by an inverse temperature dependence of the thermal conductivity of the semiconductor material. The transient solution therefore ignores this problem, although a steady-state solution is found for variable conductivity so as to evaluate the magnitude of error introduced by the linear analysis.

The resulting temperature distribution within the active layer for both single-pulse and periodic excitation has been derived for several geometries. The results are presented in graphical form using normalized variables with respect to both the thermal time constants and steady-state temperatures. The time constants associated with specific geometries, using both GaAs and Si substrates, are computed using the active layer length and relative capping layer thickness as parameters.

A more thorough familiarity with the problem, with particular regard for both material and geometrical considerations, should provide an understanding of certain device limitations that result from the temperature distribution within the active region and make it possible to minimize these effects. An improvement of the device electrical efficiency is directly associated with the package thermal efficiency and, consequently, the overall heat-sink design.

The TEO (Gunn Oscillator) as a Heat Source

A. Electrical Characteristics

The application of a uniform electric field to a bulk semiconductor results in the generation of a uniform heat density. In the active n-layer of a GaAs epitaxial device that is suitably doped to exhibit the

"two-valley" conduction band-effect,[8] a differential negative resistance appears, thus permitting this material to sustain microwave oscillations. The rf power that is removed from the material represents only a small portion of the dc power that must be supplied to the device. The remaining energy is consumed by the material, resulting in uniform heating throughout the active region.

The heat density may be represented by

$$A(s) = \frac{A_o}{s} = \frac{JE}{s}(1-\eta), \tag{1}$$

where we have introduced the Laplace transform $A(s)$ of the power density, A_o. The dc-to-rf conversion efficiency is given by η. Although this analysis considers only uniform heating, the introduction of nonuniform values of $A(s)$ could be used to analyze localized heating effects and other devices (i.e., avalanche diodes, etc.).

The usual geometry for a layered device with a rectangular cross section is shown in Figure 1. The thickness l of the active region is grown epitaxially to establish the electrical length necessary for efficient oscillator performance at a specific frequency. The transit-time frequency f_τ is inversely related to this length (i.e., $f_\tau \simeq 10^7/l$ Hz). The lower limit on frequency is determined by both the increased temperature and the difficulties in homogeneous growth associated with thick layers. A 1 GHz oscillator operating in a resonant mode near f_τ would require as much as 100 microns of active material, with a carrier concentration near 1.5×10^{14} cm^{-3}. The optimum efficiency for an oscillator tuned into a resonant mode near f_τ is obtained when the nl product is a constant (i.e., $nl \simeq 1.5 \times 10^{12}$ cm^{-2}).[9]

At the threshold field, $E_T \simeq 3000$ V/cm, the current density $J_T = E_T/\rho_o$, where ρ_o is the volume resistivity ($1/ne\mu$ ohm-cm) of the active layer, which has a carrier concentration n and electron mobility μ. The operating field, which is usually several times E_T, may be specified as βE_T, where $\beta \geq 1$. The valley current density becomes γJ_T, where $\frac{1}{2} \leq \gamma \leq 1$.

We may now express the heat power density A_o from Equation (1), in terms of the material constants* and the constant nl product specified.

---

[8] McCumber, D. E., and A. G. Chynoweth, "Theory of Negative-Conductance Amplification and of Gunn Instabilities in 'Two-Valley' Semiconductors," *IEEE Trans. ED*, Vol. 13, p. 4, Jan. 1966.

[9] Copeland, J. A., "Theoretical Study of a Gunn Diode in a Resonant Circuit," *IEEE Trans. ED*, Vol. 14, p. 55, Feb. 1967.

* In this analysis we shall neglect the temperature susceptibility of the resistivity ρ_o which is evaluated at an ambient temperature $T_o \simeq 300$°K.

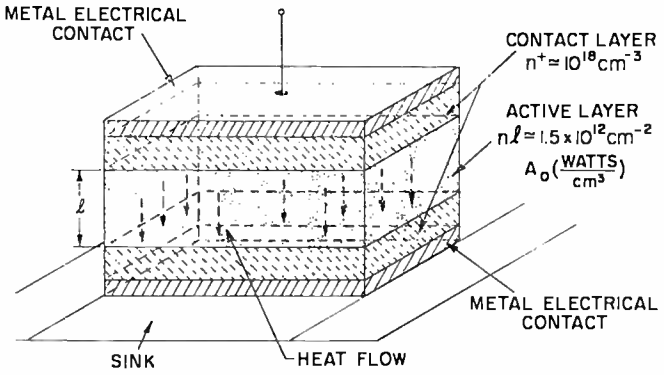


Fig. 1—Layered geometry of a TEO Gunn bulk-effect device.

$$A_o = \frac{\beta\gamma E_T^2}{\rho_o}(1-\eta) = \frac{\beta\gamma}{l}[1.28 \times 10^4] \text{ watt/cm}^3. \qquad (2)$$

Here l is in cm and a nominal value of $\eta \simeq .05$ has been used. Figure 2 shows the variation of power density A_o versus length l with $\beta\gamma$ as a parameter, obtained with Equation (2). Although the power density

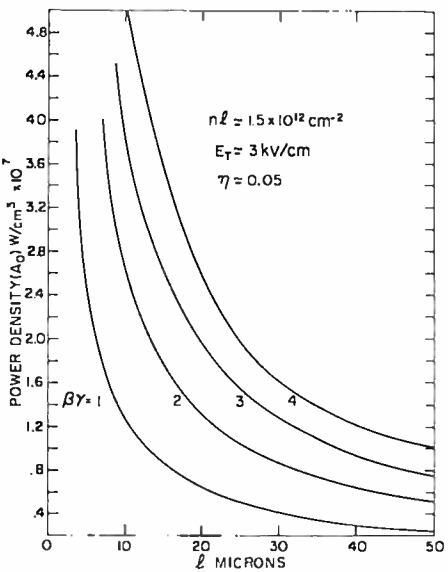


Fig. 2—Variation of input power density with layer thickness.

increases with decreasing layer thickness, the actual dissipated power is independent of l and is proportional to the cross-sectional area. The temperature rise within the active layer is affected by l and determines the limitations on performance for this device.

B. Material Considerations—Thermal Characteristics

The properties of several well known materials,[10-14] that could conceivably be used in the sink geometry are indicated in Table I.

The appearance of different thermal properties for high-resistivity and highly doped single-crystal semiconductors creates a thermal problem. The thermal conductivity of highly doped material ($n > 10^{18}$ cm^{-3}), which is necessary for good ohmic constants, is actually lower than that of lightly doped material, even though more electrons are available for conduction of heat. The decreased thermal conductance appears to result from a phonon–electron scattering process.[11,12] In this transient analysis, we have disregarded the resulting $1/T$ conduction dependence. This inverse temperature effect on K has been used in the Appendix to calculate the steady-state behavior.

The most common geometry for a TEO consists of n/n+ epitaxial GaAs, where the difficulty is that the alloying of metallic contacts directly to the already thermally poor n layer necessitates the use of highly doped capping material. The relatively poor thermal response of GaAs limits the effectiveness of heat removal through the n+ contact material, thereby requiring a minimum layer thickness. Even if a direct alloy to the active n layer were achieved (i.e., Sn, Au-Ge, etc.), a significant improvement in thermal performance would not occur. Unfortunately, the high-conductivity materials (i.e., Ag, Au) do not effectively alloy with n-GaAs ($n \leq 10^{16}$ cm^{-3}). The use of a planar geometry would separate the conduction processes (electric versus thermal) but still retain a semi-insulating GaAs thermal barrier unless another single-crystal substrate material were used. When the required electrical length l becomes large ($> 30\ \mu$), the temperature rise in a layered geometry can become excessive and a planar structure may be more desirable. The most effective method for removing heat

[10] Berman, R., *Physical Properties of Diamond*, Clarendon Press, Oxford, 1965.

[11] Maycock, P. D., "Thermal Conductivity of Si, Ge, III-V Compounds and Alloys," *Solid State Elect.*, Vol. 10, p. 161, March 1967.

[12] Holland, M. G., "Photon Scattering in Semiconductors from Thermal Conductivity Studies," *Phys. Rev.*, Vol. 134, p. A471, April 1964.

[13] Willardson, R. K. and Beer, A. C. *Semiconductors and Semimetals*, Vol. 2, "Physics of III-V Compounds," Academic Press, New York, 1966.

[14] Drabble, J. R., and Goldsmith, H. J., *Thermal Conduction in Semiconductors*, Pergamon Press, New York, 1961.

from thin n-layer material (e.g., $l \leq 20\ \mu$) is to flip-chip bond the epitaxial side to the heat sink. A minimum thickness of n+ material is desirable.

The most interesting, as well as effective, material for use in the heat-sink geometry is diamond Type IIa.[10] This material can be used to spread the heat efficiently from the concentrated area of the device over a broader base, thereby minimizing thermal flux entering the heat

Table I—Thermal Constants for Selected Materials

Material	Thermal Conductivity K (Watts/°C − cm)	Diffusivity α (cm²/sec)
GaAs (n)	0.5	.29
GaAs (n+)	0.4	.232
Si (undoped)	1.45	.825
Si (high doping)	1.15	.650
Be_2O_3	2.45	.572
Diamond (IIa)	22.0	1.01
Diamond (IIb)	13.0	.592
Silver	4.27	1.71
Gold	3.16	1.18
Copper (OFHC)	4.0	1.14
Aluminum	2.36	.86
Molybdenum	1.46	.58
Tin	0.625	.38
Brass (70:30)	1.04	.33
Lead	0.34	.25

sink. The relatively high thermal conductance of type IIa diamond is very effective in reducing the spreading resistance associated with the device-to-metal contact.[15] The allowable dissipation of the resulting package is thereby improved.

Although the termal conductivities of the various materials vary considerably, a less significant difference exists between their diffusivities. The higher specific heat and density of diamond II, Be_2O_3, etc. limit their thermal response, since the associated thermal time con-

[15] Kennedy, D. P., "Spreading Resistance in Cylindrical Semiconductor Devices," *Jour. Appl. Phys.*, Vol. 31, p. 1495, Aug. 1960.

stants are inversely proportional to the diffusivity, α. Diamond IIa has the minimum thermal resistance and therefore yields the least rise in temperature, but *silver* has the smallest time constant.

Transient Temperature Response for Several Geometries

The transient temperature build-up within the active n layer $(0 < x < l)$ of a transferred-electron-oscillator geometry is readily found as a solution to a one-dimensional thermal-conduction equation, which is described in Appendix I. Using electrical analogies, the layered geometry may be represented as a thermal network, as shown in Figure 3, consisting of a set of cascaded transmission lines. Using

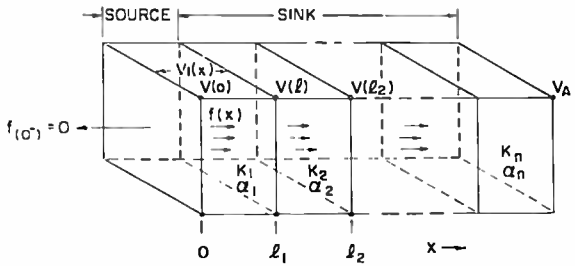


Fig. 3—Cascaded thermal network.

Laplace notation, each line is characterized by a transfer matrix relating the flux and temperature distribution between layers.

Separating the source region (i.e., active layer) from the sink geometry and introducing an input "sink" impedance, $Z(l) = V(l)/f(l)$ where $x = l_1 = l$, we may relate the passive sink parameters to a termination on the source. The sink is terminated by a finite controllable source, V_A. The use of a fixed termination V_A neglects the spreading resistance normally associated with steady spherical flow $(\partial/\partial t = 0)$ from the device into a larger sink. This model is justified during the initial transient, where the temperature rise within the device will be shown to be considerably faster than that in the sink, which has a considerably larger thermal mass. The contribution of the spreading resistance will be determined for the case of a repetitive pulse source where the system has reached an equilibrium condition and the continuity of an average thermal flux is maintained through the sink geometry.

The temperature distribution within the active n layer is derived in Appendix I and may be expressed in Laplace space as

$$V_1(x) = \bar{A}\left[1 - \frac{C_1(x)}{C_1(l) + \sigma S_1(l)}\right] + \frac{V_A}{s}\left[\frac{C_1(x)}{AC_1(l) + \sigma S_1(l)}\right], \quad (3)$$

where

$V_1(x)$ = transformed temperature in active layer as a function of the distance l measured into the device from the top of the source region.

V_A = sink-load-temperature boundary condition.

$\bar{A} = A(s)\alpha_1/(sK_1)$ = normalized equivalent source temperature for a zero initial material temperature (i.e., $t = 0$, $V_o = 0$).

$A(s)$ = source power density (watts/cm^3), with time dependance explicit in Laplace operator s.

A = variable in ABCD combined transfer matrix for cascaded sink geometry (see Equation (4)).

α_n = thermal diffusivity (cm^2/sec for n^{th} layer (e.g., $n = 1$ for source region).

K_n = thermal conductivity (watts/ $^oC - cm$) for n^{th} layer.

$C_n(x) = \cosh q_n x$

$S_n(x) = \sinh q_n x$

$q_n = \sqrt{s/\alpha_n}$ = normalized Laplace operator.

$\sigma = Z(l)/Z_1$ = dimensionless variable.

Z_n = characteristic impedance of n^{th} region.

The impedance transfer matrix for each layer may be expressed, as shown in Appendix I, as,

$$\begin{bmatrix} A_n B_n \\ C_n D_n \end{bmatrix} = \begin{bmatrix} C_n(l_n) & -Z_n S_n(l_n) \\ Z_n^{-1} S_n(l_n) & C_n(l_n) \end{bmatrix} \quad (4)$$

where $l_n = x_n - x_{n-1}$ corresponds to the length of the n^{th} subregion. The composite matrix, $\begin{bmatrix} AB \\ CD \end{bmatrix}$ for a series of cascaded lines is obtained by direct matrix multiplication. The primary-source driving point impedance $Z(l)$ may be shown to be,

$$Z(l) = \frac{V(l)}{f(l)} = -B/A, \quad (5)$$

where A, B are found from the composite matrix.

Unit-Pulse Heating of a Semi-Infinite Homogenous Material

The results derived in Appendix I and summarized above are used to examine the problem in which heat is generated in an active layer $0 < x < l$ and dissipated in similar material that extends infinitely in the positive x direction. This geometry is shown in Figure 4(a), with

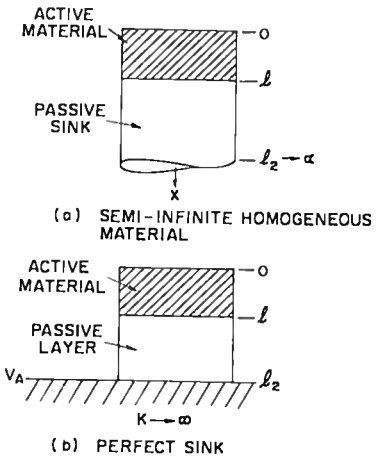


Fig. 4—Finite composite media.

$V_A \to 0$ as $l_2 \to \infty$. From Equations (3), (4), and (5), the normalized sink impedance σ is found to be

$$\sigma = \frac{Z(l)}{Z_1} = \frac{Z_2}{Z_1}\left(\frac{S_2(l_2)}{C_2(l_2)}\right)\Bigg|_{\substack{l_2 \to \infty \\ V_A \to 0}} = \frac{K_1}{K_2}\left(\frac{\alpha_2}{\alpha_1}\right)^{1/2}. \tag{6}$$

For homogeneous media, both source and sink are characterized by the same material constants, and $\sigma = 1$. For the active-layer, temperature distribution, Equation (3), becomes

$$V_1(x) = \bar{A}\left[1 - \frac{1}{2}\left(\exp\{-q_1(l-x)\} + \exp\{-q_1(l+x)\}\right)\right]. \tag{7}$$

A solution for $V_1(x, t)$ is obtained by replacing $\bar{A}$ by the suitable Laplace transform of a single pulse of width T, e.g., $\bar{A} = (\alpha/s^2)A_0(1 - e^{-sT_1})$. The resulting inverse transform of Equation (7) may be normalized to a temperature V_N

$$V_N = \frac{A_0 l^2}{K_1}. \tag{8}$$

Then,

$$\frac{V_1(x,t)}{V_N} = \frac{\alpha_1 t}{l^2}\left[1 - 2i^2\,\mathrm{erfc}\,\frac{l-x}{2\sqrt{\alpha t}} - 2i^2\,\mathrm{erfc}\,\frac{l+x}{2\sqrt{\alpha t}}\right]$$
$$\text{with } 0 < x < l \text{ and } 0 \leq t < T_1, \tag{9}$$

and

$$\frac{V_1(x,t)}{V_N} = \frac{\alpha_1}{l^2}\left\{T_1 - 2t\left[i^2\,\mathrm{erfc}\,\frac{(l-x)}{2\sqrt{\alpha t}} + i^2\,\mathrm{erfc}\,\frac{(l+x)}{2\sqrt{\alpha t}}\right]\right.$$
$$\left. + 2(t-T_1)\left[i^2\,\mathrm{erfc}\,\frac{(l-x)}{2\sqrt{\alpha}(t-T_1)} + i^2\,\mathrm{erfc}\,\frac{(l+x)}{2\sqrt{\alpha}(t-T_1)}\right]\right\} \text{ with } t > T_1. \tag{10}$$

Similarly, the average solution becomes

$$\frac{V_{av}(t)}{V_N} = \frac{\alpha_1 t}{l^2}\left[1 - \frac{2(\alpha_1 t)^{1/2}}{l}\left(\frac{1}{3\sqrt{\pi}} - 2i^3\,\mathrm{erfc}\,\frac{l}{\sqrt{\alpha t}}\right)\right] \text{ with } 0 < t - T_1. \tag{11}$$

and

$$\frac{V_{av}(t)}{V_N} = \frac{\alpha_1}{l^2}\left\{T_1 - \frac{2}{3l}\left(\frac{\alpha_1}{\pi}\right)^{1/2}\left(t^{3/2} - (t-T_1)^{3/2}\right)\right.$$
$$\left. + \frac{4\alpha_1^{1/2}}{l}\left(t^{3/2} i^3\,\mathrm{erfc}\,\frac{l}{\sqrt{\alpha t}} - (t-T_1)^{3/2} i^3\,\mathrm{erfc}\,\frac{l}{\sqrt{\alpha(t-T_1)}}\right)\right\} \text{ with } t > T_1. \tag{12}$$

The normalized temperature $V_1(x,t)/V_N$ is plotted in Figure 5, using Equations (8)—(12), for $x = 0$ and l. Figure 5 also shows the average value V_{av} and the differential $V(0) - V(L)$ normalized to V_N, as functions of the variable $\alpha_1 t/l^2$. The pulse width T_1 has been introduced as a normalized parameter $T_1' = \alpha_1 T_1/l^2$.

Also indicated are the results for a TEO with an active source-region length $l = 20$ μm, corresponding to a transit-time frequency near 5 GHz. The diffusivity averaged over n type and n+ type GaAs material has been used here ($\alpha_1 \simeq 0.265$ cm^2/sec), with $\alpha/l^2 \simeq 0.067 \times 10^6$ sec^{-1} and the appropriate material constants inserted as discussed previously (e.g., $nl \simeq 1.5 \times 10^{12}$ cm^{-2}).

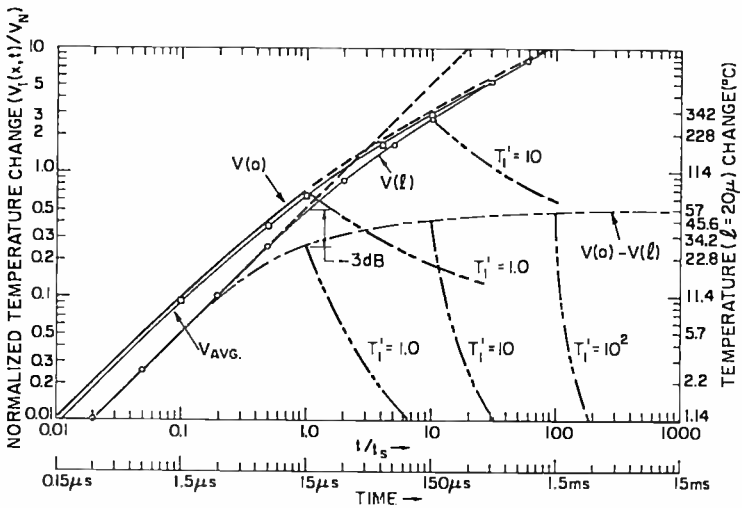


Fig. 5—Transient temperature response of n-layer terminated in infinite n+ material due to unit pulse.

From Equation (2) the normalization temperature V_N may be found for $\beta = 3$, $\gamma = 2/3$, and $K_1 \simeq 0.45$ watt/°C-cm; with a conservative device efficiency, $\eta = 0.05$, $V_N \simeq 5.70l = 114$°C. It should be noted that the thermal properties of GaAs actually vary with temperature[12,13] (e.g., $K_1 \simeq 1/T$). In addition, the thermal conductivity of GaAs is a function of the electron concentration: $K(\text{n}) \simeq 0.5$ at 300°K for $n = 10^{15}$ cm^{-3} and $\rho = 1$ ohm-cm, and $K(\text{n+}) \simeq 0.4$ at 300°K for $n = 10^{18}$ cm^{-3} and $\rho = 0.001$ ohm-cm.

The results of this example are intended primarily for illustrative purposes. The significance of the n+-layer thickness on effective heat removal is discussed more thoroughly when the GaAs material is referred to a fixed-temperature surface in the examples that follow. It is most interesting to note the "steady-state" temperature difference $V(0)$-$V(L) = 0.5V_N$, as indicated in Figure 5. The steady-state condition may be defined to occur after a time $t_s = l^2/\alpha_1$, when the temperature is within 3 dB of its steady-state value. In this example $t_s \simeq 15$

μsec. This parameter may be interpreted as the thermal time constant for this material geometry, where spreading resistance effects have been neglected since a one-dimensional heat flow is maintained.

Spreading Resistance Effect for Heat Flow into a Semi-Infinite Half-Plane

In the practical case, a small device is usually flip-chip bonded to a large post of thermal conductivity K_p as shown in Figure 4(b). The post geometry may be interpreted as a semi-infinite half-space with a thermal response time considerably larger than the time constants associated with the small device. The temperature build-up in the post may therefore be assumed to be negligible during initial heating transients. When the heat flux entering the post reaches an equilibrium value ($f(t+T) = f(t)$, where T is the period of repetitive heating pulse), an average steady-state thermal resistance to flow from the device cross-sectional area into the half-space of the post may be calculated. Several solutions for this problem have been found for a cylindrical device with a radius r. The assumptions of a variety of boundary conditions result in several values for the spreading resistance R_{sp} with a worst-case solution given by[1],

$$R_{sp} = \frac{V_{pav}}{F_p} = \frac{8}{3\pi^2 K_p r} \text{ (°C/watt)} \tag{13}$$

where,

V_{pav} = average rise in temperature at infinite plane of the post.

$F_p = \pi r^2 f_p(t)$—is the rate of heat flow across a device area of radius r into infinite half-space.

The effective average heating during pulse conditions is obtained by multiplying the heat density A_o by a duty factor a. For GaAs material with a constant doping length product for the active layer (e.g., $nl \simeq 1.5 \times 10^{12}$), the average temperature produced at the post surface, using Equations (2) and (13), is

$$V_{pav} \simeq \frac{a\beta\gamma r}{K_p} \times 10^4 = \frac{a\beta\gamma}{K_p} (25.4r \text{ mils}). \tag{14}$$

The normalized temperature is plotted in Figure 6 as a function of r, with K_p as a parameter.

Pulse Heating Of Composite Media

Unit-Step Single-Pulse Heating

The thermal time constants and temperature response for the semiconductor device shown in Figure 4(b) may be found as a function of the layer parameters. Media with different thermal characteristics are considered in this case. The source region, $0 < x < l$, with thermal

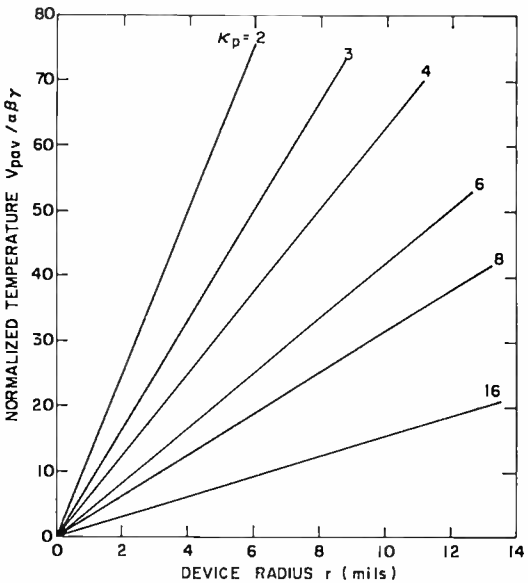


Fig. 6—Average temperature rise due to spreading resistance, with the post thermal conductivity as a parameter.

constants K_1, α_1, etc., passes heat through a subregion $l \leq x \leq l_2$, with constants $K_2\alpha_2$, etc., to an isothermal sink maintained at a temperature V_A. In this case, we shall neglect the contribution of the spreading resistance during a single pulse transient and use a super-imposed source V_A to provide a suitable isothermal reference plane. As before, the transient temperature rise is referred to zero initial conditions (i.e., $V_o(t) = 0$).

The normalized sink impedance for this geometry is found, using Equations (4)-(6), to be

$$\sigma = \frac{Z_2}{Z_1}\frac{S_2(l_2)}{C_2(l_2)}. \tag{15}$$

The resulting active-layer temperature distribution for a single pulse

source function $\bar{A}$ is found, by substituting Equations (8) and (15) into Equation (3), to be

$$V_1(x) = \frac{\alpha_1}{s^2 l^2} V_N \left[1 - \frac{N(s)}{D(s)} C_2(l_2) \right] + \frac{V_A}{s} \frac{N(s)}{D(s)}, \quad 0 < t < T_1. \tag{16}$$

The rational function $N(s)/D(s)$ is given by

$$\frac{N(s)}{D(s)} = \frac{C_1(x)}{C_1(l_1) C_2(l_2) + \sigma' S_1(l_1) S_2(l_2)}, \tag{17}$$

where

$$\sigma' = \frac{Z_2}{Z_1} = \left(\frac{\alpha_2}{\alpha_1} \right)^{1/2} \left(\frac{K_1}{K_2} \right) \tag{18}$$

Equation (17) is not readily transformed into real time by using tabulated results. The most direct approach is to apply the theorem of residues by expanding Equation (17) into partial fractions and using the Laplace inversion theorem. The poles of Equation (16) occur at $s = 0$ and $s = -\alpha_1 \beta_m^2$, where the $\pm \beta_m$ $(m = 1,2,\ldots.)$ are the roots of one of the two following equations:

$$\cos\beta l \cos k\beta l_2 - \sigma' \sin\beta l \sin k\beta l_2 = 0 \tag{19a}$$

$$\cot\beta l = \sigma' \tan k\beta l_2 \tag{19b}$$

with

$$k = \frac{q_2}{q_1} = \frac{1}{\sigma'} \left(\frac{K_1}{K_2} \right) = \left(\frac{\alpha_1}{\alpha_2} \right)^{1/2}. \tag{20}$$

When $l \neq kl_2$, there is no explicit analytic method for computing the roots of Equation (19a). The only feasible methods are either graphical construction or numerical computation. With $l = kl_2$, the roots are found from $\tan\beta l = \pm\sqrt{1/\sigma'}$.

The residues at $s = 0$ for the $\bar{A}$ source term in Equation (16) are found for the multiple poles $1/s^2$ and $1/s$ to be zero and $[(l^2 - x^2) + 2l_2 l(K_1/K_2)]/2\alpha_1$, respectively. The $s = 0$ residue in the V_A source term is unity. The remaining residues are found by solving Equation

(19) for the roots and substituting into the conventional expansion term,

$$\frac{1}{s^n}\left[\frac{N(s)}{dD(s)/ds}\right]_{s=-\alpha_1\beta_m}$$

using the correct multiplicity factor n. From Equation (16) and the appropriate tabulated transforms,

$$V_1(x,t) = V_s + V_A - \sum_{m=1}^{\infty} (V_{sm} + V_{Am})\, e^{-t/\tau_m},\ 0 < x < l,\ 0 < t < T_1 \tag{21}$$

with the steady-state primary source response, as $t \to \infty$, given as

$$V_s = V_N\left[\frac{1}{2}\left(1 - \left(\frac{x}{l}\right)^2\right) + \frac{l_2}{l}\frac{K_1}{K_2}\right], \tag{22}$$

which is essentially reached for $t \simeq 3\tau_m$, neglecting the contribution of spreading resistance during this short response time.

The steady-state temperature gradient, $\Delta V_s = V_s(0) - V_s(l)$, may be found from Equation (2) to be $V_N/2$, which is plotted in Figure 7 as a function of the length l with $\beta\gamma$ as a parameter. The average disipated power P_D versus the device area is also plotted in this figure, again with $\beta\gamma$ as a parameter. Thus, a knowledge of P_D and the device area can be used to find an effective value of the drive parameter $\beta\gamma$, which, in turn, may be used to relate V_N to l.

The thermal time constants are

$$\tau_m = \frac{1}{\alpha_1\beta_m^2} = \frac{l^2}{\alpha_1(\beta l)^2}\ ; \tag{23}$$

the transient term coefficients are

$$V_{sm} = \frac{V_N\left(\dfrac{2}{l^2\beta_m^3}\right)\cos k\beta_m l_2 \cos\beta_m x}{(l + \sigma' k l_2)\sin\beta_m l \cos k\beta_m l_2 + (k l_2 + \sigma' l)\cos\beta_m l \sin k\beta_m l_2} \tag{24}$$

$$V_{Am} = V_A \left[\frac{\left(\frac{2}{\beta_m}\right)\cos\beta_m x}{(l + \sigma' k l_2)\sin\beta_m l \cos k\beta_m l_2 + (k l_2 + \sigma' l)\cos\beta_m l \sin k\beta_m l_2} \right], \tag{25}$$

which may be slightly simplified using Equation (19). These co-

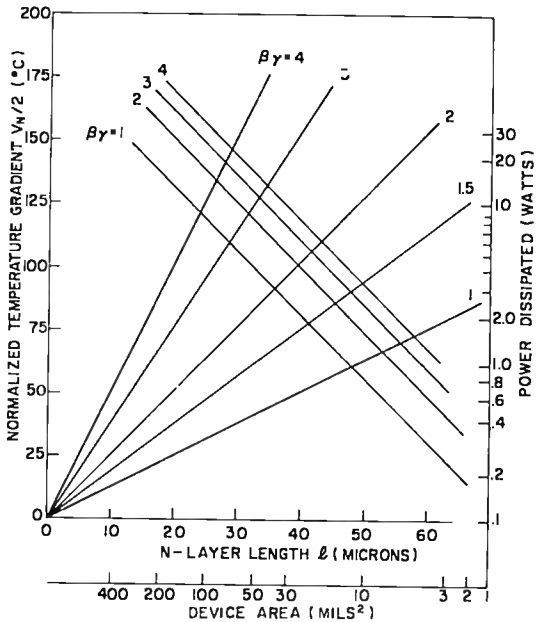


Fig. 7—Steady-state temperature gradient and average power dissipated in active n-layer with a thermal conductivity $K_1 = 0.5$ watts/°C-cm.

efficients may be related to the steady state response V_s, using Equation (21) for $t = 0$,

$$V_s = \sum_{m=1}^{\infty} V_{sm} \tag{26}$$

$$V_A = \sum_{m=1}^{\infty} V_{Am}. \tag{27}$$

The single-pulse thermal transient build-up normalized to $V_A + V_s$

is shown in Figure 8. In a similar manner, the "off" pulse result for the primary source removed at T_1 *is*

$$V_1(x,t) = V_A + \sum_{m=1}^{\infty} \left[\left(V_{sm} \exp\left\{ -\left(\frac{t - T_1}{\tau_m} \right) \right\} \right) \left(1 - \exp\left\{ -\frac{T_1}{\tau_m} \right\} \right) - V_{Am} \exp\left\{ \frac{-t}{\tau_m} \right\} \right], \quad t > T_1, \tag{28}$$

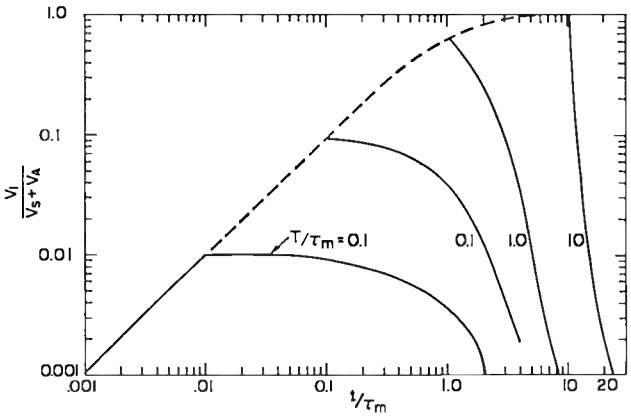


Fig. 8—Single pulse response.

wnere the isothermal surface at l_2 is maintained at V_A (i.e., unit step at $t > 0$). The use of a negative sink V_A (i.e., a thermoelectric device) would reduce the net rise in temperature, assuming sufficient compensation is made for the inevitable contribution of the spreading resistance.

The steady-state asymptotic limit of Equation (22) has been derived by assuming a constant value of thermal conductivity with respect to temperature for the different semiconductor layers. For the temperatures that would be encountered in practice, $K \sim 1/T$. A comparison may be made between the results obtained for V_s in Equation (22) and those obtained for the inverse temperature value of K as derived in Appendix II.

We may relate V_s to an actual temperature V_{so} with a reference temperature $T_o \simeq 300°\text{K}$. These results are indicated in Figure 9 for $x = 0$ as a function of the layer thickness l. Various drive conditions are indicated with $\beta\gamma$ as a parameter (i.e., $A_o \simeq 1.28 \times 10^7\ \text{W/cm}^3$ for

$\beta\gamma = 2$ and $l = 20\ \mu$m). The error involved in neglecting the inverse temperature dependance of K, relative to the steady state result, is $\leq 20\%$ for material with $l < 20\ \mu$m and $A_o \leq 2 \times 10^7$ W/cm^3. In most cases the error is significant, but in thicker layers, where the calculated results indicate abnormally high temperatures, the errors become considerable. These high temperature results are primarily illustrative and have been derived assuming perfectly linear material

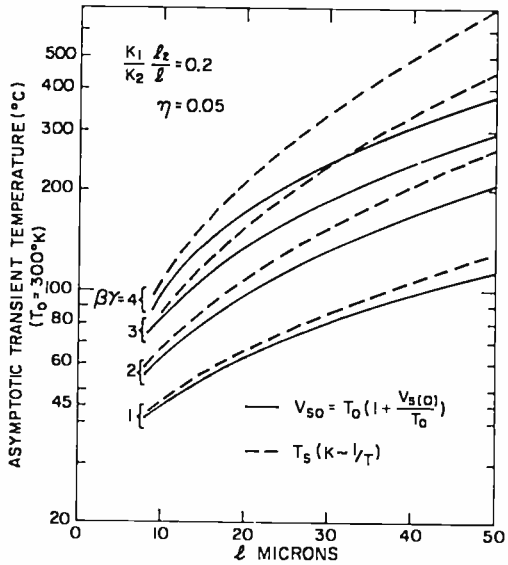


Fig. 9—Steady-state temperature rise at $x = 0$ (maximum temperature location) for optimum $nl = 1.5 \times 10^{12}$ cm^{-2} in a TEO.

parameters and are *not valid* when these parameters are greatly affected by thermal effects (i.e., n or $\rho_o = f(T)$). During cw operation, the temperature V_{pav} from Equation (14) may be added to V_s to give an estimate of the actual device temperature at $x = 0$.

Steady Periodic Heating

The primary source solution for a repetitive pulse[16] applied to the geometry as shown in Figure 4(b) is found by using the appropriate pulse transform for $A(s)$ in Equation (3) and following the procedure for inversion that lead to Equations (21) and (28). The additional

[16] Jaeger, J. C., "Pulsed Surface Heating of SemiInfinite Solid," *Quart Appl. Math.*, Vol. 11, p. 132, 1953.

poles resulting from the periodic source term are at,

$$s = 2n\pi iT,\ n = \pm 1, \pm 2, \ldots,$$

which gives a series of oscillations of periods T, $T/2, \ldots$. These new roots lead to a steady periodic temperature variation, while the $s = 0$ roots provide the steady state average value, $V_s T_1/T$. Both roots lead to a steady periodic term V_{sp}, whereas the roots $s = -\alpha_1\beta_m^2 = -1/\tau_m$ provide the transient part. The V_A unit-step-source solution has already been found in Equations (21) and (25). The transient result for a periodic source V_A is analogous to the following result for the primary source. The effect of spreading resistance can be added to the steady-state terms by replacing V_A by V_{pav} from Equation (14).

An evaluation of the residue at each pole leads to the combining result

$$V_1(x, t) = V_{sp} - \sum_{m=1}^{\infty}\left(V_{sm}\left[\frac{\exp\{-t/\tau_m\} - \exp\{-(t - T_1)/\tau_m\}}{1 - \exp\{T/\tau_m\}}\right] + V_{Am}\exp\left\{\frac{-t}{\tau_m}\right\}\right). \quad (29)$$

The previous results for a single pulse may be used to find a convenient solution for V_{sp}. Since Equation (29) is valid for all time, it must agree with Equation (21) for $0 < t < T_1$, and with Equation (28) for $T_1 < t < T$. Thus, equating Equations (29) and (21), V_{sp} for $0 < t < T_1$ is

$$V_{sp} = V_s + V_A - \sum_{m=1}^{\infty} V_{sm}\exp\{-t/\tau_m\}\left(\frac{1 - \exp\{-(T - T_1)/\tau_m\}}{1 - \exp\{-T/\tau_m\}}\right). \quad (30)$$

The periodic temperature V_{sp} given in Equation (30) is the result for each "on" pulse at any time $nT + t$ where n is an integer.

Again using Equation (28), the steady periodic temperature at a time $nT + t$, where $T_1 < t < T$, is

$$V_{sp} = V_A + \sum_{m=1}^{\infty} V_{sm}\exp\{-(t - T_1)/\tau_m\}\left(\frac{1 - \exp\{-T_1/\tau_m\}}{1 - \exp\{-T/\tau_m\}}\right). \quad (31)$$

Since the heating time T_1 is usually considerably smaller than the thermal constants τ_m, we may simplify Equation (30) by using Equation (26) and a small argument exponential expansion for low-duty-cycle operation:

$$V_{sp} \simeq V_A + a \sum_{m=1}^{\infty} V_{sm} \left(\frac{b}{\lambda_m} \right), \quad T_1 \gg \tau_m, \tag{32}$$

Similarly, for high duty cycle (i.e., $T_1/T \geq 0.1$, $T_1 < \tau_m$,

$$V_{sp} \simeq V_A + a \sum_{m=1}^{\infty} V_{sm} \left[1 + \frac{b}{\lambda_m} (1 - a) \right], \quad T \ll \tau_m, \tag{33}$$

and for medium duty ($T \simeq \tau_m$, $\lambda_m \simeq 1$)

$$V_{sp} \simeq V_A + a \sum_{m=1}^{\infty} \frac{V_{sm}}{\lambda_m} \left[\frac{1}{(e^{1/\lambda_m} - 1)} + b \left(1 - \frac{a}{\lambda_m} \frac{1}{(e^{1/\lambda_m} - 1)} \right) \right], T \simeq \tau_m \tag{34}$$

where $a = T_1/T$ (duty factor), $b = t/T_1$ (normalized heating time), and $\lambda_m = \tau_m/T$ (relative time constants). These approximate expressions have been used to plot the normalized curves in Figures 10(a), (b), and (c), which indicate the periodic build-up of the "average" temperature as a function of the parameters a, b, and λ_m. Similar results could be readily obtained from Equation (30) for wide pulse widths (i.e., $T_1 > \tau_m$). With the inclusion of an additional pulse-modulation term in the source $A(s)$, this analysis may be extended to cover the burst-heating mode, in which short fast pulses at a high duty cycle occur periodically during a gated interval.

In a practical situation, where the external source V_A is not available, one may substitute V_{pav}, the steady-state effect of the spreading resistance from Equation (14), such that $V_A \rightarrow V_{pav}$ in Equations (30) through (34). The validity of this substitution is based upon the fact that the thermal time constant of the post geometry is considerably larger than any τ_m, thereby resulting in a rising pulse transient response for the device (Figure 10) superimposed upon a *constant* temperature, $V_A = V_{pav}$.

The equations derived in this section are very general in form. In many cases considerable simplification is possible after the roots of Equation (19) are found. Their relative magnitude often permits a

first-order approximation for the constant τ_m. These roots and the associated thermal-transient response for this layered geometry are functions of the material thermal constants and the length associated with the sink interface layer $l \leq x \leq l_2$.

A computation of the thermal constants τ_m for all GaAs n/n+ material, as well as for GaAs n on Si n+ substrate, based on Equation (19) and the appropriate constants from Table I is shown in Table II.

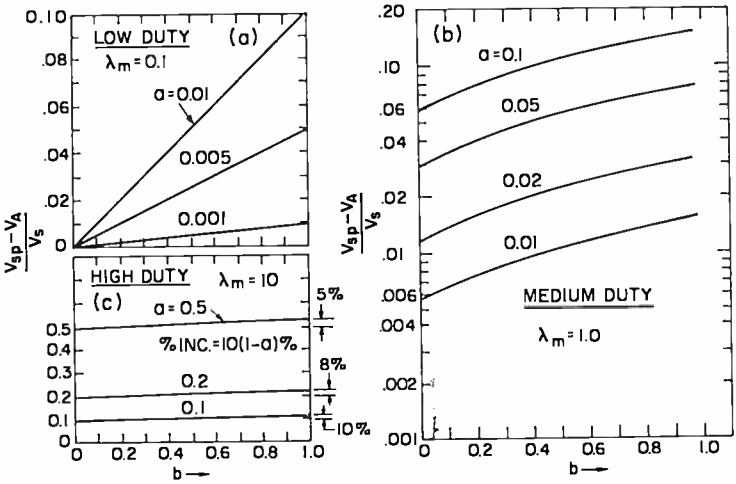


Fig. 10—Temperature rise during pulse for steady periodic heating.

Several relative lengths kl_2/l have been used as parameters. The limiting result for $l = 0$ (i.e., $\beta_m l = m\pi/2$) is also indicated. The roots extracted from Equation (19) require an active-layer length l to be converted into the time constants in Equation (23). The τ_m for both 10 μm and 20 μm samples have been computed. The result for any length is readily obtained by multiplying the 10 μm result by the square ratio of the new length relative to 10 μm. Similarly, in the case of symmetrical heat sinking from both sides of the active layer, the effective length l is halved and τ_m thereby quartered.

The presence of a contact layer ($l_2 \neq 0$) is seen to increase τ_m, although insignificantly, for $kl_2/l < 1/5$ relative to the ideal condition where $l_2 = 0$. The use of Si rather than GaAs(n+) substrate improves the relative transient response when $kl_2/l > 1/5$, but has only a small effect for thin layers. The most practical scheme for reducing the thermal rise in the contact n+ layer is to bond the thin epi-layer to the sink using suitable flip-chip techniques.

Table II—Thermal Time Constants for Specific TEO Geometries

kl_2/l	τ_1 μsec		τ_2 μsec		τ_3 μsec		τ_4 μsec	
	$l = 10$ μm	$l = 20$ μm	$l = 10$ μm	$l = 20$ μm	$l = 10$ μm	$l = 20$ μm	$l = 10$ μm	$l = 20$ μm
				GaAs n/n+, $k = \sigma' = 1.12$				
1	6.10	24.4	.610	2.44	.230	.921	.112	.450
1/2	3.36	13.5	.354	1.52	.125	.5	—	—
1/5	2.1	8.4	.23	.925	.082	.328	—	—
1/10	1.72	6.9	.192	.77	.07	.28	—	—
				GaAs n, Si substrate, $k = .670$, $\sigma' = .650$				
1	4.4	17.6	.68	2.72	.213	.86	.12	.48
1/2	2.44	9.8	.354	1.52	.134	.54	—	—
1/5	1.77	7.1	.203	.82	.078	.312	—	—
1/10	1.58	6.3	.178	.71	.064	.256	—	—
				GaAs n ($l_2 = 0$)				
0	1.40	5.6	.156	.62	.056	.224	—	—

Conclusion

A method of analyzing the temperature gradient within the active layer of a bulk-effect GaAs device during pulsed excitation has been presented. The close analogy between the layered thermal geometry of this and other semiconductor devices and a cascaded electrical transmission line permits the use of electrical duals to the thermal parameters. The Laplace transform, matrix techniques, and superposition of sources may readily be applied in analyzing the thermal network.

When a TEO device is characterized by an optimum product of the electron concentration and the active layer length, the heat density is inversely related to this length. The asymptotic steady-state temperature is *linearly* dependent on n-region thickness, while the total dissipated power remains independent of the n-region thickness and proportional to the cross-sectional area. The error involved in neglecting the inverse temperature dependence of the semiconductor thermal conductivity when calculating its steady-state asymptotic value is insignificant in most cases where $l \leq 20$ and the heat density $A_0 \leq 2 \times 10^7$ W/cm^3.

The thermal transient response and associated time constants are most seriously affected by the contact-layer thickness. The poor thermal characteristics of GaAs require both minimum active-layer and contact-layer thickness to reduce thermal rise time. The limiting response is imposed by the active layer, where the time constants are related to the square of the different relative lengths. The thickness of the contact layer becomes insignificant when it is less than 20% of the active-layer lengths. Since the semiconductor material (primarily because of its relative volume) is actually faster thermally than the larger volume of metallic sink, the transient temperatures derived in this analysis, may be referred to an isothermal surface. The spreading resistance is usually significant and the resulting equilibrium temperature may be added directly to the steady periodic or cw solutions.

Although this analysis has used the TEO as a model for studying the transient temperatures associated with a pulsed device, the mathematical approach may be readily applied to evaluating other layered devices, possibly including those with nonuniform heating.

Acknowledgment

The author is grateful to T. E. Walsh for his encouragement and independent research, which provided the basis for several constructive suggestions.

APPENDIX I—HEAT CONDUCTION IN ONE DIMENSION

A. Fundamental Relations

Consider a solid capable of generating heat energy (source) throughout its volume and supplying that heat through a process of conduction to a passive medium (sink), which is assumed to be unheated by any internal Joule effects. The entire volume is considered to be enclosed in an adiabatic environment where no heat is removed from any surface through any process other than conduction (i.e., no radiation or convection). We will now examine closely the thermal behavior of a one-dimensional geometry (see Figures 1 and 2).

The flow of heat in a one-dimensional linear medium containing a continuous source of heat ($t > 0$) is described by the following equation[1]

$$\frac{\partial^2 V}{\partial x^2} - \frac{1}{\alpha_1}\frac{\partial V}{\partial t} = -\frac{A(x,t)}{K_1}, \quad 0 < x < l, \; t > 0 \tag{35}$$

where $\alpha_1 = K_1/\rho_1 c_1$ is the thermal diffusivity, (K, ρ, and c are thermal conductivity, density, and specific heat, respectively), $A(x, t)$ is the rate of heat generation per unit time per unit volume, and the temperature $V = T\ (x, t)$. It has been assumed that the thermal properties, K, α, c, and A are all independent of the temperature as well as the position within a given medium. In the more general case,[2,5] where $K, c = f(V)$, the solution of a nonlinear partial differential equation is required. For the general case, an exact analytic transient solution for the present model would not be possible; however, a steady-state solution may be found (see Appendix II) and compared with asymptotic limit in the following linearized analysis.

Equation (35) may be transformed into Laplace s-space using

$$L\{V(x,t)\} = V(s) = \int_0^\infty e^{-st} V(x,t)\, dt. \tag{36}$$

Then,

$$\frac{\partial^2 V_1(s)}{\partial x^2} - q_1^2 V_1(s) = -\frac{A(s)}{K_1} - \frac{V_o}{\alpha_1}, \quad 0 < x < l, \tag{37}$$

and, similarly,

$$\frac{\partial^2 V_n(s)}{\partial x^2} - q_n^2 V_n(s) = -\frac{V_o}{\alpha_n}, l < x. \tag{38}$$

Here, V_o is the initial temperature of the solid at $t = 0$. Also,

$$q_n^2 = s/\alpha_n, (n = 1, 2, 3, \text{etc.}) \tag{39}$$

Substituting $V = U + \bar{A}$ into (37), where the particular source solution is

$$\bar{A} = \frac{A(s)\alpha_1}{sK_1} + \frac{V_o}{s}, \tag{40}$$

we obtain an ordinary homogeneous equation in $U(s)$. This equation is readily solved to give

$$U(x) = U(l)\cosh q_1 (l - x) + f(l)Z_1 \sinh q_1 (l - x), \tag{41}$$

where l is the length of the source (Region I) and

$$U(l) = V(l) - \bar{A}, \tag{42}$$

and where

$$Z_1 = \frac{1}{K_1 q_1}, Z_n = \frac{1}{K_n q_n} \tag{43}$$

B. Boundary Conditions

The appropriate flux boundary conditions for the source in Region I are

$$f(x) = -K_1 \frac{\partial V}{\partial x} = -K_1 \frac{\partial U}{\partial x} = \begin{cases} 0 \text{ at } x = 0 \\ f(l) \text{ at } x = l \end{cases} \tag{44}$$

corresponding to unilateral heat flow in the positive x direction. At all remaining interfaces (between each subregion), both the fluxes and temperatures are continuous (i.e., V and dV/dx are continuous).

For the purposes of this analysis, V_o, the initial temperature of the solid at $t = 0$, is selected to be equal to zero. The resulting solution of Equation (38) provides the temperature relationship between the sink

regions and the source (all initially at zero temperature) after consideration of all the necessary boundary conditions, particularly at the interface where $x = l$ in Equation (44). In most practical cases, the device can assumed to be terminated by an infinite half-space resulting in a finite spreading resistance. During thermal transients, the device response time is considerably faster than that for the half-space sink. The transient solution shall ignore the spreading resistance and refer the device to an isotherm. A constant temperature is achieved by introduring a second source (i.e., V_A/s) into this analysis by using the superposition principle. This method provides the transient response resulting from the sink termination.

The only other case of interest is when $V_A = 0$ at $x_n \to \infty$ and the sink is thereby considered to be infinite in its length. Without the mechanism of heat removal by convection or radiation, this geometry would result in a continued increase in temperature due to $A(s)$.

C. *Source Condition (Uniform Heat Generation)*

In Laplace notation, we may represent a uniform power density A_o (watts/cm^3) as

$$A(s) = \frac{A_o}{s}. \tag{45}$$

The primary source driving function $A(s)$ may consist of a periodic waveform of period T and pulse "on" width T_1. In Laplace notation,

$$A(s) = \frac{A_o(1 - e^{-sT_1})}{s(1 - e^{-sT})}, \quad \begin{matrix} A(t) = A_o,\ nT < t < nT + T_1 \\ A(t) = 0,\ nT + T_1 < t < (n+1)T \end{matrix} \tag{46}$$

In many cases, it is of significant interest to study the "heating" behavior of our sample during and after a *single* pulse. This simplified case would provide useful rise and decay transient information for a specific geometry. In this case a delayed negative unit step is superimposed upon a unit step at $t = 0$, so that Equation (46) reduces to

$$A_2(s) = \frac{A_o}{s}(1 - e^{-sT_1}) \tag{47}$$

D. *Linear Thermal Network*

The solution to Equation (37), for the temperature distribution in

the source region of the geometry shown in Figure 3 arising from $A(s)$ and with the necessary boundary conditions of Equation (44), is

$$\overline{V_1}(x) = \bar{A}\left[1 - \frac{C_1(x)}{C_1(l) + \sigma S_1(l)}\right], 0 < x < l, \tag{48}$$

where

$$\sigma = \frac{Z(l)}{Z_1} \tag{49}$$

$$C_1(x) = \cosh q_1 x$$
$$S_1(x) = \sinh q_1 x \tag{50}$$

and a driving point thermal impedance has been defined as

$$Z(l) = \frac{V(l)}{f(l)}, \tag{51}$$

where $Z(l)$ provides a direct algebraic means of introducing the passive ($V_A = 0$) heat-sink parameters into the solution of Equation (48). Equation (38) may be solved for each individual sink region in a similar way. The terminal relationship for each subregion may be written in matrix form as

$$\begin{bmatrix} V_n(l_n) \\ \\ f_n(l_n) \end{bmatrix} = \begin{bmatrix} C_n(l_n) & -Z_n S_n(l_n) \\ \\ -\frac{1}{Z_n} S_n(l_n) & C_n(l_n) \end{bmatrix} \begin{bmatrix} V_{n-1}(l_{n-1}) \\ \\ f_{n-1}(l_{n-1}) \end{bmatrix} \tag{52}$$

where $l_n = x_n - x_{n-1}$ corresponds to the length of the n^{th} subregion, and $q_n^2 = s/\alpha_n$. The form of solution in Equation (52) is readily extended to a cascaded series of subregions, each characterized by a set of constants, K_n, ρ_n, C_n, α_n, etc.

The function $Z(l)$ may be found by multiplication of the sink matrices of Equation (52), and substitution into Equation (48) with appropriate boundary conditions. For the case where a finite length of sink is terminated at a constant temperature V_A, we may use superposition. Initially, $V_A = 0$, so that the multiplication of the subregion matrices of Equation (52) gives,

$$\begin{pmatrix} 0 \\ f_A \end{pmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{pmatrix} V(l) \\ f(l) \end{pmatrix} \tag{53}$$

where it is readily found that AD-CB = 1. The primary source driving point impedance of Equation (51) for $V_A = 0$ becomes,

$$Z(l) = \frac{V(l)}{f(l)} = -B/A. \tag{54}$$

It is now necessary to let $A(s) = 0$ and solve for the temperature distribution in Region I due to a constant source V_A at $x = x_n$. Reversing the flow of heat at $x = l$ and using Equations (52) and (44),

$$V_A(l) = Z_1 \frac{C_1(l)}{S_1(l)} f_A(l). \tag{55}$$

The subscript A is used to denote dependance on source V_A with flow in the negative x direction. Introducing a finite source V_A (i.e., $0 \to V_A$, $f_A \to -f_A$, $f(l) \to -f_A(l)$) into Equation (53) we have,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} 1 & \\ -\frac{1}{Z_1} & \frac{S_1(l)}{C_1(l)} \end{bmatrix} = \begin{bmatrix} V_A & \\ V_A(l) & \\ -f_A & \\ V_A(l) & \end{bmatrix} \tag{56}$$

or,

$$V_A(l) = \frac{V_A}{A - \frac{BS_1(l)}{Z_1 C_1(l)}}. \tag{57}$$

Then, using Equations (41), (55), and (57) and noting the direction of $f_A(x)$,

$$V_{1A}(x) = V_A \left[\frac{C_1(x)}{AC_1(l) - \frac{B}{Z_1} S_1(l)} \right]. \tag{58}$$

Combining the results of Equations (47), (49), (54), and (58),

$$V_1(x) = \overline{V_1(x)} + V_{1A}(x)$$
$$= \bar{A}\left[1 - \frac{C_1(x)}{C_1(l) + \sigma S_1(l)}\right] + \frac{V_A}{s}\left[\frac{C_1(x)}{AC_1(l) + \sigma S_1(l)}\right]. \tag{59}$$

It may often be necessary to know the variation of the average temperature in the active layer (Region I). For this case, using Equation (59),

$$V_{AV}(x) = \frac{1}{l}\int_0^l V_1(x)\,dx,$$

$$= \bar{A}\left(1 - \frac{S_1(l)}{lq_1[C_1(l) + \sigma S_1(l)]}\right) + \frac{V_A}{s}\left(\frac{S_1(l)}{lq_1A[C_1(l) + \sigma S_1(l)]}\right)$$
$$\text{for } 0 < x < l$$

E. Inverse Transformation

The resulting temperature distribution, $V_1(x, t)$ is found from the Laplace transform by the use of an inversion theorem, and contour integration. Thus,

$$V_1(x, t) = \frac{1}{2\pi i}\int_{\gamma - i\infty}^{\gamma + i\infty} e^{st}V_1(x, s)\,ds, \gamma > 0. \tag{50}$$

For the present problem, where we are dealing with finite subregions, the transforms are free of branch points and usually have poles at the origin and along the negative real axis. The inverse transforms of Equation (59) reduce to a sum of residues of $V_1(x, s)$. In many cases, the transforms are readily obtained as tabulated functions.

Appendix II—The Effect of an Inverse Temperature Dependence of K on the Steady-State Solution

If the inverse temperature variation of the thermal conductivity K

is included in the thermal diffusion equation with $\partial T/\partial t = 0$ in the steady state, it is readily found[1] that for uniform heat generation throughout the active layer ($0 \leq x \leq l$),

$$\frac{d^2}{dx^2}(\ln T) = \frac{-A_o}{K_1 T_o}, \quad 0 < x < l \tag{63}$$

$$\frac{1}{T}\frac{dT}{dx} = -\frac{(A_o l)}{K_2 T_o}, \quad l \leq x \leq l_2. \tag{64}$$

T represents the variable temperature and T_o is the reference ambient sink temperature (i.e., 300°K) at which the *constant* values of K have been calculated in the preceeding analysis.

Equations (63) and (64) are readily solved[2] and may be used for comparison with the two-layer case discussed in this paper. After introducing similar boundary conditions and neglecting sink resistance[15] for $x > l_2$,

$$\frac{T_s}{T_o} = \exp\left\{\frac{V_s}{T_o}\right\} \tag{65}$$

where T_s is the steady-state asymptotic temperature characteristic of the inverse temperature variation of K, and V_s is the steady-state asymptotic temperature relative to T_o for a constant value of K (Equation (22)). Substituting Equation (22) into Equation (65) and choosing a reference temperature $T_o \simeq 300$°K, we may calculate the actual device temperatures T_s and V_{so}. These results are plotted in Figure 9 versus the layer thickness l.

DIGITAL LOGIC FOR RADIATION ENVIRONMENTS: A COMPARISON OF METAL—OXIDE–SEMICONDUCTOR AND BIPOLAR TECHNOLOGIES

BY

W. J. DENNEHY,* A. G. HOLMES-SIEDLE,* and K. H. ZAININGER†

Summary—In analyzing the impact of radiation environments on logic circuitry constructed of MOS and bipolar transistors, one must consider the magnitude and relative importance of two very different classes of phenomena—permanent effects and transient effects. This paper reviews the phenomenology of these two classes of effect, presents a detailed analysis of a typical pair of existing, competing circuit designs, and describes the advantages that are likely to result from recent improvements in device technology. These improvements include (1) the use of aluminum oxide as the MOS gate-insulator material in place of silicon dioxide and (2) the adjustment of the geometry of certain "critical junctions" in MOS devices. These advances are discussed in relation to the overall problem of hardening integrated circuits for pulsed and steady-state radiation environments. It is concluded that, although some current designs of MOS logic device may have a sensitivity to permanent damage that is a serious problem in some expected radiation situations in space, MOS technology has a high potential for producing very "hard" circuits and systems.

1. INTRODUCTION

A. General Comments on MOS Logic

THE TECHNOLOGY for the design and fabrication of Metal–Insulator–Semiconductor (MIS) devices has progressed very rapidly in the past five years. As a result, this class of device is now overtaking its competitors in certain fields of use, e.g. aerospace logic, which is the main topic of this article. The MOS field-effect transistor (MOSFET) has certain intrinsic advantages as regards size, power consumption, and cost that make it very attractive in this field.

* RCA Astro-Electronics Division, Princeton, N. J.
† RCA Laboratories, Princeton, N. J.

For the MOSFET to emerge as a competitor of bipolar transistors in aerospace logic, however, improvements in the gate-insulator-layer material were required. The insulator used initially in these devices was a thermally grown layer of silicon dioxide. This is simple to produce but difficult to perfect as a gate insulator. Many more complex insulator layers were examined as replacements and some have been adopted quite widely. As pointed out in earlier publications[1-6], the electrical demands made by the MOSFET system on this thin insulator layer are very severe. Even in conventional circuit applications, the electrical stresses put on the film can produce instabilities or even complete breakdown. Thus, research is still in progress to find a better insulator for these devices.[7] In addition, when MOS devices are used in space vehicles, missiles, and some other military equipment, a new, very unconventional "stress" must be met, namely, bombardment by high-energy particles or photons that are sometimes in the form of intense submicrosecond bursts. This type of environment results in serious degradation of the performance of this device, and has acted as a powerful deterrent to its use in the present generation of space and missile systems.

This paper describes the impact of two recent research results that could change this situation: (1) the development of a gate insulator material that appears promising for all MOS applications, and that is excellent in its tolerance to large doses of ionizing radiation (i.e., permanent radiation effects); and (2) the development of principles for topology adjustment to reduce the transient response of devices to pulsed irradiation.

The new insulator of most interest is aluminium oxide (Al_2O_3) formed by various processes, such as, e.g., plasma anodization[8, 22] or vapor deposition.[9, 22] If this gate insulator can be developed successfully and applied economically to a topologically adjusted integrated circuit, it will become possible to employ MOS devices in the wide range of aerospace logic circuits that are subjected to radiation environments. The main object of this paper is to describe the advantages of using MOS devices in these applications; in other words, to describe quantitatively what the development of a radiation-resistant MOS logic device would "buy" in system performance. In the process, the current state-of-the-art in tolerance levels of components to transient and permanent radiation effects and the physics behind these comparisons are reviewed. We will then review our own recent work on Al_2O_3 devices and the phenomenology of transient effects. The object will be to give a realistic comparison of truly parallel cases of logic circuit constructed from the three main competing devices (bipolar, complementary MOS (CMOS),

and all-p-channel MOS (PMOS) transistors) and their use in a typical high-radiation application.

B. Radiation Tolerance In Aerospace Systems

In the following discussion, the descriptive terms *hard, hardness,* and *hardening* are frequently used. The terms *soft* and *hard* have been adopted from military terminology, where an installation or system is *hard* or *survivable* if it can survive the maximum expected level of a given type of assault and is *soft* or *vulnerable* if it cannot. A branch of *survival technology* is *radiation hardening*. The *hardness* level of a component is the highest radiation rate or integrated flux that the component will withstand without undergoing a specified permanent degradation or temporary malfunction. Here, of course, the terms can be given precise physical values, whereas the military use involves a less precisely definable "figure of merit."

Among the elements that make up satellite and airborne vehicles, bipolar transistors and MOS transistors are usually the most easily damaged or disturbed by radiation. In designing such vehicles, countering the effects of radiation is complicated by the dominating need to conserve weight. Since it is rarely possible to provide enough shielding to stop the impinging radiation completely, the electronic devices must function at some acceptable level in spite of the radiation. As will be shown, most SiO_2-MOS logic systems are softer than bipolar logic systems in certain important respects. However, the MOS device is very flexible with regard to device geometry and materials used, and several modifications can be made in these devices that will bring their level of hardness up to and probably higher than that of the bipolar transistor. Such MOS devices could then be chosen by the designer for use in the many circuits where the MOS device is preferable for electrical and other considerations. Thus, the system as a whole could be hardened without penalty in weight or power.

A brief comparison of the radiation environments expected in aerospace missions, and of the radiation levels at which many solid-state devices malfunction, may serve to indicate the severity of the problem in today's hard system design. Some typical levels at which malfunction due to permanent radiation damage occurs are given below for conventional commercial (unhardened) MOS and bipolar transistors:

(1) 10^{13}n/cm^2 of fission-spectrum neutrons produces "bulk displacement defect damage" in silicon and hence destroys gain in bipolar transistors.

(2) 3×10^5 rads of gamma or X-rays produces charge trapping or "ionization damage" in the insulator layers at the surface of MOS and bipolar transistors, reducing gain, shifting turn-on voltages, and increasing surface leakages,

(3) 10^{13}e/cm^2 of 1-MeV electrons produces both ionization damage and bulk damage, the former being the dominant effect for many integrated-circuit bipolar transistors.*

Even more critical for logic operation is the nondestructive transient malfunction, often called TREE Response (Transient Radiation Effects in Electronics). TREE responses are most commonly the result of a nuclear-weapon burst. A flash of carrier-generating illumination penetrates the encapsulant around the device. This effect is usually dominated by gamma-rays. As a result, information in the form of logic voltages or stored charges may be lost. The response of the device is in the form of microsecond pulses of electrical current or disturbances of logic voltages within the integrated silicon chip. A soft bistable circuit may change state at peak-radiation intensities (dose rates) as low as 10^7 rad/sec.

Conditions in space and the current design of nuclear weapons are such that, for many space, aircraft, and missile missions, the radiation levels quoted for malfunction can be exceeded greatly, particularly with respect to the peak dose rates experienced. As an example, RCA's Relay satellite, flying for a few months through the trapped radiation belts generated by the "Starfish" space nuclear explosion that occurred in July 1962, accumulated radiation fluxes greater than 10^{14} electrons/cm^2 of energy exceeding 1 MeV. Internal doses from these particles exceeded 10^6 rads. In addition, if the satellite had been within line of sight of the Starfish explosion in July 1962, it could have experienced peak dose rate levels as high as 10^8 rad/sec, even in its relatively high orbit.* A vehicle flying near the Starfish burst could receive a fluence of neutrons of value many orders of magnitude higher than that quoted above as causing failure in some devices.

* For a heavily driven bipolar transistor, 10^{13}e/cm^2 at $E = 1$ MeV is equivalent to approximately 3×10^5 rads (Si) of cobalt-60 gamma radiation in the amount of ionization damage produced, and 10^{11} fission-spectrum neutrons/cm^2 in the amount of bulk damage produced.

* The orbit of the Relay I satellite, built by RCA for NASA and launched in December 1962, varied from 819 to 4612 statute miles at an inclination of 47.5 degrees to the equator. It was hardened by local shielding and selective replacement of components.

2. Speed and Power Consumption in Digital Circuits

In the design of any system, the designer must be concerned with all facets of the system requirements. To design a system that is hard to the radiation environment but fails to meet the other system requirements is an exercise in futility. In digital aerospace systems, in particular, the question of power consumption and switching speed are of paramount importance. This section will briefly compare the power consumption and switching speeds of both MOS and bipolar digital integrated circuits (IC's). Since most of the details of this comparison have already appeared in the literature,[10-12] we will only summarize the results.

A. MOS Integrated Circuits

Figure 1 is a schematic diagram of a complementary-pair MOS (CMOS) inverter circuit and shows a cross-section of a typical monolithic complementary pair of MOS transistors. Basically, the operation of the circuit (an n-p inverter) is as follows. When the input voltage is 0 volts, the p-transistor is biased "on" by $-V_0$ volts and the n-transistor is biased "off", so that the output voltage is the positive supply voltage ($+V_0$). When the input voltage is $+V_0$ volts, the p-transistor is biased off and the n-transistor is biased on so that the output voltage is 0 volts. In either state, there is a high impedance between the power supply and ground, and the total current flow is that of the "off" transistor, which is a few microamperes* at most. Appreciable current will flow through the device only during the switching transient, and this current will be used to charge and discharge the capacitance of the output node. Also, since the CMOS circuit, in either state, has an "off" (high-resistance) transistor between the power supply and ground, good logic-level definition is obtained when both the n-channel and the p-channel transistor have the same transconductance (g_m). Consequently, since g_m can be made reasonably large, good switching speeds can be obtained. The switching speed of commercially-available monolithic CMOS inverter circuits is between 30 and 40 nanoseconds, and for a fan-out of four, the switching speeds will be between 100 and 150 nanoseconds. For CMOS circuits fabricated using thin film silicon-on-sapphire technology, the input and output capacitances are greatly reduced, due to the elimination of an active substrate. Switching speeds between 5 and 10 nanoseconds for a fan-out factor of unity are attainable. For a fan-out factor of four, switching speeds between 20 and 40 nanoseconds are obtainable.

* Recently, a 9-bit CMOS memory circuit employing 90 silicon-on-sapphire (SOS) devices was fabricated.[11] It consumed a measured standby power of 90 microwatts, or 1 microwatt per device.

For certain applications, digital integrated circuits employing all p-channel MOS transistors (PMOS) offer very great advantages over bipolar transistors with respect to device packing density, simplicity of fabrication, and low cost.[12] Figure 2 shows a simple PMOS inverter

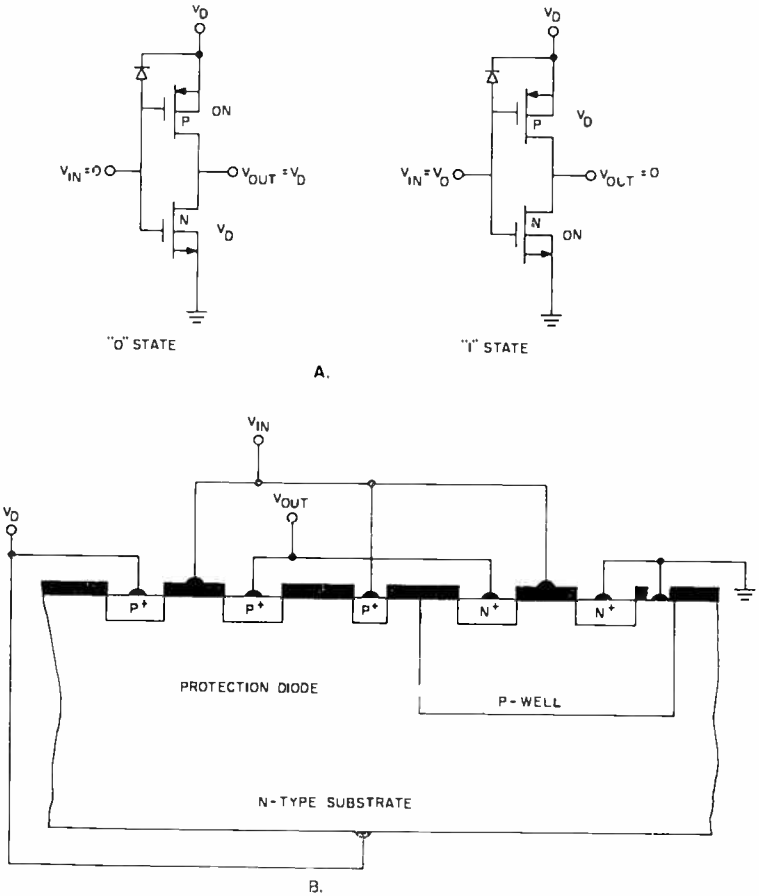


Fig. 1—(Top) Schematic diagram of CMOS inverter circuit and (bottom) cross-section of typical monolithic complementary pair of MOS transistors.

circuit. Here, the gate of one MOS device is connected to a power supply, and the "pull-up" device (Q_1) acts simply as a passive load and is always "on." Consequently, in one state of the inverter (i.e., with the input at negative voltage) there is a dc path between the power supply and ground, and the inverter will draw significant stand-by power. In PMOS circuits, unlike CMOS circuits, there is a dc path

between the power supply and ground, and it is therefore necessary to design circuits such that $g_m(Q_1) \ll g_m(Q_2)$ in order to obtain good definition in logic-level voltages. This leads to turn-off times significantly longer than turn-on times, which in turn, limits the overall switching speed of this type of PMOS circuit to about 0.5 microsecond.

Several other more complex switching schemes employing PMOS IC's are available. These schemes rely on the storage of charge on a capacitive element of the IC, and eliminate the dc path between the power supply and ground. However, since the storage capacitance does have a shunt leakage, it is necessary to recharge the capacitor quite

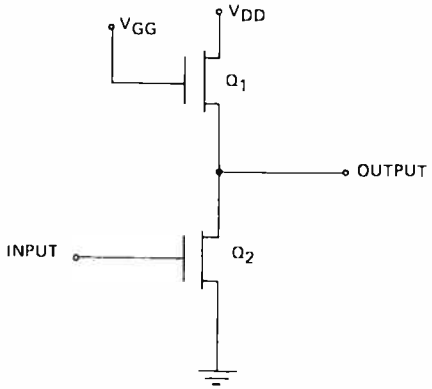


Fig. 2—Schematic diagram of basic PMOS inverter circuit.

frequently. Hence, these circuits are referred to as "dynamic" and PMOS circuits of the type shown in Figure 2 are referred to as "static". Since dynamic PMOS logic eliminates the dc path between the power supply and ground, the quiescent or stand-by power dissipation is greatly reduced and overall switching speeds are reduced. Using such dynamic schemes, switching speeds of the order of 50 nsec for a fan-out of unity are obtainable.

B. Bipolar Transistor Digital Integrated Circuits

Figure 3 shows a bipolar transistor integrated circuit that is a DTL 4-input NAND gate. Figure 4 is a schematic diagram of a typical 3-input T^2L NAND gate. Both of these types of bipolar integrated circuits fall in the class of saturated transistor logic and are used extensively in aerospace logic applications.

In Figure 3 (DTL-gate), it is easily seen that when the input voltage

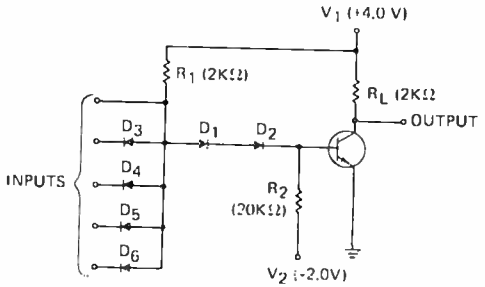


Fig. 3—Bipolar transistor integrated circuit (DTL 4-input NAND gate).

is at 0 volts, the transistor is "off" (both junctions are reverse-biased) and the output is typically at +4.0 volts. Current flows from the collector supply (V_1) to ground through resistor R_1 and the input diodes (D_3 through D_6 which are "low" or at ground). When the input voltage is raised to +4.0 volts, the transistor is biased "on", in a saturated condition (both junctions are now forward-biased) and the output falls to a potential near that of ground. The current flow from the supply is controlled by load resistor, R_L. In either state the current drain from the supply is approximately the same. For a typical DTL logic gate, the stand-by-power dissipated is thus orders of magnitude higher than that of a CMOS gate (i.e., the lowest-power DTL circuits typically dissipate 2 to 3 milliwatts per gate).

Similar considerations apply to the T²L gate shown in Figure 4. Typical stand-by power dissipation lies in the range of 5 to 20 milliwatts per gate (depending on circuit design parameters.)

The switching speeds for both DTL and T²L circuits are dominated

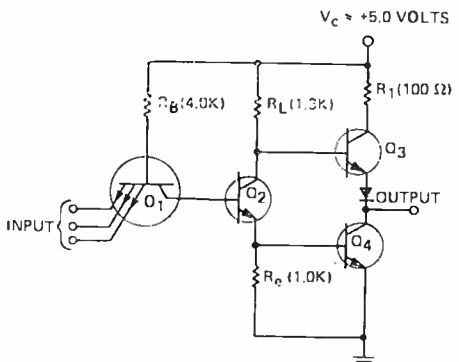


Fig. 4—Schematic diagram of a typical T²L logic gate.

by the delay associated with the storage of minority carriers in the base and collector regions that occurs when the transistor is driven into saturation. T^2L circuits are slightly faster due to transistor action at the input and the low circuit-output impedance. Typically, maximum attainable switching speeds for DTL circuits are of the order of 20 nanoseconds and T^2L circuits of the order of 10 nsec. These speeds are approximately independent of fan-out for moderate fan-out factors.

In emitter-coupled or current-mode logic (ECL), the switching transistor is always operated in the active region of the transistor characteristic (i.e., the circuit conditions are adjusted so that the collector-base junction never becomes forward-biased). Consequently, the storage of charge in the collector junctions, which occurs in saturated transistor logic, does not enter into the speed term, and switching times of the order of 2 to 4 nsec per gate are thus obtainable. The minimum stand-by power dissipation of these circuits is again, approximately 15 mW/gate.

C. *Summary of Advantages of MOS Devices in Aerospace Logic*

Table I lists the stand-by power dissipation per gate for the various forms of digital IC's. As can be seen from this table, the stand-by power consumption of a CMOS circuit is lower than that of the best bipolar circuit (i.e., DTL) by three orders of magnitude. Most types of PMOS circuit also consume considerably less stand-by power than bipolar circuits. It might be suggested that this table presents a somewhat unfair comparison, since the transient power consumption of CMOS circuits increases at high repetition rates, negating the low stand-by power advantage of the CMOS circuit.* However, in most large digital systems, only a small fraction of all circuits switch during each cycle, so that the total power consumption of the CMOS system will still be considerably less than that of a bipolar system.

A bar chart summarizing the switching speeds of the various forms of digital logic is shown in Figure 5. As can be seen, the types of monolithic, single-crystal CMOS circuits that are already widely available compete with saturated bipolar logic over a significant range of switching speeds. The most advanced silicon-on-sapphire CMOS arrays

* Burns[10] has shown that for a square-wave switching signal of frequency f_o, the dynamic power consumption of a CMOS gate is given by $P = C_o V_o^2 f_o$, where C_o is the output node capacitance of the gate and V_o is the supply voltage. Therefore, for $C_o = 10\ pf$, $V_o = 10\ V$, and $f_o = 10$ MHz, $P = 10$mw/gate.

Table I—Comparison of Standby Power Consumption of the Various Forms of Digital Logic (Typical Biasing Conditions)

Digital Logic Type	Standby Power Consumption/Gate
CMOS	1 μW
PMOS (Static)	0.5 mW
PMOS (Dynamic)	50 μW
DTL	4 mW
T^2L	15 mW
ECL	15 mW

compete with saturated bipolar logic over almost the entire range of speeds required. However, if very high switching speeds (2 to 3 nsec) are required, only nonsaturated bipolar-transistor logic (e.g., ECL logic) can be used.

In summary then, CMOS integrated circuits will, in the future, be able to perform over approximately the same speed range as saturated bipolar transistor logic. Moreover, a digital system employing CMOS integrated circuits can function with considerably less power consumption than can an equivalent system employing bipolar transistor logic. This reduced power consumption is an extremely attractive feature for aerospace applications. In addition, in many aerospace applications, speed is not a critical requirement, but low power consumption is.

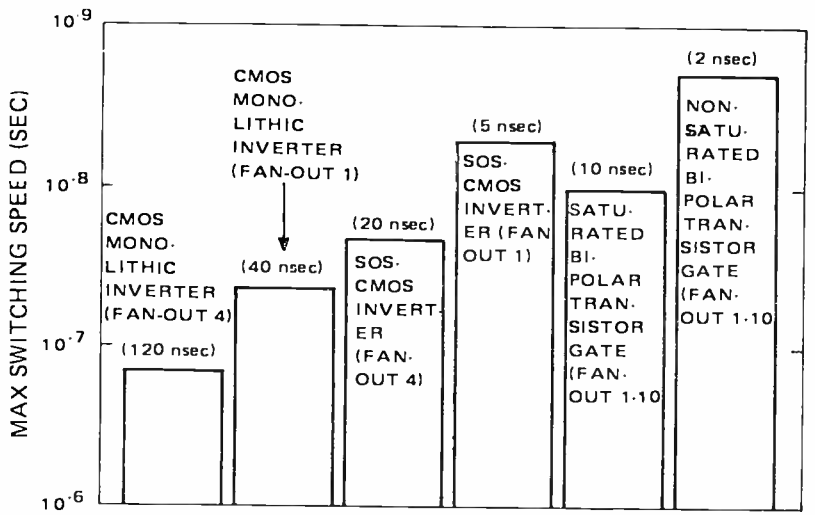


Fig. 5—Comparison of the maximum switching speeks of MOS, bipolar, and other micrologic.

3. Permanent Degradation of MOS and Bipolar Devices in a Radiation Environment

In the paragraphs that follow, the permanent degradation of device characteristics for MOS and bipolar integrated circuits in a nuclear radiation environment will be discussed. As in the previous section, a detailed comparison will be drawn between MOS inverter circuits and a DTL gate. The cause of circuit failure for these two types of digital circuits is quite different. For the MOS circuits, failure will be caused by threshold voltage shifts that increase smoothly in proportion to the ionizing dose received from the gamma-ray component of the radiation:

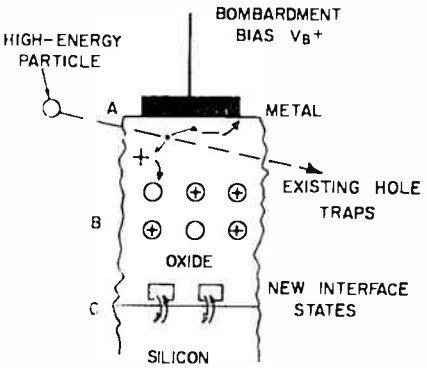


Fig. 6—Charge-trapping mechanism in MOS transistors.

for the bipolar DTL gate, failure will be caused by degradation of gain and increases in saturation voltage caused by bulk displacements, produced by the neutron component. Consequently, the failure mechanisms as well as the available hardening techniques will be discussed before a comparison is made.

A. Permanent Radiation Effects in MOS Devices

The primary mechanism by which MOS transistors degrade in a radiation environment is the build-up of strongly trapped positive charges in the gate insulator. The nature of this trapping is indicated in Figure 6. The authors have jointly studied this phenomenon since 1964 and an extensive review of its significance can be found in Reference (4). The result of this charge-trapping is a shift, ΔV_T, in the threshold voltage, V_T, of both n- and p-channel MOS transistors toward more negative voltages. To a first approximation, this charge buildup is directly related to the number of ion pairs created (i.e., the energy deposited) in the oxide and, hence, depends on the total dose received.

The presently accepted model for the radiation-induced charge buildup process in MOS structures is as follows. The incident radiation creates electron–hole pairs in the silicon dioxide. The holes are relatively immobile and get trapped almost immediately; the electrons, however, are mobile, and will drift under the action of an applied or built-in electric field until they either recombine with a trapped hole or escape from the oxide, leaving a trapped hole behind. The larger the electric field in the oxide, the more electrons escape. In this process, illustrated in Figure 6, the trapped positive charge builds up primarily at the silicon–silicon-dioxide interface. This positive charge induces a negative image charge in both the silicon and the metal electrode. The negative image charge in the silicon implies that the threshold voltage, V_T, of an MOS transistor shifts toward more negative voltages. Since a negative voltage large enough to counteract this negative charge must be applied to the gate before an n-channel device will turn off or a p-channel device will turn on, there is a shift in the threshold voltage. From this model, it can be seen that the important parameter that will affect this charge-buildup process is the net density of traps in the oxide, N_t.

The magnitude of the effect in a variety of devices and the increase of the effect with dose is shown in Figure 7. While this figure is for p-channel devices, the situation is very similar for n-channel devices, under the same bias conditions. It can readily be appreciated that the effect of a shift such as 5 volts in the turn-on voltage of the component MOS devices could be catastrophic for many common MOS circuits; for the worst type of device shown in Figure 7, this value of ΔV_T can be attained at dose levels as low as 1.2×10^4 rads. For some MOS circuit designs, which, for reasons of electrical performance, have been designed with little tolerance for shift in threshold voltage or increased power drain, a ΔV_T value as small as 2 volts could be sufficient to degrade logic performance seriously. Thus, for aerospace logic circuitry, this problem of threshold shift is obviously of importance, although, as described in References (13) and (14), design engineering approaches are already available for moderating the effect in currently available devices and accommodating circuit designs to minimize the effect. Nevertheless, in the long term the best approach is the development of insulators with reduced tendency to charge buildup. The curve marked Al_2O_3 indicates the authors' findings that, on the laboratory scale, radical improvements can, indeed, be made in this way.

It should be noted here that the effects of ionization-induced shift in threshold voltage and interface-state creation are the strongest permanent effects observed in MIS devices. However, MOS transistors

would also undergo some shift in threshold voltage and loss of transconductance (g_m) in a pure neutron environment, due to the introduction of carrier-removal sites alone. In most real-life radiation environments, and MIS devices, these carrier-removal effects will be secondary compared to the effect of the ionizing radiation components. The problem of hardening MOS transistors to the effects of ionizing radiation

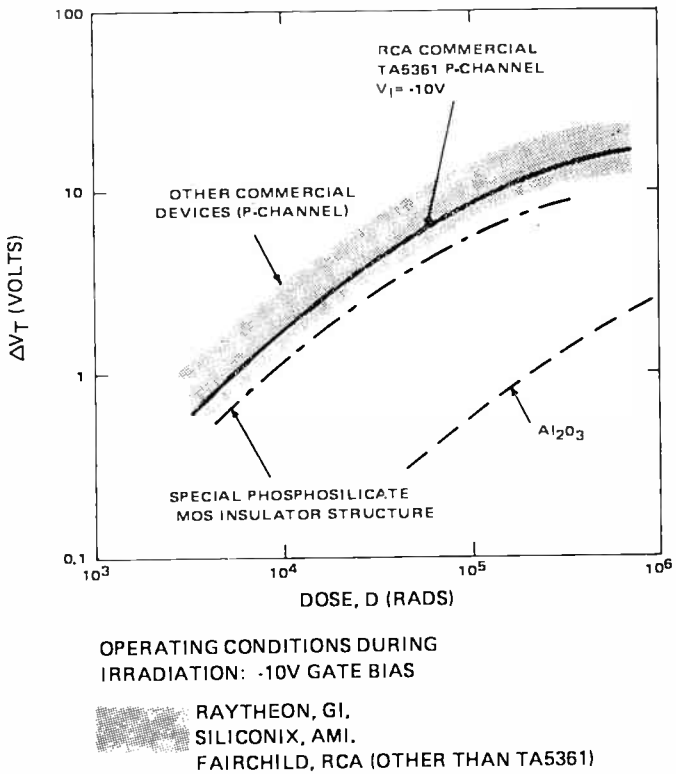


Fig. 7—Typical radiation-induced threshold voltage shifts.

is mainly one of developing an insulator layer that does not show appreciable net trapping-charge buildup and survives the radiation without the generation of interface instabilities.

B. Effects of Threshold Shift on Logic Operation

If the radiation-induced threshold shift, ΔV_T, is large enough, an MOS inverter can lose its inverting capability. This is illustrated in Figure 8, where the inverter input–output characteristic of a CMOS

pair is plotted for an operating potential of 10 volts, with radiation dose, D, as a parameter. From this figure it can be seen that, initially, the CMOS inverter changes state at an input voltage that is approximately half the operating voltage (e.g., 5.0 volts). Figure 8 further shows that the circuit can suffer a voltage shift of at least 4.0 volts without impairing logic operation. Thus, in analyzing the problem, a shift of 40% of the operating voltage has been taken as a conservative criterion for the onset of malfunction in a series of CMOS logic gates.

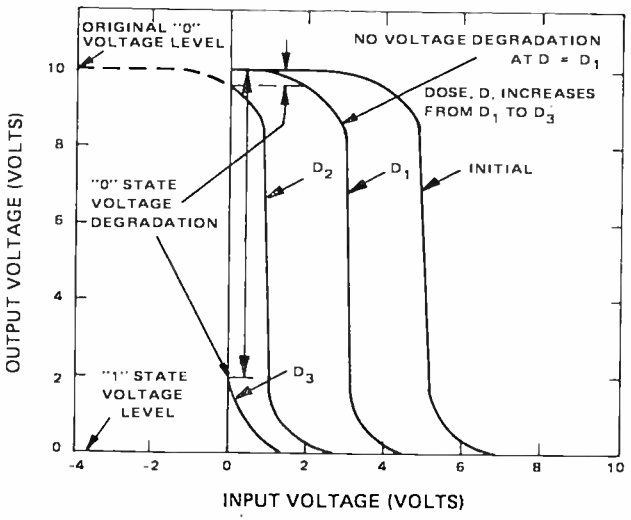


Fig. 8—Permanent radiation-induced shifts in the transfer characteristics of a complementary-symmetry MOS inverter.

In the case of a PMOS inverter, the "pull-down" device will require a larger negative voltage to bias it "on" after irradiation. A point may come at which the logic "1" voltage supplied (say −10 volts) will not open a sufficiently low-impedance electrical path to ground. While increasing the logic input voltages would restore operation, this may not be a practical countermeasure in a spacecraft.

Table II gives a set of worst-case logic malfunction dose levels that would apply for CMOS logic devices fabricated from several different types of gate insulators to be described later. Also shown are dose levels that produce malfunction in some MOS circuits that are more sensitive than CMOS devices are. The significance of these figures can be appreciated better if it is realized that a tenfold increase in the dose leading to malfunction implies a tenfold increase in life span for a subsystem operating in a given orbit in space; a 3.3-times reduc-

tion in the range to which a system can approach a point source of nuclear radiation (power reactor, nuclear burst, etc.) and still survive; or a potential reduction in shielding of a large sub-system amounting to many tens of pounds. Clearly, such magnitudes of improvement could have radical effects on the designs adopted for nuclear-powered satellites, satellites orbiting near the Van Allen belts, and missiles required to survive nuclear countermeasures.

Table II—Worst-Case Values for the Doses Producing Logic Malfunction in CMOS Inverters Fabricated from Several Forms of Gate Insulator

ΔV_T* (volts)	Dose D required to produce ΔV_T values shown		
	Typical Thermally-Grown SiO_2	Composite Layers $P_2O_5(SiO_2) - SiO_2$	Plasma-Anodized Al_2O_3
−4.0‡	4×10^4 rads	1.3×10^5 rads	2×10^6 rads
−2.0†	10^4 rads	4×10^4 rads	6×10^5 rads

* The data are for an operating potential, V_{DD}, of 10 volts and hence include the effect of a 10 volt gate bias, applied by way of the preceding MOS inverter.

‡ This condition could constitute onset of logic malfunction in typical commercial CMOS inverters. Note that some increase in the quiescent power drain in these devices, due to incomplete turn-off of the n-channel device in the "0" state, would be observed at lower dose levels than shown here.

† This condition could produce a malfunction in some very shift-sensitive MOS logic circuits.

C. *Hardening of Field-Effect Devices for Permanent Damage Effects*

The following approaches have been used in studying the improvement of radiation resistance of insulated-gate field-effect transistors:

(a) Replacing the SiO_2 gate-insulator layer with silicon nitride, Si_3N_4.

(b) The use of composite or "sandwich layer" insulator films.[3,15]

(c) Doping of the SiO_2 films.[22]

(d) Replacing the SiO_2 insulator layer with Al_2O_3.[8]

1. *Silicon Nitride as Gate Insulator*

The first material, pyrolytically deposited silicon nitride, originally created interest because it showed promise of radiation resistance.[6] However, if the nitride film extends to the silicon interface, severe interface instability problems usually arise. Unless these problems can be eliminated, such devices will not be useful. However, a newer class

of device, containing composite $Si_3N_4 - SiO_2$ layers, is now appearing commercially. Preliminary results indicate that these devices have much in common with the composite-layer group described below.[15]

2. Composite Layers as Gate Insulators

The second approach, the use of composite layers of pyrolytically formed phosphosilicate glasses or other insulator layers, deposited over a layer of thermally-grown SiO_2, was originally adopted as a means of reducing the temperature-bias instability in MOS transistors due to the drifting of sodium ions in the silicon dioxide.[16] However, the authors[3] have observed that gate insulators formed by the pyrolytic decomposition of organophosphorus, organosilicon compounds, and oxygen [hereafter described as $(P_2O_5)SiO_2$] on a thermally grown SiO_2 surface are less sensitive to ionizing radiation by about a factor of three than strictly comparable thermally grown silicon dioxide. At present, it is thought that this reduction in sensitivity is due to compensatory electron trapping at the insulator–insulator interface. In effect, a reduction in the effective density of trapped charge has been accomplished. The composite MIS structure, unlike the continuous nitride layer, can yield devices with good insulator–silicon interface characteristics. MOS circuits using these materials are now routinely fabricated by many manufacturers. The improvements observed by the authors for phosphorus-doped layers are indicated in Table II and Figure 7.

3. Doping of SiO_2 Films

It was found[22] that by doping of the SiO_2 films with aluminum or phosphorus, the radiation resistance of MIS capacitors could be increased. However, the degree of hardening that could be achieved was not comparable to that obtainable by using a different insulator material.

4. Aluminum Oxide as Gate Insulator

The third material, Al_2O_3, has shown the highest degree of hardness to ionizing radiation yet achieved. Compare with grown silicon dioxide films, Al_2O_3 films can be a factor of 10 to 30 times less sensitive to ionizing radiation. It is thought that this is due to a different trap structure of the Al_2O_3 films. In addition, these films have shown threshold voltage shifts in the unusual positive direction when irradiated under negative bias. This indicates that, under this oxide field condition, electrons, rather than holes, are being trapped, and immediately suggests that a reversal of the conventional charge buildup

might be accomplished by suitable biasing during irradiation. Table II indicates the degree of radiation tolerance that the use of Al_2O_3 as gate insulator would confer on a CMOS circuit. Because of the importance of this insulator, some of the details of two different fabrication techniques, as well as irradiation testing of laboratory Al_2O_3-MOS devices, are included in Appendix I.

5. Junction Field-Effect Transistors

It is well known that Junction Field Effect Transistors (JFET) are extremely insensitive to radiation.[17] Because of the complexity of fabrication of these devices, however, there has been no widespread effort to integrate JFET devices into large-scale logic arrays. In addition, the JFET has disadvantages similar to those of the bipolar transistor as regards transient radiation effects, namely large junction size. However, there are undoubtedly many circuit functions in aerospace systems where the judicious use of a JFET device in place of a bipolar transistor would impart higher overall hardness to the system.

D. Damage to Bipolar Transistors

In digital integrated circuits employing bipolar transistors, the primary cause of permanent degradation of logic operation is the loss of transistor gain. Gain degradation is caused both by ionizing radiation (gamma rays), which produces a surface effect analogous to the effect in MOS transistors, and by heavy-particle irradiation (neutrons), which produces defects in the silicon that act as recombination centers in the base and emitter–base depletion regions. The impact of both the ionizing radiation effects[18, 19] and the neutron effects[20] in discrete bipolar transistors has been studied, and rules have been devised for predicting these effects in small-signal-device applications for both types of radiation. These studies indicate that the surface effects saturate, so that, for digital applications and for the neutron gamma ratios of interest, the neutron effects will be the actual cause of failure.

Messenger[21] has studied the MC-201 gate under neutron bombardment and has shown that the gain degradation is governed by the standard equation,

$$\frac{1}{\beta} - \frac{1}{\beta_o} \approx \frac{0.2\Phi}{f_T K}, \tag{2}$$

where

β is current gain after bombardment,

β_o is current gain before bombardment (typically about 60 for the MC-201 transistor),

f_T is transistor gain bandwidth product in Hz,

Φ is neutron fluence in cm^{-2}, and

K is a constant that depends on base resistivity and the energy of the bombarding particles.

The above equation predicts that the gain of the MC-201 will degrade to about 2 for a neutron fluence of 10^{15} cm^{-2}.

Defining a failure level for bipolar logic is more difficult than for the case of CMOS circuits. This is because, as the transistor gain

Table III—Neutron Failure Fluence of DTL Logic Gate for Various Fan-Out Factors

Fan-Out Factor	Neutron Failure Fluence (cm^{-2})	β
10	1.1×10^{14}	~21
5	2.7×10^{14}	~11
4	3.3×10^{14}	~9
3	4.6×10^{14}	~7.5
2	5.5×10^{14}	~4.8
1	10^{15}	~2.1

Data derived from Reference (21).

decreases, the maximum fan-out capability of the gate decreases. This effect does not exist in CMOS circuits. Hence, in defining a failure level for bipolar logic, a fan-out factor must also be specified. This will be dictated by the particular digital system in which the circuit is required to operate. Table III is a list of the neutron fluences producing failure of the MC-201 gate for various fan-out factors. Taking a fan-out factor of 4 as being representative, the neutron failure fluence is 3.3×10^{14} cm^{-2}.

E. Hardening of Bipolar Devices for Permanent Damage Effects

The problem of hardening against the permanent effects of radiation in bipolar transistors is considerably different from that of MOS transistors. In bipolar transistors, hardening against permanent (neutron) damage is achieved primarily by keeping base widths small (i.e., keeping the value of f_T as large as possible). Hence, hardening is essentially a problem of controlling diffusions very closely. It is not likely that further radical improvements will be possible in this field.

F. *Summary of Comparison of MOS and Bipolar Logic Circuits*

To make a comparison of the steady-state radiation failure mechanisms of both the CMOS and bipolar digital circuits, a meaningful neutron-to-gamma ratio must be specified. An n/γ ratio of $2 \times 10^8 \text{cm}^{-2}\text{rad}^{-1}$ has been chosen as representative. Using this ratio and the foregoing considerations, Figure 9 has been constructed, giving the expected logic failure levels of CMOS and bipolar logic circuits in a typical mixed neutron–gamma environment. The important point to note in this figure is that the technological breakthrough represented by Al_2O_3 can result in a CMOS circuit that will function to higher radiation levels than bipolar-DTL logic, working with a fan-out of four. Since the failure level of CMOS circuits is independent of fan-out, systems that require large fan-out capabilities and employ CMOS circuits will be far superior to bipolar transistor logic under irradiation.

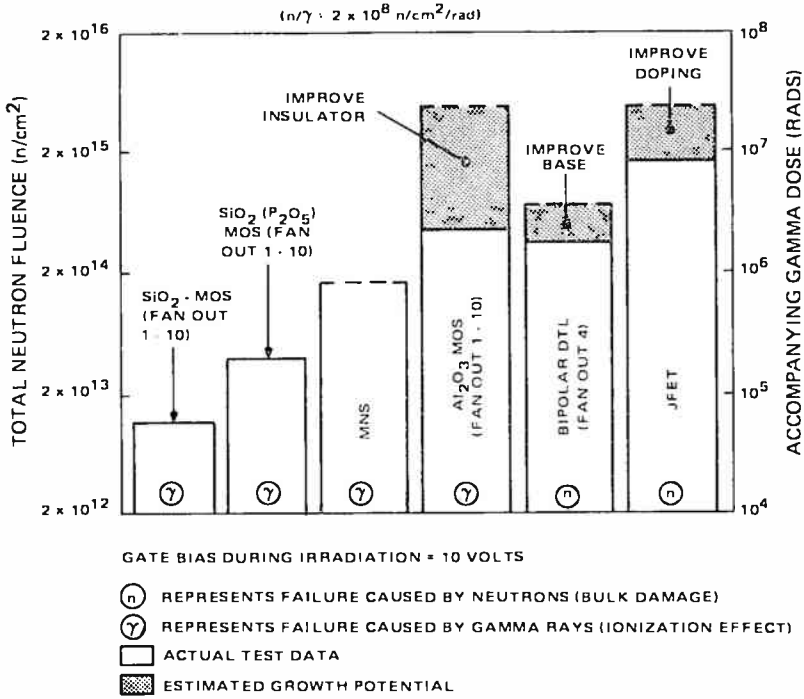


Fig. 9—Comparison of expected failure levels of various forms of digital logic due to permanent damage by neutrons and gamma rays combined.

4. TREE EFFECTS IN DIGITAL CIRCUITS

Many aerospace digital logic systems will be required to survive the effects of a nuclear burst in space within several hundred miles. Military vehicles may be required to operate during the burst without any loss of control or of stored data. Even a temporary lapse may be unacceptable. Even civilian satellites, as a matter of economy, must be designed so that, if they are innocent "bystanders" of a nuclear burst in space, they are not unintentionally switched off or irreparably degraded by the radiation from the burst. A nuclear burst in space has associated with it a sub-microsecond burst of high-intensity, high-energy ionizing radiation in the form of gamma rays. When this ionizing radiation is incident on a semiconductor circuit, photocurrent will flow in the circuit. This may result in two types of damage to the circuit. Firstly, if the photocurrents are sufficiently large, physical damage (due to heating) may occur with the resultant permanent failure of the system. Secondly, for digital systems, photocurrents can cause some digital gates to change state in a spurious manner and cause information to be lost or spurious commands to be given in control circuits. Thus, even though the system is still operable after the burst, some critical piece of information may have been lost, or a command carried out that could have catastrophic results for the system. It is this second type of failure of a system due to "TREE malfunction," rather than a permanent device degradation, with which we are concerned in this section.

A. Definition of TREE Failure Levels

Before discussion the TREE failure levels of digital circuits, we must first define what we mean by failure of a digital circuit from a burst of ionizing radiation. In a digital system, these circuits will be interconnected in such a manner that the output of one inverter stage drives the inputs of one or more following inverter stages. Hence, a failure criterion should be based on what a transient output pulse does to the stages immediately following it, i.e., on the noise immunity of the following stage. If the radiation burst causes the stages following it to change state, this can constitute a failure of the digital system. The noise immunity of a digital circuit can be determined from the voltage transfer characteristic. Figure 10 shows the voltage-transfer characteristics, for typical operating conditions, of a CMOS inverter and of the bipolar-transistor DTL inverter previously described. As can be seen from the figure, the noise immunity of CMOS inverter is approximately the same in either the "on" or the "off" state and is

slightly less than half of the drain supply voltage, which is typically 10 volts. On the other hand, the DTL gate is more susceptable to a noise pulse in the "off" state, the noise immunity being between 1.0 and 1.3 volts. For a CMOS inverter then, a radiation-induced transient output pulse of 4.0 volts will probably *not* change the state of the next inverter stage that it is driving. On the other hand, for the DTL gate,

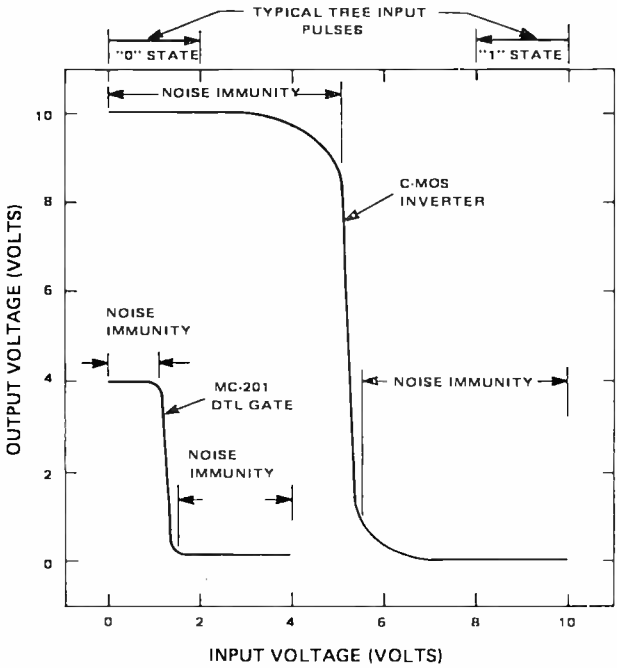


Fig. 10—Comparison of noise imunity of an CMOS inverter and bipolar-transistor DTL inverter.

a radiation-induced transient output pulse of 1.3 volts has a 50% chance of changing the state of the next inverter stage. This basically high immunity of the CMOS circuit to noise gives such circuits a three-fold advantage over bipolar gates with regard to TREE hardness. Of course, if the CMOS inverter characteristic is shifted toward zero volts as the result of steady-state ionizing radiation, (see Section III), then the noise immunity in the "off" state of the inverter is decreased. This is one of the several instances where TREE effects and permanent radiation damage effects interact.

B. TREE Effects in MOS Integrated Circuits

In discussing the magnitudes of TREE responses in MOS and

bipolar-transistor IC's, and the mechanism of their generation, we will start by considering the transient response of discrete MOS transistors to a burst of ionizing radiation. Several accounts of the transient response of discrete MOS transistors have appeared in the literature.[22, 25] These accounts indicate that the important transient responses of discrete MOS transistors occur in the drain circuit. This response is the result of three separate mechanisms:

(1) Transient photocurrent of the normally reverse-biased drain-junctions.

(2) Modulation of the conductivity of the channel of a transistor that is "on".

(3) Charge-emission and gaseous-ion currents influence the drain current by gain action in the MOS transistor. Effectively, electrons scattered out of the gate electrode leave a net positive charge on the gate electrode. This positive charge can turn an n-channel device on and a p-channel device off.

The relative magnitudes of the various contributions are different for different device types. Two of the authors[22] working with semi-integrated MOS transistors as well as CMOS inverters[26] have shown that gate-replacement-current effects are essentially negligible. In addition, the results indicate that the transient drain photocurrent is due primarily to the transient response of drain–substrate junction. The conductivity modulation of the channel is only a secondary effect. In addition, since channel-conductivity-modulation effects can only occur when the device is on, and since this effect tends to lower the on impedance of the device, it can be considered a beneficial effect in digital circuits.

The variation of the transient photocurrent with device type, operating conditions, and dose rate has also been studied by the authors.[22] The results show that the transient drain current increases with increasing drain bias, dose rate, and drain-substrate junction area and that the response tends to be larger for n-channel transistors than for p-channel transistors. With a knowledge of these effects in MOS transistors, we can begin to analyze TREE effects in MOS integrated circuits.

Figure 11 shows a simple schematic model that is helpful in analyzing the transient response of the CMOS circuit. In this model, the photocurrents generated in the drain diodes of both the n-channel and the p-channel transistors, ΔI_1 and ΔI_2 respectively, will flow through the inverter circuit, and it is these photocurrents that will be mainly responsible for the transient output voltage that is observed when the

integrated circuit is exposed to a burst of ionizing radiation. Note that the photocurrent of the n-channel transistor, ΔI_4, flows out of the output node, and the photocurrent of the p-channel transistor, ΔI_2, flows into the output node so that *diode compensation* is inherent in the* CMOS *structure*. This is a very attractive basic feature of CMOS circuits.

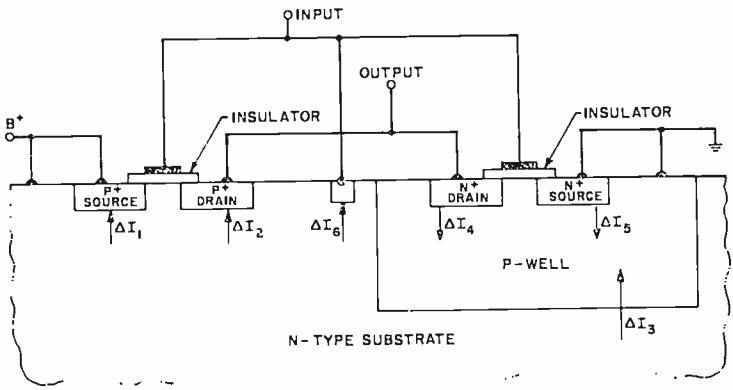


Fig. 11—Model for analyzing transient response of CMOS circuit.

Neglecting capacitive loading at the output, we can very simply estimate the peak transient voltage to be expected at the inverter output. It is equal to the current that flows through the resistance of the "on" transistor so as to balance the charge that is being driven out of the output node by the photocurrents of the drain diodes.

$$\Delta V_{out} = R_{on} (\Delta I_4 - \Delta I_2). \tag{3}$$

Consequently, the transient voltage observed at the inverter output will be positive or negative according to whether $\Delta I_4 > \Delta I_2$ or $\Delta I_4 < \Delta I_2$, respectively. The values of ΔI can be calculated from a knowledge of the drain junction parameters, namely area, depth, and the minority-carrier lifetime on the high-resistivity side of the junction.

Equation (3) has been found to describe adequately the results of experimental test on CMOS integrated inverter arrays up to a dose

* *Diode compensation* is a technique being developed to reduce the very high sensitivity of bipolar logic circuitry to transient photocurrents. An extra diode is placed in the circuit with the sole object of introducing current flows into output nodes that oppose those generated by the transistor junctions of the circuit. Some success has been achieved, but production devices are not yet available.

rate of about 8×10^8 rads(Si)/sec.[26] For a typical CMOS inverter in the "1" state (i.e., input high) operating with a 10.0-volt drain supply (V_D), a TREE output voltage pulse of about +0.4 volts is observed at this dose rate. The voltage varies approximately linearly with dose rate. In the "0" state, (i.e., input low) the TREE voltage pulse is very small indeed, indicating good diode compensation. If we extrapolate the "1" state response to higher dose rates, we would predict a TREE failure level of approximately 7×10^9 rads(Si)/sec. However, at a threshold in the region of 8×10^8 rads(Si)/sec, a new response mechanism comes into play. Output voltages become much larger, and other

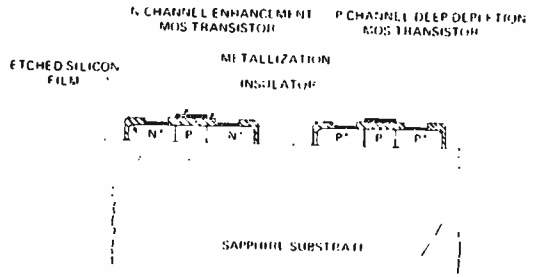


Fig. 12—Typical CMOS silicon-on sapphire circuit construction.

temporary disturbances occur in the circuit that would be sufficient to make a CMOS memory cell lose stored information. No permanent damage occurs, however, such as the "latchup" phenomenon, an equivalent parasitic response that sometimes occurs in diode-isolated bipolar integrated circuits.[27] This parasitic response is not fully understood, but the results indicate that its occurence is related to the presence of the p-well isolation diode and/or the gate-protection diode. The results also suggest that it may be eliminated by appropriate changes in diffusion profiles and by eliminating the p-well isolation diode, using techniques such as dielectric isolation or a thin-film approach such as silicon-on-sapphire.

The thin-film silicon-on-sapphire (SOS) technique offers very attractive features for transient radiation hardening. Figure 12 shows a cross section of an SOS CMOS inverted presently being fabricated.[11] As seen from the figure, there are no isolation junctions. The insulating sapphire isolates the transistor elements. In addition, source and drain junctions are diffused completely through to the isolating substrate, so that the effective volume of depletion region that supplies carriers

contributing to the photocurrent response of the drain junction is greatly reduced. Note also that the p-channel transistor is a deep depletion transistor.[28] The use of this technique reduces the number of diffusions required to fabricate a CMOS array.

The expression for the TREE output voltage pulse of the CMOS inverter, given earlier (Equation (3)), suggests certain measures that could be taken to minimize these voltages and hence harden the device. Since the response is dominated by the photocurrent response of the drain junctions, hardening becomes a problem of keeping these photocurrents as small as possible, and matching the photocurrents flowing in the n-channel and the p-channel transistors as closely as possible (i.e., taking advantage of the inherent diode compensation). The junction photocurrents may be kept small by keeping junction areas small and keeping the minority-carrier lifetime in the high-resistivity substrate as small as possible. Minority-carrier lifetime can probably be made very low without seriously impairing device operation. In addition, bounded collection-volume techniques such as dielectric isolation, or thin-film silicon on sapphire or spinel (SOS), reduce the photocurrents by reducing the photocurrent current collection volume. These techniques, especially the SOS, are the most effective techniques to use with MOS devices and, if they could be made readily and repeatably, should produce the ultimate in radiation hardening.

Equation (3) also suggests that, in addition to minimizing photocurrents, minimizing the "on" resistance of the transistors acting as inverter loads should also be effective in reducing the transient output voltage. This could be accomplished primarily by changing the geometry of the transistors in question. In their first studies of transient effects on MOS arrays, the authors devoted a majority of their attention to the CMOS device. This was because of the clear basic advantages of this configuration for withstanding TREE effects. However, the PMOS configuration must still be of high interest. The prospects are that PMOS arrays, while ranking below CMOS, can be made considerably more resistant to TREE than competing bipolar arrays and, of course, have some of the advantages for use in space which were mentioned in Section 2. Moreover, at present, PMOS techniques are better established than CMOS, and a large variety of PMOS logic designs are available on the market. Of special note are the "dynamic" memories, which can achieve very high densities of bit storage per cubic centimeter; moreover, PMOS processing will probably always be less complex and costly than CMOS. Thus, some close comparative studies of PMOS arrays are being carried out by two of the authors (W. Dennehy and A. Holmes-Siedle). The preliminary results of experiments con-

ducted to date confirm that, as would be expected, "static" PMOS schemes will be harder to TREE effects than "dynamic" or capacitance-storage schemes.

The ultimate TREE logic failure levels of PMOS circuits have not yet been finally ascertained, but it appears that appropriate chip design should place the failure level in excess of 10^9 rads(Si)/sec.

The greatest advantage of PMOS circuits is in the small device geometries that can be achieved by virtue of the simple processing involved. This feature is particularly attractive for TREE effects since the junction sizes at all critical nodes can be kept extremely small. In addition, since these circuits use a single conductivity type of transistor, isolation junctions, which produce a large photocurrent, are not required. Also, because of the absence of isolation junctions, we can be reasonably sure that parasitic effects will not occur. On the negative side, PMOS circuits have less noise immunity than CMOS circuits and do not have diode compensation inherent in these structures.

C. *TREE Effects in Bipolar Integrated Circuits*

The transient radiation response of discrete bipolar transistors is reasonably well understood.[29] Theory and experimental results have shown that the transient radiation response of bipolar transistors is dominated by the response of the reverse-biased collector–base junction. In the active region of bipolar-transistor operation (collector–base junction reverse biased and emitter–base junction forward biased), two elements of current can be identified. For an n-p-n transistor, holes created in the collector region diffuse into the base region. This charge flow is the conventional photovoltaic effect and the resultant current is termed the primary photocurrent, I_{pp}. However, the resultant temporary charge imbalance has a secondary effect; the base potential is raised and the device conducts more heavily. The magnitude of this secondary photocurrent will depend on the rate at which the charge stored in the base can escape through the external base lead, and hence its magnitude will depend on the base resistance. The maximum magnitude of this secondary photocurrent is the produce of the current gain (β) of the device and the primary photocurrent and occurs for large base resistances. Hence, the transient collector photocurrent (ΔI_c) lies in the range

$$I_{pp} \leqq \Delta I_c \leqq I_{pp}\,(1+\beta).$$

It should be noted that a transistor biased in the cutoff region prior to a transient radiation pulse may be turned on by the peak primary

photocurrent flowing through the base resistance. If the transistor is biased on by the radiation pulse, then the secondary photocurrent mechanism comes into play. As discussed below, this is very often the cause of failure in bipolar digital microcircuits.

In an integrated circuit made from bipolar transistors, the various circuit elements (resistors, transistors, etc.) in the silicon chip must be electrically isolated from each other. The usual technique for accomplishing this is through the use of p-n isolation junctions. These junctions are the ones with the largest area on the chip. Thus, when a burst of ionizing radiation is incident on the structure, they will produce the largest photocurrents. The electrical arrangement of these isolation junctions is usually such that the photocurrents generated from them will flow through the load resistors and cause the logic failure of the circuit before the primary and secondary photocurrent failure mechanisms discussed above.

The only way of eliminating the isolation junction response is to use one of the several dielectric isolation techniques. Production versions of dielectrically isolated digital integrated circuits are presently becoming available* and, as is discussed below, are proving to be considerably harder to TREE effects then their diode-isolated counterparts. The remaining effort to harden against transient radiation effects in bipolar digital integrated circuits, therefore, must be directed toward reducing I_{pp} and, where possible, eliminating secondary photocurrent effects. The peak primary photocurrent of the collector–base junction is determined primarily by area, doping levels, and minority-carrier lifetimes in both the collector and base regions. Since little change can be made in doping levels and minority-carrier lifetime in the base region without degrading the device characteristics, the primary means of achieving hardening against primary photocurrent must be by area reduction and by minimizing minority-carrier lifetimes in the collector region. Hardening against secondary photocurrents is accomplished by keeping base resistances as small as possible, consistent with the operating characteristics of the circuit.

From the work of Messenger,[21] Raymond et al,[30] and Bowman,[31] the failure levels of a typical diode-isolated DTL logic inverter circuit (type MC-201) can be estimated. Using Bowman's data for the transient output voltage of this device when subjected to 200-nsec-wide pulses of electrons from a LINAC,* the failure level can be estimated as approximately 10^8 rads(Si)/sec. Raymond, using 20-nsec-wide pulses

* These circuits cost a premium price, typically six times that of the equivalent diode-isolated circuit.

* Linear electron accelerator.

from a flash x-ray machine, finds somewhat higher failure levels. The measurements on CMOS inverters, which were discussed in the previous section, were performed using 100-nsec-wide pulses of electrons from a LINAC, and hence it is felt that Bowman's data provides a closer comparison with the CMOS data. The failure of the MC-201 gate is dominated by the transient photocurrents of the isolation junctions

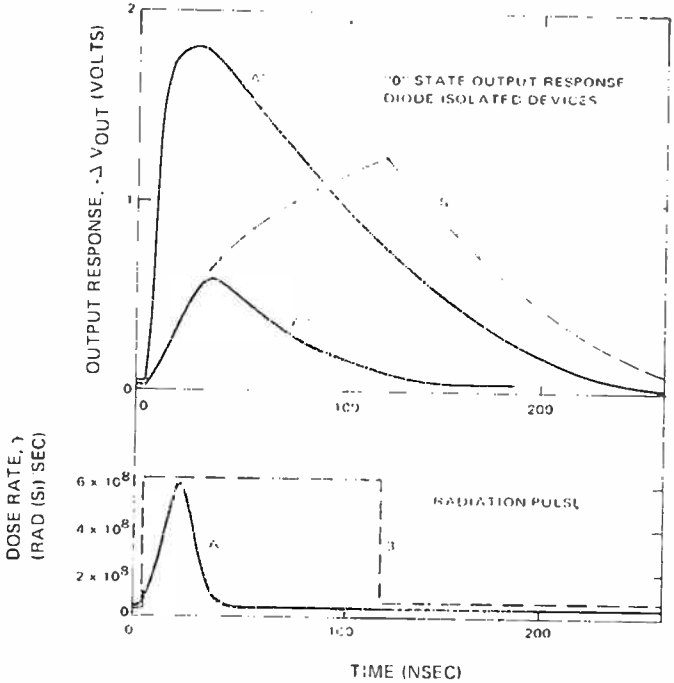


Fig. 13—Comparison of transient output responses of bipolar and CMOS integrated circuits under pulses of ionizing radiation: Curve A′ is experimental data by Chang and Raymond on the response of bipolar IC when exposed to pulse of form A; curve A″ is predicted response of RCA type CD4007 gate under an identical pulse; curve B′ is experimental data of the authors from which curve A″ was derived (B′ was obtained using a longer radiation pulse of form B).

and by the secondary photocurrent mechnism discussed in the beginning of this section. Indeed, the type XC-201, which has basically the same circuit and geometry as the MC-201, but which employs dielectric isolation instead of p-n junction isolation, was also tested in the LINAC by Bowman. His experimental results for this circuit indicate a failure level of approximately 1 to 2×10^9 rads(Si)/sec., an increase of at least an order of magnitude over the junction-isolated version, confirming the importance of isolation-diode responses.

Figure 13 shows a comparison of the actual TREE wave forms obtained for a MC-201 gate and a CMOS inverter, both for a peak dose

rate of approximately 6×10^8 rads(Si)/sec. While comparing these responses it is important to reiterate the fact that the MC-201 gate has a lower noise immunity than the CMOS inverter (1.3 versus 4.0 volts, respectively). The response of the MC-201 gate is taken from Raymond's data[30] (20-nsec flash x-ray pulse). The transient output voltage is seen to increase to approximately 1.8 volts at the end of the radiation pulse and then return to zero in about 200 nsec when the radiation pulse is removed. The response of the CMOS inverter is from data obtained using 120-nsec-wide pulses from a LINAC. The transient output voltage is seen to increase during the pulse to approximately 1.2 volts and then to start to return to zero when the radiation pulse is removed. We have also shown in Figure 13 an estimate of what the response of a CMOS inverter would look like for a 20-nsec flash x-ray pulse. In this case we would expect the peak transient output voltage to be about 0.5 volt and to occur near the end of the radiation pulse.

Summarizing the above discussion:

1. The junction areas of an MOS transistor are, in general, smaller than the area of the collector–base junction of a bipolar transistor by about a 5-to-1 ratio.[12] Hence, the peak primary photocurrent is correspondingly smaller.
2. Depending on operating conditions, the bipolar transistor can multiply the collector–base junction photocurrent by the common-emitter current gain. Indeed, this secondary photocurrent is frequently the cause of failure of bipolar integrated circuits. Such a secondary photocurrent mechanism is not dominant in MOS transistors.
3. The ready adaptability of the MOS configuration to a thin-film approach is a significant factor toward ultimately achieving high transient radiation hardness in logic arrays. When a thin-film approach is used, the effective collection volume is reduced by orders of magnitude, so that the peak primary photocurrents of the controlling drain junction diodes are reduced.
4. For the CMOS configuration, diode compensation is intrinsic to the nature of the circuit.
5. One of the few points against MOS devices is that bipolar transistors normally have considerably lower "on" impedances than do MOS transistors.

D. Summary of TREE Effects in Digital I.C.'s

The conclusions of this section are summarized in Figure 14 in which the peak dose rates for onset of logic failure in various forms

of digital logic are compared. Some words of explanation are in order, since several of the bars present on this chart represent estimates based on extrapolation of experimental results, rather than a direct measurement. Shaded areas indicate projections, and the open areas indicate

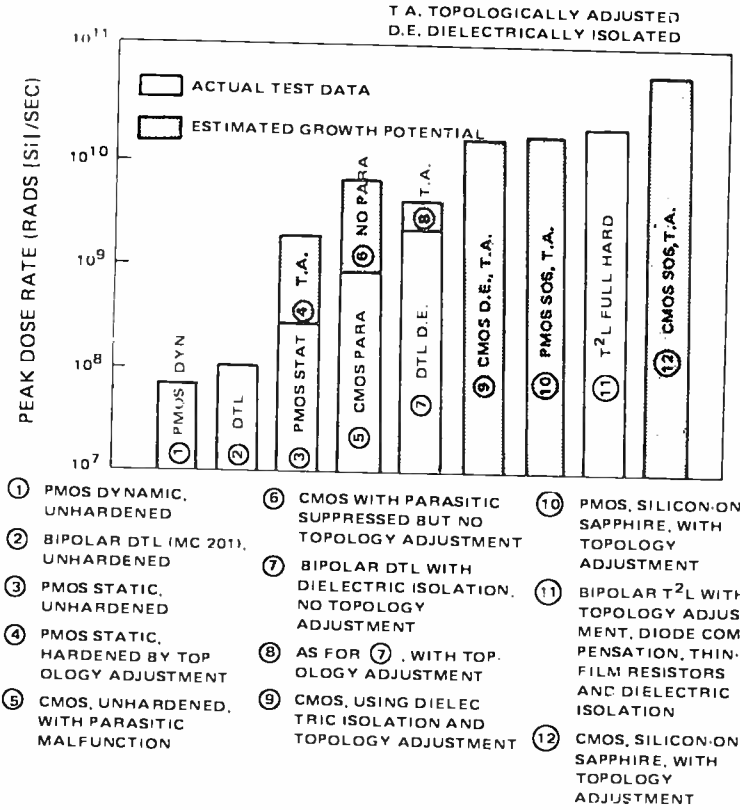


Fig. 14—Bar chart indicating the TREE malfunction levels of the various types of digital logic circuits.

experimentally determined logic failure levels. The columns "PMOS dynamic" and "PMOS static" (first and third columns) were derived from some preliminary experiments on two related RCA arrays containing two-phase shift registers and static shift registers, respectively. The shaded portion in the column "PMOS static" (third column) is an estimate of the hardening potential of PMOS-static logic. This hardening would be accomplished mainly through topological redesign of the chip. The two columns "bipolar DTL" (second and fifth column) are self-explanatory, since they are based on the experimental results as

discussed above. The shaded area in the column "CMOS unhardened" (fourth column) is an estimate of what the logic failure level would be if the observed parasitic response of the CMOS inverter could be eliminated without adjusting other circuit features. The shaded area in the column "bipolar DTL with dielectric isolation" (fifth column) is an estimate of the level of increased hardening that might be obtained by improving the circuit design. These design changes would include area reduction of the collector–base junction and appropriate changes in resistance values. These changes could only be minor before circuit performance would be affected. Not included is the increased hardening that might be accomplished by using diode compensation.

The column "CMOS hardened, etc." (sixth column) represents an estimate of the failure level of the CMOS inverter that could be attained by using dielectric isolation instead of p-n junction isolation of n- and p-channel device groups and by making the appropriate topological adjustments in the chip layout. The use of dielectric isolation would, of course, eliminate the parasitic response mechanism found in diode-isolated CMOS circuits and, in addition, could reduce the junction photocurrents by reducing the effective collection volume. Appropriate topological rearrangements could minimize both the "on" resistance and the drain junction area within the constraints of good logic operation. It is worth noting that these topological adjustments for TREE hardness are also advantageous in that they produce faster switching speeds.

The column "PMOS-SOS" (seventh column) represents an estimate of what the increased hardness will be if one turns to PMOS arrays fabricated in epitaxial silicon-on-sapphire films. This estimate is based on the fact that preliminary experiments have found the transient drain photocurrent of a MOS-SOS transistor to be about a factor of 8 times smaller than that of an equivalent monolithic MOS transistor. Theoretically, even higher factors should be obtainable with further effort. It is reasonable to assume that this decrease in photocurrent will result in a corresponding decrease in the transient output voltage of the inverter, with a resultant increase in the hardness level, as indicated in the figure.

The column "Bipolar T^2L" (eighth column) represents results recently reported on experimental integrated bipolar T^2L logic circuits.[33] These circuits employed all of the experimental techniques that can be used to obtain TREE hardening of bipolar integrated circuits, namely,

(1) Dielectric isolation,

(2) Gold-doping,

(3) Thin-film nichrome resistors,

(4) Diode compensation,

(5) Smallest possible junction areas,

(6) Shallow-junction diffusions.

The degree of increased TREE hardness that one can obtain in practical logic circuits by application of such techniques is still a subject of considerable discussion. However, the above data demonstrates what may be the ultimate in hardening technology for bipolar-transistor devices.

The column "CMOS SOS" (ninth column) represents an estimate of what the increased hardness will be if one turns to CMOS arrays fabricated in epitaxial silicon-on-sapphire films, using deep-depletion transistors. This estimate is based on the same rationale as that for PMOS-SOS.

It is worth noting that the technology of silicon-on-sapphire is making rapid advances*, so that the use of SOS-CMOS circuits in military systems in the near future is not merely an academic question. It is also worth noting that silicon-on-sapphire technology is also very adaptable to the PMOS approach.

6. Conclusions

In this analysis an attempt has been made to describe and relate the two distinct types of effects of radiation—permanent and transient—in the light of work on MOS structures, and to indicate measures for achieving tolerance to these two effects by proper device design. A word should be said here about the interaction of the two effects of radiation and the two techniques required to harden against them. The two problems, charge buildup and photocurrent generation are, of course, basically different and affect different regions of the device. The hardening techniques, insulator modification and topology adjustment, can thus be applied independently. However, it seems fairly certain that when measures for hardening against both effects are combined in the same chip, some interaction will take place, although it probably will not be large. For example, design constraints imposed by the need for small junctions could make the formation of aluminum oxide layers in the desired geometry more difficult because of problems in controlling oxide etch rates. However, it is clear that it will always be possible to tailor the hardness levels of integrated circuits with

* See Reference (11) for the discussion of a SOS-CMOS array using 90 transistors that has been designed, fabricated, and operated satisfactorily.

respect to permanent and transient effects independently, within these limits, to fit a particular environmental demand. Many space vehicles, for example, may not demand hardening to transient effects, since the dose rates experienced when oribiting through space radiation belts are many orders of magnitude lower than those commonly produced near nuclear events, while the total dose received in the two cases are commensurate.

Looking to the future, it is likely that the ability to apply "hardening" techniques to state-of-the-art integrated circuitry will become a common and desirable skill of the designer of semiconductor devices for aerospace use. This skill may prove of general benefit in that, to date, most of the methods developed for improving tolerance to radiation in devices has also generally produced improvement in the electrical performance of the device, while the phenomenological work done on radiation effects has substantially increased our total understanding of device physics.

Acknowledgements

We gratefully acknowledge the constant support and encouragement of Martin Wolf, Paul Rappaport, and John Boning in the studies described above.

Appendix I—Aluminum Oxide as an MOS Gate Insulator

Various methods exist for the fabrication of aluminum oxide. Here, we shall briefly describe the two methods employed in our studies, namely plasma anodization and vapor deposition.

A. Plasma Anodization Process

The technique of plasma anodization of a metal to form the metal oxide is a relatively new one and, to date, has been primarily used to form oxides on metals as dielectrics for thin-film capacitors. The most frequently studied materials have been Al_2O_3 and Ta_2O_5. While the plasma-anodization technique is easily implemented and has been used for the fabrication of thin-film capacitors, an understanding of the fundamental mechanism controlling oxide growth and maximum obtainable oxide thickness has not yet been achieved.

For the fabrication of Al_2O_3 films on a silicon substrate,[8] the anodization of the aluminum is carried out in a bell jar as shown in

Figure 15. The aluminized silicon wafer is placed in a pressure-contact jig that makes electrical contact to the back side of the wafer. The system is then pumped down and subsequently back-filled with dry oxygen. A glow discharge is ignited between anode and cathode, and the sample, which is in the glow or "positive column" portion of the discharge, is biased positively with respect to the anode. The applied bias is set sufficiently high so that all the deposited aluminum film is converted to Al_2O_3. The oxide thickness is determined mainly by the

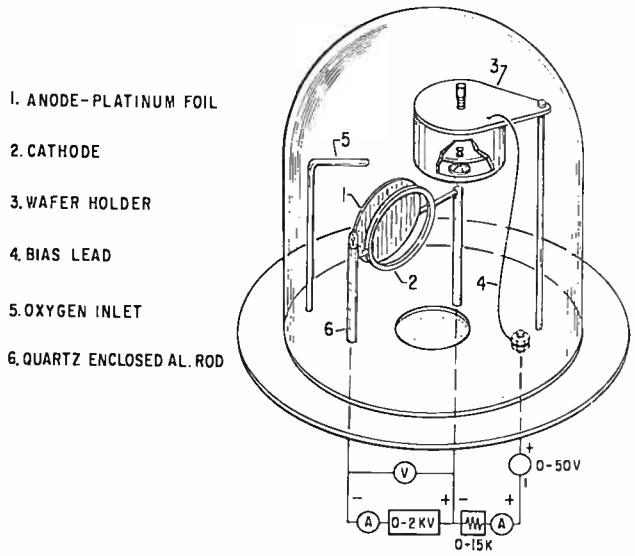


Fig. 15—Method of fabricating AL_2O_3 films by plasma anodization.

initial aluminum film thickness. Typically, oxides of 500-1000 Å have been fabricated. Experimentally it was found that the ratio of Al_2O_3/Al thickness is 1.41, in good agreement with calculated values using published values of densities for Al and Al_2O_3. Both n and p-type wafers with various resistivities have been used with no apparent effect on the anodization of the aluminum.

Device Characteristics

In Figure 16, *C-V* and *G-V* data are shown for a typical Al_2O_3 MOS capacitor (10 ohm-cm p-type silicon). The aluminum oxide thickness is 550 Å. Immediately after anodization, a large positive oxide charge is found in the samples. Typically, it is of the order of $5 - 7 \times 10^{12}$ charges/cm^2. At this point, the capacitors also exhibit a small hystere-

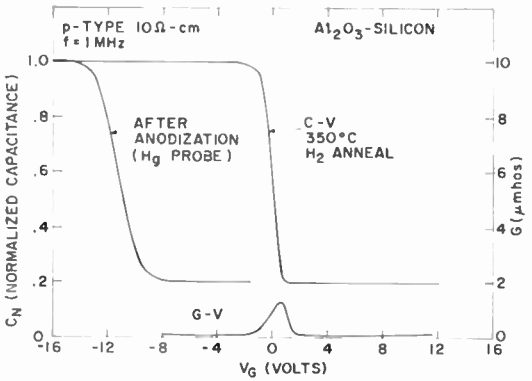


Fig. 16—*C-V* and *G-V* data for a typical Al_2O_3 MOS capacitor (10 ohm-cm *p*-type silicon).

sis under bias at room temperature. The devices are then annealed in hydrogen or forming gas. This results in a removal of both oxide charges and interface states. This low surface-state density is found to be independent of whether n or p type wafers are used. Effects of silicon orientation have also been investigated but little difference between orientations was observed.

A typical Al_2O_3 MOS transistor characteristic is shown in Figure 17. The transconductance is excellent and the field-effect mobility is 600 cm^2/volt-sec, indicating again a high degree of perfection of the Al_2O_3-Si interface.

Effects of Ionizing Radiation

As with MOS devices using SiO_2 or other oxides as the gate insulator, the bombardment of Al_2O_3-MOS devices with 1-MeV electrons

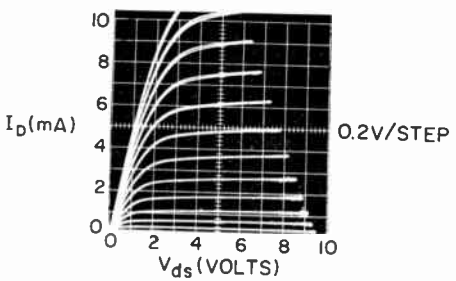


Fig. 17—Typical Al_2O_3 MOS transistor characteristic.

results in the build-up of a space charge in the insulator with a corresponding shift of the device characteristics along the voltage axis.[8] However, in contrast to results obtained with SiO_2 and many other insulators, irradiation does not always cause a shift of the *C-V*, *G-V*, or *I-V* characteristics towards more negative bias. This can be seen from the typical 1 MeV electron bombardment result shown in Figure 18. Note that for negative bombardment bias the *C-V* curve is displaced toward the right, indicating the introduction of a negative space charge into the oxide and/or interface states. It should also be noted

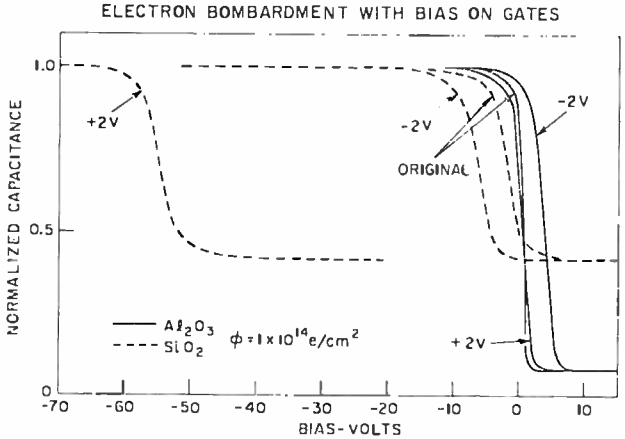


Fig. 18—Effect of 1 MeV bomardment with different gate biases.

that the generation of interface states is not completely negligible for this oxide–semiconductor system, as indicated by the slight change in slope of the *C-V* characteristics. Up to 10^{13} e/cm^2, under positive or negative bias, no oxide charge build-up and no interface state generation is detectable. Above these fluences only slight changes are observed, as indicated in Figure 18. For comparison purposes, radiation damage data for the SiO_2-Si system, typical for rf-grown dry-oxygen SiO_2 films, is also included. For the same bombardment conditions, the shifts for an SiO_2-Si device can typically reach −60 volts, whereas the Al_2O_3-Si device will generally shift less than a few volts. This radiation behavior is better than that observed for metal-nitride-semiconductor, or hardened, SiO_2 devices. The relative behavior of the Al_2O_3 devices under bias-bombardment has been examined in detail, and the results are shown in Figure 19. In this graph the range of bombardment-induced charge for dry-grown SiO_2 films, as well as results obtained

for the "best" steam oxides, is also indicated. The comparison between Al_2O_3 and SiO_2 is quite striking. It must be remembered here that Al_2O_3 devices would also be operated at half the applied voltage as compared with SiO_2 devices because of the value of the dielectric constant of Al_2O_3. Notice also that for negative values of bombardment bias, negative oxide charge is induced, i.e., the curves shift to the right. This is most probably due to predominant trapping of electrons in the Al_2O_3 under these conditions.

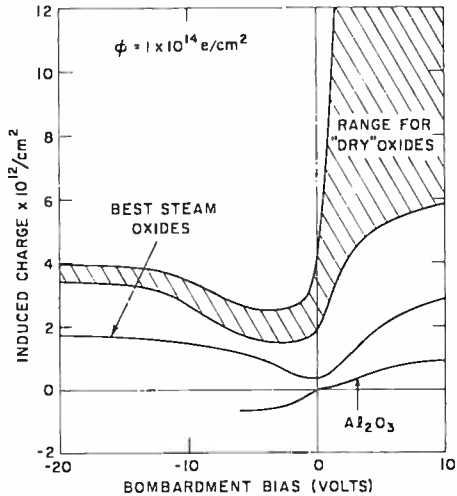


Fig. 19—Relative behavior of Al_2O_3 devices under bombardment bias.

The authors have also carried out irradiations of the Al_2O_3 MOS transistor structures with 1-MeV electrons, mixed neutron/gamma fluxes from a reactor, gamma-rays and x-rays (the latter two irradiation experiments were done in colloboration with P. Poch). The results are generally consistent with the capacitor data and support the large superiority of Al_2O_3 over SiO_2 and (P_2O_5/SiO_2)-SiO^2 as expressed in Figure 7.

Even though these results are, as yet, on laboratory devices, they indicate that if the technology of plasma-grown Al_2O_3 can be developed successfully, then MOS arrays can be developed that have an adequate degree of hardness for aerospace use. The anodization process is compatible with planar integrated-circuit technology, and has enabled RCA to fabricate MOS transistors that have near-zero threshold voltages and that are stable to temperature-bias stress without resorting to ultra-clean technology.

B. Vapor Deposition Process

The deposition process[33] is essentially the pyrolytic decomposition of Al-isopropoxide. The deposition is generally carried out in either a resistance-heated or rf-heated furnace. In either case, helium gas is bubbled through the liquid alkoxide. This gas is then mixed with oxygen and, depending on the He/O_2 ratio, the Al content of the gas mixture can be varied. Hence, the deposition rate at a given temperature is controllable. The Al_2O_3 film forms on the substrate in the temperature range from 400°C to 900°C. At the lower temperature, the oxide is noncrystalline, but with increasing temperature there is a

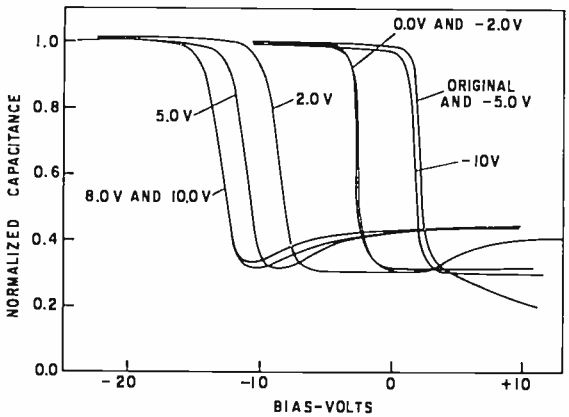


Fig. 20—Effect of 1 MeV bombardment on Al_2O_3 MOS capacitor formed by vapor deposition.

tendency toward crystallinity. The films are smooth and uniform, and the index of refraction varies between 1.60 and 1.70, depending on the processing conditions.

To date, only MOS capacitors have been fabricated with this process. The process steps are very simple. After standard cleaning of the Si chip the wafer is placed in the reaction chamber. Film deposition can then follow immediately. After a suitable post-deposition heat treatment the gate electrode and the back contact are deposited.

For the case of a properly prepared single Al_2O_3 film about 1000 Å thick, the flat-band voltage is around +2.0 V and the interface state density is typically below 1×10^{10} states/cm^2–eV.

MOS capacitors with a single Al_2O_3 layer, when bombarded by 1-MeV electrons, showed a large spread in their irradiation behavior. However, the best results obtained are similar to those for plasma-grown units. Such a result is shown in Figure 20.

References

1. K. H. Zaininger, "Electron Bombardment of MOS Capacitors," IEEE Silicon Interface Specialists Conf., Las Vegas, Nov. 1965, and *Appl. Phys. Letters*, Vol. 8, p. 140, 15 March 1966.
2. K. H. Zaininger, "Irradiation of MIS Capacitors with High-Energy Electrons," IEEE Trans. GNS, Vol. NS-13, p. 237, Dec. 1966.
3. W. J. Dennehy, A. G. Holmes-Siedle and K. H. Zaininger, "Process Techniques and Radiation Effects in MOS Structures," IEEE Trans. GNS on Nuclear Science, December 1967.
4. K. H. Zaininger and A. G. Holmes-Siedle, "A Survey of Radiation Effects in Metal-Insulator-Semiconductor Devices," *RCA Review*, Vol. XXVIII, No. 2, June 1967.
5. E. H. Snow, A. S. Grove, and D. J. Fitzgerald, "Effects of Ionizing Radiation on Oxidized Silicon Surfaces and Planar Devices," *Proc. IEEE*, Vol. 55, p. 1168, July 1967.
6. A. G. Stanley, "Comparison of MOS and Metal-Nitride-Semiconductor Insulated Gate Field-Effect Transistors Under Electron Irradiation," *IEEE Trans. GNS*, Vol. NS-13, p. 248, Dec. 1966.
7. K. H. Zaininger and C. C. Wang, "Thin-Film Dielectric Materials for Microelectronics," *Proc. IEEE*, Vol. 57, p. 1564, Sept. 1969.
8. A. Waxman and K. H. Zaininger, "Al_2O_3-Silicon Insulated Gate Field-Effect Transistors," *Appl. Phys. Letters*, Vol. 12, No. 3, Feb. 1968; and K. H. Zaininger and A. S. Waxman, "Radiation Resistance of Al_2O_3 MOS Devices," *IEEE Trans. Electron Devices*, Vol. ED-16, p. 333, April 1969.
9. M. T. Duffy and A. G. Revesz, "Interface Properties of Si-(SiO_2)-Al_2O_3 Structures," *Jour. Electrochem. Soc.*, in press.
10. J. R. Burns, "Switching Response of Complementary Symmetry MOS Transistor Logic Circuits," *RCA Review*, Vol. 25, No. 4, Dec. 1964.
11. J. F. Allison, F. P. Heiman and J. R. Burns, "Silicon-on-Sapphire Complementary MOS Memory Cells," *IEEE Journal of Solid State Circuits*, Vol. SC-2, No. 4, Dec. 1967.
12. R. L. Petritz, "Current Status of Large Scale Integration Technology," *IEEE Journal of Solid State Circuits*, Vol. SC-2, No. 4, Dec. 1967.
13. A. G. Holmes-Siedle and K. H. Zaininger, "The Physics of Failure of MIS Devices Under Ionizing Radiation," *IEEE Trans. Reliability*, Vol. R-17, No. 1, March 1968.

14. A. G. Holmes-Siedle and K. H. Zaininger, "Designing MOS Systems for Radiation Environments," *Solid State Technology*, Vol. 12, No. 5, May 1969.
15. C. W. Perkins, K. G. Aubuchon, and H. G. Dill, "Radiation Effects in Modified Oxide Insulators in MOS Structures," *IEEE Trans. GNS*, Dec. 1968.
16. D. R. Kerr et. al., "Stabilization of SiO_2 Passivation Layers with P_2O_5," *IBM Jour. Res. and Development*, Vol. 8, No. 4, Sept. 1964.
17. B. Buchanan, R. Dolan and S. Roosild, "Comparison of the Neutron Radiation Tolerance of Bipolar and Junction Field-Effect Transistors," *Proc. IEEE*, Dec. 1967.
18. G. J. Brucker, W. J. Dennehy and A. G. Holmes-Siedle, "High-Energy Radiation Damage in Silicon Transistors," *IEEE Trans. GNS*, Vol. NS-12, July 1965.
19. G. J. Brucker, W. J. Dennehy and A. G. Holmes-Siedle, "Electron-Induced Surface Damage in Silicon Transistors," *Proc. IEEE*, Vol. 53, p. 1800, Nov. 1965.
20. G. Brucker, "Correlation of Radiation Damage in Silicon Transistors Bombarded by Electrons, Protons, and Neutrons," Proc. on Radiation Effects in Semiconductor Components, Toulouse, France, Vol. B7 (Mar. 1967).
21. G. C. Messenger, "Radiation Effects on Microcircuits," *IEEE Trans. GNS*, Vol. NS-13, No. 6, Dec. 1966.
22. K. H. Zaininger, W. Dennehy and A. Holmes-Siedle et. al., "Stability and Radiation Effects in Complementary Symmetry MOS Transistor Circuits," Contract AF33615-C1140, Quarterly Reports 1 through 8 (March 1967-Oct. 1968) and Final Report (March 1969).
23. D. Sullivan, "Transient Radiation Induced Response of MOS Field Effect Transistors," *IEEE Trans. GNS*, NS-13(6), (Dec. 1965).
24. J. P. Raymond, E. Steele, and W. Chang, "Radiation Effects in Metal-Oxide-Semiconductor Transistors," *IEEE Trans.* GNS Feb. 1965.
25. D. M. Long, "A Radiation Effects Large Signal Equivalent Circuit for MOS Transistors," *IEEE Trans. GNS*, Vol. NS-14, No. 6, Dec. 1967.
26. W. J. Dennehy, A. G. Holmes-Siedle and W. F. Leopold, "Transient Radiation Response of Complementary Symmetry Integrated Circuits," *IEEE Trans. GNS*, Vol. NS-16, No. 6, p. 114, Dec. 1969.

27. G. Kinoshita, C. T. Kleiner, and E. D. Johnson, "Radiation Induced Regeneration through P-N Junction Isolation in Monolithic I/C's," *IEEE Trans. GNS*, Oct. 1965.
28. S. R. Hofstein, "An Analysis of Deep-Depletion Thin-Film MOS Transistors," *IEEE Trans. GED*, Vol. ED-13, No. 12, Dec. 1966.
29. J. L. Wirth and S. C. Rogers, "The Transient Response of Transistors and Diodes to Ionizing Radiation," *IEEE Trans. GNS*, Nov. 1964.
30. J. P. Raymond, W. W. Chang and R. E. Budris, "Lumped-Model Analysis of Microcircuit Vulnerability," (to be published) *IEEE Trans. GNS*, Dec. 1968.
31. W. C. Bowman, W. Beezhold and A. H. Johnston, "Models for Predicting Transient Radiation Response in Microcircuits," *IEEE Trans. GNS*, NS-15(6), (Dec. 1968).
32. R. L. Russel, R. L. Skavland, G. G. Spehar and K. R. Walker, "Photocurrent Compensation Techniques for Radiation Hardening of Integrated Circuits," (Final Report) Rome Air Development Center, RADC-TR-68-366.
33. J. A. Aboaf, "Deposition and Properties of Aluminum Oxide Obtained by Pyrolytic Decomposition of an Aluminum Oxide," *Jour. Electrochem. Soc.*, Vol. 114, p. 948, 1967.

PARALLEL PROCESSING FOR PHASED-ARRAY RADARS

By

J. E. Courtney and H. M. Halpern

RCA Missile and Surface Radar Division
Moorestown, N.J.

Summary—Some results of a study of a large radar–computer system are reported. Emphasis is on phased-array radars and parallel-processing computers in a high-target-density environment. Communication is a critical problem in parallel processing, and a modular-array computer configuration was found to be efficient in this area. The particular problems of input–output, array steering, and target-data storage and handling are analyzed with respect to parallel processing. The importance of stating each problem in a form that takes the greatest advantage of parallel processing is emphasized.

Introduction

THIS PAPER REPORTS some results of a study of the characteristics and requirements of a large radar-computer system, with special emphasis on phased-array radars and parallel-processing computers. A survey of radar systems requirements was made, with emphasis on the high-density target environment. A detailed analysis of the various system tasks was performed as they were related to digital processing of signals and data. A study of various parallel-processing computer configurations was also carried out, and a particular configuration was chosen for evaluation against the radar problem. This configuration was a modular-array computer. A number of radar tasks were applied to this configuration in order to determine its applicability under various conditions.

A major problem of parallel processors is to find a configuration that is both flexible and efficient. Flexibility is desired to enable the system to expand in an orderly manner when the task grows, so that the same concept may be applied to a wide variety of radar-computer systems. Efficiency is used here in the sense that internal housekeeping and executive functions should be small compared to the total computational

capacity. The modular-array system was found to be both efficient and flexible.

Nature of the Radar Problem

Operation of a large phased-array radar requires a great many data processing tasks, including target search, target track, array steering, and beam scheduling. Beam scheduling alone is a considerable task; the radar may have multiple beams, all of which must be scheduled in accordance with target priorities and transmitter-power constraints. Thus, the radar imposes a very large data-processing load that must be performed in real time.

Considerable flexibility is needed in the computer configuration due to the manner in which large radar systems evolve. A radar may take three to five years to develop and have an operational life of another ten years. During that time, requirements may change and new technical developments become available. Thus, modifications are frequently made after a system becomes operational. This need for long-term flexibility has a direct bearing on whether a special-purpose or general purpose computer configuration is chosen. While a special-purpose system may be more efficient for a specific computational task, changes in the system require changes in the hardware and are very expensive. The general-purpose computer, on the other hand, permits changes to be made by changing the software. While changes in the software are far from trivial, they do allow the system to respond to changing requirements.

Another requirement is the ability to increase capacity in order to accommodate the increased processor load that results from most changes in the radar system. This is very difficult to do with a conventional general-purpose computer. For this reason, a modular configuration employing parallel processors was chosen.

The use of parallel processors, however, introduces problems in efficiency, i.e., the applicability of the software to the required task. Some radar tasks are naturally parallel in nature. An example is the array-steering problem, in which settings must be computed simultaneously for a large number of phase shifters or time-delay units. In other tasks, however, the method of application of parallel processing is not so obvious, and some study is required in order to efficiently fit the problem to the computer.

Some Parallel-Processor Configurations

The efficiency of a computer becomes low when too much of the capacity of the computer is devoted to internal housekeeping. House-

keeping includes such jobs as shifting data from point to point inside the computer and assigning tasks to the various processors. If the proper configuration is not used, it is difficult to keep all the processors busy at one time. This is related to the problem mentioned in the previous section of fitting tasks to the computer. Three possible configurations were considered: the multi-processor, the array processor, and the modular array.

In the multi-processor configuration, a number of processors and controllers are all inter-connected so that any controller has access to any processor. The problem with the multi-processor is communications. The amount of housekeeping required increases rapidly with the number of processors employed, so that a point is eventually reached, beyond which the addition of more processors does not result in an increase in overall capacity.

In an array processor, a single controller issues commands to a number of processors, and these all perform the same task simultaneously. The array processor avoids some of the housekeeping problems by using a single controller. However, the array processor tends to become limited in its capability as additional processors are added, since it is not able to keep all the processors working all the time. Since many of the tasks that the radar must perform will not require the full capacity of the computer at any given time, the efficiency of the computer begins to decline as the number of processors grows.

A solution to some of these problems lies in a combination of the array processor and the multi-processor. This is the modular array computer, the subject of this paper.

The Modular-Array Computer

The modular-array computer may contain from 1 to 16 modules, as shown in Figure 1. Each module in turn may contain from 1 to 16 processors and one controller, arranged as in Figure 2. Each module contains a Next Instruction Bus, a Data Bus, and an Input–Output Bus. Each processor has associated with it a 32-bit, 4,096-word memory. The instructions for the controller are stored in the processor memories.

Communication is the key to success in parallel processing. Each module is interconnected to the others via a data exchange bus and an instruction exchange bus, but each processor also communicates directly with its immediate neighbors via the memory buffer registers. A processor communicates with its neighbors within the module by the "East–West" transfers shown in Figure 2. In addition, each processor communicates with its counterpart in the two adjacent modules by

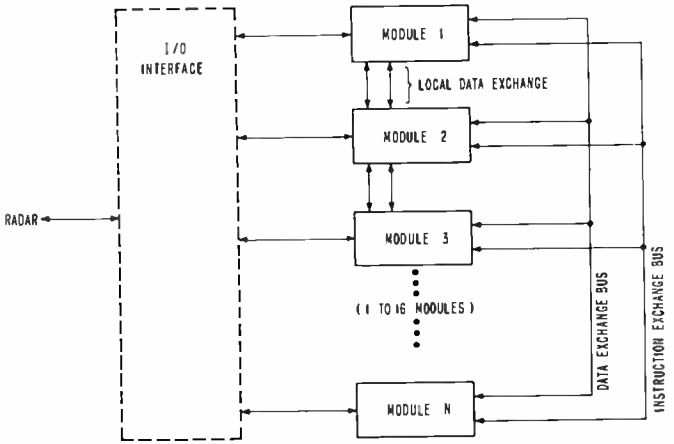


Fig. 1—Modular array computer configuration.

means of the "North–South" transfer lines. Not only may data be transferred by this means, but data stored in the memory of one processor may be directly accessed by a neighboring arithmetic unit, without actually transferring the data to the neighbor's memory.

Figure 3 shows the relative performance of the multiprocessor, the array processor, and the modular-array processor. Here the effects of housekeeping and processor utilization are seen. By selecting the proper number of modules for the radar task at hand, most of the processors

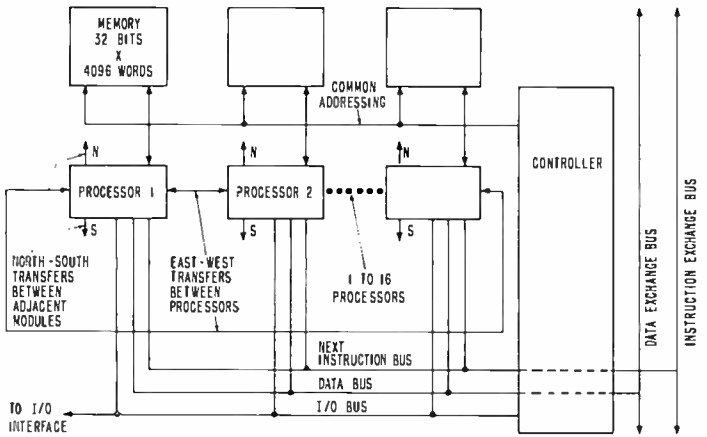


Fig. 2—Computer module.

may be utilized most of the time, so that the efficiency of the modular-array processor is kept high.

The processor assumed for the basic unit in each module has capabilities that might be expected to obtain in a system being designed today, i.e., a memory cycle time of 200 nanoseconds, an add time (including the memory cycle) of 250 nanoseconds, and a multiply time of 650 nanoseconds. Extensive use of Large-Scale Integration (LSI) was assumed. Use of LSI is important since the power of the computer may

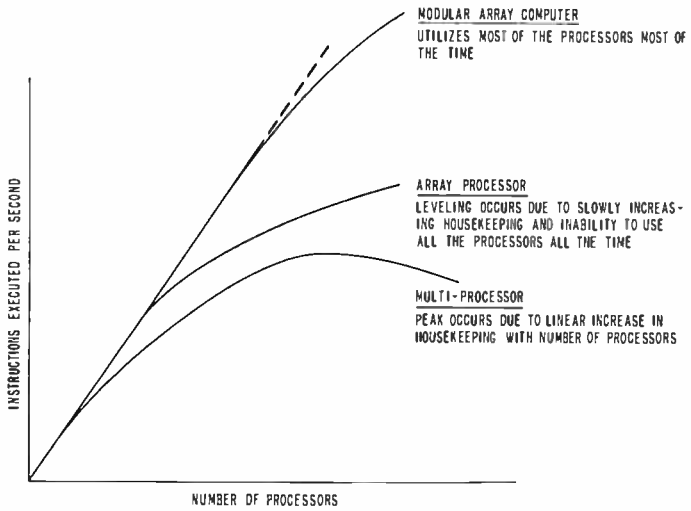


Fig. 3—Relative performance of computer configurations.

be rapidly dissipated if the individual components are so large that delays are encountered in transferring data from one component to another. Each processor or controller is to fit on a single 19 × 19-inch platter.

Examples of Radar Problems

This section gives a description of some typical radar problems, and how the modular-array computer was applied to these problems.

Data Input

The first problem considered is that of data input. This problem is very important because no other processing may take place until the radar data has been loaded into the computer. The problem of loading

radar data hinges on the special nature of this data. The output of a radar receiver is characterized by short bursts at very high rates, followed by relatively long quiescent periods during which processing may take place. An example of one of these bursts of data is the contents of a range gate. The problem was aggravated in the particular case studied, because the radar had several parallel receivers, all operating independently, with more than one range gate in each receiver. We will now show how the particular structure employed in this computer is useful in loading radar data.

Referring to Figures 1 and 2, it should be remembered that each module has its own input–output bus. This bus connects directly to the processors in each module. Since each processor within the module has the capability of loading a 32-bit word in one memory cycle and since there are 8 processors in each module, 256 bits may be loaded into each module during a single memory cycle.

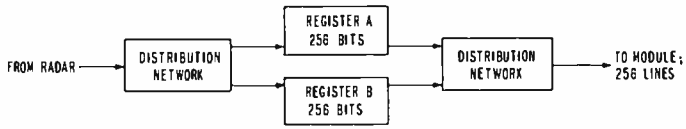


Fig. 4—I/O interface for one module.

Figure 4 shows the input–output interface for each module. Each radar receiver has an output of relatively few bits in each sample, and the buffer registers are used to collect enough of the short samples to make one large 256-bit word. Two buffer registers are used, so that one may be loading into the module while the other is accumulating data from the radar receivers. When enough data has been transmitted from the radar to fill one of the buffer registers, an interrupt signal is sent to the computer. Upon reception of the interrupt signal, computation is suspended and the contents of the buffer registers are transferred into the modules.

Figure 5 shows how the modules are used as buffers for each other during a high rate burst from the radar. For most of the processing problems, the data to be processed is needed in one module. However, due to the method of high-speed loading, the data that has just been received is scattered through a number of modules. The sequence of operations shown in Figure 5 is used to shift the data into the module in which the processing is to be done. This procedure makes use of the North–South transfer connections between modules. During each

step of the sequence, the data in each module is shifted to the module just below, i.e., the data in module 1 is transferred to module 2, while the data that was in module 2 goes to module 3, and so on. During this first step, 256 bits are transferred from each module to the module just below. These steps are repeated until all the data required has been transferred into module 4. (The data exchange bus may also be used to transfer data between modules that are not adjacent.) The number of actual steps or memory cycles required to complete the transfer depends on how many words of data were contained originally in each module.

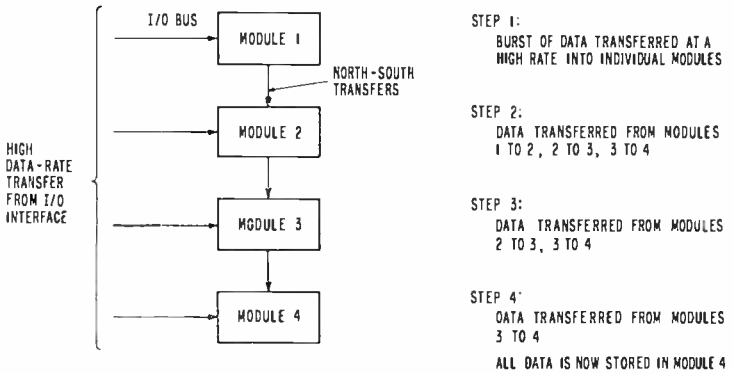


Fig. 5—Loading of burst of radar data into one module.

Computation of Phase Steering Commands

A second radar problem to be illustrated for application to the computer is the one of generating the steering commands for a phased array. This is an example of a problem with a large amount of parallelism; a single beam command will give rise to hundreds of phase-shifter commands.

The geometry is illustrated in Figure 6. Commands for positioning a beam are received in radar (spherical) coordinates—azimuth and elevation of the beam. Some radars have a variable carrier frequency, and if this is the case, the frequency to be used must be supplied with each beam order. Other parameters that enter the computation concern the structure and orientation of the array. The phase shifters for the array being considered are arranged on rectangular (row and column) coordinates. The array also has a fixed azimuth and elevation, as shown in Figure 6, that must be taken into account.

The command for each phase shifter is computed in two components

that are added together to obtain the setting for the phase shifter. The reason for this is that all the phase shifters in a row, say, will have the same y component, and the x components will be integer multiples of one phase increment. Hence the basic quantities to be

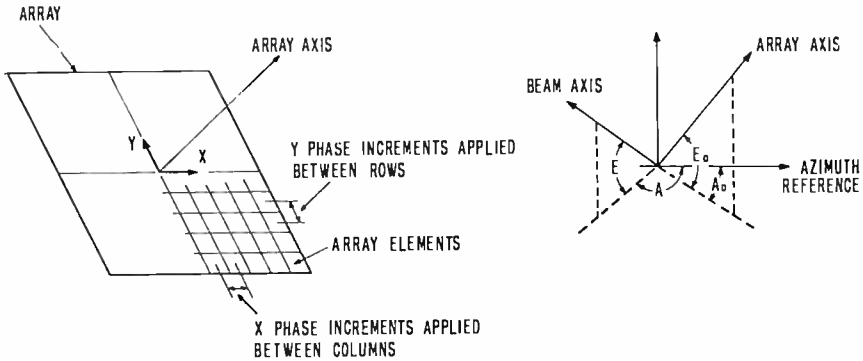


Fig. 6—Beam steering geometry.

computed are x and y phase increments, given by the following expressions.

x phase increment:

$$\phi_x = 360\, f\, D_x \left(\frac{1}{c}\right) \cos E \sin\,(A_o - A),$$

y phase increment:

$$\phi_y = 360\, f\, D_y \left(\frac{1}{c}\right) [\sin E \cos E_o - \cos E \cos\,(A_o - A) \sin E_o],$$

where

f is the transmitter carrier frequency,
D_x is the horizontal array element spacing,
D_y is the vertical array element spacing,
c is the velocity of light,
A is the beam azimuth,
E is the beam elevation,
A_o is the array azimuth, and
E_o is the array elevation.

Not only is there parallelism, in that many phase shifter commands are to be computed at once, but in the particular system studied, commands for up to 16 different beams are to be computed simultaneously. For each beam, separate values of A, E, and f are given. The sequence of computations and the amount of parallelism for this problem is shown in Figure 7. At each stage of the process, the maximum number of processors that may be used simultaneously is given. The sequence shown in Figure 7 is for a single module. Notice, however, that steps

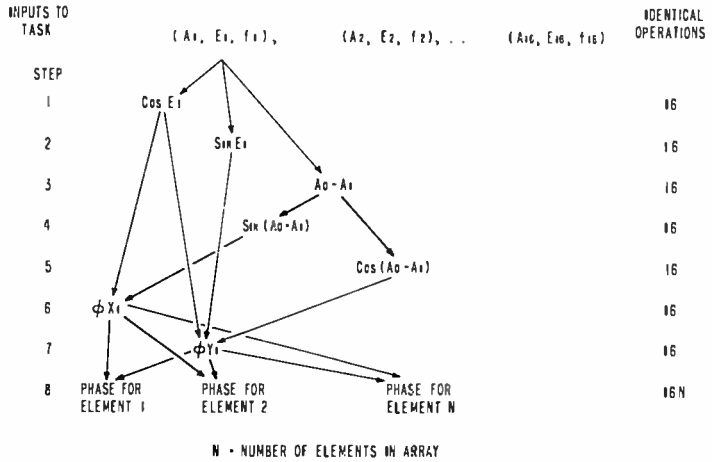


Fig. 7—Processing of phase steering commands.

1, 2, and 3 could be performed simultaneously using three adjacent modules, since the values of A, E, and f need be stored only in the middle one; the other two access the A, E, and f values in the middle memory directly.

As an example of how the problem may be arranged to take maximum advantage of parallel processing, note that at some stages the cosine of an angle is needed, while at others the sine of the same angle is needed. The inclusion of a short addition step will allow the paralleling of the much longer sine or cosine computation, as shown in Figure 8. If the processors are available to perform this parallel operation, a considerable amount of time may be saved by *adding* operations.

Problems of Target-Data Storage and Handling

In illustrating the problems of target data storage and handling, it

is helpful to consider the radar-signal processing task as a sequence of filters. The earlier in the sequence, the more "targets" there are to be manipulated. After each step in the process, the number of targets is reduced and, consequently, more sophisticated filtering can be applied. Typical filtering operations require association of the targets resulting from successive transmissions to permit doppler filtering, drag estimation, scintillation tests, etc.

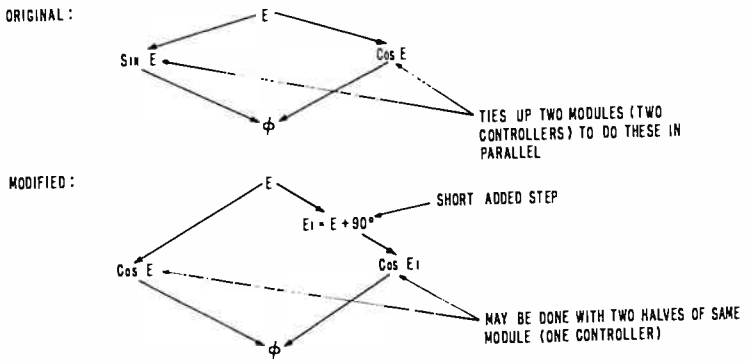


Fig. 8—Change of program to increase parallelism.

One can approach the association problem in two ways. The first, which is more commonly used, becomes extremely complicated for large numbers of targets. In this method, each target is assigned a fixed location in memory that contains position and coordinates. The target priority and expected position versus time is updated periodically in a scheduler, which groups the targets by beams and expected range and orders radar pulses in accordance with the priority ranking. As the number of targets increases, the continuous shifting of order and establishment of range gates greatly increases the communication time between processors, particularly if the targets are distributed among several processors.

A second method of handling the target association task is to place the target data in memory by some associative scheme, so that the location in memory depends in some fashion upon the location of the target in space.* (All the target data need not be stored at this location, but a system of pointers may be used, with some of the data, such as target priorities, stored at some other location designated by the

* The location in space may be relative to the earth, or relative to some point within a large group of targets ("cloud-centered coordinates").

pointer.) A very simple method of doing this is to assign a separate memory location to each resolvable point in space. This of course is extremely uneconomical of memory. A more efficient approach is to establish bins in memory capable of holding many pointers. The pointers are placed in a bin by selecting the bin address through a random number computed from the target position. This method has proved to be very economical in computing time and memory require-

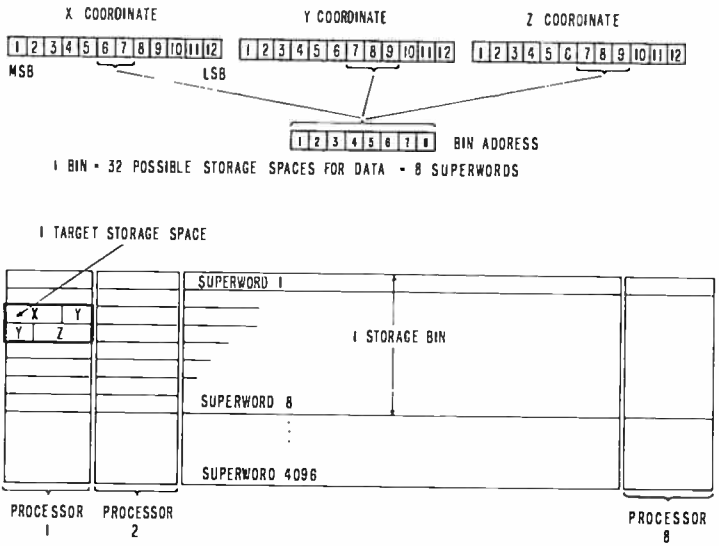


Fig. 9—Memory organization for scintillation test.

ments. Randomizing the address has a distinct advantage in that targets are more uniformly distributed among the processors, and a bunching of targets in space does not place an undue burden on any one processor.

Figure 9 illustrates the organization of one such memory used in a scintillation test. For this test, two beams scan the same moving volume. The two beams maintain a fixed time interval, τ. A second pair of beams are swept over the volume about 10τ later. Targets that pass the first comparison (i.e., those that persist for τ seconds) are associated with the targets that pass the second comparison. Only those targets that persist through the two successive examinations are passed by the scintillation filter.

The leading beams establish the contents of the memory bins. The figure shows a 12-bit word for each of the x, y, z coordinates of the

relative target position.* The intermediate bits of the coordinate words are used to generate a random address. As shown in Figure 9, all the words in memory that are called out by the same address are lumped together as a superword. All the processors in a module have a common controller, and the controller manipulates data as superwords.

Each piece of target data that is stored is located in two superwords, one containing x and half of y, and the other containing the rest of y and z. For the scintillation filter studied, each bin consisted of 32 possible storage spaces, which required eight superwords. A total of 256 storage bins were provided.

The system operates in the following fashion. For the leading beam, a bin address is computed for each target, and the target coordinates are placed in the next available storage space in the bin. For the trailing beam, a bin address is also computed for each incoming target. As each trailing beam target arrives, comparisons are made simultaneously by all processors going successively through superwords containing x coordinates until a match is found. If x and half of y compare, then the remaining superword is called for a comparison of the remaining y and z. A successful comparison is passed to the next store, to await the passage of the second pair of beams. A scintillating "target" that does not compare at any stage is discarded. Figure 10 shows a three-module implementation with the assignment of superwords of memory in each module. For the case shown, 12,000 bits of memory were fully occupied. The processors were utilized 500 milliseconds/second of the time. Less than 8 milliseconds/second of that time was involved in data transfers.

In any randomization scheme such as the one just described, the possibility of having an overflow in one of the memory bins must be considered. When the targets have been randomly distributed among storage bins, the probability of any bin containing a given number of targets follows a Poisson distribution, and from this the probability of overflow may be obtained. For the filter just described, about 3000 targets were to have been distributed among 256 bins, giving an average of 12 targets per bin. For the probability of overflow to be less than 0.00001, the number of storage spaces per bin must be 32. On the average, $\frac{3}{8}$ of the cells in memory will be filled. This means that processing steps (comparisons) have been saved at some cost in memory utilization, raising the question of how these two items are to be traded off. Figure 11 compares the results of the three mechanizations. Note that the randomization method just described greatly re-

* Cloud-centered coordinates are assumed here.

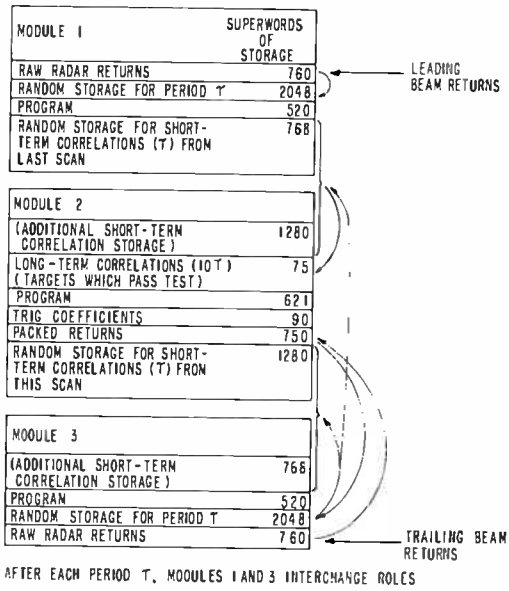


Fig. 10—Parallel processing of scintillation test.

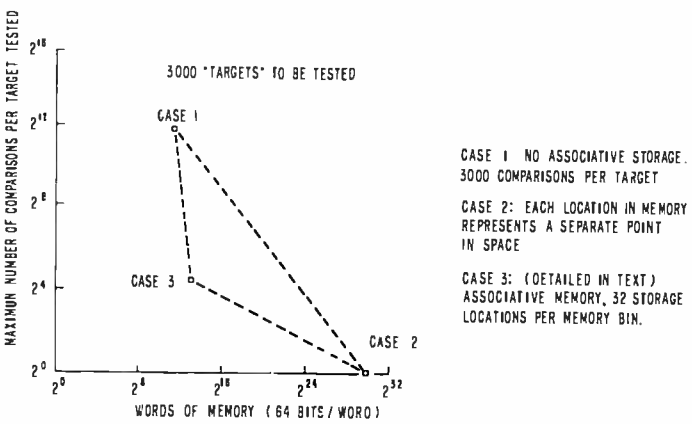


Fig. 11—Comparison of mechanizations for scintillation test.

duces the number of computations without a substantial increase in the amount of memory required.

In addition to the problem of overflow, another problem arises when associating targets from successive scans. This is the problem of an ambiguous location in storage arising from uncertainties in the position of the target. There is always a certain amount of error associated with measuring the position, due to noise in the radar system, interference from jamming, uncertainties in predicting positions, etc. For this reason, the least-significant bits of the coordinates are disregarded when making comparisons, or when computing an address in memory from a given target position. As a result, the storage bin corresponds to a finite volume in space, (or a set of disconnected volumes) rather than a point as small as the best resolution of the radar. The ambiguity problem arises when the target is near the boundaries of this volume. When a second scan is made, the target may very well appear in a neighboring volume, and hence be referred to a different storage bin, where it will not compare with the data stored from the first scan.

In the case of the scintillation test the number of lost targets is not a serious consequence, since the prime concern is to reduce the number of targets to be examined later by more refined methods. However, for filters that occur later in the chain, where the targets are subjected to more demanding criteria, each of the objects being tracked has a higher probability of being an actual threat, and it is important not to lose any through processing deficiencies.

For the case of storage-bin overflow, target loss can be very nearly eliminated by providing overflow bins. These bins hold the overflow from any of the associatively addressed bins. Since the number of overflows is small compared to the total number of returns being handled, the extra processing load is relatively small. For the case of addressing ambiguities, a provision can be made to place a given return in more than one storage bin if a possibility of an ambiguity arises.

The problem of addressing ambiguities may be greatly alleviated by computing bin addresses from radar coordinates (range and angle) rather than by rectangular (x, y, z) coordinates. (Target data will still be *stored* in rectangular coordinates in many cases, since use of these coordinates simplifies other aspects, such as tracking algorithms.)

The reason that addressing by radar coordinates alleviates the ambiguity problem has to do with the relative accuracy with which range and angle are computed, and the fact that a target falling outside the beam will be missed in any event, in which case addressing ambiguities are immaterial. The target position in space may be measured more accurately in the range dimension than in the angle

dimension. The bin in memory should therefore correspond to a volume in space that is pancake shaped—thin in the range dimension and wide in the angle dimension.

By thus suitably tailoring the volume represented by a bin, the frequency of ambiguous jumps from bin to bin will be minimized. When an ambiguity does arise, it will apply only in the range dimension, so that a target will usually only have to be stored in two bins to avoid ambiguity losses.

If an error in angle measurement or prediction occurs, it usually results in the target being missed because it was not illuminated, rather than from some ambiguity of storage.

ACKNOWLEDGMENTS

We wish to acknowledge the contributions of R. Klein, and those of the RCA-Defense Communications Systems Division who worked on the ILLIAC IV Project. This work was supported by the Air Force under Contract F30602-67-C-0211.

INTEGRATED-CIRCUIT METALIZED PLASTIC SYMMETRICAL MILLIMETER TROUGH WAVEGUIDE SYSTEM WITH NONRECIPROCAL ELEMENTS

By

C. P. Wen

RCA Laboratories
Princeton, New Jersey

Summary—From a fabrication standpoint, size reduction is no longer desirable in millimeter-wave integrated circuits. Dielectric loading becomes a liability due to radiation and loss considerations. Plastic symmetrical trough waveguides with metalized surfaces that are easily adaptable to batch fabrication techniques are most suitable for hybrid millimeter-wave integrated-circuit systems. Practical applications of the symmetrical trough waveguide for nonreciprocal devices have been experimentally demonstrated by the fabrication of a ferrite resonant isolator with a figure of merit over 50 and of a differential phase shifter at X-band frequencies.

NUMEROUS DIFFICULTIES arise when the applications of conventional microwave integrated-circuit systems are extended to the millimeter-wave region. Transmission lines fabricated on dielectric substrates, such as the strip line, the microstrip line, and the coplanar waveguide,[1] all suffer from high dielectric loss and the possibility of mode conversion or radiation unless the transverse dimensions are kept substantially smaller than one-half of a wavelength. Dielectric loading becomes a liability in this high-frequency portion of the microwave spectrum, because it is no longer desirable, from the fabrication standpoint, to further reduce the dimensions of the circuits. Precision hollow waveguides, which are not adaptable to integrated-circuit techniques, have been the sole acceptable alternative. In radar and communication systems where large numbers of identical modules are required, nonconventional open-waveguide circuits that can be batch fabricated become highly desir-

[1] C. P. Wen, "Coplanar Waveguide, A Surface Strip Transmission Line Suitable for Nonreciprocal Gyromagnetic Device Applications," *IEEE International Microwave Symposium Digest*, Dallas, Texas, May 1969.

able. In the following, an integrated-circuit metalized-plastic symmetrical millimeter trough-waveguide system is proposed. The performance of some of the passive nonreciprocal gyromagnetic elements in such a system is described.

Symmetrical trough waveguides[2,3] (shown in Figure 1) were found to be practical transmission lines in miscrowave systems. The wavelength corresponding to the lower cutoff frequency is approximately four times the depth of the slots. Since the trough waveguide is an open system, plastic casting techniques similar to those employed in phonograph record fabrication, which promise high precision and reproducibility, may be employed to form integrated-circuit systems.

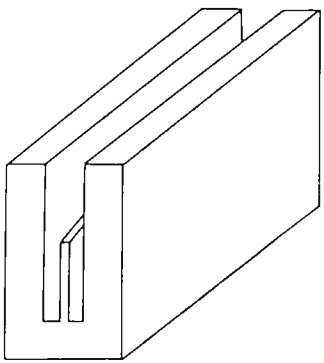


Fig. 1—Symmetrical trough waveguide.

In the millimeter-wave range, where the skin depth is very small, a few microns of highly conducting metal film (i.e., gold, silver, copper, etc.) evaporated or plated on the surfaces of the plastic blocks and the slots will be sufficient for low-loss propagation of rf signals in the troughs. Matching of the symmetrical trough waveguide into coaxial systems may be achieved by placing the center conductor of a coaxial line directly in contact with the ridge of a trough waveguide. The system will operate best if all discontinuities are made symmetrical to the center ridge. As with any open transmission-line system, a small amount of radiation leakage is unavoidable, even though its severity drops off rapidly with increasing height of the side walls. This

[2] K. S. Packard, "The Cutoff Wavelength of Trough Waveguides," *IEEE Trans. PGMTT*, Vol. 6, No. 4, p. 455, Oct. 1958.
[3] W. Rotman and A. A. Oliner, "Asymmetrical Trough Waveguide Antenna," *IEEE Trans. PGAP*, Vol. 7, No. 2, p. 153, April 1959.

undesirable leakage may be completely eliminated if a metallic cover-plate is placed across the top of the trough waveguide.

A symmetrical trough waveguide may simply be considered as an analog of a rectangular waveguide where one broad side wall has been opened along its center and then, with the center of the opposite wall as an axis, each half has been rotated 90° (i.e., folded back). The magnetic vector of the rf signal is nearly circularly polarized one-half way down the slots, with the sense of polarization depending on the propa-

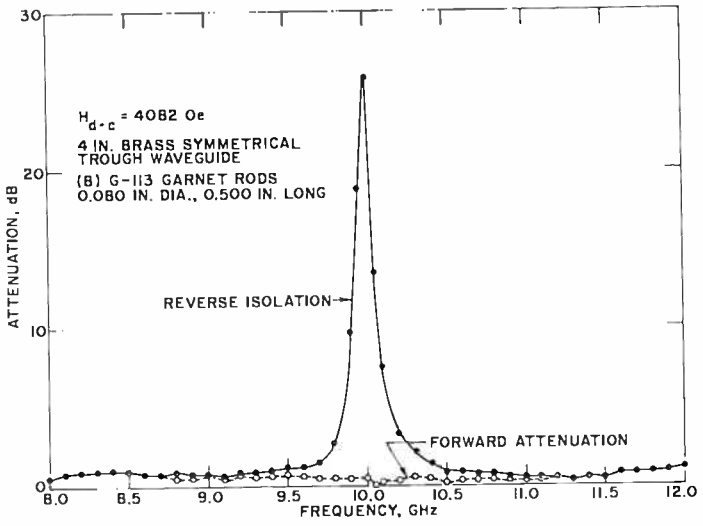


Fig. 2—Performances of a symmetrical trough waveguide ferrimagnetic resonant isolator.

gation direction of the signal. As a result of the existence of the direction-dependent circularly-polarized magnetic vector, nonreciprocal gyromagnetic elements can be built into the transmission system.

A symmetrical trough waveguide section has been built by cutting two parallel grooves in a brass block. The width of the ridge is 0.050 inch; the slots are 0.100 inch wide by 0.450 inch deep; and the side walls are 0.180 inch taller than the center ridge. Broadband matching to a 50-ohm coaxial system with OSM connectors has been achieved between 8.0 and 12.0 GHz with better than 1.5 VSWR. Little effect on the transmitted power has been observed when a metal object is placed across the top of the side walls.

Experimental results for a symmetrical trough-waveguide ferrite isolator are shown in Figure 2. Over 25.8 dB isolation accompanied by

0.5 dB forward attenuation, or a figure of merit of 51.6, has been achieved at 10 GHz when eight pieces of G-113 ferrite rods (0.020 inch in diameter by 0.500 inch long) are attached with double-sided tapes to the side walls of the slots approximately half way up from the bottom. Maximum attenuation has been measured to be less than 1 dB from 8.0 to 12.0 GHz.

In Figure 3, the configuration and the performance of a broadband differential phase shifter are shown. Minimum differential phase shift

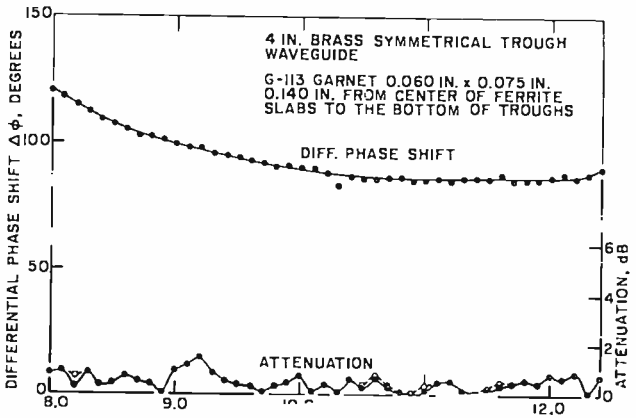


Fig. 3—Performances of a symmetrical trough waveguide ferrimagnetic differential phase shifter.

is slightly over 85°, while the insertion loss in either direction is less than 1.0 dB between 9.2 GHz and 12.4 GHz. The amount of differential phase shift varies little with frequencies beyond 10 GHz. Performance of the nonreciprocal devices can probably be improved by adapting the dielectric loading techniques.[4] Nearly identical performance has been achieved on a gold-plated plastic trough-waveguide section with physical dimensions the same as the previously described brass one. The waveguide is first formed with Castolite liquid casting plastic in a teflon mold. Surfaces of the plastic block and the slots are then plated with electroless copper followed by a few microns of gold. Total thickness of the composite metallic film is estimated to be 0.001 inch. These preliminary results have demonstrated the nonreciprocal gyromagnetic device capability of metal-coated plastic symmetrical trough waveguides and their potential applications in millimeter-wave integrated circuits.

[4] B. Lax, K. J. Button, *Microwave Ferrites and Ferrimagnetics*, McGraw-Hill Book Co., Inc., New York, 1962.

Powders of ferromagnetic and ferrimagnetic materials may be mixed in the plastic before casting so that the body of the symmetrical trough waveguide will serve as part of the closed magnetic loop required for proper operation of latching nonreciprocal gyromagnetic devices (as, e.g., the differential phase shifter shown in Figure 4). The

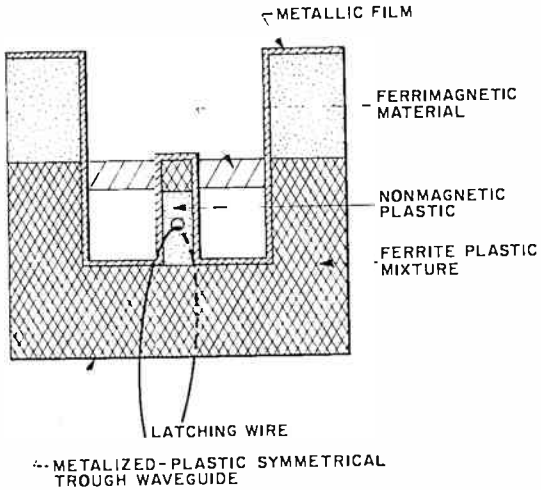


Fig. 4—A latching metalized-plastic symmetrical trough waveguide ferrimagnetic differential phase shifter.

saturated magnetization of the ferrite-plastic mixture is equal to the product of the saturation magnetization of the ferrite employed and the fraction of the ferrite powder in the mixture (by weight). Permanent magnets may be incorporated in the system in a similar fashion.

In summary, plastic symmetrical trough-waveguide networks with metalized surfaces that are easily adaptable to low-cost mold-casting techniques are found to be suitable for high-frequency microwave hybrid integrated-circuit systems. Practical applications of the symmetrical trough waveguide for nonreciprocal devices have been experimentally demonstrated by the fabrication and the operation of a ferrite resonant isolator and differential phase shifter at X-band frequencies.

ACKNOWLEDGEMENTS

The author wishes to thank H. Davis for his assistance during experimentation and W. Dimitruk, R. DeStefanto, and W. Reichert for fabricating and plating the plastic waveguides. Valuable suggestions from B. Hershenov are also acknowledged.

PREDICTION OF THE DIELECTRIC BEHAVIOR OF TITANATE DISPERSIONS

By

J. M. Guiot

RCA Consumer Electronics Division
Indianapolis, Ind.

Summary—The rapid development of thick-film technology has intensified the search for multiphase dielectric materials having optimized electric properties. This communication gives some predictions about the dielectric behavior of such heterogeneous systems based on relation between the volume fraction of dispersed phase and the dielectric constants of the mixture and of the continuous phase.

THE NEED FOR PREDICTING electrical properties of multiphase mixtures has become increasingly important because of the current interest in using such heterogeneous materials as dielectrics. Many theories have been proposed. For a comprehensive discussion of this subject, the reader is referred to the work of Van Beek.[1] An analytical solution of the general problem has not yet been obtained, since the relations of the electrical properties of such mixtures to the properties of the pure phases is very complex. Among approximate treatments involving no empirical parameters, Bruggeman's approach[2] provides a valuable and convenient method to calculate, for instance, the dielectric constants of heterogeneous mixtures consisting of dielectric powders embedded in an homogeneous and continuous medium. The equation derived by Niesel[3] (for isotropic crystals with shapes approximating spheres) and Bruggeman's formula have the same form:

$$V_2 = 1 - \frac{E_2 - E_0}{E_2 - E_1}\left(\frac{E_1}{E_0}\right)^{1/3}, \qquad (1)$$

[1] L. K. H. Van Beek, *Progress in Dielectrics*, Vol. 7 (1965).
[2] D. A. G. Bruggeman, "Berechnung Verschiedener Physikalischer Konstanten von Homogenen Substanzen," *Ann. der Physik*, Vol. 5, p. 636, 1935.
[3] W. Niesel, "Die Dielektrizitatskonsten Heterogener Mischkorper aus Isotropen und Anisotropen Substanzen," *Ann. der Physik*, Vol. 6, p. 336, 1952.

where V_2 = volume fraction of the dispersed phase, E = dielectric constant used with subscripts 0, 1, and 2 to refer to the mixture, the continuous phase, and the dispersed phase, respectively. Bruggeman used a model of regularly arranged spherical particles, and his equation should be applied to dispersions for which the particle-size distribution is broad. Only in this way may close-packing volume fractions approaching $V_2 = 1$ be achieved without contact between adjacent spheres. Satisfactory experimental confirmation of Equation (1) has been obtained by several authors.[3-6] However, disagreements can be encountered and essentially explained on the basis of a poor correspondence between Bruggeman's model and the actual configuration of experimental systems. For instance, Equation (1) does not adequately describe the experimental results in the case of random arrangements of nonspherical particles of a narrow range of sizes, especially when $E_2 \gg E_1$ and at high values of V_2. Very often porosity is present and has to be considered as an additional phase. In this respect, wetting of the dispersed phase by the surrounding continuous medium is an important factor. Limited solubility of the dispersed phase into the matrix may exist and introduce further deviation from the theoretical treatment. Despite these limitations, Bruggeman's derivation gives a good representation of experimental results in the case of titanates (either dispersed in a medium—like glass or polymer—or in the form of a sintered body, porosity being the dispersed phase up to about $V_2 = 0.3$).

The object of this short communication is to show that Bruggeman's equation represents a valuable tool in the understanding of the complex systems that are multiphase dielectric mixtures. Equation (1) has been used to calculate optimum values of E_0 achievable when titanate powders of given E_2 are dispersed in media of given E_1. The cases considered here are depicted in Figures 1, 2, and 3. Some interesting information can be derived from these results concerning the procedure for achieving high-dielectric-constant dispersions.

The interesting feature of Figure 1 is that a decrease in the amount of titanate results in a sharper decrease of the resultant dielectric constant of the mixture when $E_2 = 14{,}000$ and $E_1 = 3.0$ than in the case when $E_2 = 7{,}000$ and $E_1 = 6.0$. It should be pointed out that, in practice, below 10% by volume the amount of the "continuous phase" becomes

[4] C. A. R. Pearce, "The Electrical Conductivity and Permittivity of Mixtures, with Special Reference to Emulsions of Water in Fuel Oil," *Brit. Jour. Appl. Phys.*, Vol. 6, p. 113, 1955.

[5] B. R. Eichbaum, "Dielectric Behavior of Solids Embedded in a Homogenous Medium," *Jour. Electrochem. Soc.*, Vol. 106, p. 804, 1959.

[6] R. M. De LaRue and C. W. Tobias, "On the Conductivity of Dispersions," *Jour. Electrochem Soc.*, Vol. 106, p. 827, 1959.

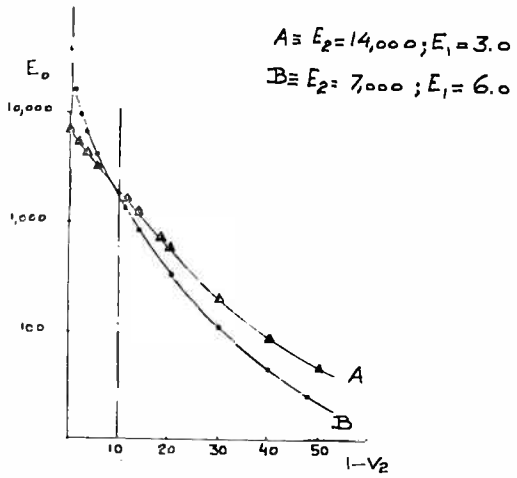


Fig. 1

too small to fill the space between the spheres (this situation is particularly critical when the range for the particle size is narrow). Therefore, in the search for mixtures with high dielectric constant, the interesting range of volume fraction of titanate lies about between 0.9 and 0.8. As shown in Figure 1, in that interval, any dispersion made of titanate $E_2 = 7{,}000$ presents a higher E_0 than dispersions obtained from titanate for which $E_2 = 14{,}000$.

The overwhelming influence of E_1, characterizing the continuous

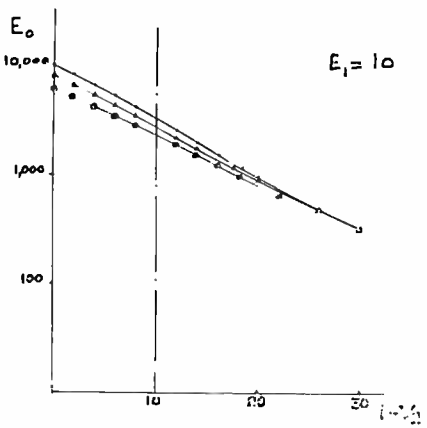


Fig. 2

phase, on the resultant dielectric constant E_0 is also shown in Figure 3. However, its influence becomes less marked when the values of V_2 are decreased. As can be seen from the graph of Figure 2, beginning at about 20% by volume continuous phase, various mixtures having an identical E_1, would differ only slightly in terms of E_0, despite large differences in E_2. A common and erroneous assumption is to consider a multiphase dielectric to behave as if it were made of capacitors in series (as many capacitors as there are phases). The importance of the

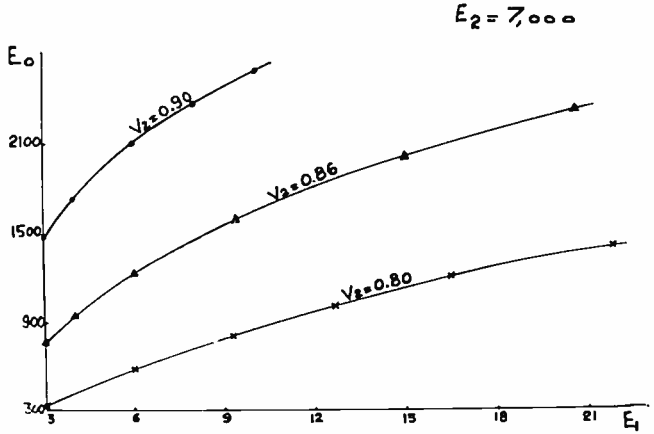


Fig. 3

theoretical results plotted in Figures 1, 2, and 3 is to show the manner in which E_0, E_1, and V_2 are related. Moreover, they provide a means of selecting a suitable system from among many possible mixtures; in this respect, Figure 1 offers a striking example.

Two Fortran IV programs were written to calculate resultant dielectric constants from Bruggeman's equation for spheres and from Niesel's equation for needles. The cubic equation for spheres was solved in the usual way by synthetic substitution. The resultant for needles was calculated using the bisection method. An attempt was made to calculate the resultant for needles from a third-degree series representation of Niesel's equation, but the resultant calculations were found very unreliable for a reasonably lengthy series. Both of the working programs were written for a time-sharing computer and both provide graphic output, either linear or semilog to the RCA Model-T display.

ACKNOWLEDGMENT

I wish to thank L. L. Simon, who developed the computer programs and the mathematical technique involved.

RCA Technical Papers

Third Quarter, 1969

Any request for copies of papers listed herein should be addressed to the publication to which credited

"Beam Shape Loss and Surveillance Optimization for Pencil Beam Arrays," P. M. Hahn and S. D. Gross, *Trans. IEEE GAES* (Correspondence) (July) 1969

"Current Nonlinearity in Electroluminescent Layers," T. N. Chin, *Jour. Appl. Phys.* (Communications) (July) 1969

"Delay Modulation," M. Hecht and A. Guida, *Proc. IEEE* (Letters) (July) 1969

"Determination of e/h, Using Macroscopic Quantum Phase Coherence in Superconductors: Implications for Quantum Electrodynamics and the Fundamental Physical Constants," B. N. Taylor and Coauthors, *Rev. Mod. Phys.* (July) 1969

"Eletronically Expanding the Citizen's World," J. Hillier and Coauthors, *IEEE Spectrum* (July) 1969

"Extremely High Power Silicon p-i-n Diode Switch," V. Stachejko, *Proc. IEEE* (Letters) (July) 1969

"A Hybrid 100-watt Linear Audio Amplifier," F. S. Kamp and J. C. Sondermeyer, *Trans. IEEE GBTR* (July) 1969

"Light-Emitting Diodes," R. S. Myers and J. O'Brien, *Electronics World* (July) 1969

"A Monolithic Integrated Circuit for Stereo Preamplifiers," L., Kaplan, *Trans. IEEE GBTR* (July) 1969

"Optimum Design of Thin-Layer GaAs Amplifiers," R. H. Dean, *Proc. IEEE* (Letters) (July) 1969

"The Tiros Decade," A Schnapf, *IEEE Spectrum* (July) 1969

"Transistor Amplifiers," D. R. Carley, *Microwave Jour.* (July) 1969

"What to Do *Before* the Lights Go Out," I. Berner, *Electronic Engineer* (IC Ideas) (July) 1969

"Zero Delay Video Systems," R. J. Butler, *Jour. SMPTE* (July) .. 1969

"Magnetophotoconductive Effects in Copper Phthalocyanine," S. E. Harrison, *Jour. Chem. Phys.* (Letters to Ed.) (1 July) 1969

"Electrical Studies of Electron-Irradiated Lithium-Containing n-Type Silicon," G. J. Brucker, *Phys. Rev.* (15 July) 1969

"Data Rate Requirements for Relay Links in Vehicular Locator Systems," G. S. Kaplan and L. Schiff, *Trans. IEEE VT* (August) 1969

"DIPNET: A General Distributed Parameter Network Analysis Program," W. N. Parker, *Trans. IEEE MTT* (August) 1969

"Effect of Substrate Imperfections on GaAs Injection Lasers Prepared by Liquid-Phase Epitaxy," H. Kressel, H. Nelson, S. H. McFarlane, M. S. Abrahams, P. LeFur, and C. J. Buiocchi, *Jour. Appl. Phys.* (August) 1969

"Electrical and Optical Properties of n-Type Si-Compensated GaAs Prepared by Liquid-Phase Epitaxy," H. Kressel and H. Nelson, *Jour. Appl. Phys.* (August) 1969

"An Electrostatic Tool for Handling Very Small Samples," F. H. Nicoll, *Rev. Sci. Instr.* (Notes) (August) 1969
"The Feasibility of a Heterode Pressure Transducer," R. M. Moore and C. J. Busanovich, *Proc. IEEE* (Letters) (August) 1969
"Growth and Characterization of Spinel Single Crystals for Substrate Use in Integrated Electronics," C. C. Wang, *Jour. Appl. Phys.* (August) 1969
"Keynote Remarks," C. A. Meyer, *Trans. IEEE EWS* (August) .. 1969
"A New Mobile Communications Systems Standard," J. R. Neubauer, *Trans. IEEE VT* (August) 1969
"Remote Color Genlock," R. J. Butler, *Jour. SMPTE* (August) 1969
"Stresses in Heteroepitaxial Layers: $GaAs_{1-x}P_x$ on GaAs," M. S. Abrahams, L. R. Weisberg, and J. J. Tietjen, *Jour. Appl. Phys.* (August) 1969
"Summary of the Case for Use of Video Tapes for Technical Training," J. W. Wentworth, *Trans. IEEE EWS* (August) 1969
"Where Is Technology Leading Communications?—A Review of the February 1969 EWS Symposium," J. Phillips, *Trans. IEEE EWS* (August) 1969
"Electric-Field-Induced Cholesteric-Nematic Phase Change in Liquid Crystals," G. H. Heilmeier and J. E. Goldmacher, *Jour. Chem. Phys.* (Letters to Ed.) (1 August) 1969
"Anisotropy of the Surface Impedance in the Surface-Sheath Regime as Evidence for Fluctuations of the Superconducting Order Parameter," G. Fischer and Coauthor, *Phys. Rev.* (10 August) 1969
"Capillary Flow of Cholestric and Smectic Liquid Crystals," W. Helfrich, *Phys. Rev. Letters* (18 August) 1969
"Effect of Fluctations in the Superconducting Order Parameter on the Tunneling Density of States," R. W. Cohen, B. Abeles, and C. R. Fuselier, *Phys. Rev. Letters* (18 August) 1969
"The Acoustoelectric Effects and the Energy Losses by Hot Electrons —Part IV Field and Temperature Dependence of Electronic Transport," A. Rose, *RCA Review* (September) 1969
"Assessment of Focused Microwave Systems for Plasma Diagnostics Against Electrostatic Probes," C. Richard and A. I. Carswell, *Jour. Appl. Phys.* (September) 1969
"Bistatic Clutter in a Moving Receiver System," E. G. McCall, *RCA Review* (September) 1969
"Characteristics of a Sealed-Off He^3-Cd^{114} Laser," J. R. Fendley, Jr., I. Gorog, K. G. Hernqvist, and C. Sun, *RCA Review* (September) 1969
"Color Correction Techniques for Television Cameras," J. F. Monahan and R. A. Dischert, *Jour. SMPTE* (September) 1969
"The CRT in Phototypesetting Systems," R. J. Klensch and E. D. Simshauser, *IEEE Spectrum* (September) 1969
"Electrical Characteristics of a Cryoelectric Memory System," A. R. Sass, W. C. Stewart, and L. S. Cosentino, *Trans. IEEE MAG* (September) 1969
"High-Efficiency Transferred Electron Oscillators," J. F. Reynolds, B. E. Berson, and R. E. Enstrom, *Proc. IEEE* (Letters) (September) 1969
"Ionospheric Phase Distortion and Faraday Rotation of Radio Waves," T. Murakami and G. S. Wickizer, *RCA Review* (September) 1969
"Low-Radiation-Noise He-Ne Laser," K. G. Hernqvist, *RCA Review* (September) 1969
"Material and Process Considerations for Monolithic Low-1/f-Noise Transistors," H. Khajezadeh and T. T. McCaffrey, *Proc. IEEE* (September) 1969
"Miniature Microstrip Circulators Using High-Dielectric-Constant Substrates," B. Hershenov and R. L. Ernst, *RCA Review* (September) 1969

"A Model for Superconductive Magnet Size Effect," (Digests), E. R. Schrader, *Trans. IEEE MAG* (September) 1969
"A Modular Three-Section 150-kG Magnet for Laboratory Use," (Digests), H. C. Schindler, *Trans. IEEE MAG* (September) .. 1969
"Note on the Discharge of Electrophotographic Material by Light," H. Kiess, *Jour. Appl. Phys.* (September) 1969
"Radar Reflectivity of Metallic Bodies of Revolution," D. A. De Wolf, *Trans. IEEE GAP* (Communications) (September) 1969
"Read-Noise Considerations in Cryoelectric Memory Systems," S. A. Keneman and L. S. Cosentino, *Trans. IEEE MAG* (September) 1969
"Semiconducting Ferromagnetic Spinels," (Digests), P. J. Wojtowicz, *Trans. IEEE MAG* (September)..... 1969
"SHF Tactical Satellite Earth Terminals," E. F. Bailey, *Signal* (September) 1969
"Sonic-Film Memory," (Digests), R. Shahbender, P. Herkart, K. Karstad, H. Kurlansik, and L. Onyshkevych, *Trans. IEEE MAG* (September) 1969
"Spectral Analysis of Turbulent Wakes," D. A. DeWolf, *RCA Review* (September) 1969
"A Temperature-Stable, Fail-Safe, Latching Ferrite TR Switch," W. W. Siekanowicz, D. J. Blattner, T. E. Walsh, and R. W. Paglione, *Trans. IEEE MTT* (September) 1969
"A Theory for the High-Efficiency Mode of Oscillation in Avalanche Diodes," A. S. Clorfeine, R. J. Ikola, and L. S. Napoli, *RCA Review* (September) 1969
"Thin Film Dielectric Materials for Microelectronics," K. H. Zaininger and C. Wang, *Proc. IEEE* (September) 1969
"Thin-Film Silicon: Preparation, Properties, and Device Applications," J. F. Allison, D. J. Dumin, F. P. Heiman, C. W. Mueller, and P. H. Robinson, *Proc. IEEE* (September) 1969
"Threshold Gate Building Blocks," S. Cohen and R. O. Winder, *Trans. IEEE GC* (September) 1969
"Size Effects in Quasiparticle Lifetime and Phonon Generation in Superconductors," A. Rothwarf, *Phys. Rev. Letters* (1 September) 1969
"Excitonic Effects in the Electroreflectance of Lead Iodide," Ch. Gähwiller and G. Harbeke, *Phys. Rev.* (15 September) 1969
"Orientation Pattern of Domains in Nematic p-Azoxyanisole," W. Helfrich, *Jour. Chem. Phys.* (Notes) (15 September) 1969
"The RCA-DME Four Phase P-MOS Universal Array and Its Application to a 6 Megahertz Digital Subsystem," M. M. Kaufman and G. E. Skorup, *Wescon Technical Papers*, Part 5, Components and Microelectronics, Session 1 1969

Patents Issued To RCA Inventors

Third Quarter 1969

July

C. H. Anderson, *Incoherent Broadband Circularly Polarized Maser Optical Pumping*, 3,454,885.

J. R. Burns and J. J. Gibson, *Complementary Field-Effect Transistor Transmission Gate*, 3,457,435.

E. W. Curtis and T. C. Jobe, *Phase Shifting Circuits for Color Television Receivers*, 3,454,708.

R. H. Dawson, N. H. Ditrick, and M. M. Mitchell, *Method of Fabricating Insulated-Gate Field-Effect Devices*, 3,455,020.

G. J. Dusheck, Jr., *Circuit That Analyzes Transient Signals in Both Time and Frequency Domains*, 3,453,540.

G. J. Gilbert, *Method of Fabricating Semiconductor Device*, 3,453,724.

G. W. Gray, *Vehicle Detector*, 3,457,547.

W. Henn, *Logic Circuit*, 3,457,434.

B. Hershenov, *Single Ground Plane Junction Circulator Having Dielectric Substrate*, 3,456,213.

K. R. Kaplan, *Threshold Gate Logic and Storage Circuits*, 3,456,126.

K. R. Kaplan, *Shift Registers Employing Threshold Gates*, 3,458,734.

J. B. Kelly, *Magnetic Tape Transport Head Assembly with Azimuth Adjustment*, 3,457,556.

M. B. Knight, *Phase Comparison Circuit*, 3,456,075.

W. F. Kosonocky, *Hybrid Transistor-Negative Resistance Diode Circuits Including Feedback*, 3,458,733.

S. Larach, R. E. Shrader, and P. N. Yocom, *Luminescent Image Device and Combinations Thereof with Optical Filters*, 3,454,715.

S. Larach, *Luminescent Image Device and Combinations Thereof with Optical Filters*, 3,454,716.

W. L. Oates, *Probe Assembly for Testing Semiconductor Wafers Including a Wafer Vibrator for Effecting Good Connections*, 3,453,545.

J. N. Pratt, *Keyed Burst Separator*, 3,454,709.

A. G. Samusenko, *Ferroelectric Control Circuits*, 3,457,455.

H. C. Schindler, *Metallic Laminated Superconductors*, 3,458,293.

F. Sterzer, *Cubic Crystal Light Modulator*, 3,454,771.

G. A. Swartz, *Grooved Bulk Semiconductor Oscillator*, 3,453,560.

August

J. Avins, *Frequency Modulation Detector Circuit Providing Balanced Detection Over a Wide Range of Signal Levels*, 3,462,694.

J. H. Avins and J. C. Miller, *Rolling Anvil Member Control Means for Serial Printer*, 3,459,287.

L. J. Bazin, *Remote Control for Deflection System of a Television Camera*, 3,463,962.

P. S. Carnt, *Color Correction System for Video Tape Recorders*, 3,461,226.

C. R. Corson, *Halftone Image Generator System*, 3,463,880.

P. E. Crookshanks and R. D. Altmanshofer, *Service Aid for Color Television Receiver*, 3,461,225.

P. Delivorias, *Complementary MOS Transistor Integrated Circuits with Inversion Layer Formed by Ionic Discharge Bombardment*, 3,461,361.

F. U. Everhard and R. E. Hartwell, *Gas Sampler*, 3,461,727.

E. Fatuzzo and W. J. Merz, *Insulating Ferroelectric Gate Adaptive Resistor*, 3,463,973.

B. F. Floden, *Slide Projector Including Two Light Paths and One Slide Magazine*, 3,462,215.

P. K-C Hsieh, *Priority Circuits*, 3,460,043.

I. Kudman, *Thermoelectric Lead Telluride Base Compositions and Devices Utilizing Them*, 3,460,996.

S. Larach and P. N. Yocom, *Phosphor and Method of Preparation Thereof*, 3,459,667.

A. Lichowsky, *Random Access Card Memory System*, 3,460,120.

A. Mayor and J. A. Dean, *Television Deflection Circuit*, 3,459,993.

P. G. McCabe, *Electron Beam Convergence Apparatus*, 3,459,989.

H. S. Miiller, R. J. Linhardt, and R. D. Sidnam, *Computer System Adapted to be Constructed of Large Integrated Circuit Arrays*, 3,462,742.

R. L. Pryor, *Integrated Memory System*, 3,460,094.

A. R. Sass and E. K. Lohner, *Circuits for Reducing Electrical Noise*, 3,460,101.

J. H. Scott, *Semiconductor Junction Device*, 3,460,007.

J. F. Slusarski and J. A. Konkel, *Continuous Video Peaking Control Circuit*, 3,461,234.

E. J. Sperber, *Preset UHF Tuning Mechanism*, 3,459,055.

C. R. Turner and P. Schiff, *Switching Type Voltage and Current Regulator, Load Therefor, and Voltage Doubling Means to Start the Load*, 3,462,643.

C. M. Wine and J. C. Miller, *Semi-Permanent Memory*, 3,462,747.

E. J. Wittmann, *Single Ended and Differential Stabilized Amplifier*, 3,460,049.

September

J. Avins and J. Craft, *Integrated Amplifier Circuit Especially Suited for High Frequency Operation*, 3,467,909.

H. R. Beelitz, *Differential Amplifier Single Ending Circuit*, 3,470,486.

J. P. Beltz, and D. E. Phelps, *Reliability Check Circuit for Optical Reader*, 3,465,130.

H. W. Bertram, O. P. Hart, and R. E. Kleppinger, *Housing and Lead Assembly for High Frequency Semiconductor Devices*, 3,465,210.

K. K. N. Chang, *Semiconductor Delay Line*, 3,466,575.

K. K. N. Chang, *High Frequency Semiconductor Systems Using Electric Fields Perpendicular to the Direction of Wave Propagation*, 3,470,375.

R. H. Cornely and W. F. Kosonocky, *Constant-Gain Low-Noise Light Amplifier*, 3,467,906.

R. Denning and C. L. Tollin, *Semiconductor Devices and Methods of Manufacture Thereof*, 3,465,209.

D. A. Donovan, *Automobile Control Manipulating Apparatus*, 3,465,577.

D. P. Dorsey and R. W. Bruce, *Synchronizing System for Television Receivers*, 3,469,032.

W. W. Evans, *Automatic Frequency Control System*, 3,469,025.

F. U. Everhard, *Electrically Operated Throttle Device*, 3,465,790.

C. E. Granger, *Current Pulse Driver with Means to Steepen and Stabilize Trailing Edge*, 3,470,391.

M. J. Grimes and H. R. Meisel, *Heat Dissipator*, 3,465,212.

G. D. Hanchett, *Switching Circuit Embodying Parallel Pair of Controlled Rectifiers*, 3,469,113.

L. A. Harwood, *Detector and AGC Circuit Stabilization Responsive to Power Supply Changes*, 3,469,195.

K. C. Hudson, *Laser Recorder with Optical Filter*, 3,465,347.

W. P. Iannuzzi, *Combination Chrominance Amplifier, Burst Amplifier, and Burst Gate Circuit for a Color Television Receiver*, 3,469,022.

S. W. Johnson, *Automatic Color Electrophotographic Apparatus*, 3,467,468.

W. E. Kinslow, *Check-Out Counter or Similar Article*, D215,380.

S. Y. Levy, *Multiprocessing Computer System with Special Instruction Sequencing*, 3,470,540.

M. H. Lewis, *Analog Position to Binary Number Translator*, 3,466,646.

J. C. Miller and C. M. Wine, *Read-Only Memories*, 3,466,625.

J. A. Milnes, *Electrical Switching System Which Defeats Automatic Fine Tuning Control Through One Switch Contact, Actuated During Manual Channel Change or Manual Fine Tuning*, 3,466,549.

R. G. Olden, *Electromagnetic Pumping Apparatus for Use in Electrophotography*, 3,469,911.

D. J. Poitras, *Amplifier Control System*, 3,469,203.

A. D. Robbi and J. W. Tuska, *Switching System for Driving Read-Write Lines in a Magnetic Memory*, 3,469,245.

E. A. O. Rutishauser, *Delay Lines*, 3,466,574.

R. B. Schilling and C. M. Wine, *Flexode Crosspoint Adaptive Matrix Circuits*, 3,465,292.

F. E. Shashoua, *Color Image Display System Utilizing a Light Valve*, 3,470,310.

B. P. Silverman and H. B. Currie, *Character Reader*, 3,465,288.

E. D. Simshauser, *Electronic Halftone Image Generator*, 3,465,199.

S. L. Starger, *Encapsulated Semiconductor Device Having Internal Shielding*, 3,469,017.

AUTHORS

JACQUES M. ASSOUR attended the Technion in Haifa, Israel, in 1955 where he completed two years of studies in the Electrical Engineering Department. In February 1958, he entered the Polytechnic Institute of Brooklyn where he received his degree of Bachelor of Electrical Engineeing (cum laude) in 1960. Upon graduation he joined RCA Laboratories in Princeton, New Jersey where he worked on the design and synthesis of antenna radiation patterns. He received his M.S.E.E. degree in 1962 and his Ph.D. degree in electrophysics in 1965 from the Polytechnic Institute of Brooklyn. From 1961 to 1966 he was engaged in the study of crystalline organic semiconductors and specifically in the area of paramagnetic resonance. Since 1966, he has been working in the field of solid-state devices. In particular, he designed and fabricated silicon PIN microwave diodes, ultrafast silicon PIN photodetectors, photosensor arrays for computer applications, and microwave avalanche diodes.

Dr. Assour is a member of the honorary societies Eta Kappa Nu and Tau Beta Pi.

JOHN E. COURTNEY received an E.E. degree from the University of Cincinnati in 1957, and an M.S.E.E. degree from the University of Pennsylvania in 1964. Since 1957 he has been employed by RCA, working in the area of sonar and radar systems. Most recently he has been concerned with the application of special-purpose and general-purpose computers to digital signal processing and radar control.

He is a member of IEEE, the Acoustical Society of America, and Eta Kappa Nu.

WILLIAM J. DENNEHY received his B.S. degree in Physics from St. Peter's College in 1963, and the M.S. degree in Physics from the Polytechnic Institute of Brooklyn in 1968. He joined RCA Laboratories in 1963, as a member of the Technical Staff assigned to the Research Training Program. In 1964, Mr. Dennehy transferred to the Radiation Physics section of the RCA Astro-Electronics Division where he worked on radiation effects in bipolar and MOS transistors. For his contributions to this work he received a NASA Award. In 1967, he initiated work in transient radiation effects within the Radiation Physics group, and since that time he has been actively engaged in this area. In addition, he is presently active in the area of the design and development of radiation hardened integrated circuits.

Mr. Dennehy is a member of the American Physical Society.

A. C. GRIMM received the BSEE from the Polytechnic Institute of Brooklyn in 1940. Upon joining RCA Electronic Components in 1940, Mr. Grimm was assigned to the development of receiving and small power tubes. In 1951, he became engineering leader, directing the advanced development effort on the RCA color kinescope. He became Product Manager for the color kinescope product group in 1954. In 1955, he was named Manager of Advanced Development in the Power Tube Group. In this capacity he was responsible, among other programs, for the development of the RCA 8568 klystron used to power the Stanford two-mile linear accelerator. From 1960 to 1967, he was Manager of engineering and Operations Manager for RCA super power tubes.

He is a member of Tau Beta Pi and Eta Kappa Nu, a Senior Member of the IEEE, and a member of the American Institute of Aeronautics and Astronautics. He is a registered professional engineer in the State of Pennsylvania.

JEAN M. GUIOT received the "Licencié en Sciences Chimiques" degree and the Sc. D. degree from the University of Brussels in 1955 and 1964, respectively. From 1964 to 1967 he worked as a postdoctoral fellow at the University of California (Lawrence Rad. Lab., Berkeley). In 1967 he joined RCA in Indianapolis, Indiana. His work has involved solid-state chemistry, high temperature electrochemistry, and kinetics of metal oxidation.

HENRY M. HALPERN received a B.E.E. degree from the City College of New York in 1950 and an M.S.E.E. degree from the University of Pennsylvania in 1962. Except for a two year period of military service at the Frankford Arsenal he has been employed by RCA, since 1951. His work has been primarily concerned with system analysis and special and general purpose computer applications to sonar, radar and communications systems. He is a member of the IEEE, and NSIA.

A. G. HOLMES-SIEDLE is a B.A. of Trinity College, Dublin (1954), and a Ph.D. of Cambridge University (1958). He conducted post-doctoral research at the Cambridge University Chemical Laboratories from 1958 to 1960, studying the transfer of energy within biological and chemical systems. From 1960 to 1962, he worked at Hawker-Siddelay Aviation on communication satellite systems and lunar vehicles. Dr. Holmes-Siedle joined RCA in 1962, and has been involved in studying both the fundamental and practical problems of radiation environments in regard to basic damage processes in materials, modification of materials for radiation hardness, the effect of the space radiation on space systems and components, and studies of the effects of intense neutron-gamma pulses on semiconductor devices. He is manager of the Radiation Effects group of the Astro-Electronics Division. Dr. Holmes-Siedle is a Fellow of the British Interplanetary Society, a member of the Institute of Electrical and Electronics Engineers, a member of the American Physical Society, and is listed in American Men of Science.

R. E. KLEPPINGER received a BS in Electrical Engineering from the University of Nebraska in 1951 and a MS in mathematics from Stevens Institute of Technology in 1955. He joined RCA in 1951 as an engineer. In this capacity he developed germanium alloy transistors for different frequencies and power and low power germanium drift transistors. In 1960 he became an Engineering Leader. As such he has been responsible for the design and processing of small signal germanium and silicon transistors, rectifiers, and MOS transistors. Mr. Kleppinger became Engineering Leader of module processing for microwave microelectronics in 1968. He is responsible for developing processes for integrated passive circuits, physical circuit layouts, and assembling both integrated and hybrid microwave circuits.

Mr. Kleppinger is a Senior member of the IEEE.

ROBERT D. LOHMAN received a B.S. degree in Electrical Engineering from Norwich University in 1949 and an M.S. degree in Electrical Engineering from North Carolina State College in 1951. In June 1951, he joined the RCA Laboratories. From 1951 to 1956, he was engaged in research in the areas of basic semiconductor noise phenomena, transistor circuit development, color television display systems, and information theory. In 1956, he joined the RCA Semiconductor and Materials Division as an applications engineer. In 1957, he became Manager of Applications for computer devices. In 1960, Mr. Lohman was promoted to Engineering Manager, Computer Products, and in 1963 was made Manager of Integrated Circuit Engineering. In June 1966, he returned to RCA Laboratories as Group Head, Integrated Electronics Research. His present position is Head, Information Sciences Systems.

BARRY S. PERLMAN received the B. E. E. from the City College of New York and the M. S. E. E. from Brooklyn Polytechnic Institute in 1961 and 1964, respectively. He is presently pursuing a Ph. D. in Electrophysics at Brooklyn Polytechnic Institute, with emphasis on solid state physics and advanced network theory. During 1961 he was a design and development trainee with RCA Defense Electronic Products. His work included signal processing, radar receivers, and parametric devices. From 1961 to 1968 he was a member of the Advanced Communications Laboratory in New York City, becoming a senior member of the technical staff in 1966. He was primarily concerned with advanced receiver techniques such as the development of all-solid-state microwave troposcatter and relay subsystems, airborne X-band receivers, integrated circuits, high level parametric circuits, and dielectric and superconducting microwave filters. In June of 1968 he joined the Advanced Technology Laboratory at the RCA David Sarnoff Research Center, Princeton, New Jersey. He is presently working on transferred electron oscillators and amplifiers.

Mr. Perlman is a member of the IEEE and NYSSPE.