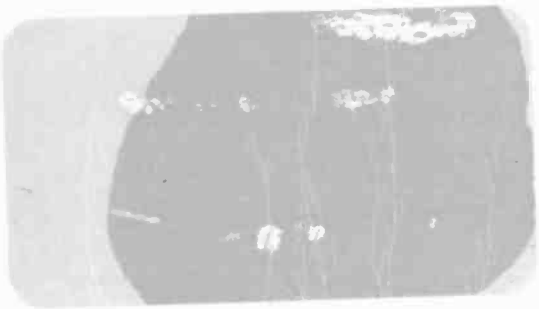
RCA Review

September 1970 Volume 31 No. 3



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.

RCA Review

A technical journal published quarterly by RCA Research and Engineering in cooperation with the subsidiaries and divisions of RCA.

Contents

September 1970 Volume 31 No. 3

Operational Dependence of the Direct-Tunneling Mode MNOS Memory Transistor on the SiO_2 Layer Thickness

E. C. Ross, A. M. Goodman, and M. T. Duffy

RCA Laboratories, Princeton, N. J.

Abstract—The results of the work presented in this paper show conclusively that the operating characteristics of the direct-tunneling mode MNOS memory transistor depend critically on the thicknesses of the Si_3N_4 and SiO_2 layers. In terms of the threshold separation obtained for a given applied gate voltage and a fixed thickness of Si_3N_4, a thickness of 20 Å for the SiO_2 layer is near optimum. The operating characteristics obtained are shown to be relatively insensitive to the value of the ambient temperature in the range from room temperature to 125° C.

Finally, it is shown that the rate of loss of stored charge decreases as the SiO_2 thickness is increased. This result supports the model in which the mechanism responsible for charge loss in the direct-tunneling mode device is tunneling between the traps and the silicon through the SiO_2 barrier.

Introduction

The MNOS (Metal-Nitride-Oxide-Silicon) memory transistor is an insulated-gate field-effect transistor in which the gate insulator has the ability to store charge.[1-3] The mechanism responsible for charging and discharging the gate insulator from the silicon terminal is dependent on the thickness of the oxide layer. If the oxide layer is made sufficiently thin (≤ 40 Å), the charge transfer is accomplished by direct-tunneling through the oxide barrier.[4] Since tunneling probabilities are strongly dependent on the thickness of the tunneling barrier, it is to be expected that the oxide thickness will have a significant influence on the operating characteristics of these devices. This paper presents some experimental results showing the effect of varying the oxide layer thickness on the operating characteristics of direct-tunneling mode MNOS memory transistors.

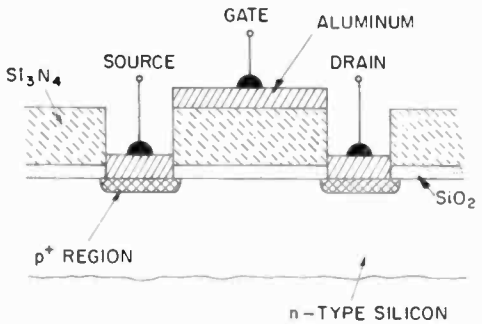


Fig. 1—Schematic diagram of a MNOS memory transistor.

Direct-Tunneling-Mode MNOS Memory Transistor

Fig. 1 is a schematic representation of an MNOS memory transistor. This transistor is a standard insulated-gate field-effect transistor, except for the gate insulator, which is a double-layer structure of silicon nitride (Si_3N_4) and silicon dioxide (SiO_2).

The MNOS structure exhibits charge storage effects that are associated with traps at or near the interface between the SiO_2 and the Si_3N_4. The amount and sign of the charge stored in the traps affects the surface potential of the silicon. The charge state of the traps, and therefore the threshold voltage of the MNOS transistor, is electronically alterable by the application of large gate fields.

Typical transfer characteristics for a direct-tunneling mode MNOS memory transistor are illustrated in Fig. 2. If a negative voltage pulse is applied to the gate with respect to the source, drain, and substrate, the transfer characteristic will be shifted to more negative values of gate voltage and will exhibit a threshold voltage V_T^-. Similarly, if a

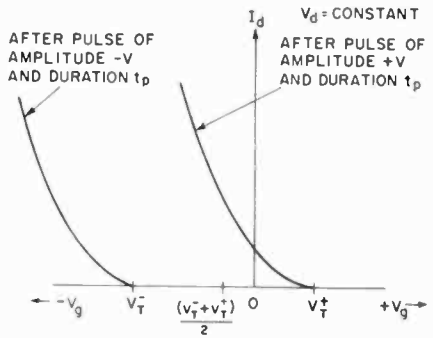


Fig. 2—Illustration of typical transfer characteristics for a direct-tunneling mode MNOS memory transistor.

positive voltage pulse is applied to the gate with respect to the source, drain, and substrate, the transfer characteristic will be shifted to more positive values of gate voltage and will exhibit a threshold voltage V_T+.

The amount of threshold voltage shift that results from the application of a given gate voltage is a function of the thickness of both the SiO_2 and the Si_3N_4 layers. Therefore, to meaningfully compare the performances of MNOS memory transistors with different SiO_2 layer thicknesses, the experimental results must be scaled in an appropriate way to eliminate the effects associated with various Si_3N_4 layer thicknesses on the experimental samples. The expressions required in order

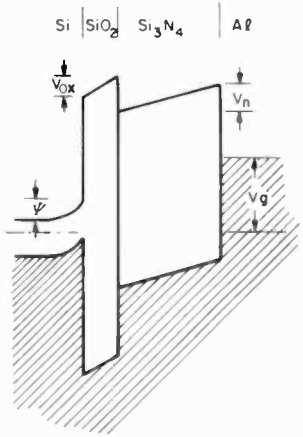


Fig. 3—Energy-band diagram of the MNOS system.

to arrive at a suitable scaling procedure can be derived from the energy-band diagram of the MNOS system shown in Fig. 3.[5] From Fig. 3,

$$V_g = V_{ox} + V_n - \phi_{ms} + \psi, \qquad [1]$$

where V_g is the applied gate voltage, V_{ox} and V_n are the voltage drops across the SiO_2 and the Si_3N_4, respectively, ψ is the surface potential of the silicon, and ϕ_{ms} is the contact potential between the metal and the silicon given by $\phi_{ms} = (\Phi_s - \Phi_m)/e$. Here, Φ_s and Φ_m are the semiconductor and metal work functions, respectively.

By application of Gauss' law at the silicon–oxide interface

$$\epsilon_{ox}E_{ox} - \epsilon_s E_s = eQ_{ss}, \qquad [2]$$

where ϵ_{ox} and ϵ_s are the dielectric constants of the oxide and the silicon,

respectively, E_{ox} and E_s are the electric fields in the oxide and silicon at the interface, respectively, e is the magnitude of the electronic charge, and Q_{ss} is the surface state charge density at the interface.

Similarly, at the Si_3N_4–SiO_2 interface

$$\epsilon_n E_n - \epsilon_{ox} E_{ox} = eQ_I, \qquad [3]$$

where ϵ_n is the dielectric constant of the Si_3N_4, E_n is the electric field in the Si_3N_4 and Q_I is the charge density at the SiO_2–Si_3N_4 interface.

In the absence of space charge in the bulk of the individual insulators

$$V_{ox} = -l_{ox} E_{ox} \qquad [4]$$

and

$$V_n = -l_n E_n. \qquad [5]$$

By combining Eqs. [1] through [5], we obtain

$$Q_I = \frac{1}{e}\left\{\left[1 + \frac{\epsilon_n l_{ox}}{\epsilon_{ox} l_n}\right][-\epsilon_s E_s - eQ_{ss}] + \frac{\epsilon_n}{l_n}[-V_g + \psi - \phi_{ms}]\right\}. \qquad [6]$$

Eq. [6] can be evaluated at the threshold condition if the value of E_s is known at the time the applied gate voltage is the threshold voltage V_T. In general, the value of E_s is[6]

$$E_s = \left(\frac{2en_i L_D}{\epsilon_s}\right) F(U_s, U_B), \qquad [7]$$

where

$$L_D = \left(\frac{\epsilon_s kT}{2e^2 n_i}\right)^{1/2}. \qquad [8]$$

n_i is the intrinsic carrier density, k is Boltzmann's constant, T is absolute temperature, and $F(U_s, U_B)$ is a function whose value depends on the Fermi potential of the bulk silicon and the Fermi potential at the silicon surface.

A practical definition of the threshold condition for an insulated-

gate field-effect transistor is when the concentration of minority carriers at the surface equals the impurity concentration in the bulk of the semiconductor.[7] For a p-channel transistor, this becomes

$$p = N_D, \qquad [9]$$

where p is the hole concentration at the surface of the silicon and N_D is the donor concentration in the bulk of the silicon. For a given substrate material, the required Fermi potentials can be obtained and the value of $F(U_s, U_B)$ can be calculated. The value of E_s at the threshold condition can then be computed, and for our experimental devices is found to be $E_s = 7.6 \times 10^3$ V/cm. Using this value of E_s, the amount of charge stored in the MNOS memory transistor, Q_I, can be determined from Eq. [6] and the experimentally determined transistor threshold voltage.

One scaling procedure that has practical significance is to scale the experimental results to a constant Si_3N_4 thickness. The threshold voltage scaled to any particular Si_3N_4 thickness for a given Q_I is readily obtained from Eq. [6]. The gate voltage that must be applied to obtain a given Q_I can be scaled to any particular Si_3N_4 thickness by requiring that E_n be the same before and after scaling. An expression for E_n can be obtained from Eqs. [1] through [5], and is given by

$$E_n = \frac{(-V_g - \phi_{ms} + \psi) + \dfrac{eQ_I l_{ox}}{\epsilon_{ox}}}{l_n\left(1 + \dfrac{\epsilon_n}{\epsilon_{ox}}\dfrac{l_{ox}}{l_n}\right)}. \qquad [10]$$

Preparation of Experimental Transistors

The starting material for the experimental memory transistors is bulk, n-type, (100) oriented, 8-12 ohm-cm silicon.

The source and drain regions (see Fig. 1) are defined by standard photoresist techniques and are degenerately doped p-type with boron in the usual fashion.

The silicon surface is then carefully cleaned and the thin SiO_2 layer thermally grown at 600°C in a water vapor ambient. The preparation of the SiO_2 has been described in detail elsewhere.[8]

Silicon nitride is deposited from the vapor phase reaction[9] of ammonia (NH_3) and silane (SiH_4). The deposition is performed in a resistance-heated furnace at 700°C with an ammonia-to-silane ratio

greater than 10^4:1. The resulting Si_3N_4 layer has been previously shown to be near optimum for a direct-tunneling mode MNOS transistor.[10]

Contact pad openings are made with buffered HF and standard photoresist techniques. Aluminum metallization is evaporated and defined for the gate, source, and drain contacts. A gold evaporation is performed to make contact to the substrate material. The contacts are alloyed at ~550°C to provide good ohmic characteristics.

A photomicrograph of the finished transistors is shown in Fig. 4. The experimental devices have a channel width of 5 mils and a channel length of 0.3 mil.

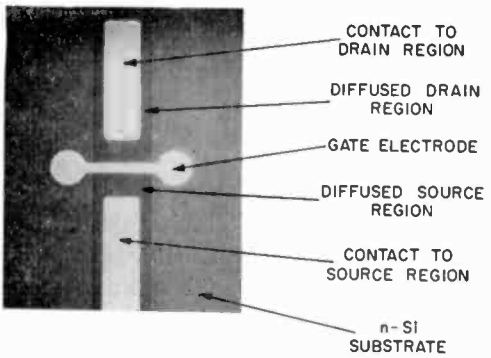


FIG. 4 PHOTOMICROGRAPH OF AN EXPERIMENTAL MNOS MEMORY TRANSISTOR.

Fig. 4—Photomicrograph of an experimental MNOS memory transistor.

Experimental Results

A. Room Temperature

A useful manner of presenting some of the experimental results is in terms of V_{TS}, the magnitude of the threshold shift, defined by

$$V_{TS} \equiv V_T^+ - V_T^-$$

and V_c, the center voltage, defined by

$$V_c \equiv \frac{V_T^+ + V_T^-}{2}.$$

Experimental results of V_{TS} as a function of applied gate voltage V_g scaled to a Si_3N_4 thickness of 300 Å are given with the SiO_2 thick-

ness as a parameter in Fig. 5. The data was obtained from positive and negative gate pulses of equal amplitude and duration 1×10^{-3} sec. For a given V_g, V_{TS} is a maximum for a SiO_2 layer thickness of 20 to 22.5 Å, and decreases for an SiO_2 layer that is either thicker or thinner than that range. The observed decrease in V_{TS} for a given V_g as the SiO_2 thickness is increased occurs because the tunneling barrier width is being increased. Therefore, the tunneling transition probabilities between the traps and the silicon are decreased. The decrease in V_{TS}

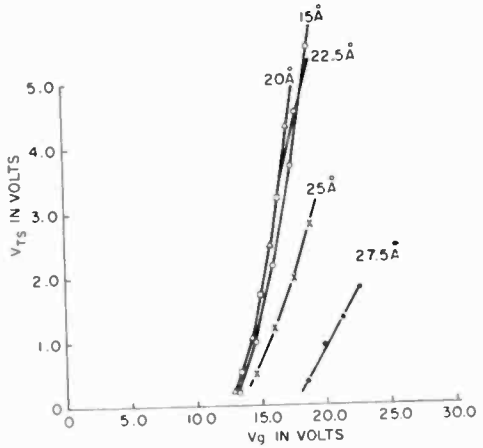


Fig. 5—V_{TS} versus V_g with the silicon dioxide thickness a parameter for a pulse duration of 1×10^{-3} second and the silicon nitride thickness scaled to 300 Å.

for a thinner SiO_2 layer is thought to be a result of the decrease in the potential drop across the SiO_2. This decreases the shift in the energy of the traps with respect to the silicon substrate for a given V_g and, therefore, decreases the tunneling probabilities, since fewer traps will be energetically permitted to communicate with the silicon.

In addition, the results presented in Fig. 5 show that for a given oxide thickness, the charge transfer characteristics are a very sensitive function of the applied voltage. Look, for example, at the results for a 20 Å layer of SiO_2. For a V_g of 15 V, V_{TS} is 1.7 V. If V_g is increased to 16.5 V, V_{TS} becomes 3.2 V. That is, in the operating region, a 10% increase in V_g nearly doubles the value of V_{TS}.

In Fig. 6, experimental results are given for V_c, the center voltage, as a function of V_g with the SiO_2 thickness as a parameter. Again, the pulse duration is 1×10^{-3} sec and the results have been scaled to a constant Si_3N_4 thickness of 300 Å.

For a given value of V_g, V_c becomes monotonically more positive for

increasing SiO_2 thickness. The importance of this result is that a device with 20 Å of SiO_2 can be operated such that in both digital threshold states the device is of the enhancement type, while a transistor with a thicker SiO_2 layer may have one or both of the threshold states such that the device is of the depletion type. Also, for a fixed SiO_2 thickness, V_c becomes more positive with increasing V_g. This means that as V_g is increased, not only does the threshold separation increase, as shown in Fig. 5, but there is a change in the "dc" level of the threshold voltages as well. These considerations are of great importance to the application of the MNOS memory transistor to memory systems.

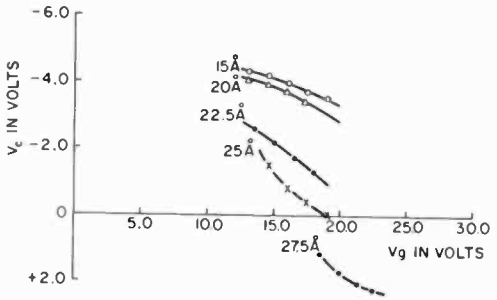


Fig. 6—V_c versus V_g with the silicon dioxide thickness a parameter for a pulse duration of 1×10^{-3} second and the silicon nitride thickness scaled to 300 Å.

The results presented in Figs. 5 and 6 indicate very clearly that a prerequisite for reproducibility of the MNOS memory transistor is a means for accurate control of the thicknesses of the gate insulators. Such control has been shown to be technologically feasible in the case of the SiO_2 layer. Using techniques that have been previously described,[8] SiO_2 layers can be repeatedly grown to within ±1 Å of the desired thickness. At present, evaluation of the reproducibility of the Si_3N_4 thickness is being performed. It is expected that techniques will be evolved to ensure the desired degree of reproducibility of this layer as well.

In order to explicitly demonstrate the effects of varying the Si_3N_4 layer thickness, V_{TS} and V_c are shown as a function of V_g in Figs. 7 and 8, respectively, with the Si_3N_4 layer scaled to thicknesses of 200 Å, 300 Å, 400 Å, and 500 Å.

The effect of pulse duration variation on the experimental results is shown in Fig. 9 for a transistor with 20 Å SiO_2 and 433 Å Si_3N_4.

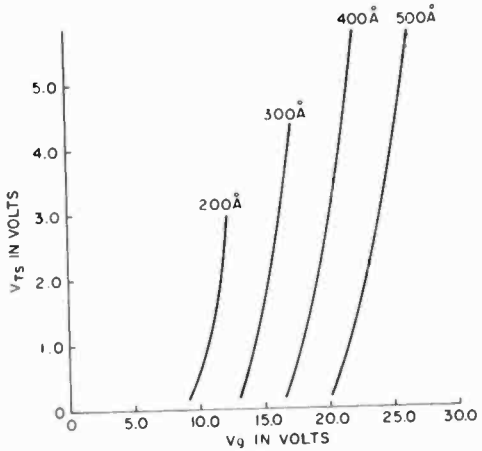


Fig. 7—V_{TS} versus V_g for a silicon dioxide thickness of 20 Å, a pulse duration of 1×10^{-3} second and the silicon nitride scaled to various thicknesses.

In this figure, $V_T{}^+$, $V_T{}^-$, and V_c are shown as a function of V_g with the pulse duration as a parameter.

First, consider the results of a fixed pulse duration of 1×10^{-3} sec. As we have seen previously, the threshold separation increases with increasing V_g and simultaneously V_c moves to a more positive value. Now, as the pulse width is reduced, we find the following for a constant V_g:

(a) The amount of charge transferred in switching from one threshold state to the other increases with increased pulse width, but is not a sensitive function of the pulse width.

(b) The center voltage, V_c, of the threshold states is a function of

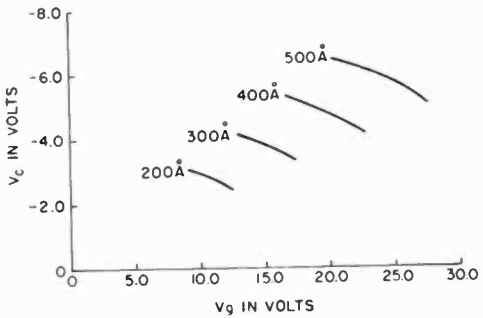


Fig. 8—V_c versus V_g for a silicon dioxide thickness of 20 Å, a pulse duration of 1×10^{-3} second and the silicon nitride scaled to various thicknesses.

V_g alone and not the pulse duration. As before, we find that V_c becomes monotonically more positive as V_g is increased.

The results of Fig. 9 graphically point out that we are not dealing with a digital element in the case of the MNOS memory transistor, but with an analog storage device that we are attempting to utilize for a digital application.

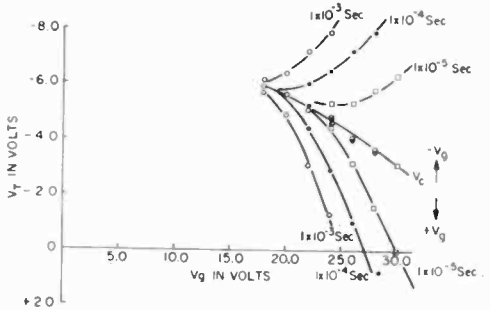


Fig. 9—V_T-, V_T+ and V_c versus V_g for an experimental transistor with 433 Å of silicon nitride and 20 Å of silicon dioxide. Pulse duration is the parameter.

B. Elevated Temperatures

The operating characteristics of a direct-tunneling mode MNOS memory transistor with a 20 Å SiO_2 layer as a function of temperature in the range from room temperature to 125°C are presented in Fig. 10. The results show that the switching characteristics are a relatively insensitive function of the ambient temperature in this experimental range. For instance, the threshold voltage separation decreases by only 12% throughout the temperature range.

Also important to the utilization of the MNOS transistor in practical applications are the charge-retention characteristics once a threshold condition has been selected. Fig. 11 shows the normalized charge stored as a function of time with ambient temperature as a parameter. The data are taken by setting the transistor into the more negative threshold state with a gate pulse of −24 V, 1×10^{-3} sec duration, then grounding the gate to the source, drain, and substrate and periodically measuring the threshold voltage as a function of time. The charge associated with the threshold voltage is then normalized by dividing it by its initial value Q', which is determined from a measurement taken a few seconds after storage.

The data shown in Fig. 11 are from the same device at 27°C and 125°C, and are found to be independent of temperature throughout this

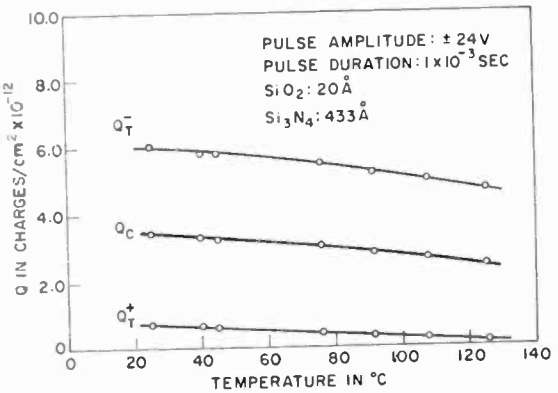


Fig. 10—Operating characteristics of a direct-tunneling mode MNOS memory transistor as a function of temperature.

range. In addition, after a brief transient period, the amount of charge stored is proportional to the logarithm of the time that has elapsed since storage, with a rate of charge loss of 2.5% per decade of time. From this rate of charge loss, the amount of the initial charge remaining can be calculated to be 85% five *weeks* after storage and 80% ten *years* after storage.

The nature of the data in Fig. 11 strongly implies that the dominant mechanism responsible for the loss of charge in the direct-tunneling mode transistor is tunneling back through the SiO_2 layer. In order to examine this hypothesis, the effect upon the rate of charge loss of

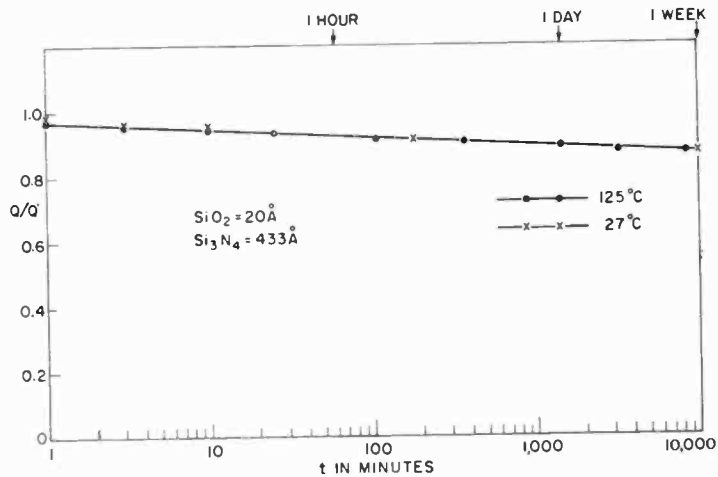


Fig. 11—Normalized charge stored versus time with ambient temperature as a parameter.

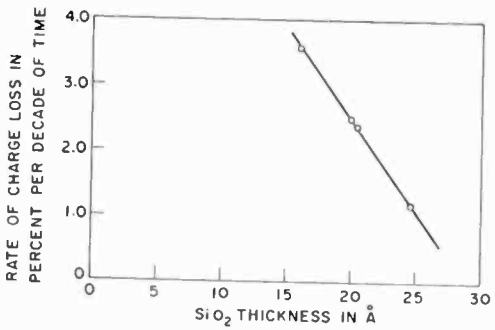


Fig. 12—Rate of loss of stored charge versus thickness of silicon dioxide.

varying the SiO_2 thickness was measured. The results are presented in Fig. 12, where the rate of loss of stored charge in per cent per decade of time is given as a function of the SiO_2 layer thickness. As can be seen, the rate of charge loss decreases with increasing SiO_2 layer thickness. These results support the concept that as the SiO_2 thickness increases, the tunneling barrier width between the traps and the silicon is increased and the tunneling probability, therefore, decreases.

Acknowledgments

The authors wish to express their appreciation to J. M. Breece, D. J. Maurizi, and R. A. Soltis for their assistance in preparing the experimental samples.

References:

[1] H. C. Pao and M. O'Connell, "Memory Behavior of an MNS Capacitor," **Appl. Phys. Letters,** Vol. 12, p. 260, 1968.

[2] J. T. Wallmark and J. H. Scott, Jr., "Switching and Storage Characteristics of MIS Memory Transistors," **RCA Review,** Vol. 30, p. 335, 1969.

[3] D. Frohman-Bentchkowsky and M. Lenzlinger, **Jour. Appl. Phys.,** Vol. 40, p. 3307, 1969.

[4] E. C. Ross and J. T. Wallmark, "Theory of the Switching Behavior of MIS Memory Transistors," **RCA Review,** Vol. 30, p. 366, 1969.

[5] A. M. Goodman, "Electron Energy-Band Diagrams of Insulators Determined from Internal Photoemission Measurements," **Proc. Third International Conf. on Photoconductivity,** Stanford University, Aug. 12-15, 1969 (in press).

[6] R. H. Kingston and S. F. Neustadter, "Calculation of the Space Charge, Electric Field, and Free-Carrier Concentration at the Surface of a Semiconductor," **Jour. Appl. Phys.,** Vol. 26, p. 718, 1955.

[7] H. K. J. Ihantola and J. L. Moll, "Design Theory of a Surface Field-Effect Transistor," **Sol. State Electron.,** Vol. 7, p. 423, 1964.

[8] A. M. Goodman and J. M. Breece, "Thin Tunnelable Layers of Silicon Dioxide Formed by Oxidation of Silicon," **Jour. Electrochem. Soc.,** Vol. 117, p. 982, 1970.

[9] G. A. Brown, W. C. Robinetter, Jr., and H. G. Carlson, "Electrical Characteristics of Silicon Nitride Films Prepared by Silane-Ammonia Reaction," **Jour. Electrochem. Soc.,** Vol. 115, p. 948, 1968.

[10] A. M. Goodman, E. C. Ross, and M. T. Duffy, "Optimization of Charge Storage in the MNOS Memory Device," **RCA Review,** Vol. 31, p. 342, 1970.

Characteristics of Coupled Microstrip Lines*

L. S. Napoli and J. J. Hughes

RCA Laboratories, Princeton, N. J.

Abstract—A measurement technique is described whereby one can accurately determine the parameters of edge-coupled microstrip lines by separately exciting one or the other mode of propagation. The parameters determined from these measurements, the odd- and even-mode characteristic impedance and guide-wavelength, correlate well with performance data on directional couplers. The measurements lead to a set of empirical design curves so that directional coupler dimensions are readily obtained by knowing only the desired coupling coefficient and center frequency.

Introduction

The microstrip transmission line has had widespread use as a result of recent advances in microwave integrated circuits. Edge-coupled microstrip lines have been used as in-line resonators, band-reject filters, and directional couplers. The measurements reported here lead to a set of quasi-empirical design curves, so that directional coupler dimensions can be readily obtained and an example of a −6-dB directional coupler design is outlined.

There are two parameters associated with the two modes of propagation in edge-coupled microstrip transmission lines—the odd- and even-mode characteristic impedances and the odd- and even-mode guide wavelengths. Knowing these parameters, one can design edge-coupled resonators and directional couplers. Under ordinary conditions, both modes of propagation are excited, and it is difficult to sort out from the data the odd- and even-mode parameters.

This paper describes a technique for determining the coupled microstrip parameters of each mode of propagation independently of

* The work reported here was sponsored in part by the U.S. Army Electronics Command, Fort Monmouth, New Jersey, Contract No. DAAB07-68-C-0296.

the other. That is, only one mode of propagation is excited at the time that its parameters are determined.

The experimental results are compared to the analytical results of Bryant and Weiss[1], and Johnson[2] and are found to differ from the theory by a small, but significant, amount. The guide-wavelength difference of the odd- and even-modes is greater, and therefore much more detrimental to coupler directivity, than the theories[1,2] predict.

The four-port scattering matrix (i.e., directional coupler characteristics) as a function of odd- and even-mode coupled-line parameters, is derived in Appendix 1. It is found to agree well with the relationship between experimentally determined scattering parameters of the lines used as a directional coupler, and the measured values of the odd- and even-mode parameters.

Appendix 2 details the method by which design data on matched directional couplers are obtained from information on mismatched couplers.

Odd- and Even-Mode Impedance Measurements

The measurement techniques described in this paper provide a means of experimentally determining the propagation parameters of the odd- and even-mode impedances, and phase velocities of edge-coupled microstrip lines. A drawing of the coupled lines with their associated fields for the two modes of propagation is shown in Fig. 1.

Data on impedance were obtained by separately exciting the odd or even mode of propagation. This technique is illustrated in Fig. 2. Power to the network analyzer is split equally into two transmission lines. The transmission path length of one line is adjustable, so that the two waves leaving the transmission test unit are of equal amplitude and of controllable relative phase. The relative phase and amplitude are measured using the harmonic frequency converter. The accuracy of the harmonic frequency converter is checked by reversing the connections to it.

The coupled microstrip line (directional coupler) is inserted between the transmission test unit and harmonic frequency converter. Either the odd mode or the even mode of propagation is individually excited by controlling the phase, i.e., by adjusting the line stretcher. The odd or even mode is excited when the harmonic frequency converted indicates 180° or 0°, respectively. A slotted line is inserted in one arm between the harmonic frequency converter and the directional coupler, and an equal length of air line is inserted in the other arm. Care is taken to ensure that each arm is essentially the same electrically; that

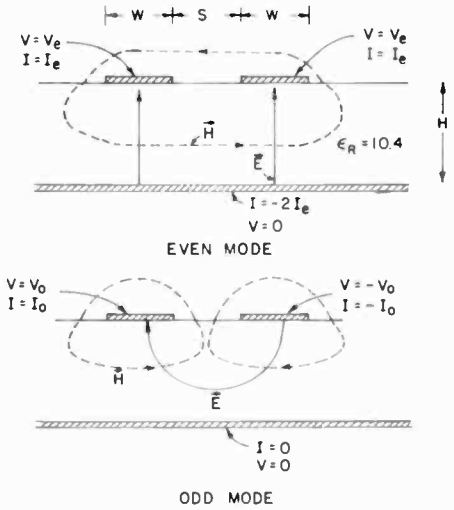


Fig. 1—Electric and magnetic fields in coupled microstrip.

is, if one arm has a length of semi-rigid coaxial line, the other arm has a nearly identical length. Small differences in length of lines are compensated for to ascertain the excitation of the odd mode at $180° \pm 2°$ or the even mode at $0° \pm 2°$. The experiment was performed near 4 GHz, where the coupling length is approximately one-quarter wavelength, because at this wavelength the effect of the odd- or even-mode characteristic impedance is maximized, and the effect of the discontinuity capacitance is minimized in the measurement of VSWR. The inversion from relatively high characteristic impedance, Z_{oe}, to relatively low characteristic impedance, Z_{oo}, is noted by a shift of the

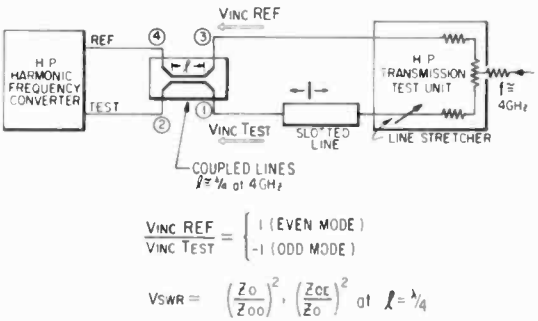


Fig. 2—Odd- and even-mode impedance measurement.

minimum voltage on the slotted line by one quarter wavelength. The impedance is determined from the VSWR measurement and corrected for transmission-line losses. Table 1 summarizes the results for directional couplers with various values of W/H and S/H, where S is the spacing between lines, W is the line width, H is the substrate thickness,* and Z_{oo} and Z_{oe} are the odd- and even-mode characteristic impedances, respectively.

Table 1—Odd- and Even-Mode Characteristic Impedances ($\epsilon_r = 10.4$)

W/H	0.96	0.863	0.785	0.368	0.176	0.107	0.1	0.254
S/H	1	0.528	0.304	0.12	0.084	0.022	0.03	0.02
Z_{oo}	40	38	34	35	37.5	28.9	33	24.9
Z_{oe}	52	58	64	86	118	142	142	108.5

Odd- and Even-Mode Guide-Wavelength

The technique employed in measuring the guide wavelength of odd and even modes for coupled microstrip lines was that of measuring the resonant frequency of lines excited in these modes. In particular, the directional couplers were severed, as shown in Fig. 3, so that two coupling gaps existed between the pair of coupled lines and the four uncoupled lines. The length of the resonator was nearly one-half wavelength at approximately 8 GHz.

The coupled pair acts as a single resonator, lightly loaded, and has two resonant frequencies associated with a half wavelength. It is apparent that the odd-mode resonant frequency should always be higher than the even-mode frequency since, upon odd-mode excitation, there is a larger proportion of the electric field energy outside the dielectric than during even-mode excitation. Any doubt can easily be overcome by purposely exciting the odd-mode resonances. This is accomplished by forcing the two inputs indicated by 1 and 3 to be out of phase for odd-mode excitation. The outputs, at 2 and 4, are detected to be out of phase for this condition. For even-mode resonance, the in-phase condition is used. End-fringe capacitance foreshortens the half-wavelength resonators by the fractional amount

* The substrate used is Alsimag 772 from American Lava. It was polished to a thickness of 0.025" and its measured dielectric constant at 1 MHz is 10.4.

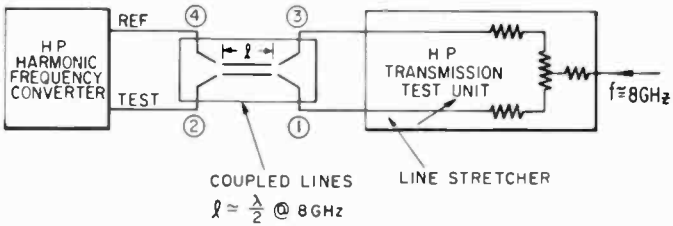


Fig. 3—Odd- and even-mode phase velocity measurement.

$$\frac{\Delta l}{l} \cong 4fC_F Z_o. \qquad [1]$$

The guide wavelength is

$$\lambda_g \cong 2l\left(1 + \frac{\Delta l}{l}\right). \qquad [2]$$

where

l = length of resonant line,

Δl = foreshortening due to end-fringe capacitance

Z_o, f = odd- or even-mode characteristic impedance or resonance frequency

C_F = experimentally determined end-fringe capacitance

Table 2 is a tabulation of the normalized guide wavelengths for the various pairs of lines tested. It is interesting to note that although

Table 2—Normalized Guide Wavelengths for Odd and Even Modes of Propagation

S/H	1	0.528	0.304	0.12	0.084
W/H	0.96	0.863	0.785	0.368	0.176
λ_{go}/λ	0.403	0.414	0.418	0.424	0.433
λ_{ge}/λ	0.355	0.356	0.357	0.372	0.383
$\lambda_{go}/\lambda_{ge}$	1.135	1.16	1.17	1.17	1.13
* $\dfrac{\lambda_{go} + \lambda_{ge}}{2\lambda_{g\infty}}$	.992	1.008	1.01	1.023	1.035
* $\lambda_{g\infty} = \lambda_g$ for $S = \infty$		$\epsilon_r = 10.4$			

the difference in the guide wavelengths is as much as 17%, the average is nearly equal to the guide wavelengths of a single uncoupled line of the same width as one of the coupled lines. The values of guide wavelengths for uncoupled lines were obtained from the paper by Caulton, Hughes, and Sobol.[3]

Comparison of Experiment and Theory

The measured impedances are somewhat lower, and the ratio of odd- to even-mode guide wavelength is higher than that predicted by Bryant and Weiss[1] and Johnson.[2] Their theory is for a zero-thickness line, whereas the experimental lines are 3 microns thick. It is doubtful, however, that the finite thickness accounts for the difference. Table 3 illustrates what seems to be a general trend. The experimental results

Table 3—Comparison of Theoretical and Experimental Results for Lightly Coupled and Heavily Coupled Lines

Lightly Coupled Line $\simeq -10$ dB $W/H = .785$, $S/H = .304$				Heavily Coupled Line $\simeq -5$ dB $Z_{oo} = 37.5$, $Z_{oe} = 118$			
	Z_{oo}	Z_{oe}	$\lambda_{go}/\lambda_{ge}$		W/H	S/H	$\lambda_{go}/\lambda_{ge}$
Bryant & Weiss	36.6*	73.2*	1.07	Bryant & Weiss**			
Johnson	36.5*	66.5*	1.12	Johnson	0.27*	0.1*	1.1
Experiment	34	64	1.17	Experiment	0.176	0.084	1.13

* The published values are for $\epsilon = 9$[1] and $\epsilon = 9.6$[2], but the assumption is made that the impedance varies as $(\epsilon + 1)^{-1/2}$, and asterisked values have been adjusted for a dielectric constant of $\epsilon = 10.4$ (measured value).

** Data not available from published curves for S/H less than 0.2.

of lightly coupled lines tend to be in closer agreement with the theory than the more heavily coupled lines. The theoretical results of Johnson are in better agreement with the experiment than those of Bryant and Weiss.

One glaring discrepancy, however, is that of the ratio of odd- to even-mode guide wavelength for the lightly coupled line. The experiment indicates a 17% difference, Bryant and Weiss predict a 7% difference, while Johnson predicts a 12% difference. As is shown in Appendix 1, differences in guide wavelength strongly affect the directivity of a directional coupler, and are therefore of great importance.

Comparison of Experiments

The analysis of Appendix 1 results in a relationship between the odd- and even-mode impedances and wave numbers and the directional coupler characteristics. At center frequency, the coupling coefficient, κ, is

$$\kappa = (Z_{oe} - Z_{oo})/(Z_{oe} + Z_{oo}), \qquad [3]$$

and the power to the directed port, relative to the power of the isolated port, or the directivity, is

$$\frac{|S_{31}|}{|S_{41}|} = \frac{4}{\pi} \frac{|S_{31}|}{\Delta(1 - |S_{31}|^2)}, \qquad [4]$$

where

$$\Delta = \frac{\lambda_{go}}{\lambda_{ge}} - 1. \qquad [5]$$

Independent measurements have been made of the coupling coefficient κ, and the odd- and even-mode impedances, Z_{oo} and Z_{oe}. Similarly, independent measurements have been made of the odd- and even-mode guide wavelength, λ_{go} and λ_{ge}, and directivity, $|S_{31}|/|S_{41}|$. The comparison is illustrated in Figs. 4 and 5. Fig. 4 shows the measured coupling coefficient versus its relationship to the measured values of odd- and even-mode impedance. The two sets of data correlate well as is evidenced by the straight-line behavior of the data points in Fig. 4. Fig. 5 is a plot of Eq. [4] with Δ as a parameter. The dashed curve of Fig. 5 is obtained from directivity measurements on directional couplers. For example, a coupler with a −10-dB coupling coefficient and 10 dB directivity indicates a 15 to 16% odd- to even-mode guide wavelength difference (Fig. 5). The same pair of lines had a measured guide wavelength difference of 17% ($S/H = 0.304$, $W/H = 0.785$, Table 2). The correlation is surprisingly good.

Directional Coupler Design Example

The average electrical length of the coupling region is, from Eq. [47], Appendix 1,

$$\left(\frac{\beta_{oo} + \beta_{oe}}{2}\right) l = \pi/2, \qquad [6]$$

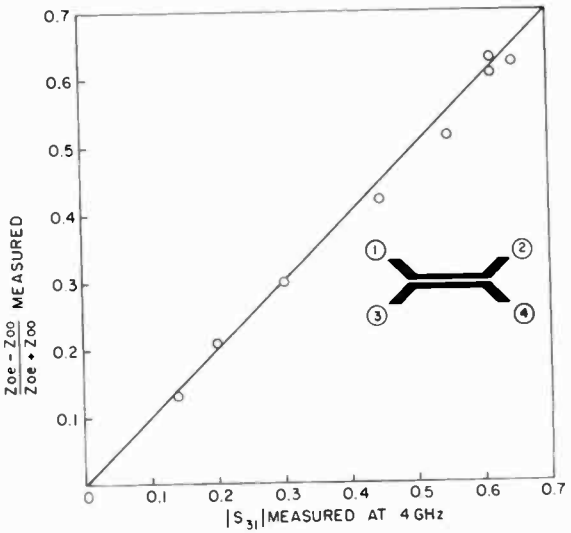


Fig. 4—Comparison of measured impedance and measured coupling coefficient.

or

$$\left(\frac{1}{\lambda_{go}} + \frac{1}{\lambda_{ge}}\right) l = 1/2. \qquad [7]$$

The guide wavelength of an uncoupled line, $\lambda_{g\infty}$, is greater than that of the even mode and less than that of the odd mode. That is,

$$\lambda_{ge} < \lambda_{g\infty} < \lambda_{go}, \qquad [8]$$

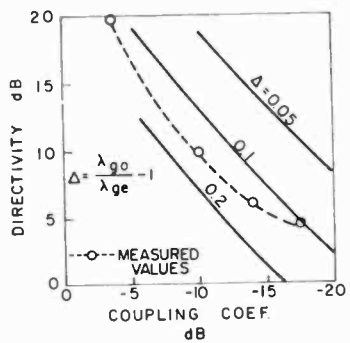


Fig. 5—Effect of phase velocity difference on coupler directivity.

where the subscript "∞" denotes the same line width as the subscript "e" or "o", but a large distance from another line. Eq. [7] reduces to

$$l \approx \frac{\lambda_{g\infty}}{4}. \qquad [9]$$

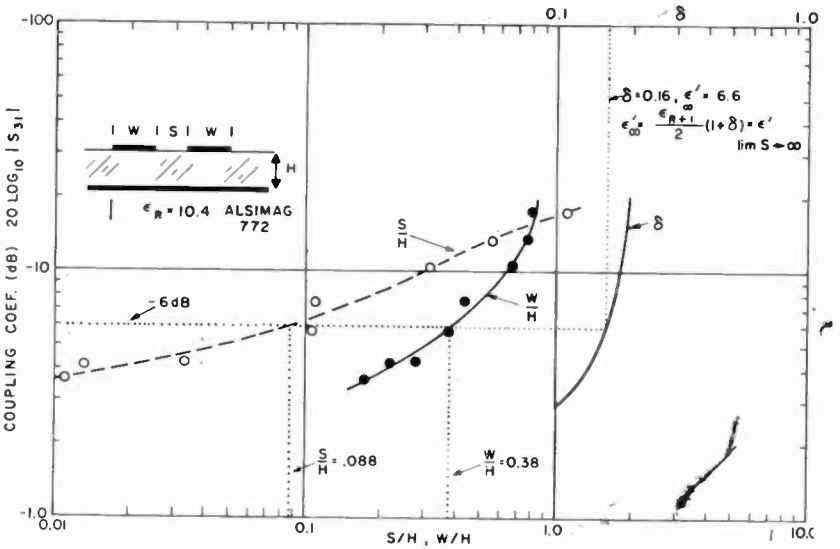


Fig. 6—Empirical design data for edge-coupled microstrip.

The dotted line on Fig. 6 for −6 dB coupling coefficient defines the normalized spacing S/H, normalized line width W/H, and fractional increase, δ, in average dielectric constant $(\epsilon_r + 1)/2$.

$$20 \log_{10} |S_{31}| = -6 \text{ dB},$$
$$S/H = 0.088,$$
$$W/H = 0.38,$$
$$\delta = 0.16.$$

The apparent dielectric constant is, for $\epsilon_r = 10.4$,

$$\epsilon' = \left(\frac{\epsilon_r + 1}{2}\right)(1 + \delta) = 6.6. \qquad [10]$$

The length of the coupling region for a center frequency of 4 GHz is

$$l = \frac{\lambda_{g\infty}}{4} = \frac{1}{4f\sqrt{\mu_0\epsilon_0\epsilon'}} = 0.728 \text{ cm}. \qquad [11]$$

Therefore for a substrate thickness of 0.025 inch,

$$H = 0.025 \text{ inch},$$
$$l = 0.287 \text{ inch},$$
$$S = 0.0022 \text{ inch},$$
$$W = 0.0095 \text{ inch}.$$

Fig. 7 details the suggested construction including the transition from an uncoupled line.

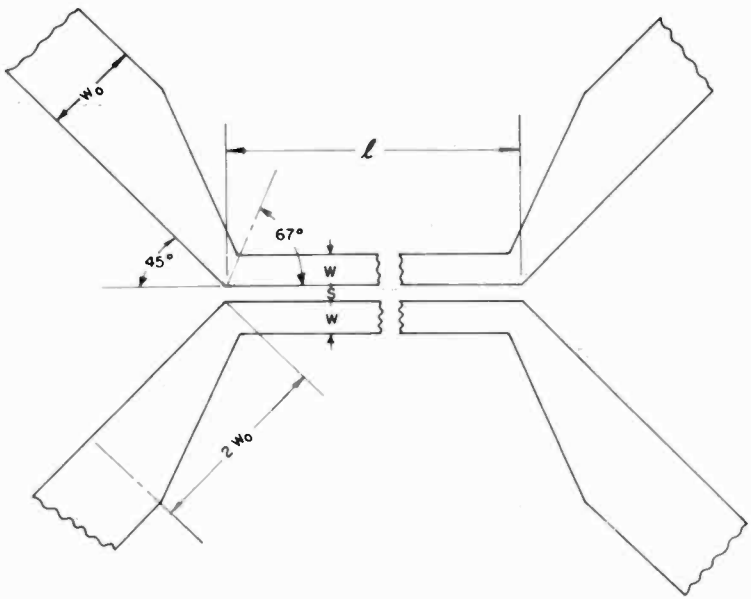


Fig. 7—Construction of microstrip coupler.

Summary

A measurement technique has been described that allows one to accurately determine the parameters of coupled microstrip lines by separately exciting either one or the other mode of propagation. The

parameters determined from these measurements, the odd- and even-mode characteristics impedance and guide wavelength, correlate well with performance data on directional couplers. The impedance data on lightly coupled lines are in fair agreement with the analytical results of Johnson, and are in accord, to a lesser degree, with the results of Bryant and Weiss. The most glaring discrepancies seem to be in the differences between the odd- and even-mode guide wavelengths. The experimental data indicate a difference between odd- and even-mode guide wavelengths of 15 to 17%, while Bryant and Weiss predict a 7% difference and Johnson, a 12% difference. In addition, for more tightly coupled lines, the data indicate that the lines should be as much as 50% wider than the analysis of Johnson predicts. The experimental lines, of course, are finite in thickness, and the theories[1,2] are for zero thickness. It is not likely, however, that a 3-micron line thickness can significantly affect the parameters of a 110-micron-wide line.

Acknowledgments

The authors would like to express their indebtedness to Alfred Schwarzmann and Bernard Hershenov for many stimulating discussions.

Appendix 1—Coupled Transmission Lines

The two identical transmission lines shown in Fig. 8 are described by the following set of equations

$$\frac{\partial V_I}{\partial z} = (M - L)\frac{\partial I_I}{\partial t} - M\frac{\partial I_{II}}{\partial t}, \qquad [12]$$

$$\frac{\partial V_{II}}{\partial z} = -M\frac{\partial I_I}{\partial t} + (M - L)\frac{\partial I_{II}}{\partial t}, \qquad [13]$$

$$\frac{\partial I_I}{\partial z} = -(C + C_c)\frac{\partial V_I}{\partial t} + C_c\frac{\partial V_{II}}{\partial t}, \qquad [14]$$

$$\frac{\partial I_{II}}{\partial z} = C_c\frac{\partial V_I}{\partial t} - (C + C_c)\frac{\partial V_{II}}{\partial t}, \qquad [15]$$

where L and C are the self-inductance and self-capacitance of the individual lines, and M and C_c are the mutual inductance and coupling capacitances between the lines.

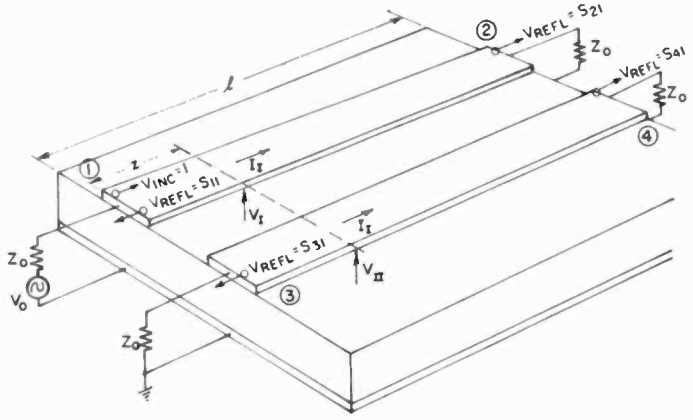


Fig. 8—Coupled microstrip lines.

If $V_I = V_{II} = V_e$ for all z, then it follows that $I_I = I_{II} = I_e$ for all z. What follows is a set of transmission line equations, Eqs. [16] and [17], describing the even mode of propagation where the wave number and characteristic impedance are $\beta_{oe} = \omega\sqrt{LC}$, $Z_{oe} = \sqrt{L/C}$. (The subscripts o and e refer to the odd and even modes.)

If $V_I = -V_{II} = V_o$, it follows that $I_I = -I_{II} = I_o$ and a similar set of transmission line equations, Eqs. [18] and [19], result with the wave number and characteristic impedance defined by, $\beta_{oo} = \omega\sqrt{(L-2M)(C+2C_c)}$, $Z_{oo} = \sqrt{(L-2M)/(C+2C_c)}$.

$$\frac{\partial V_e}{\partial z} = -L\frac{\partial I_e}{\partial t} \qquad [16]$$

$$\frac{\partial I_e}{\partial z} = -C\frac{\partial V_e}{\partial t} \qquad [17]$$

$$\frac{\partial V_o}{\partial z} = -(L-2M)\frac{\partial I_o}{\partial t} \qquad [18]$$

$$\frac{\partial I_o}{\partial z} = -(C+2C_c)\frac{\partial V_o}{\partial t} \qquad [19]$$

Assuming excitation of the form exp {jωt} and combining Eqs. [12] through [19] gives,

$$\frac{\partial^2 V_I}{\partial z^2} = -\frac{1}{2}(\beta_{oo}^2 + \beta_{oe}^2)\, V_I - \frac{1}{2}(\beta_{oo}^2 - \beta_{oe}^2)\, V_{II}, \qquad [20]$$

and

$$\frac{\partial^2 V_{II}}{\partial z^2} = -\frac{1}{2}(\beta_{oo}^2 - \beta_{oe}^2)\, V_I - \frac{1}{2}(\beta_{oo}^2 + \beta_{oe}^2)\, V_{II}. \qquad [21]$$

If solutions of the form exp $(j\Gamma z)$ are assumed, then $\Gamma = \pm \beta_{oo}, \pm \beta_{oe}$, which are simply the odd- and even-mode reflected and incident wave numbers.

The voltage on each line is a linear combination of the odd and even incident and reflected waves, and in particular,

$$V_I = A \exp\{j\beta_{oo}z\} + B \exp\{-j\beta_{oo}z\} + C \exp\{j\beta_{oe}z\} + D \exp\{-j\beta_{oe}z\}. \qquad [22]$$

From Eqs. [12]-[15], it follows that

$$V_{II} = -A \exp\{j\beta_{oo}z\} - B \exp\{-j\beta_{oo}z\} + C \exp\{j\beta_{oe}z\} + D \exp\{-j\beta_{oe}z\}. \qquad [23]$$

$$I_I = -\frac{A}{Z_{oo}} \exp\{j\beta_{oo}z\} + \frac{B}{Z_{oo}} \exp\{-j\beta_{oo}z\} - \frac{C}{Z_{oe}} \exp\{j\beta_{oe}z\} + \frac{D}{Z_{oe}} \exp\{-j\beta_{oe}z\}, \qquad [24]$$

and

$$I_{II} = \frac{A}{Z_{oo}} \exp\{j\beta_{oo}z\} - \frac{B}{Z_{oo}} \exp\{-j\beta_{oo}z\} - \frac{C}{Z_{oe}} \exp\{j\beta_{oe}z\} + \frac{D}{Z_{oe}} \exp\{-j\beta_{oe}z\} \qquad [25]$$

Both lines are of length l and are terminated in a load Z_o at both ends. The open circuit voltage at port 1 is V_o as shown in Fig. 8. From this set of boundary conditions, the voltages at the four ends are

$$V_2 = V_I(z = l) = \frac{V_o}{4}$$

$$\left[\frac{(1+\rho_o)(\rho_o-\exp\{2j\beta_{oo}l\})}{\rho_o^2-\exp\{2j\beta_{oo}l\}}+\frac{(1+\rho_e)(\rho_e-\exp\{2j\beta_{oe}l\})}{\rho_e^2-\exp\{2j\beta_{oe}l\}}\right], \quad [26]$$

$$V_2 = V_I(z=l) = \frac{V_o}{4}$$

$$\left[\frac{(\rho_o^2-1\exp\{j\beta_{oo}l\}}{\rho_o^2-\exp\{2j\beta_{oo}l\}}+\frac{(\rho_e^2-1)\exp\{j\beta_{oe}l\}}{\rho_e^2-\exp\{2j\beta_{oe}l\}}\right], \quad [27]$$

$$V_3 = V_{II}(z=0) = \frac{V_o}{4}$$

$$\left[\frac{-(1+\rho_o)(\rho_o-\exp\{2j\beta_{oo}l\})}{\rho_o^2-\exp\{2j\beta_{oo}l\}}+\frac{(1+\rho_e)(\rho_e-\exp\{2j\beta_{oe}l\})}{\rho_e^2-\exp\{2j\beta_{oe}l\}}\right], \quad [28]$$

$$V_4 = V_{II}(z=l) = \frac{V_o}{4}$$

$$\left[\frac{(1-\rho_o^2)\exp\{j\beta_{oo}l\}}{\rho_o^2-\exp\{2j\beta_{oo}l\}}-\frac{(1-\rho_e^2)\exp\{j\beta_{oe}l\}}{\rho_e^2-\exp\{2j\beta_{oe}l\}}\right], \quad [29]$$

where

$$\rho_o = -\left(1-\frac{Z_{oo}}{Z_o}\right)\Big/\left(1+\frac{Z_{oo}}{Z_o}\right)$$

$$\text{and } \rho_e = -\left(1-\frac{Z_{oe}}{Z_o}\right)\Big/\left(1+\frac{Z_{oe}}{Z_o}\right).$$

The scattering coefficients relating the reflected voltages to the incident voltages, normalized to the characteristic impedance Z_o, are as follows:

$$S_{11} = \frac{\rho_e}{2}\left(1-\frac{1-\rho_e^2}{\exp\{2j\beta_{oe}l\}-\rho_e^2}\right)+\frac{\rho_o}{2}\left(1-\frac{1-\rho_o^2}{\exp\{2j\beta_{oo}l\}-\rho_o^2}\right), \quad [30]$$

$$S_{21} = \frac{(1-\rho_e^2)\exp\{j\beta_{oe}l\}}{2(\exp\{2j\beta_{oe}l\}-\rho_e^2)}+\frac{(1-\rho_o^2)\exp\{j\beta_{oo}l\}}{2(\exp\{2j\beta_{oo}l\}-\rho_o^2)}, \quad [31]$$

$$S_{31} = \frac{\rho_e}{2}\left(1 - \frac{1-\rho_e^2}{\exp\{2j\beta_{oe}l\} - \rho_e^2}\right) - \frac{\rho_o}{2}\left(1 - \frac{1-\rho_o^2}{\exp\{2j\beta_{oo}l\} - \rho_o^2}\right), \quad [32]$$

$$S_{41} = \frac{(1-\rho_e^2)\exp\{j\beta_{oe}l\}}{2(\exp\{2j\beta_{oe}l\} - \rho_e^2)} - \frac{(1-\rho_o^2)\exp\{j\beta_{oo}l\}}{2(\exp\{2j\beta_{oo}l\} - \rho_o^2)}. \quad [33]$$

The entire matrix is defined from these four terms because of the symmetry of the network.

One of the more common uses for edge-coupled lines is the directional coupler. An ideal directional coupler has the same wave number, or phase velocity, for both the odd- and even-mode excitations. The electrical length of the coupling region is 90°, i.e., $\beta_{oo}l = \beta_{oe}l = \pi/2$. Therefore, Eqs. [30] to [33] reduce to

$$S_{11} = \frac{\rho_e}{1+\rho_e^2} + \frac{\rho_o}{1+\rho_o^2}, \quad [34]$$

$$S_{21} = \frac{-j}{2}\left(\frac{1-\rho_e^2}{1+\rho_e^2} + \frac{1-\rho_o^2}{1+\rho_o^2}\right), \quad [35]$$

$$S_{31} = \frac{\rho_e}{1+\rho_e^2} - \frac{\rho_o}{1+\rho_o^2}, \quad [36]$$

$$S_{41} = -\frac{j}{2}\left(\frac{1-\rho_e^2}{1+\rho_e^2} - \frac{1-\rho_o^2}{1+\rho_o^2}\right). \quad [37]$$

The optimum coupling condition is that the system be matched when terminated in 50 ohms; i.e., the condition $S_{11} = 0$. Eqs. [34] through [37] reduce further to

$$S_{11} = 0, \quad [38]$$

$$S_{21} = -j\frac{1-\rho^2}{1+\rho^2}, \quad [39]$$

$$S_{31} = \frac{2\rho}{1+\rho^2}, \quad [40]$$

$$S_{41} = 0, \quad [41]$$

where $\rho = \rho_e = -\rho_o$ for a matched system or $Z_o^2 = Z_{oo}Z_{oe}$.

The coupling coefficient in dB to port 3 is

$$20 \log_{10} S_{31} = 20 \log_{10} \frac{2\rho}{1+\rho^2} = 20 \log_{10} \frac{Z_{oe} - Z_{oo}}{Z_{oe} + Z_{oo}} \text{ dB} \qquad [42]$$

Therefore,

$$Z_{oo} = Z_o \left(\frac{1 - 10^{\text{dB}/20}}{1 + 10^{\text{dB}/20}} \right)^{1/2}, \qquad [43]$$

$$Z_{oe} = Z_o \left(\frac{1 + 10^{\text{dB}/20}}{1 - 10^{\text{dB}/20}} \right)^{1/2}. \qquad [44]$$

For example, if one designs a −10 dB coupler ($Z_o = 50$ ohms), then $Z_{oe} = 69.5$ ohms and $Z_{oo} = 36$ ohms.

Eqs. [43] and [44] are plotted in Fig. 9.

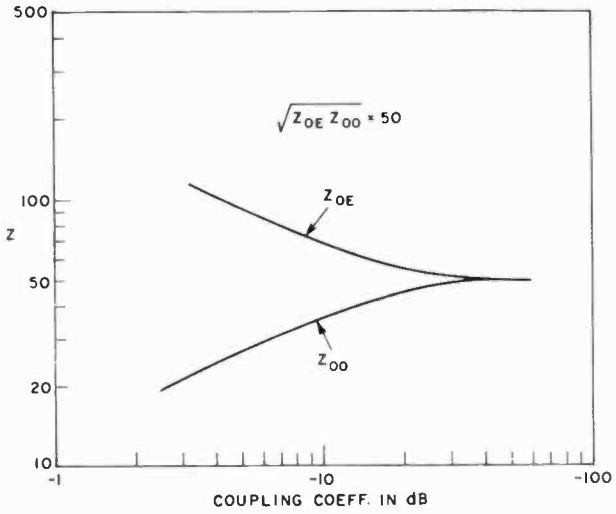


Fig. 9—Odd- and even-mode impedance for matched couplers.

If on the other hand, the odd- and even-mode phase velocities differ, the relative voltage at port four is not zero, while the voltage at port three is essentially unmodified from that shown in Eq. [40]. The relative voltage at port four, S_{41}, to a first approximation is

$$S_{41} = \frac{\pi}{4}\left(\frac{1-\rho^2}{1+\rho^2}\right)\Delta, \qquad [45]$$

where

$$\Delta = \frac{\beta_{oe}}{\beta_{oo}} - 1 = \frac{\lambda_{go}}{\lambda_{ge}} - 1, \qquad [46]$$

and

$$(\beta_{oo} + \beta_{oe})\, l = \pi. \qquad [47]$$

The directivity is then

$$\frac{|S_{31}|}{|S_{41}|} = \left(\frac{4}{\pi(1-|S_{31}|^2)}\frac{|S_{31}|}{\Delta}\right)^2. \qquad [48]$$

The directivity (the power to port three relative to that at port four) is plotted as a function of coupling coefficient in Fig. 5. The relative difference in guide wavelength, or phase velocity for the odd- and even-mode, Δ, is the parameter.

Fig. 10 shows plots of the scattering parameters for four directional couplers. Three are theoretical and the fourth is experimental. The −6 dB and −16 dB couplers are assumed to satisfy Eqs. [37] and [38] and are therefore matched, but the −10 dB coupler is assumed to be mismatched so a comparison could be made to the experimental −10 dB coupler. Note that even though the "matched" condition, $Z_o^2 = Z_{oo}Z_{oe}$, is met, there is a finite reflection coefficient for both the −6 and −16 dB couplers. This comes about because of the finite difference in odd- and even-mode phase velocities. The experimental and theoretical curves for the −10 dB couplers are in fairly good agreement. The wiggles in the experimental curves are probably due to transition discontinuities from the uncoupled line to the coupled pair of lines.

Appendix 2—Extension of Measured Data to the Case of the Matched Coupler

In order that a pair of edge coupled lines be matched at all ports, the geometric mean of the odd- and even-mode impedance should be 50 ohms. The tested couplers did not meet this requirement as can be seen from Table 4.

$$\Delta \frac{W}{H} = \frac{\partial W/H}{\partial Z_{oo}} \Delta Z_{oo} + \frac{\partial W/H}{\partial Z_{oe}} \Delta Z_{oe} + \cdots . \qquad [50]$$

Where the partial derivatives are obtained from Cohn's nomograph and evaluated at the experimental values of S/H and W/H. The change in impedance is

$$\Delta Z_{oo} = Z_{oo}' - Z_{oo}, \qquad [51]$$

$$\Delta Z_{oe} = Z_{oe}' - Z_{oe}. \qquad [52]$$

The primed numbers represent the desired values, and the unprimed are the measured values.

The corrected values of odd- and even-mode impedances are related to the coupling coefficient by

$$Z_{oo}' = 50 \left(\frac{1 - \kappa}{1 + \kappa} \right)^{1/2}, \qquad [53]$$

and

$$Z_{oe}' = 50 \left(\frac{1 + \kappa}{1 - \kappa} \right)^{1/2}, \qquad [54]$$

where κ, the voltage coupling coefficient, is defined as

$$\kappa = \kappa' = \frac{Z_{oe} - Z_{oo}}{Z_{oe} + Z_{oo}}. \qquad [55]$$

The corrected values of S/H and W/H are plotted in Fig. 6 as a function of coupling coefficient.

References:

[1] T. G. Bryant and J. A. Weiss, "Parameters of Microstrip Transmission Lines and of Coupled Pairs of Microstrip Lines," **IEEE Trans. Microwave Theory and Techniques,** Vol. MTT-16, p. 1021, Dec. 1968.

[2] K. M. Johnson, "X-Band Integrated Circuit Mixer with Reactively Terminated Image," **IEEE Jour. Solid-State Circuits,** Vol. SC-3, p. 50, June 1968.

[3] M. Caulton, J. J. Hughes, and H. Sobol, "Measurements on the Properties of Microstrip Transmission Lines for Microwave Integrated Circuits," **RCA Review**, Vol. 27, p. 377, Sept. 1966.

[4] S. B. Cohn, "Shielded Coupled-Strip Transmission Line," **IRE Trans. Microwave Theory and Techniques,** Vol. MTT-3, p. 29, Oct. 1955.

On The Fabrication of High-Efficiency Silicon Avalanche Diodes

J. M. Assour, J. Murr, Jr., and D. Tarangioli

RCA Laboratories, Princeton, N. J.

Abstract—The properties of each region of p^+-n-n^+ silicon avalanche devices were characterized using etching techniques, optical microscopy, IR transmission, and C-V measurements. Crystal defects and inhomogeneous impurity distributions were found to occur frequently in the n^+ substrate, in addition to work damage due to improper polishing procedures. The epitaxial n-layer was found to exhibit nonuniformities in thickness and resistivity, and excessive dislocations and surface defects. These difficulties were minimized by careful examination and selection of both substrates and epitaxial layers, and the use of improved polishing procedures. To avoid creating defects when introducing the p^+ layer by diffusion, a doped-oxide diffusion source was employed. A diffusion bonding technique that provides good heat-sinking and the proper electric field profile at the device surface is described in detail.

1. Introduction

A study[1] of the impurity distributions in avalanche regions of high-efficiency silicon oscillators (HEO) has shown that most of the doping profiles routinely achieved in p-n junctions result in high-efficiency performance. Indeed, L-band narrow-base p^+-n-n^+ diodes designed with abrupt, hyperabrupt, graded, and linearly graded junctions with doping densities ranging from 10^{14} to $2 \times 10^{15}/cm^3$ were operated at 40% efficiency levels. S- and C-band devices designed[2] with abrupt and linearly graded junctions were also operated at 30% efficiency levels. A comparison[1] between the experimental parameters and those derived from theory[3] has shown that the physical design parameters are not critical. Thus from the point of view of device fabrication, high-power HEO devices are simple. However, this statement is true only if good-quality epitaxial silicon layers are processed.

To operate in the high-efficiency mode, the diodes are negatively biased beyond their breakdown voltage to critical field values of the

order of 3×10^5 Vcm^{-1}. These critical fields must be sustained during the period of the plasma formation throughout relatively wide depletion regions. In L-band devices, for example, the depletion region width extends to 10 μm. Moreover, these devices must be able to respond to peak current densities of the order of 10,000 A/cm^2. In order to support these electric fields and current densities and to prevent premature avalanche breakdown and filament formation, the devices must be free of crystallographic and chemical defects and must adequately dissipate the heat generated within the junction.

In this paper, the proper characterization of the silicon active layers and the influence of their quality on the fabrication of high-power HEO devices are discussed. First, the structure of these devices and methods of realizing them using both epitaxial layers and bulk silicon wafers are presented in Section 2, followed by a detailed discussion on the properties of the silicon n+ substrate wafers in Section 3. The characteristics of the lightly doped n layers grown by epitaxy on n+ substrates is then examined in Section 4 in terms of the uniformity of their thickness, profile of their impurity distributions, and density of their chemical and structural defects. In Section 5, the properties of p+ regions made by high-temperature diffusion and their effects on the current–voltage characteristics of the devices are described. Finally, a method for processing and bonding high-power silicon avalanche devices is detailed in Section 6.

2. Structure of HEO Devices

The structure of HEO devices mostly studied by several workers[1-9] is a narrow-base p+-n-n+ diode. Our devices have a p+ region typically 3 to 5 μm wide with an average resistivity of 0.01 ohm-cm. Since the width of this region contributes to the total thermal resistance of flip-chip bonded devices, it must be minimized. The width of the n region ranges from 0.7 μm in diodes designed for 10-GHz fundamental oscillation frequency to 10 μm for 1-GHz oscillators. The average resistivity of this most active region has been varied from 0.5 to 15 ohm-cm. The n+ region is about 20 μm wide and its average resistivity is of the order of 0.01 ohm-cm. Thus the total thickness of the p+ and n layers in these devices does not exceed 15 μm. The p+-n-n+ structures have been fabricated by diffusion into epitaxial n on n+ layers and by double diffusion into bulk silicon wafers. The use of epitaxial n on n+ layers represents the most convenient and economical fabrication technique since it requires the least processing steps. Bulk silicon wafers, on the other hand, are available with excellent crystal qualities that lead to uniform and high-yield devices.

One method consists of growing *both* the p+ and n layers by epitaxy on n+ substrates. In principle, this method is the simplest and most efficient; unfortunately, in practice it requires precisely controlled processes that are not yet universal or reproducible. Another method consists of diffusing the p+ impurity into n layers grown by epitaxy on n+ substrates. A cross-section of HEO mesa diodes fabricated with the latter technique is shown in Fig. 1(a). A sketch of a typical impurity distribution in these diodes is shown in Fig. 1(b). It is

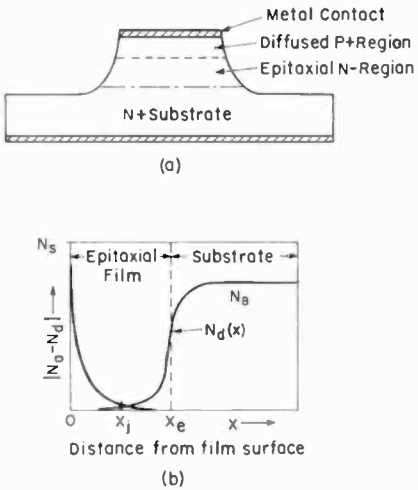


Fig. 1—(a) Typical structure and (b) impurity distribution in a p+-n-n+ mesa avalanche diode fabricated by diffusing boron into epitaxial n layers.

noteworthy that a doped-oxide source is used for these diffusions. This technique has been chiefly applied for the fabrication of our devices. Since both methods, however, require an epitaxial n layer whose characteristics are largely determined by the properties of the n+ substrate, our investigation has started with the low-resistivity n+ substrate wafers.

HEO devices have also been fabricated with bulk silicon wafers by double-diffusion techniques. The heavily doped p+ and n+ regions are diffused simultaneously on both sides of an n-type bulk silicon wafer. To achieve a thin n region, a cavity is first etched in the bulk wafer before diffusion. The cross section of the double-ended structure is shown in Fig. 2. The diameter of this experimental device is 0.025 inch. The cavity is 50 μm deep. The physical dimensions of the pertinent semiconductor regions are shown in Fig. 2(b) for a cross-sectioned

diode designed to operate at 750 MHz fundamental oscillation frequency. The double-ended devices have also a mesa configuration, similar to that shown in Fig. 1(a). The fabrication of double-diffusion bulk devices requires more critical handling procedures in comparison to those prepared with epitaxial layers.

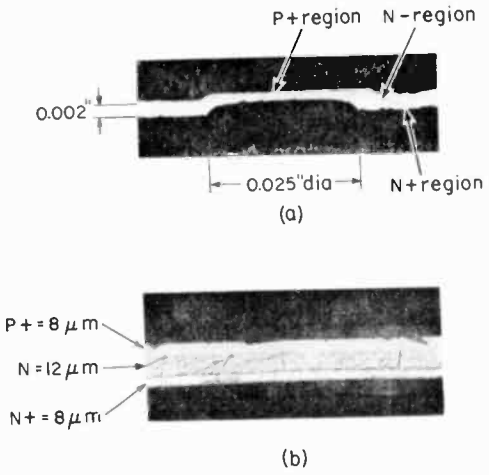


Fig. 2—(a) Typical structure and (b) width of the active silicon layers in a p+-n-n+ avalanche diode fabricated by double diffusion in bulk silicon wafers.

3. Properties of n+ Substrates

The silicon n+ substrates are routinely doped with antimony or arsenic dopants with average concentrations of approximately $5 \times 10^{18}/cm^3$ (0.01 ohm-cm). These dopants have lower diffusion constants than the boron or phosphorus atoms and, consequently, are expected to minimize autodoping and outdiffusion from the substrate during the high-temperature growth of higher-resistivity epitaxial layers. Although typical doping concentrations of the order of $5 \times 10^{18}/cm^3$ are necessary to decrease the parasitic resistance of the n+ contacting region in microwave devices, these doping levels are often the origin of several defects.

First, crystallographic imperfections due to dislocations are readily identified by a high density of etch pits as shown in Fig. 3(a) for <111> silicon substrates. The density of these dislocations might vary from $500/cm^2$ to $10^6/cm^2$. Second, chemical defects due to antimony or arsenic cones are shown in Fig. 3(b), and defects due to SiO_2 and other metallic precipitates are shown in Fig. 3(c). Spiral rings

caused by dopant concentration variation in the wafer during growth from the melt are shown in Fig. 3(d). These spiral rings are geometrically similar to those observed for oxygen swirls in higher-resistivity material. Other imperfections present in n+ substrates are defects originating from work damage during routine cutting and lapping processes applied to the silicon crystal. Invariably, the surfaces of the silicon wafers are left under stress and tension after mechanical lapping. Thus, dislocations and cracks caused by work damage are introduced in the semiconductor layers during high-temperature treatments.

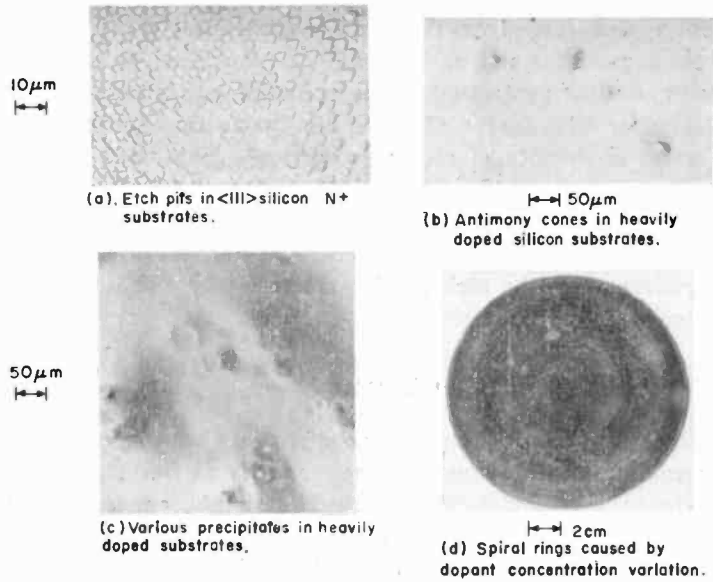


Fig. 3—Crystallographic and chemical imperfections created in heavily doped silicon n+ substrates.

All of the chemical and crystallographic imperfections existing in the n+ substrates ultimately destroy the homogeneity of the high-resistivity n layers grown by epitaxy on these substrates. Defects caused by mechanical abrasion are usually eliminated by chemical polishing of both faces of the silicon slices. On the other hand, no reliable cures are known for wafers containing the other defects. The structural and chemical imperfections observed here are by no means the only defects that exist or have been previously detected in heavily doped silicon substrates. It must be noted, however, that the proper characterization and choice of n+ substrates with relatively low den-

sity of defects have ultimately resulted in good quality epitaxial layers and high-performance HEO devices.

4. Characteristics of n Layers Grown by Epitaxy

Good-quality epitaxial layers* must satisfy the following criteria: (a) uniform thickness across the wafer, (b) uniform resistivity, and (c) crystalline perfection. Data compiled for epitaxial layers purchased from several semiconductor manufacturers have revealed the following capabilities in epitaxial technology. Layers with thickness ranging from 6 to 40 μm show a $\pm 10\%$ uniformity across each silicon wafer and $\pm 15\%$ between wafers in a given batch. Epitaxial layers with an average resistivity from 0.5 to 15 ohm-cm show a $\pm 100\%$ uniformity across each wafer and up to $\pm 300\%$ between wafers in a given batch. The crystalline perfection of the epitaxial layers varied from poor to reasonable. The degree of crystalline perfection is commonly defined in terms of density of etch pits per cm^2. This definition is believed here to be too generalized for silicon HEO and other avalanche devices. There exist a number of imperfections such as fast growing protrusions, polycrystalline silicon islands, and various precipitates that are not characterized by etch pit decorations. Their presence, however, is too important to be neglected.

A. Thickness of Epitaxial n Layers

The first step towards the characterization of epitaxial n layers is the determination of the layer thickness and its uniformity by the infrared interference method.[10] This method is simple, nondestructive, and is routinely applied to measure the layer uniformity throughout the wafer. The intensity of the fringes in the infrared reflectance spectrum yields valuable information on outdiffusion effects at the n-n+ interface. Angle-lapping and staining techniques were also applied for thickness measurements. Although these techniques accurately delineate p-n junctions, we have found them to be ineffective and inaccurate when applied to n-n+ or p-p+ boundaries. Therefore, a combination of both the above methods has been used with consistent and reasonably accurate results. The thickness of the initial epitaxial n on n+ layer is first measured by the infrared interference method. After diffusion, the width of the p+ region is determined by angle lapping and staining. The difference between the two measured values is equal to the final thickness of the n layer. In the case of deep p+ diffusion a small correction for outdiffusion from the n+ substrate is included.

* This discussion applies to both n on n+ and p on p+ epitaxial layers.

B. Resistivity of n Layers

The resistivity of n on n+ epitaxial layers is commonly determined from four-point probe measurements done on n on p+ control wafers processed with the group of substrates to be evaluated. The epitaxial n layer grown on the p+ control wafer is effectively isolated from the low-resistivity substrate by the p-n junction, thus making possible the resistivity measurement. Consequently, significant autodoping and outdiffusion effects are not easily identified in the processed n on n+ epitaxial layers unless a detailed impurity profile is determined.

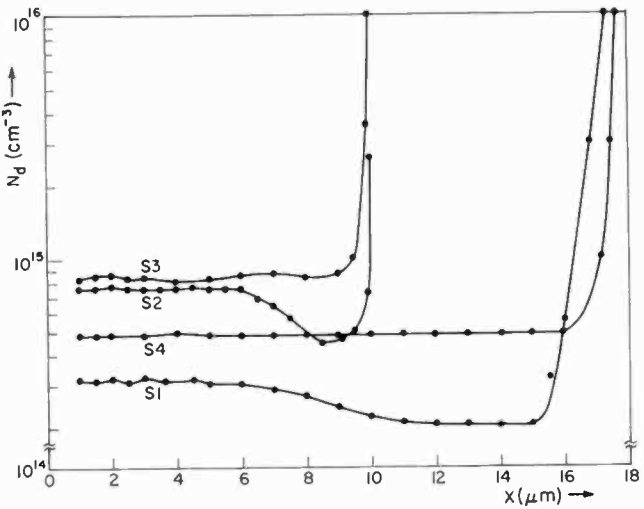


Fig. 4—Impurity distribution in n layers grown by epitaxy on n+ substrates.

The resistivity and its uniformity throughout the width of the n layers has been determined here from capacitance–voltage measurements. This measurement is performed on a control wafer selected from a batch of silicon slices. Experimental samples are prepared by forming a Schottky-barrier diode with evaporated titanium and gold films on the n layer. The *C-V* technique described by Hilibrand and Gold[11] for the determination of the impurity distribution in abrupt junction devices is applied. In general, any arbitrary impurity distribution can be evaluated with reasonable accuracy from the *C-V* data provided one boundary of the space charge region is fixed. Recently, a more convenient method has been reported by Copeland[12] for directly plotting the doping profile of semiconductor wafers. Both methods were found in reasonable agreement. Examples of impurity distributions determined from *C-V* data are shown in Fig. 4 for four different

n layers grown by epitaxy on n+ substrates. Samples S3 and S4 show excellent uniformity throughout the n layer and little or no autodoping effects at the n-n+ interface. The resistivity and width of the n layer determined from the impurity distribution are compared to those specified by the manufacturer in Table 1. The profiles of samples S1 and S2 display a higher-resistivity region near the n+ boundary. This nonuniformity in doping level of epitaxial films is quite common in many samples. Its origin might be related to oxygen contamination at the start of the epitaxial growth. These samples also show no autodoping effects at the n-n+ interface.

Table 1—Resistivity and Thickness of Epitaxial n layers.

Wafer	Manufacturer's Specifications		Impurity Distribution Data				
	Thickness (μm)	Resistivity (ohm-cm)	Thickness (μm)			Resistivity (ohm-cm)	
			Infrared	Profile	Error %	Profile	Error %
S1	15	5.6	15.2	15.5	3.3	15	170
S2	8.25	4.5	9.0	10	21	6.5	45
S3	8.8	4.7	9.35	9.5	8	6.0	28
S4	15	5.1	15.7	16	6.7	8.0	57

The evaluation of the impurity distribution in epitaxial layers yields valuable information on the resistivity, uniformity, thickness, autodoping, and outdiffusion effects at the n-n+ interface. This data, when later compared to data measured for HEO devices, reveals pertinent details that are necessary in planning a controlled fabrication process. More importantly, these valuable data allow the direct verification of proposed theories and design parameters for high-efficiency devices.

C. Defects in Epitaxial Layers

The origins of crystallographic and chemical imperfections in epitaxial silicon layers are varied and most complex. Since the cause of these defects and their prevention fall in the domain of material technology, they are outside the scope of the present study. In most cases, crystal defects in epitaxial layers are not visible to the naked eye except in extremely poor layers. Consequently, selective chemical etching[13] and optical microscopy must be used to delineate the different types of

defects according to their etch figures. Four predominant types of imperfections are shown in Fig. 5. These include stacking faults, dislocations, polycrystalline silicon islands, and fast growing protrusions.[14] The relative densities of these defects might vary from $10^2/cm^2$ to $10^5/cm^2$ and determine in large part the mechanical and electrical properties of the epitaxial silicon layers. Efforts to correlate the various types of defects with their influence on the performance of the HEO devices are underway.

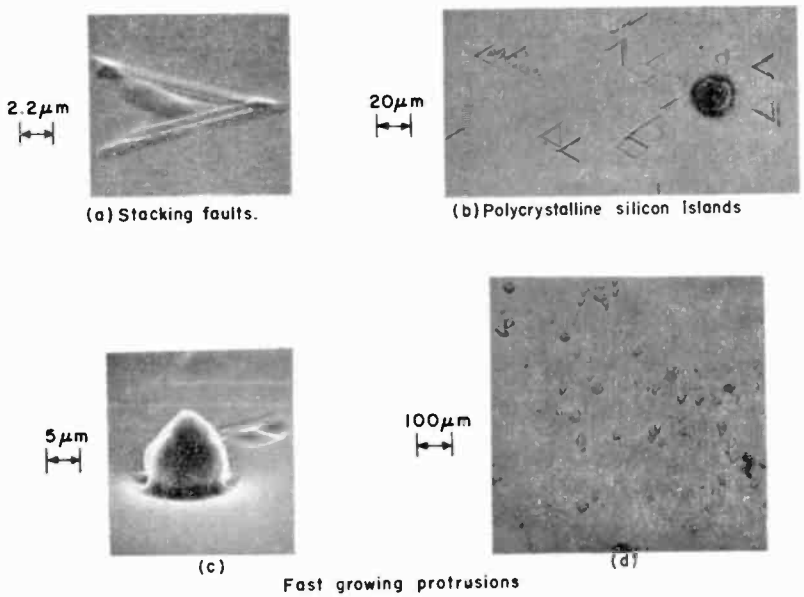


Fig. 5—Crystallographic imperfections in (111) n layers grown by epitaxy on n+ substrates.

Our results have shown that the yield of processed silicon wafers that have been properly characterized and contain a density of defects less than $500/cm^2$ has increased about 90%. Moreover, the yield of uniform avalanche devices from a single processed wafer has been maintained above 75%. These results are extremely encouraging when compared to a yield of 10% for starting silicon wafers with a density of defects exceeding $500/cm^2$.

5. Characteristics of Diffused p+ Regions

Diffused p+ regions are routinely achieved by a two-step process involving the deposition and redistribution of the p+ dopant. During the

deposition step (1000°-1100°C), a shallow layer heavily doped with boron atoms is formed at the surface of the wafer. During the redistribution step, which is generally carried out at a higher temperature of about 1200°C, the boron concentration at the surface decreases as the atoms diffuse in the active n layer. The boron source used here for the fabrication of HEO devices is a silicon dioxide film heavily doped with boron atoms. This film is formed by mixing silane and diborane gases, which are then deposited as an oxide film on silicon surfaces. The heavily doped oxide film is finally capped with a clean silicon dioxide film.

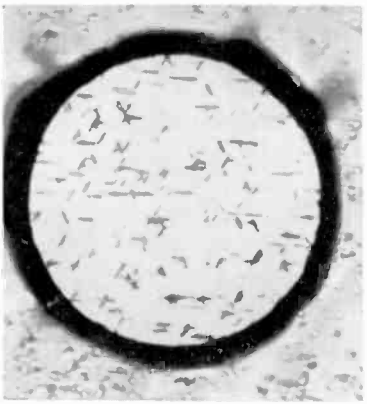


Fig. 6—Defects due to heavy boron precipitates in silicon devices subjected to a boron nitride diffusion source.

This diffusion source has several important advantages when compared to other well known p+ sources such as boron nitrides and boron silicates. The deposition step of heavily doped oxide layers is a low-temperature process which does not exceed 400°C. The purity of the oxide layers decreases considerably the amount of contaminants introduced in the starting silicon material during high-temperature treatments and thus preserves the resistivity of the n layers. High surface concentrations of dopants (of the order of $10^{20}/cm^3$) are achieved without causing surface defects in silicon due to precipitates. This is a significant advantage during the formation of shallow p+ regions. An example of the typical imperfections seen in silicon avalanche devices when subjected to a heavily doped boron nitride diffusion source is shown in Fig. 6. The defects are due to boron atoms precipitates at the silicon surface. Other defects that might not be visible to the naked eye are present in the p+ region and are deliniated by selective etching. Finally, since the heavily doped oxide layer is in

intimate contact with the silicon surface during the redistribution step, we believe that it effectively performs a gettering action of metallic impurities such as Au, Cu, etc.

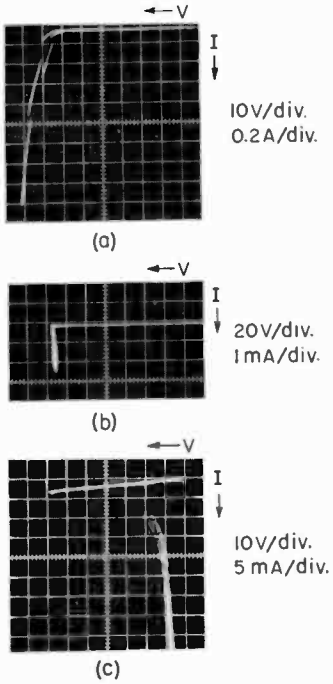


Fig. 7—Current–voltage characteristics of negatively biased p+-n-n+ silicon avalanche devices. (a) Slope discontinuities due to imperfections at the periphery of the mesa. (b) Low-frequency microplasma oscillations. (c) Current filaments in localized defects.

The influence of imperfections introduced during p+ diffusions on the *I-V* characteristics of HEO diodes and other p-n junction devices is illustrated in three examples shown in Fig. 7. Slope discontinuities in the breakdown voltage due to defects at the periphery of mesa diodes are shown in Fig. 7(a). Low-frequency microplasma oscillations shown in Fig. 7(b) are usually accompanied with light emission in discrete spots. Current filaments, shown in Fig. 7(c), result from a highly localized thermal runaway condition that is initiated at localized defects. These current filaments are predominant in nonuniform diffusion fronts at p-n interfaces. Devices displaying the above *I-V* characteristics cannot sustain the high current densities required to trigger the high-efficiency avalanche oscillations. A high-density of imperfec-

tions in epitaxial n layers will also give rise to discontinuities in the *I-V* characteristics similar to those initiated by defects in p+ regions.

6. Bonding of High-Power Silicon Avalanche Devices

The application of HEO devices as high-power oscillators and amplifiers depends on the adequate dissipation of the heat generated within the p-n junction. HEO devices operating at 40% efficiency levels dissipate 60% of the input power within the semiconductor layers, thus resulting in junction temperatures above 200°C. These excessive temperatures deteriorate the output performance and ultimately cause device failures. A generally applied technique for effective heat dissipa-

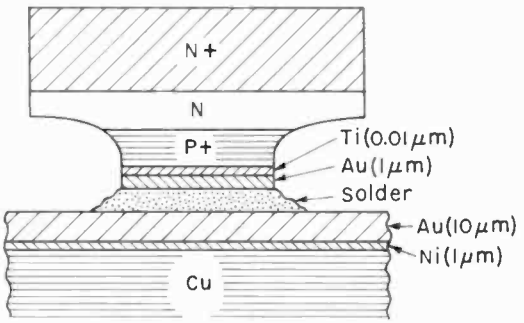


Fig. 8—Typical flip-chip mounting configuration for a high-power avalanche device.

tion consists of flip-chip bonding, i.e., for p+-n-n+ diodes the p+ surface is in contact with the heat-sink as shown in Fig. 8. The total thermal resistance of this structure is approximately $R_T = R_{Si} + R_{Ti} + R_{Au} + R_{solder} + R_{Ni} + R_{Cu}$ (°C/W); for simplicity we have assumed an ideal thermal contact exists between all metallic layers. One can readily show that for the typical structure in Fig. 8, $R_{Cu} > R_{solder} > R_{Au} > R_{Si}$. The devices in Fig. 8 are individually bonded in semiconductor packages after they have been scribed, separated, and electrically tested. After bonding, the low-frequency parameters are once more tested for reliability before determining the microwave performance of these devices.

Below we describe an extremely simple and economical flip-chip bonding technique[15,16] that includes the proper control of the electric field profile at the surface of the p+-n junction. It has been long established that in mesa devices, where the junction is exposed to the surrounding ambient, both the leakage current and breakdown voltage are

greatly influenced by surface fields. It is desirable, therefore, to design devices with electric fields at the surface significantly less than those in the bulk. Davies and Gentry[17] have calculated the electric profile at the surface of mesa devices with different bevelled junctions. A positively bevelled p+-n-n+ junction as shown in Fig. 9(d) has a linearly decreasing area going from the heavily doped p+ side of the junction to the lightly doped n side. The electric field at the surface of the mesa has its peak value near the n-n+ interface rather than the p+-n junction. Thus, a positively bevelled p+-n-n+ device is likely to show extremely low leakage current and high breakdown voltage characteristic of the bulk semiconductor. The bonding technique applied here provides both a good thermal contact and proper contour of the electric field at the surface of mesa devices.

Good thermal contact is achieved by diffusion welding between two Au layers. This technique is superior to ultrasonic bonding, which is the most widely applied technique in the semiconductor industry. In the latter method, the ultrasonic energy applied in the plane of the weld shatters the contaminating layers on the metal surfaces to be bonded together. This mechanical disruption does not create uniformly clean metal surfaces and results in decreased area of contact between the surfaces. In diffusion welding, on the other hand, the applied pressure and temperature decrease the yield strength of the materials and facilitate local plastic flow and a diffusion of Au atoms across the joint interface of the two Au layers. Thus the primary bonding mechanism in diffusion welding is the exchange of atomic sites between similar metals. The proper contour of the electric field at the surface is achieved by selective chemical etching of the silicon layers.

To fabricate p+-n-n+ diodes, p+ impurities are first diffused into n layers grown by epitaxy on n+ substrates. A thin Ti film is evaporated onto the p+ surface and it is followed by an evaporated Au film. Ti is used primarily for adherence between Si and Au. Recently, a palladium film has been evaporated on the Ti layer to minimize chemical interactions at the Ti-Au interface. Although evaporated Cr has also been used instead of Ti, we have found that Cr often interacts with heated Si surfaces and forms silicides with dendritic structures. During etching, these dendritic structures create pinholes in the silicon layers and result in device failures. The substrate used for heat dissipation is Cu electroplated with Ni and Au layers.

The silicon wafer and the Cu substrate are then placed face-to-face so that the Au layers are in contact. This assembly is heated to 150°C and pressed under 40,000 psi. The assembly is then cooled while under pressure. The silicon wafer is mechanically lapped and chemically

polished to a final thickness of about 25 μm. The silicon n+ surface is then metallized as described above. A standard photoresist technique is applied to the n+ surface for the definition of the avalanche devices. After selective delineation of the metal films, the silicon layers are etched into mesa structures. The processing sequences for a positively-bevelled p+-n-n+ HEO device are shown in Fig. 9. A photograph of a 1.25-inch-diameter silicon wafer containing 0.025-inch-diameter HEO devices processed by diffusion bonding is shown in Fig. 10.

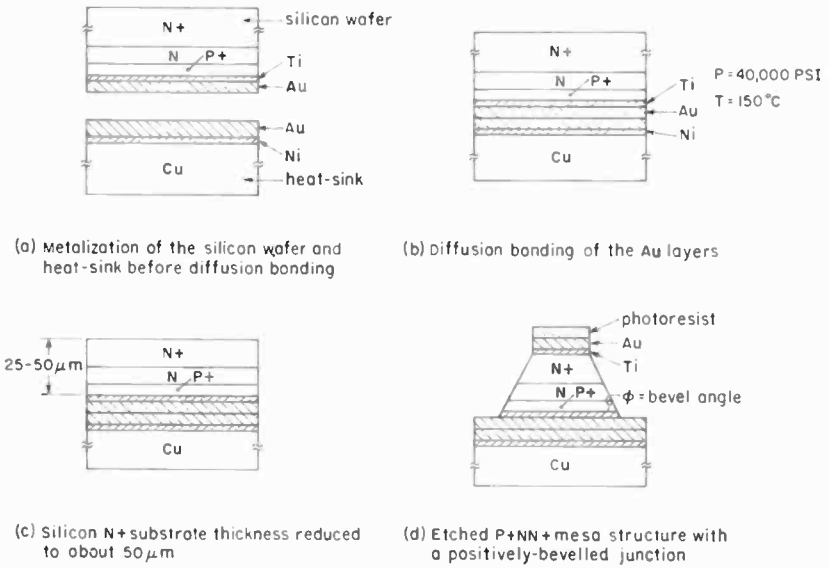


Fig. 9—Pertinent processing sequences for diffusion bonding of high-power avalanche devices.

The quality of diffusion welding between two Au layers is shown in Fig. 11. This photograph represents a cross-section of a diffusion bonded p+-n-n+ diode. The pertinent metallic layers are clearly shown. The sharpness of the Si-Au interface illustrates the uniform contact achieved throughout the area of the diode. Moreover, the absence of an interface between the evaporated Au layer on silicon and the electroplated Au layer on the copper substrate confirms the fact that diffusion welding has been achieved, i.e., Au atoms have exchanged their atomic sites by diffusion across the joint interface of the two Au layers. This condition is not achieved if pressures below 25,000 psi and temperatures below 150°C are applied during the welding process. These latter variables lead to poor thermal contact and the adherence of the Au layers.

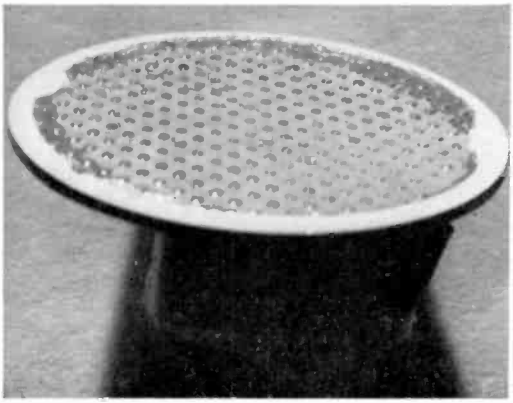


Fig. 10—Photograph of a 1.25-inch-diameter silicon wafer processed by diffusion bonding.

Bevel angles, ϕ_s, ranging from 20° to 80° in HEO p+-n-n+ mesa structures were achieved by controlling the thickness of the n+ substrate and the etch rate of the silicon layers. A profile of a 0.025-inch-diameter HEO device with a bevel angle $\phi = 55°$ is shown in Fig. 12(a). The reverse bias characteristics of this device display a breakdown voltage of 150 V and a reverse current of 10^{-8} A as shown in Fig. 12(b).

The superior thermal contact between the p-n junction and heat sink achieved by diffusion bonding was confirmed with thermal resistance data. The heat-flow resistance of avalanche diodes mounted face-up (n+ side in contact with heat sink) and flip-chip (p+ side in contact with heat sink) was determined from temperature measurements with an infrared radiometer. The diode is first heated to 100°C with zero bias. Temperature increases are then recorded as a function of dc input power. In Fig. 13, the temperature plotted as a function

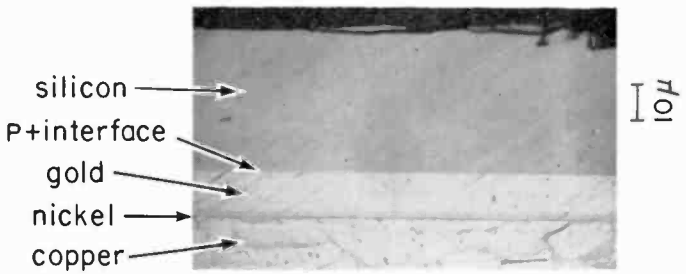
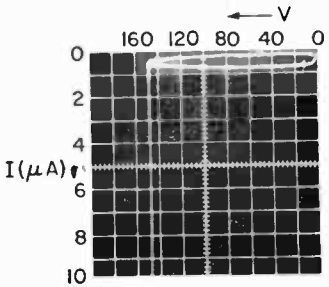


Fig. 11—Quality of diffusion bonding at the silicon-gold interface.

(a)



(b)

Fig. 12—(a) Profile of a p+-n-n+ mesa device with a bevel angle $\Phi = 55°$. (b) Reverse bias characteristics of a 0.025-inch-diameter HEO device.

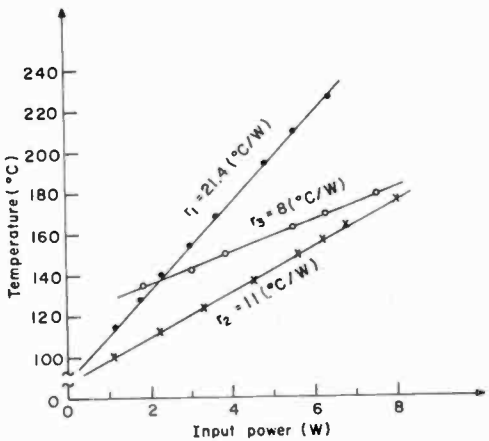


Fig. 13—Surface temperature as a function of input power as measured by an infrared radiometer for face-up and flip-chip mounted avalanche diodes.

of input power for three samples is seen to increase linearly with a slope equal to the heat flow resistance. The value $r_1 = 21°C/W$ is the thermal resistance of a 0.030-inch-diameter diode mounted face-up. A similar diameter diode mounted flip-chip by ultrasonic bonding has a lower resistance $r_2 = 11°C/W$. Finally, a 0.020-inch-diameter diode mounted flip-chip by diffusion bonding has shown the lowest value of thermal resistance, $r_3 = 8°C/W$. Further reduction in the value of r_3 can be achieved by decreasing the thickness of the gold metal layers sandwiched between silicon and the copper heat sink. HEO devices mounted by diffusion bonding were operated in the high-efficiency mode at 1% duty cycle at room temperature. Devices pulsed with a "burst" mode at 1% duty cycle were successfully operated from 25° to 100°C.

In addition to good thermal contact and proper shaping of the electric field profile at the surface of the HEO devices, the diffusion bonding technique described above offers important economical advantages. First, the whole process has been applied to bond silicon wafers with diameters up to 1.25 inches. There are no reasons to believe that larger diameter wafers could not be similarly processed. Since the Cu heat-sink represents a support for the silicon wafer, no special procedures are required to handle these wafers especially after they have been thinned down to 0.001 inch. Second, after the definition of the mesa structures, the devices are separated by scribing the Cu substrate. In other words, the silicon devices are not directly handled and, consequently, the chances of damaging the devices are considerably decreased. This technique, therefore, results in high yield and reliable devices. Third, since the n+ substrate is thinned down to about 25 μm, this surface can be effectively used as a second thermal contact for heat dissipation.

7. Conclusion

The proper characterization of the silicon n+ substrates, epitaxial n layers, and diffused p+ regions has been found to be a necessary step towards the fabrication of uniform, high-power, and high-efficiency silicon avalanche devices. Detailed measurement of the impurity distribution in the epitaxial n layers yields valuable information on the resistivity, width, impurity profile, autodoping and outdiffusion effects at the n-n+ interface. This data when later compared to those measured for HEO devices has revealed pertinent details that were found quite useful in planning a controlled fabrication process as well as direct verification of the proposed theories. Low-temperature boron-doped oxide layers were used as a diffusion source to obtain highly doped and shallow p+ regions without defects at the p-n junction.

A diffusion bonding technique that gives a good thermal contact between the silicon and a copper heat-sink and allows the proper contour of the electric field profile at the surface of the p+-n-n+ mesa devices has been successfully applied. This technique offers several economical advantages when compared to the widely accepted ultrasonic bonding methods. The diffusion bonding has been applied to silicon wafers 1.25 inches in diameter but it can accommodate larger wafers.

Acknowledgment

The authors gratefully acknowledge the assistance of C. Leuthauser for the infrared radiometer measurements.

References:

[1] J. M. Assour, "Impurity Distributions in High-Efficiency Silicon Avalanche Diode Oscillators," **IEEE Trans. Electron Devices,** Oct. 1970.

[2] J. Assour and R. V. D'Aiello, "Silicon Avalanche Diodes as Oscillators and Power Amplifiers in S-Band," **1970 ISSCC Digest.**

[3] A. S. Clorfeine, R. J. Ikola, and L. S. Napoli, "A Theory for the High-Efficiency Mode of Oscillation in Avalanche Diodes," **RCA Review,** Vol. 30, p. 397, (1969).

[4] H. J. Prager, K. K. N. Chang, and S. Weisbrod, "High-Power, High-Efficiency Silicon Avalanche Diodes at UHF and L-Band," **Proc. IEEE,** Vol. 55, p. 586 (1967).

[5] C. P. Snapp and B. Hoefflinger, "High Efficiency Avalanche Resonance Pumped Amplification," RADC-TR-69-184 (1969).

[6] R. S. Ying, R. G. Mankarious, and D. L. English, "High-Efficiency Anomalous Mode Oscillation From Silicon Impatt Diodes at 6 GHz," **1969 ISSCC Digest.**

[7] P. A. Levine and S. G. Liu, "Tunable L-band High-Power Avalanche Diode Oscillator," **1969 ISSCC Digest.**

[8] J. F. Dienst, R. V. D'Aiello, and E. E. Thomas, "The Use of a High-power, High-efficiency Avalanche Diode as a Self-pumped Parametric Amplifier," **Electronic Letters,** July 1969.

[9] S. G. Liu, "Stacked High-power Avalanche Diode Oscillators," **Proc. IEEE,** Vol. 57, p. 707 (1969).

[10] M. P. Albert and J. F. Combs "Thickness Measurement of Epitaxial Films by the Infrared Interference Method," **J. Electrochem. Soc.,** Vol. 109, p. 709 (1962).

[11] J. Hilibrand and R. D. Gold, "Determination of the Impurity Distribution in Junction Diodes From Capacity-voltage Measurements," **RCA Revlew,** Vol. 21, p. 245 (1960).

[12] J. A. Copeland, "A Technique for Directly Plotting the Inverse Doping Profile of Semiconductor Wafers," **IEEE Trans. Electron Devices,** Vol. ED-16, p. 445 (1969).

[13] Research Triangle Institute, "Integrated Silicon Device Technology, Vol. IX," ASD-TDR-63-316 (1965).

[14] J. M. Green, "Fast Growing Protrusions from Epitaxial Semiconductor Surfaces," Conference on Defects in Electronic Materials for Devices, August 1969, Boston, Mass.

[15] K. G. Hambleton, "Fabrication of Avalanche Diodes Directly on Copper Heat Sinks," IEEE workshop on avalanche devices, Dec. 1969, N. Y.

[16] R. A. Zettler and A. M. Cowley, "Batch Fabrication of Integral-Heat-Sink IMPATT Diodes," **Electronic Letters,** Vol. 5, p. 693 (1969).

[17] R. L. Davies and E. F. Gentry, "Control of Electric Field at the Surface of PN Junctions," **IEEE Trans. Electron Device,** Vol. ED-11, p. 313 (1964).

Electronically Generated Halftone Pictures*

R. J. Klensch

RCA Graphic Systems Div., Dayton, N. J.

Dietrich Meyerhofer and J. J. Walsh†

RCA Laboratories, Princeton, N. J.

Abstract—When pictorial information for printing or duplication is generated, transmitted, or manipulated electronically, it is advantageous to also transform continuous-tone images into halftones electronically. The techniques and possibilities of electronic halftone generation are the subject of this paper. Novel kinds of screening systems, made possible by the serial manipulation of the data, are discussed and compared with conventional halftones. Particular attention is paid to increasing the resolution of the final printed output, to quantizing of the grey levels, and to overcoming Moiré and color-shift problems in multicolor printing. An electronic screening system was built to test the predictions, using a modified mechanical scanner. It changes a continuous-tone picture into one consisting of array of equal-sized square dots. By using clusters of such dots, conventional-looking halftones can also be generated. Examples are shown of pictures prepared by this technique. They show that improvements over conventional halftone pictures can be obtained with particularly striking results for low-resolution printing systems.

Introduction

Many printing and duplicating systems include steps in which pictorial information is generated and processed electronically. There are examples in photocomposition, color separation, and facsimile transmission. When a picture is to be produced electronically by mechanical scanning or on a cathode-ray tube, it is practical to convert the picture into a halftone at the time the electronic information is changed into optical information, and thereby eliminate at least one photographic step.[1,2] At the same time, it becomes possible to employ halftone techniques that differ from conventional ones and obtain better looking printed pictures. Some of the possibilities are discussed in this paper.

* Presented at the 1970 GATA Annual Meeting in Boston, Mass.
† Consultant to RCA Laboratories.

Halftone Pictures

Most printing and reproduction techniques do not reproduce continuous-tone pictures well, so that halftones must be used. The reason for this is that the printed page generally exhibits only two levels of optical density: the presence or absence of ink. Wherever there is ink, its density is relatively uniform (an exception is gravure, where the density can be varied over a limited range). The most commonly used halftone is based on a square screen of uniformly spaced dots whose size varies with the picture density. It is the simplest kind of halftone to generate photographically. It also produces very pleasing pictures, as the eye integrates out the effect of the discontinuities as long as they occur in a very regular pattern.

The amount of detail that can be reproduced in a screened picture is related to the screen size or the spacing between dot centers. The smaller this spacing is, the higher will be the resolution. In turn, the dot spacing is regulated by the smallest size dot that can be printed reproducibly by the system. A regular array of such minimal dots must produce the lowest density grey level. A good picture typically requires that these dots cover 6 to 10% of the total area (optical density 0.04 above background),[3] so that the dot spacing must be 3 to 4 times minimal dot diameter.

Additional problems due to conventional screening develop in multicolor printing. When screened images are printed on top of one another, Moiré patterns result unless the screens are very nearly parallel. In conventional 4-color printing, the Moiré problem may be minimized if necessary by locating the screen directions of the cyan, magenta and black images at 60° angles to each other, and the direction of the yellow image halfway between two of them. Even if the screens were held exactly parallel, variations in the printing process would cause one the images to shift with respect to the others by a small amount. This produces varying amounts of overlap of the different inks. Since the inks are not purely subtractive colors, this causes changes in hue of the colors from sheet to sheet. This problem is reduced by having large angles between screens, since then the hue changes are averaged out over the picture.

Electronically Generated Pictures

Electronically generated pictures[1] are characterized by the fact that the picture is produced serially, one bit of picture information at a time (there may, of course, also be a relatively small number of parallel operations, each by itself serial). This is in contrast to photographic

picture reproduction which is done in parallel productions of all bits at the same time. Electronic generation is required if a picture is to be stored, transmitted, or modified electronically. It may be performed on a mechanical or other facsimile scanner or on a random-access generator, such as a cathode ray tube.

It is advantageous to be able to perform screening at the same time as the picture is being scanned so as to save processing time and, possibly, storage requirements. Since the information is processed sequentially, the screening may also be done sequentially rather than all at once. It becomes feasible to use screening arrangements that produce patterns entirely different from the conventional photographic ones. These may result in improved pictures.

Much of the electronic picture-generation equipment is based on digital processing of the information. This implies a quantization of the density values in the continuous-tone image. If this quantization is fine enough, the picture is indistinguishable from the corresponding continuous-tone (analog) picture. Storage or processing time limitations may not allow this for some applications. Also, the particular screening system itself may require quantization. These questions must be taken into consideration in designing a particular system.

Systems of Electronic Screening

The conceptually simplest kind of system is merely a duplication of the conventional photographic procedure. In certain hardware applications, the information can even be handled in analog fashion so that an exact replica of the *conventional halftone picture* is obtained. This is, for example, the case in the laser-writing color scanner application,[2] where the dot size is a function of the laser exposure time. This system has all the advantages and disadvantages of the conventional halftone including limited resolution. This also presents special problems in multicolor printing since screen rotation is not readily accomplished on a scanner.

An alternative technique of obtaining a conventional looking halftone is to "write" each halftone dot with a much smaller unit of constant size. The dot can either be formed all at once, if a random-access system is used, or it can be built up in succeeding scans on a scanner type of arrangement. Both spiral scans and parallel strokes have been proposed for this "writing." We will show examples of this technique below. For high quality work, 32 to 64 grey levels will be required in a quantized system.

Here, we are concerned mainly with a system that makes maximum use of the flexibility available in electronic screening. We call it a

constant-dot-size halftone. The image is built up out of a series of dots, each of which is formed exactly the same way in the electro-optical part of the process. This does not necessarily imply that the final printed image consists only of equal sized dots, since non-linearities may be introduced in the various steps mainly due to partial overlap of adjacent dots.

The first such system was proposed for rotogravure.[4] In it the density of dots was varied in analog fashion along the scanning direction only. This system would lead to increased reproducibility and linearity of grey scale rendition, because all gravure cavities would be of equal size. However, the random dot location produces noise in the picture and the scan lines are very obvious because the two picture directions are not handled in the same fashion.

In the present constant-dot system, the dots are all located on points of a regular raster which may be rectangular, triangular or possibly some other symmetrical arrangement. The description of the picture is then given by noting for each point whether or not it is occupied by a dot. The point raster location of the dots allows one to treat the two picture dimensions in the identical manner and, at the same time, eliminates the noise in the picture. An enlarged section of a constant-dot halftone is shown in Fig. 1. It is based on a 5×5 subarray shown by the dotted line. It includes a transition from one grey level to another. For comparison, a "conventional" halftone is shown on the same scale.

It is obvious that for maximum resolution, the dot spacing should be as small as possible. The minimum value is related to the smallest possible dot size that still produces a reliable, reproducible image on the printed output. This depends on the printing or reproduction process used, as well as on the electro-optical system that forms the dot. The dot is about the size of a highlight dot in the conventional halftone system. The dot shape will depend on the optical properties of the system and may be round, square, hexagonal or some combination. The dot spacing must be such that the maximum desired density in the printed image is reached when all points are occupied with dots, just as in the conventional system. For the typical printing process, these requirements lead to a dot spacing of about $\frac{1}{4}$ of the spacing required for a conventional screen.

In any region of the picture, the density of dots is determined in such a way that the printed image has the desired average density in that region. This is done by assigning each point in the subarray a threshold value. If the desired density at the point is higher than the threshold value of that point, then a dot is written there. The threshold

values are distributed in such a way that the screened print most closely approximates the desired picture.

One major advantage of this screening system over the conventional one is an *increase in resolution* due to the reduction in dot spacing. Consider the extreme case of a line separating a black area where all the dots are present or of maximum size from a white area without dots.

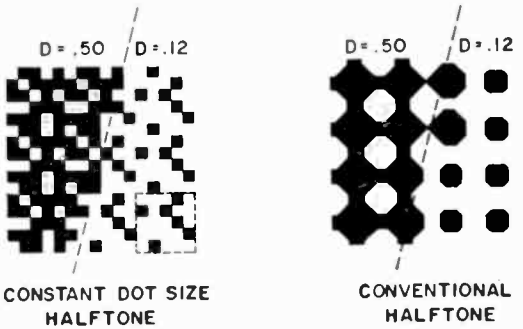


Fig. 1—Comparison of a constant-dot size halftone system with a conventional halftone. In both cases, an enlarged picture of a sharp transition between two grey levels is shown. The constant-dot size halftone is based on a 5×5 subarray shown by the dotted line.

As the change then takes place over one dot spacing, the constant dot system will have about 4 times the resolution of the conventional system. For smaller changes of density, the resolution of the constant-dot system decreases while that of the conventional system does not change and the advantage of the former diminishes. Overall, there is a resolution increase of a factor of 2 to 3.

The increase in resolution of the picture is obtained by making better use of the inherent resolution of the printing process. The appearance of the picture is also, however, affected by the *patterns formed by the dots*. The reason the conventional halftone is so successful is that the dots are all located on a regular array. Even though the dot spacing may be subliminal at the usual viewing distance of the picture, the eye would still be very sensitive to irregularities in the array. The constant-dot system used a much finer array of points but it is the patterns formed by groups of dots that determine the overall appearance. To make these as regular as possible, it is necessary to divide the point array into an equally shaped subarray which may be square, rectangular, rhombic or of some other shape. The size of the subarray should be subliminal or approximately the same as the conventional screen spacing. The subarrays are defined by the threshold

levels assigned to the points within them, which are the same in each subarray. A region of uniform density will then have an identical dot pattern in each subarray, thus optimizing the overall appearance. In the example of Fig. 1, the subarray contains 5×5 points as indicated.

The difficulty of rotating screen angles in *multicolor work* can be circumvented in the constant-dot system. The patterns of the subarrays can be chosen in a manner to minimize both the Moiré and the color shift problems even though all the screen angles are the same. Each color must have different patterns, but they are all based on the same basic array of points. Even more flexibility can be introduced with a small increase in electronic complexity by changing the array size for each color, e.g., 3×5, 4×4 and 5×3. Finally, it is even possible to vary the spacing of the basic array from color to color for some of the halftone writing systems.

A difficulty arises in the constant-dot system due to the *quantizing* of the grey scale. The number of quantized levels cannot exceed the number of points in the subarray. It may actually be considerably less, if, e.g., it is desirable to divide the grey scale into steps of approximately equal density on the finished print. To demonstrate this point, Fig. 2 shows measured densities as functions of the number of points in a subarray occupied by dots. Two experimental curves are shown of stabilization paper exposed on a color scanner as described below. By changing the processing conditions, very different curves are obtained, neither of which approaches the simple theoretical one. We show below that an 8×8 point subarray comes close to satisfying the quantizing criterion for a high quality reproduction but a somewhat larger array may still be necessary.

If the size of the subarray is larger than a subliminal area, undesirable patterns may appear in the picture. The problem can be overcome by going to a *mixed system*. In one system, a smaller subarray is used (say 5×5 points) and the subpatterns are arranged, so that each quantized grey level has a pattern which contains all the dots of the next lower level, plus one or more new dots. It is then only necessary to modulate the added dot or dots, e.g., by changing the light-pulse length which exposes this dot. The rest of the dots remain constant size. A transition can be made from one level to the next as gradually as desired by subdividing the interval into the required number of digital steps.

A different kind of mixed system can be designed which combines some of the advantages of both the conventional screen and the constant-dot size screen. It is particularly suited for systems where the minimal dot is large, so that a conventional picture uses a very coarse

screen. A constant-dot system does not improve the appearance. The mixed system for this case has the same coarse arrangement of minimal dots in the lightest highlights as the conventional screen. However, as the density increases, new dots are added halfway between the original ones, followed by another doubling of the number of dots. From this point on, all the dots increase uniformly in size. The result is that in the middle tone and shadow areas, there are four times as many dots as in the conventional halftone with correspondingly doubled resolution. The detailed behavior in the highlight region can be adjusted to make the density changes as smooth as required.

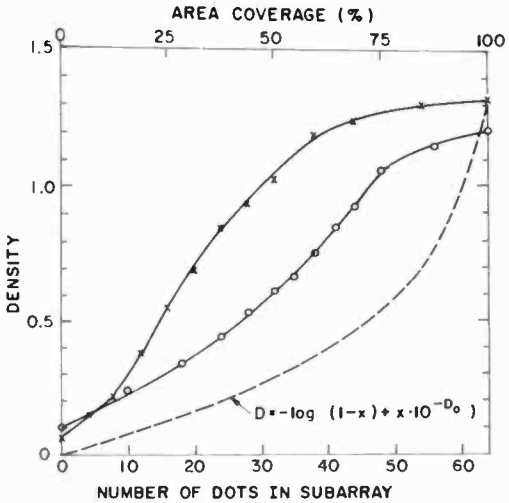


Fig. 2—Two examples of the densities obtained on stabilization paper exposed on the color scanner. The abscissa is the number of points in an 8×8 subarray that are occupied by dots. The dashed curve is the calculated density if all the dots were sharp and not overlapping.

Another problem that may arise in electronic screening is a *restricted information storage* or transmission capability. In an analog reproduction, this would require a reduction in resolution. In the simple constant-dot system, the restriction expresses itself as the fact that there is not a bit available for each dot to be written. This, however, does no imply that one should make the dots larger and the pattern coarser but only that the "video" signal should change more slowly. As an example, consider the situation where there is just enough bandwidth available for a conventional halftone, namely 6 bits for one dot. The constant-dot system might have 16 array points in the area of one conventional halftone dot, and so require 16 bits for full

specification. If only 6 are available, it is best to consider the 16 locations together as having a single grey level and assign dots to these locations according to which of the 64 levels it is. The resulting picture resolution is no different from the conventional halftone but the advantages of the finer pattern and the possibilities of effective multicolor reproduction remain. The same amount of information is simply distributed in a different manner.

Experimental Constant-Dot-Size System

To test the predictions made in the previous discussion, a constant-dot system was built. The image-generating equipment was an RCA Color Scanner, a facsimile machine with reading and writing cylinders and heads mechanically locked together. It produces a 1:1 reproduction of the original on the reading cylinder onto photographic film on the writing cylinder. The video signal is passed through an analog computer to make the desired modifications in the reproduced image (such as making a color separation). The scanning movement takes place in a spiral line with 50 μm pitch. For these experiments, the scanner was modified in two ways as shown in Figs. 3 and 4. First, the video signal was extracted from the computer after some preliminary processing. In the dot generator, it is converted to equal size pulses which drive the glow lamp that writes on the film. To determine the location of the dots and timing of the pulses, two counters are added—one counts the cylinder revolutions and the other a series of lines on a special timing track attached to the rotating cylinder. These lines are spaced 50 μm apart. The two counting techniques, therefore, define a square-point raster with 50 μm point spacing (500 lines/inch) over the entire picture surface. The dot generator causes dots to be written on the raster points only, as required by the screening system.

The components of the dot generator are shown schematically in Fig. 4. The counters divide the point array into the desired subarrays and assign each point a number corresponding to its location in the subarray. A threshold level is assigned to each subarray point, so that if the video signal is above this level, a dot is written at that point. This means that the video level at each point determines the decision at that particular point. This produces the high resolution of the resulting image. The considerably larger subarray, with a certain dot pattern at a given grey level, does not limit the resolution but only the appearance of the image.

In the present experiments, the subarray size was fixed at 8×8 points. The area (0.4×0.4 mm) is thus larger than most conventional

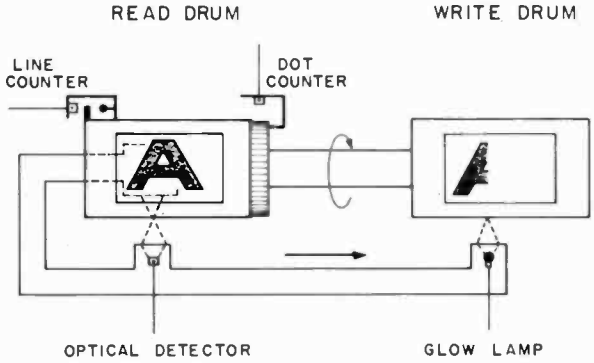


Fig. 3—Schematic diagram of the mechanical parts of the color scanner with modifications for halftone generation. The transparent Read drum contains the transparency that is being scanned while the unexposed film is located on the Write drum.

halftone spacings and a random pattern in that area would produce an undesirably large repeat pattern. However, by arranging the patterns symmetrically within the area, a finer overall appearance is produced.

The equipment used allowed for only 13 different threshold levels that could be assigned to the 64 points in the subarray. This will, of course, produce noticeable quantizing effects in some pictures with slowly varying grey levels. A straightforward modification of the electronics could produce more threshold levels. The present implementation of the system requires that each grey level contain all the dots of the preceding (lighter) levels.

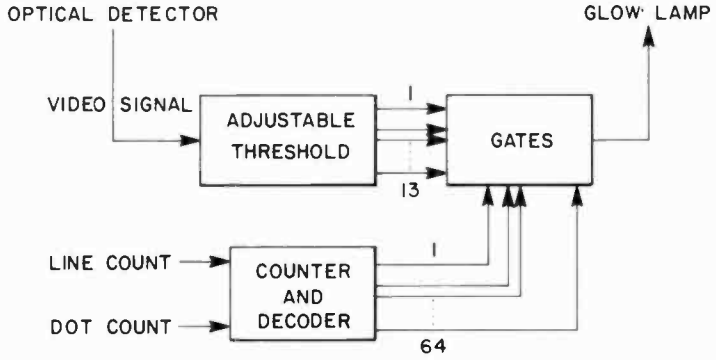


Fig. 4—Schematic diagram of the electronic configuration of the constant-dot size halftone system.

Results

Various classes of patterns were designed and tried out. A class of patterns consists of a pattern for each of the grey levels, all of which have been designed according to a certain criterion. The pattern classes were tested first by printing each pattern over a large area and by studying its appearance and the transitions from one pattern to another. As an example, the prints of one pattern class are shown enlarged in Fig. 5. The number of dots in each pattern was adjusted in such a way as to have all 13 grey levels approximately evenly spaced in density, so that the density steps between successive levels become rather large.

To determine how many grey levels are needed for a high-quality picture, we printed complete sets of pattern classes with 64 grey levels, each one having one more dot than the previous one. These

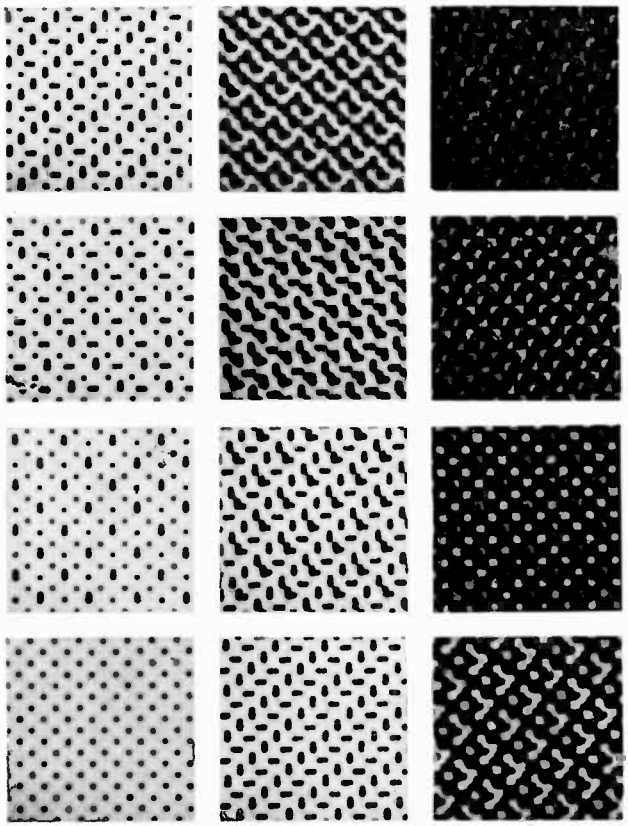


Fig. 5—The 13 grey levels of one of the pattern classes studied (enlarged 3 times).

64 level prints had to be produced in 5 separate groups of 13 levels because of the equipment limitations. A linearly changing video signal was used to generate each group. The resulting print represents the most difficult kind of picture to reproduce, similar to what is encountered in fleshtones. Fig. 6 shows a part of the print of one pattern class in actual size. It demonstrates that 64 or more grey levels are required for good reproduction under these conditions since steps from one grey level to the next are visible in the lighter regions. The steps can be eliminated and the number of required grey levels reduced by

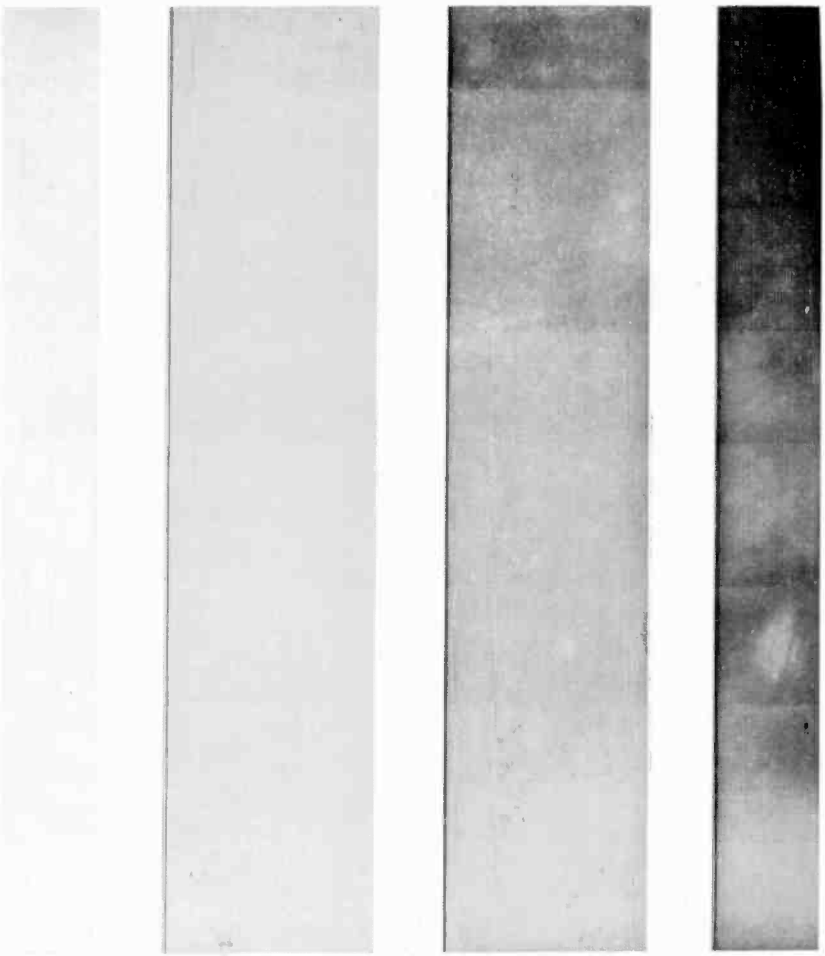


Fig. 6—A pattern class of 64 grey levels. It was produced in 5 different steps. In each case, the video signal was a linearly increasing voltage synchronized with the rotation of the drum. This simulates a picture with a very slowly changing grey level.

Fig. 7—Two reproductions of the same picture using different pattern classes of constant dots.

choosing different patterns or by adding a small amount of noise (less than half the step between two levels) to the video signal.

Reproductions of a picture are shown in Fig. 7 using two different pattern classes. The grey scale was adjusted so that there are more steps in the highlight region, where there is more detail. The thresholds were set to give the desired densities for the best appearance of the

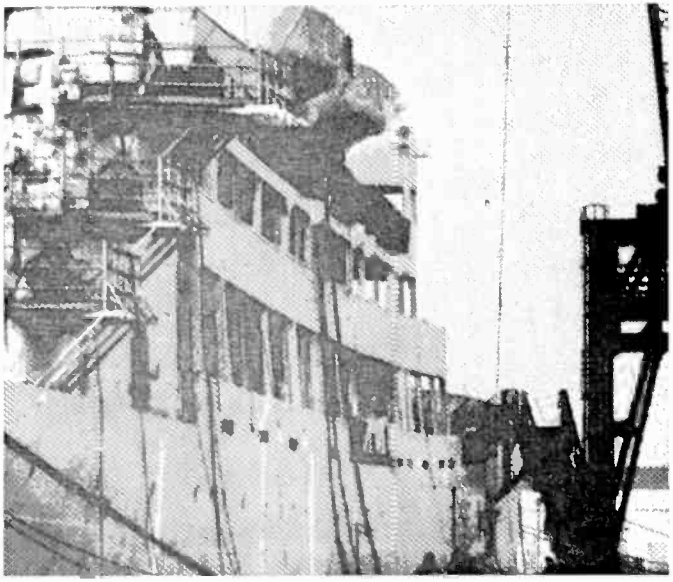


Fig. 8—An enlargement of the film from which the first print of Fig. 7 was made.

picture. Because of time limitations, the grey scales of the various pictures could not be made exactly equal and the results in Fig. 7 show this variation. It was found that the various pattern classes produce pictures that look about the same at the normal viewing distance as long as the dots are uniformly distributed within the sub-array. All patterns produce prints with higher resolution than conventional halftone pictures.

To show the dot pattern in a picture region in more detail, an enlargement of one of the films from which Fig. 7 was printed is shown in Fig. 8. The loss in resolution and contrast of the individual dots is due to the limitations of the enlarging process as the original film had hard square dots.

To show that this system can also produce conventional halftone pictures, such patterns were designed and produced the picture of Fig.

9. It is a 45° screen with 88 lines/inch. The resulting halftone dots are not round because they are composed of small square dots but the printed result approximates that of a conventional halftone.

In the same way, it was possible to simulate the mixed system described earlier. To demonstrate its possibilities, a comparison was made between the conventional and the mixed system in Fig. 10. In both cases, the smallest dot size was arbitrarily limited to a 100×100 μm

Fig. 9—The same picture as Fig. 7 with a simulated conventional halftone patterns of 88 lines/inch.

square (composed of four dots on our system). By using a slight overexposure, the dots were also made less sharp as is more commonly encountered in practice. The detailed treatment is shown in Fig. 11 by two enlargements of a highlight to middle-tone region of the picture. It is clear that in the example, the mixed system produces a much higher-resolution picture than the conventional system.

Also investigated were the problems of multicolor printing from electronically generated halftones. Because of experimental limitations, the tests were restricted to studying the overlap of black and white halftones. A series of films were prepared containing sizeable areas of each pattern of a class. Two or three such films were then super-

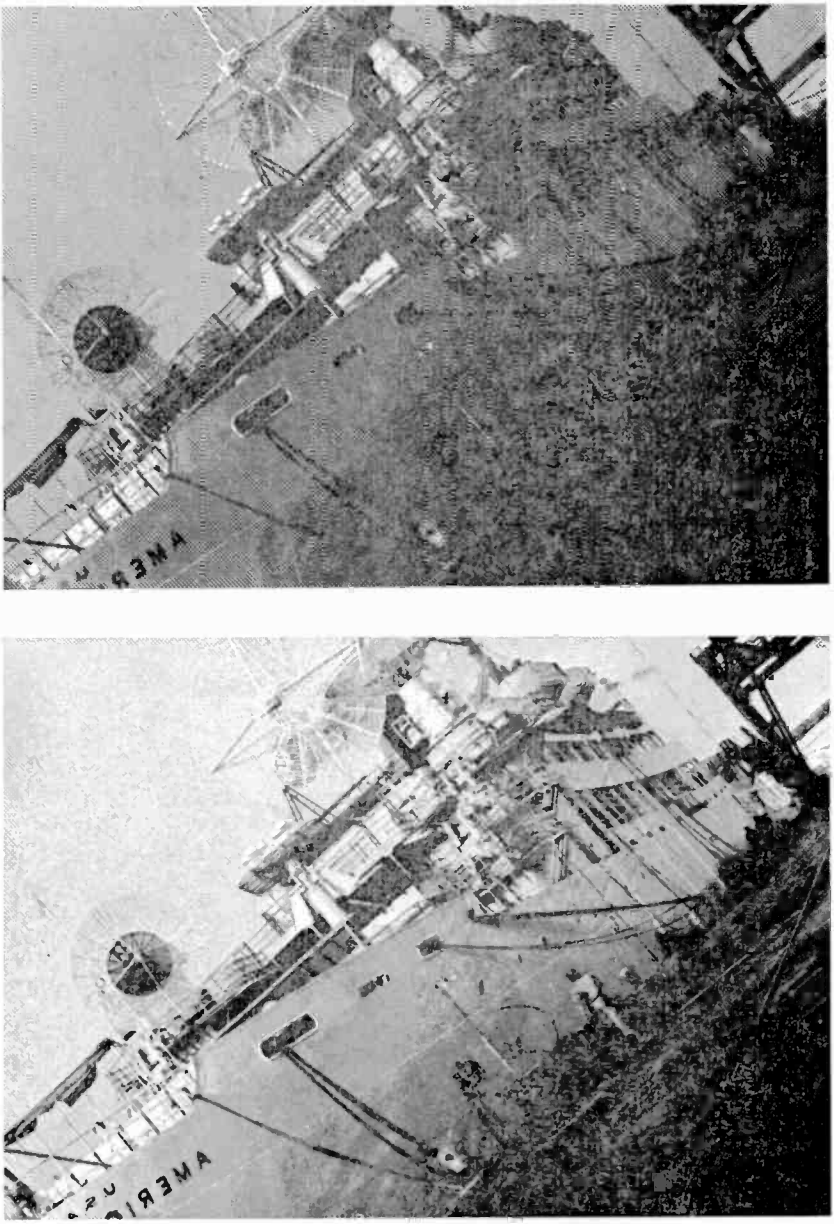


Fig. 10—A comparison of two pictures made by a conventional and a mixed halftone system. In both cases, the highlight areas are the same with 62 line/inch square pattern. Because of the limitations in the logic, the patterns had to be parallel to the scanning direction (screen angle of 0°). The pictures were therefore rotated to produce a more conventional looking 45° screen angle.

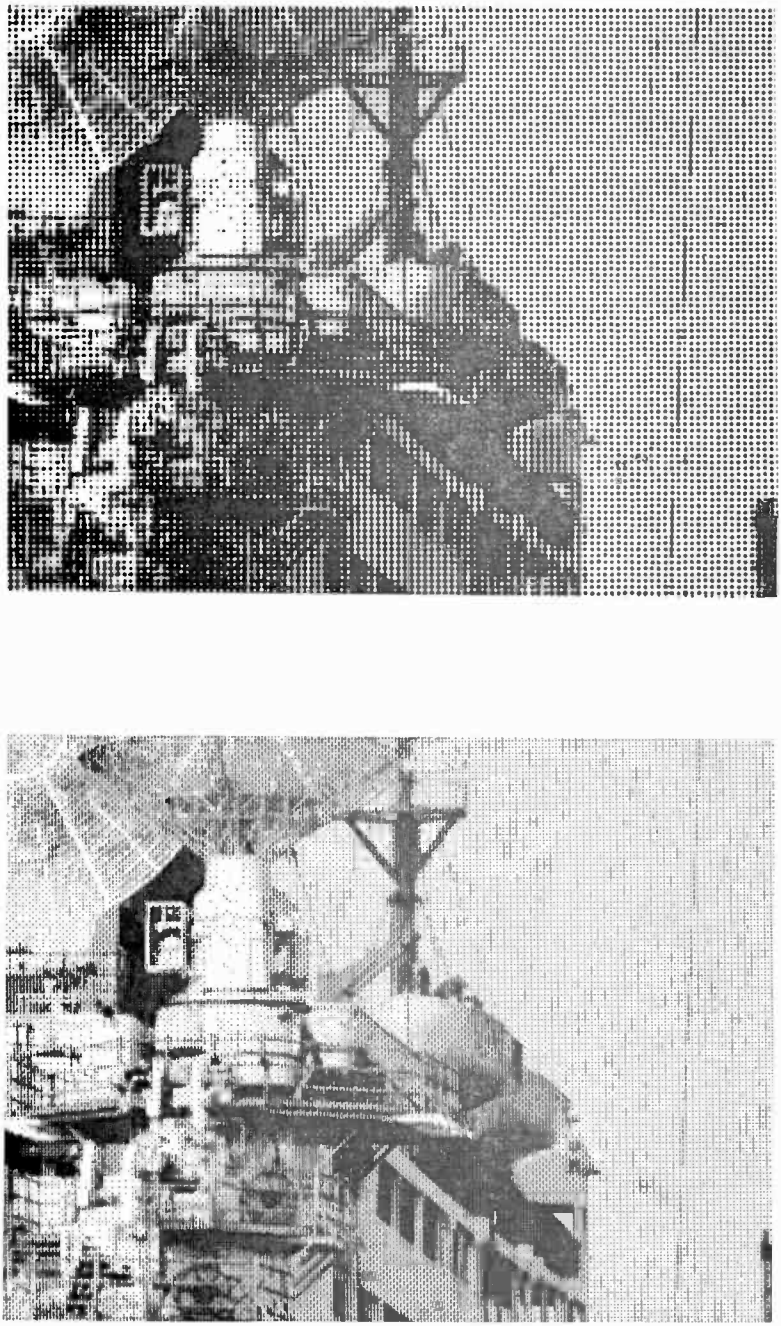


Fig. 11—Two enlarged sections of Fig. 10.

imposed, and a photograph taken of the composite. This is actually a much more severe test both for Moiré effects and color shifts than would be encountered in multicolor printing. This is because areas of uniform tone were studied and because the density shifts in the black and white photo are much more noticeable than color shifts.

As expected, the undesired effects could be minimized or eliminated for correctly chosen combinations of patterns. This must be done for each combination between the two classes of patterns, at least in the middle tone region where the problem is most severe. It could readily be accomplished for two classes of patterns, but became more difficult in the case of three classes. (Usually, at most, three of the colors will give problems as yellow has a weaker effect.) We found that the overlap problem could be studied analytically by taking two patterns and calculating the number of dots that overlap in a subarray for each of the 64 possible translations of the two subarrays relative to each other. If each translation produced the same number of dot overlaps, then the two patterns also showed no density shifts or Moiré patterns visually. The more the number varied with translation, the worse the visual appearance became.

Further refinements of varying the subarray size and the spacing between dots in the different colors have not yet been studied. This would require considerable modification of the experimental equipment.

Conclusions

It has been shown that it is feasible to include screening during reproduction or processing of electronically generated pictures and, in particular, that novel kinds of screening techniques can be used which produce better results than conventional halftones. The constant-dot system is particularly suited for high quality reproduction. The hybrid system produces advantages for reproduction system where the smallest dot that can be reliably printed on the output is large enough to require an undesirably course conventional halftone screen.

References

[1] R. L. Hallows and R. J. Klensch, "Electronic Halftones," IEEE Spectrum, p. 64, Oct. 1968.

[2] D. Meyerhofer, A. W. Stephens and J. J. Walsh, "Electronic Color Separation With Laser Light Sources," **IEEE Trans. Communication Technology,** Vol. COM-18, p. 361 Aug. 1970.

[3] E. Jaffe, **Halftone Photography,** Graphic Arts Technical Foundation, Inc., Pittsburgh, 1964.

[4] E. W. Herold and K. H. Fischbeck, private communicatlon.

Chromatic-Aberration-Limited Image-Transfer Characteristics of Image-Tube Lenses of Simple Geometry

I. P. Csorba, RCA Electronic Components, Lancaster, Pa.

Abstract—The Chromatic-aberration-limited image-transfer characteristics of image-tube lenses of simple geometry are analyzed by calculation of the image-current-density distribution of disk and line images and the modulation-transfer function.

Introduction

The three most commonly used electron lenses in image tubes are (1) the electromagnetic lens, consisting of homogeneous axial electric and magnetic fields;[1] (2) the point-symmetric-type electrostatic lens, made of two concentric spheres,[2,3] the outer sphere serving as an electron source (cathode) and the inner pierced sphere as the anode; and (3) the biplanar electron lens, consisting of a homogeneous axial electric field.

The combined action of the electric and magnetic fields of an electromagnetic lens produces an erect image, unity magnification, and flat object and image fields. This lens is free of geometrical aberrations and the only aberration present is caused by the variation of axial emission energy of the photoelectrons. Because of this variation, the image of a point of the photocathode is a small disk at the plane of the phosphor screen. Specifically, in the case of a monoenergetic Lambertian electron emitter, if the phosphor screen is placed in the image plane of the electrons emitted at an angle θ_f, the transverse focusing error Δr may be given by the following equation:[1]

$$\Delta r = 2\frac{V_i}{E}(\cos\theta_f - \cos\theta)\sin\theta, \qquad [1]$$

where V_i is the accelerating voltage required for a photoelectron at

rest to acquire its emission energy, E is the electric field, and θ is the angle between the emission direction and the axis of the electron lens.

In a point-symmetric system, the central field and the anode-aperature lens focus an inverted image of the cathode on a sphere concentric with the cathode, as shown in Fig. 1. The transverse

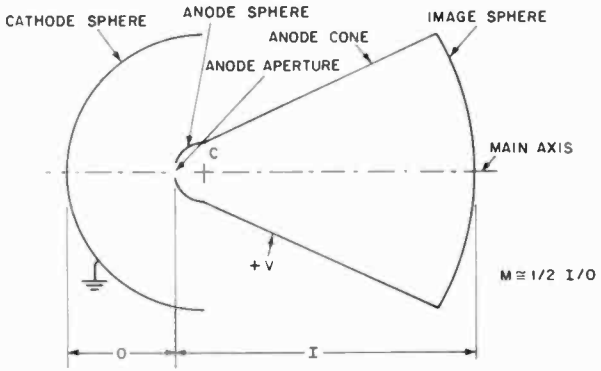


Fig. 1—A diagram of two-concentric sphere-type image-inverting electrostatic diode.

focusing error Δr produced by the chromatic aberration in the Gaussian image plane can be described by[3,4]

$$\Delta r = 2M \frac{V_i}{E_c} \sin\theta \cos\theta, \qquad [2a]$$

where E_c is the electric field at the cathode and M is the magnification of the point-symmetric lens. (It is interesting to note that Eq. [2a] is also valid for cathode lenses of axial symmetry.[4])

If the phosphor screen is placed in the image plane of the electrons emitted at an angle θ_f, the transverse focusing error Δr becomes

$$\Delta r = 2M \frac{V_i}{E_c} (\cos\theta_f - \cos\theta) \sin\theta \qquad [2b]$$

For unity magnification, Eqs. [1] and [2b] are identical. For simplicity, therefore, in the following discussion M is assumed to be unity.

In the case of a biplanar electron lens, the transverse focusing error may be obtained by multiplication of the radial emission velocity of electrons v_r with the transit time t. For a monoenergetic Lambertian electron emission, the transverse focusing error is described by

$$\Delta r = tv_r = 2L\sqrt{\frac{V_i}{V_s}}\sin\theta. \qquad [3]$$

In Eq. [3], the first-order transit time t is equal to $L\sqrt{2m/(eV_s)}$, and radial-emission velocity v_r is equal to $\sqrt{2eV_i/m}\sin\theta$, where V_s is the screen-to-photocathode voltage, L is the screen-to-photocathode separation, and e and m are the charge and mass of an electron, respectively.

Image-Current-Density Distribution of a Disk of Uniform Emission Current

In essence, Eqs. [1], [2], and [3] describe the radial distance in a point image where the electrons emitted at an emission angle θ land at the phosphor screen. The image-current density is then derived by the assumption that ΔA is an infinitesimal element of area A on the surface of the photocathode. It is further assumed that a monoenergetic emission of electrons is produced from ΔA with an angular distribution obeying Lambert's Law. Based on these assumptions, the incremental current di_θ emitted between angles θ and $\theta + d\theta$ from ΔA is given by

$$di_\theta = 2I\Delta A \sin\theta\cos\theta\, d\theta, \qquad [4]$$

where I is the emission-current density.

The incremental current di_θ strikes the phosphor screen over an annular element having a radius Δr. The current density j in this annular element may be obtained by division of di_θ by the area of the annular element, as follows:

$$j = \frac{I\Delta A \sin\theta\cos\theta\, d\theta}{\pi\Delta r d(\Delta r)}, \qquad [5]$$

where $d(\Delta r)$ is the width of the annular element.

For an extended Lambertian emitter at the cathode, the current-density distribution in the plane of the phosphor screen may be obtained by integration of j over the extended current-emitting surface. Therefore, the image-current-density distribution of a sine-wave resolution pattern and, subsequently, the modulation transfer function (MTF) of an electron lens may be derived by integration of Eq. [5] over the resolution pattern. Because a closed-form solution for the lenses in consideration is not known, the integration must be carried

out numerically. The computation is straightforward for a biplanar lens. However, for the electromagnetic or the two concentric-sphere type lenses, it is more complicated to find the optimum image plane where the MTF has a maximum value. The computation of the MTF for these lenses may be simplified by calculation of the image-current-density distribution of disks of uniform emission current and then by transformation of these data into MTF. Fig. 2 shows the geometrical outline for derivation of the image-current-density distribution of a disk having a uniform emission current. By use of the

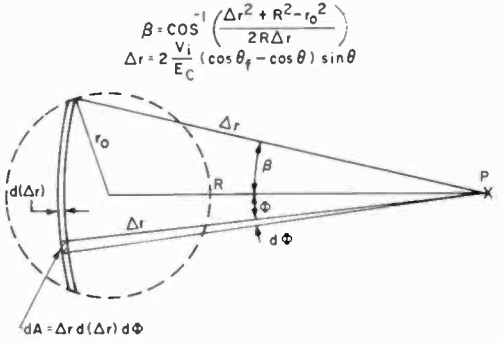


Fig. 2—Projection of a circular object on the plane of the phosphor screen (P is at a radial distance R from the center of the disk).

notations in Fig. 2, the image-current density $J(R)$ at a point P of the phosphor screen may be written

$$J(R) = \int_A \frac{I \sin\theta \cos\theta \, d\theta}{\pi \Delta r d(\Delta r)} dA$$

$$= \int_{R-r_o}^{R+r_o} \int_{\Phi=-\beta}^{+\beta} \frac{I \sin\theta \cos\theta \, d\theta \, \Delta r d(\Delta r) d\Phi}{\pi \Delta r d(\Delta r)}$$

$$= \frac{2I}{\pi} \int_{R-r_o}^{R+r_o} \beta \sin\theta \cos\theta \, d\theta, \qquad [6]$$

where $\beta = \cos^{-1}\left(\frac{\Delta r^2 + R^2 - r_o^2}{2\Delta r R}\right)$.

Substitution for β in Eq. [6] modifies the image-current $J(R)$ as follows:

$$J(R) = \frac{2I}{\pi} \int_{R-r_o}^{R+r_o} \cos^{-1}\left(\frac{\Delta r^2 + R^2 - r_o^2}{2\Delta r R}\right) \sin\theta \cos\theta \, d\theta \qquad [7]$$

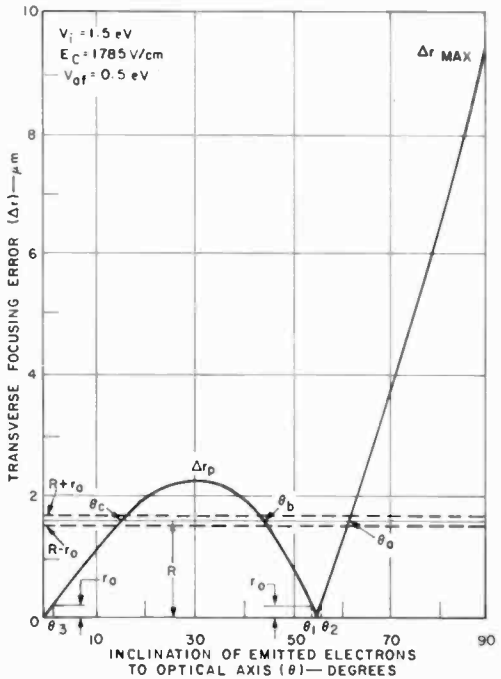


Fig. 3—Transverse focusing error as a function of the emission angle.

In general, for the condition $V_i > V_{af}$, where V_{af} is the axial emission energy of photoelectrons in focus at the phosphor screen, each radial distance is associated with three emission angles up to Δr_p. These emission angles are shown in Fig. 3 as θ_a, θ_b, and θ_c. Consequently, the total current density is the sum of current densities obtained at three different emission angles.

In the center of image ($R = 0$), β equals π; therefore, the current density $J(0)$ is

$$J(0) = 2I \int_0^{r_o} \sin\theta \cos\theta \, d\theta$$

$$= I \left[\sin^2\theta \right]_{\theta_1(r_o)}^{\theta_2(r_o)} + I \left[\sin^2\theta \right]_0^{\theta_3(r_o)}, \qquad [8]$$

where θ_1, θ_2, and θ_3 are emission angles corresponding to $\Delta r = r_o$ transverse focusing error at the screen, as shown in Fig. 3.

For a polyenergetic emitter, the image-current density may be found by integration of Eq. [7] (or in the center Eq. [8]) over the emission-energy distribution curve.

Calculation of Image-Current Density of Disk and Line Electron Emitters

In the following calculations it is assumed that the electron source has an emission-energy distribution with a maximum emission energy of eV_{im} at 2 electron-volts. The fraction of electrons dN emitted in a voltage interval $d(V_i/V_{im})$ is then

$$dN = \frac{\pi^2}{2(\pi - 2)} \frac{V_i}{V_{im}} \cos\left(\frac{\pi}{2} \frac{V_i}{V_{im}}\right) d\left(\frac{V_i}{V_{im}}\right). \qquad [9]$$

The distribution given by this equation is a normalized energy distribution similar to that measured for typical photocathodes. The focusing conditions are adjusted to yield a maximum peak current density in the image of a small disk.

The image plane where the maximum current density occurs may be found by successive placement of the phosphor screen in the image planes of various axial-emission energy groups and by calculation of the sum of current density at each plane produced by the different monoenergetic groups, as described in the Appendix.

Fig. 4 shows the image-current density as a function of axial emission energy, V_{af}, that produces a sharp focus at the phosphor screen. In Fig. 4, a maximum occurs in the image plane of 0.5-volt axial-emission-energy electrons. In the remaining discussion, therefore, the phosphor screen is assumed to be in the image plane of 0.5-volt axial-emission-energy electrons.

By use of Eqs. [7] and [8], and the above assumed emission-energy distribution, the image-current density was calculated for disks of uniform emission current. Figs. 5 and 6 show the results of the calculation.

For transformation of the image-current-density distribution of disks into MTF, it is necessary to form a line image. By movement of a disk of uniform emission current along a straight line in the object plane, a line may be formed with an elliptic emission-current-density distribution in the cross section of the line. The image of this line may be generated by division of the image of the stationary disk into disk sections of uniform image current, as shown in Fig. 7, and, as in the object plane, by transformation of these disk sections into lines with elliptic current-density distribution.

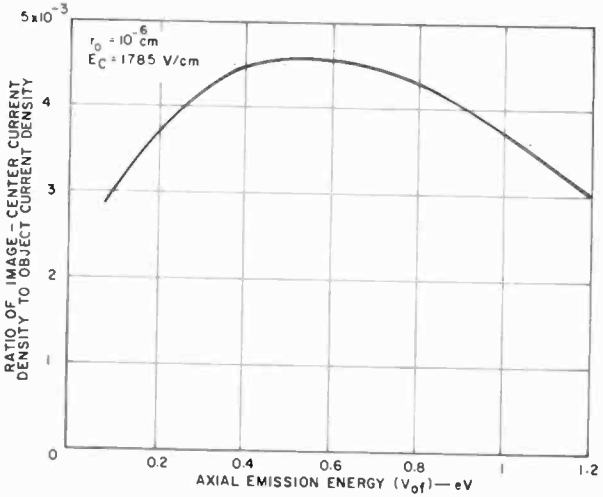


Fig. 4—Ratio of image-center current density to object current density of a disk as a function of axial emission energy of electrons in sharp focus at the phosphor screen.

The line image is obtained by summation of the current density produced by the individual disk segments. However, the current density of each segment must be multiplied by the ratio of the segment diameter to the object diameter for proper transformation of the disk image into a line image. Fig. 8 shows the image-current-density distributions of lines with elliptic emission-current-density distribution.

Elliptic and Sine-Wave Responses

An elliptic-wave pattern at the photocathode of an image tube may be generated by superimposition of alternating black and white lines of amplitude B_1 over a uniform illumination B_1 with the elliptic distribution of light in the cross section, as shown in Fig. 9. At each point of

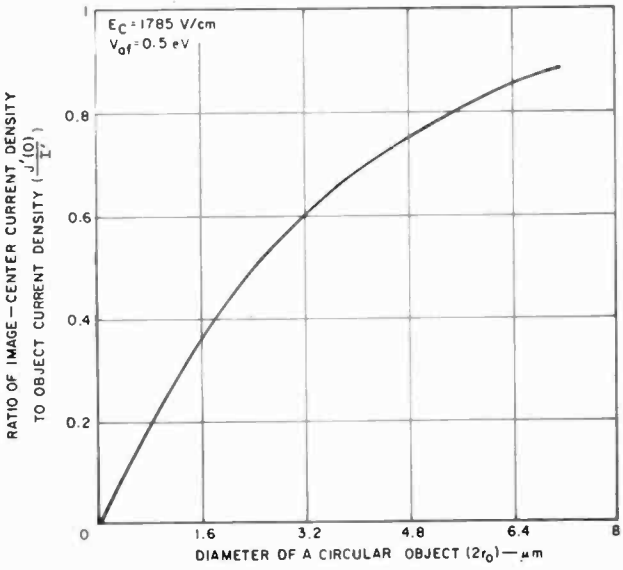


Fig. 5—Image-current density of a disk as a function of diameter. An emission energy distribution of Eq. [9] is assumed.

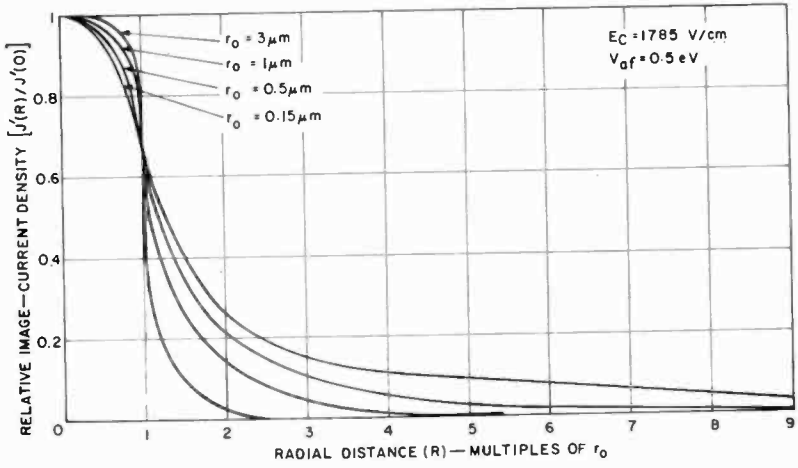


Fig. 6—Zonal image current-density distribution of disks as a function of radial distance. An emission energy distribution of Eq. [9] is assumed. $J'(R)$ and $J'(0)$ designate the image current density at the radial distance R and in the center, respectively.

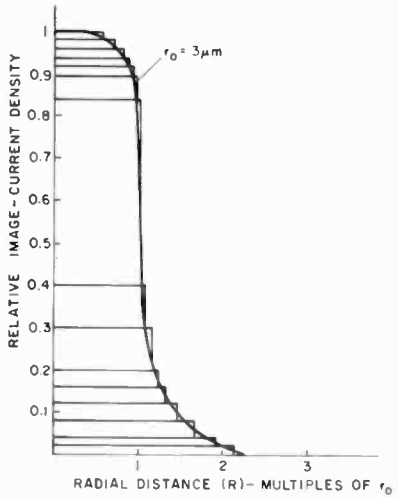


Fig. 7—Zonal image current-density distribution of a disk as a function of radial distance. The image is divided into disk sections of uniform image current for calculation of the line image.

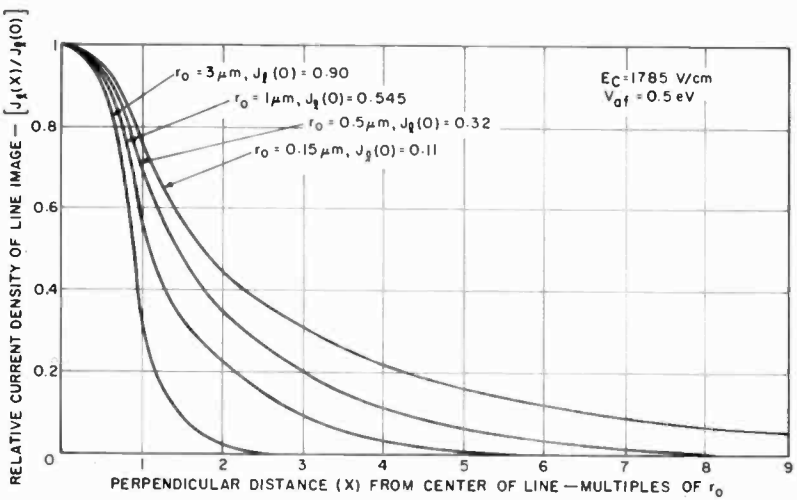


Fig. 8—Current-density distribution in the cross section of the image of a line having eliptic emission current-density distribution and $2r_o$ width. The emission current density in the center of the line is unity. $J_i(0)$ is the image current density in the center of line $(X = 0)$.

a line in the image plane, the image-current density is jointly established by the image current of the particular line plus the current contributions of all neighboring lines that are within the circle of maximum transverse-focusing error. The elliptic-wave response r, therefore, is simply

$$r = \Delta B / B_2, \qquad [10]$$

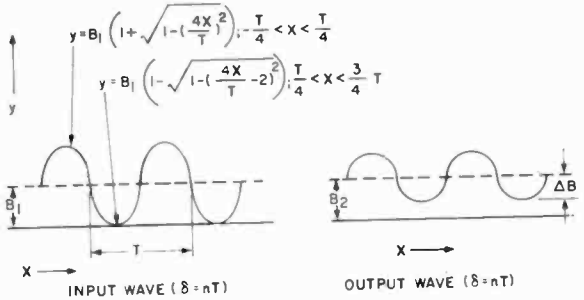


Fig. 9—Input and output waves of an eliptic wave pattern.

where ΔB is the amplitude of the wave pattern at the image plane. Fig. 10 shows the calculated elliptic-wave response curve obtained by use of the line-image curves of Fig. 8.

The sine-wave response may be obtained from the elliptic response by Fourier analysis. The Fourier series $F(x)$ of the input elliptical wave pattern of Fig. 9 is given by*

$$F(x) = B_1 + B_1 \frac{2\sqrt{3}}{\pi}\left[\cos\left(2\pi\frac{nx}{\delta}\right) - \frac{0.51}{3}\cos 3\left(2\pi\frac{nx}{\delta}\right) + \frac{0.38}{5}\cos 5\left(2\pi\frac{nx}{\delta}\right) - \frac{0.32}{7}\cos 7\left(2\pi\frac{nx}{\delta}\right) + \frac{0.25}{9}\cos 9\left(2\pi\frac{nx}{\delta}\right) - \cdots\right], \qquad [11]$$

* An exact form of Eq. [11] (as suggested by E. G. Ramberg) is

$$F(x) = B_1\left[1 + 2J_1\left(\frac{\pi}{2}\right)\left(\cos 2\frac{\pi nx}{\delta}\right) + \frac{2}{3}J_1\left(\frac{3\pi}{2}\right)\left(\cos 6\frac{\pi nx}{\delta}\right) + \frac{2}{5}J_1\left(\frac{5\pi}{2}\right)\left(\cos 10\frac{\pi nx}{\delta}\right) + \cdots\right].$$

where n is the number of cycles in a unit distance δ.

The output wave may be obtained by multiplication of each term of the input wave by the sine-wave response $R(n)$ that corresponds to the line frequency n. At the maximums and minimums of the output wave, all cosines are either unity or -1; therefore, the elliptic response factor $r(n)$ is given by

$$r(n) = \frac{2\sqrt{3}}{\pi}\left[R(n) - \frac{0.51}{3}R(3n) + \frac{0.38}{5}R(5n) - \frac{0.32}{7}R(7n) + \frac{0.25}{9}R(9n) - \cdots\right] \quad [12]$$

By use of similar series for $r(3n)$, $r(5n)$ $\cdots r(kn)$, the above series can be solved for $R(n)$ as follows:

$$R(n) = \frac{\pi}{2\sqrt{3}}\left[r(n) + \frac{0.51}{3}r(3n) - \frac{0.38}{5}r(5n) + \frac{0.32}{7}r(7n) - \cdots\right] \quad [13]$$

The calculated sine-wave response is shown in Fig. 10 for a cathode field strength E_c equal to 1785 volts per centimeter. The square-wave response shown in Fig. 10 was obtained from the sine-wave response by the well-known transformation process described in the literature.[5] Because the transverse-focusing error is inversely proportional to E_c, the response for a different value of cathode field strength E_c' may be found by multiplication of the abscissa (spatial-frequency axis) with the ratio E_c'/E_c. Similarly for magnification values of M other than unity or for maximum emission energy values of V_{im} other than 2 electron-volts, the abscissa is multiplied by $1/M$ or $2/V_{im}$, respectively. An inspection of the square-wave response curve shown in Fig. 10 reveals that at response values lower than 0.9, the response is inversely proportional to the spatial frequency and directly proportional to the cathode field strength (anode voltage). The square-wave response, therefore, may be expressed by

$$C \approx 0.1\frac{E_c}{nMV_{im}},$$

where n is the spatial frequency in line-pairs per millimeter.

The electric-field strength used to calculate the sine-wave response curve provides an accurate description of the conditions found in high-quality magnetic image-intensifier tubes of 38-millimeter useful diameter size. However, electrostatic-type image-intensifier tubes of the same useful diameter usually have about five times smaller electric-field strength at the photocathode under normal operating conditions.

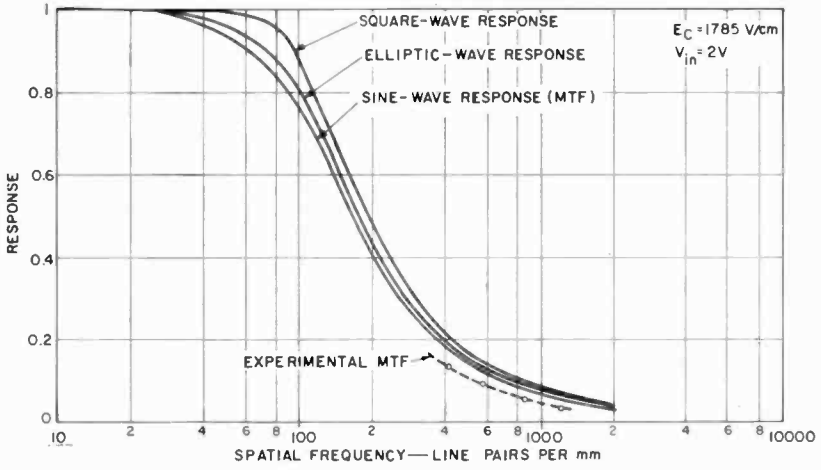


Fig. 10—Response curves of electromagnetic and two-concentric-sphere type electron lenses.

As described in the literature,[1] measurements in a developmental magnetic image-intensifier tube showed 1300 lines per millimeter electron optical resolution for a field strength E equal to 1785 volts per centimeter. This resolution figure was obtained with the assumption that the electron-optical MTF has a Gaussian distribution. A somewhat better correlation between the experimental data and calculated data may be obtained by expressing the MTF at the spatial frequency of the observed limiting resolution points as the product of the MTF of the electron optics and the combined MTF of other image-transmission elements (glass optics and phosphor screen). The square-wave response is 0.02* at the spatial frequency of the observed limiting resolution. Further, the combined MTF of the phosphor screen and glass optics may be approximated by the function $r = \exp\{-kn^2\}$ (Gaussian distribution). In addition, if we consider that the MTF of electron optics is proportional to the anode voltage and inversely proportional to

* For a square-wave resolution pattern, the threshold contrast of the observer was measured at 0.02 (for monocular vision).

the spatial frequency in the neighborhood of limiting resolution, enough information is available to deduce the MTF of the electron optics from the data of Fig. 11. The electron-optical MTF deduced from Fig. 11 is shown in Fig. 10. At each measured point, the spatial frequency has been translated to correspond to an electric field strength of 1785 volts per centimeter.

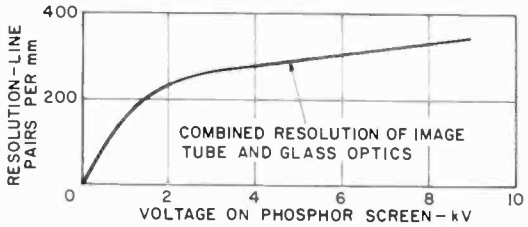


Fig. 11—Resolution of a magnetic image intensifier as a function of applied voltage.

In addition to the above experimental data, it is noted that, in a two-concentric-sphere-type developmental image tube, at a cathode field strength E_c equal to 60 volts per centimeter, 72 line-pairs per millimeter limiting resolution was observed, which corresponds to $72 \times 1785/60$ or 2142 line-pairs per millimeter electron-optical resolution for a cathode field strength of 1785 volts per centimeter.

The Biplanar Electron Lens

For a monoenergetic Lambertian electron emission, the point-image-current density of a biplanar electron lens may be given by substitution of Eq. [3] and the differential $d(\Delta r)$ into Eq. [5],

$$j = \frac{2I\Delta A \sin\theta \cos\theta \, d\theta}{2\pi 2L\sqrt{V_i/V_s} \sin(\theta)\, 2L\sqrt{V_i/V_s} \cos(\theta)\, d\theta} = \frac{I\Delta A V_s}{4\pi L^2 V_i} \qquad [14]$$

The value of $d(\Delta r)$ in Eq. [14] was obtained by differentiation of Eq. [3].

In the case of a polyenergetic electron emitter having an emission-energy distribution described by Eq. [9], the incremental image-current density dJ is as follows:

$$dJ = jdN = \frac{\pi \Delta A I'}{8(\pi - 2)L^2} \frac{V_s}{V_{im}} \cos\left(\frac{\pi V_i}{2V_{im}}\right) d\left(\frac{V_i}{V_{im}}\right), \qquad [15]$$

where I' is the emission-current density of the polyenergetic electron emitter. The point-image-current-density distribution $J(\Delta r)$ is given by

$$J(\Delta r) = \int_{V_i}^{V_{im}} \frac{\pi \Delta A I'}{8(\pi - 2)L^2} \frac{V_s}{V_{im}} \cos\left(\frac{\pi V_i}{2V_{im}}\right) d\left(\frac{V_i}{V_{im}}\right) \qquad [16]$$

$$= \frac{\Delta A I'}{4(\pi - 2)L^2} \frac{V_s}{V_{im}} [1 - \sin(\pi/2\ V_i/V_{im})]$$

In the center of the point image, the current density is

$$J(0) = \int_0^{V_{im}} dJ = \frac{\Delta A I'}{4(\pi - 2)L^2} \frac{V_s}{V_{im}} \qquad [17]$$

The relative image-current density distribution of the point image is, then,

$$\frac{J(\Delta r)}{J(0)} = 1 - \sin(\pi/2\ V_i/V_{im}) \qquad [18]$$

In this equation $\Delta r = 2L\sqrt{V_i/V_s}$. Eq. [18] then may be expressed in terms of Δr as

$$\frac{J(\Delta r)}{J(0)} = 1 - \sin(\pi/8\ V_s/V_{im}\ \Delta r^2/L^2) \qquad [18a]$$

Eq. [18a] is plotted in Fig. 12. The square-wave response may be obtained by integration of $J(\Delta r)$ over the square-wave resolution pattern. From Fig. 13, the image-current density i_w in the center of a white line is

$$i_w = I' - 4\int_{w/2}^{\Delta r_m} J(\Delta r)\Delta r \cos^{-1}\frac{w}{2\Delta r}\, d(\Delta r) + 4\int_{3/2w}^{\Delta r_m} J(\Delta r)\Delta r \cos^{-1}\frac{3w}{2\Delta r}\, d(\Delta r) \qquad [19]$$

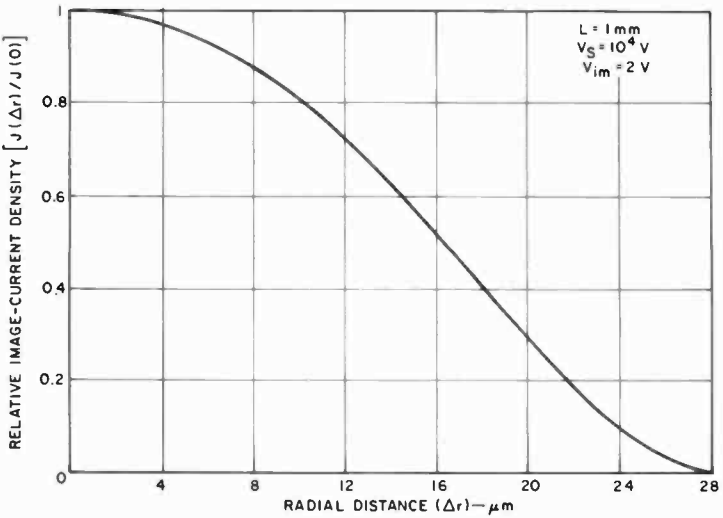


Fig. 12—Zonal image current-density distribution of a point at the photocathode.

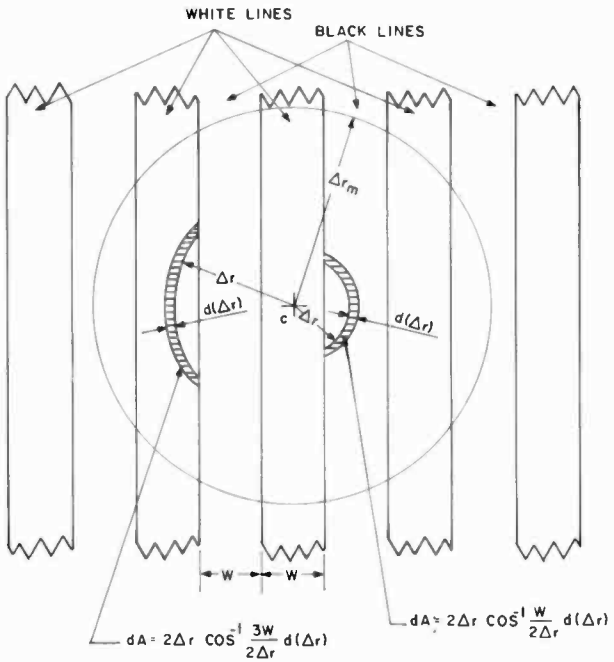


Fig. 13—Schematic diagram for deriving the image current-density distribution of a line pattern at the photocathode with uniform emission current in the white lines. The diagram is indicative for the calculation of image current density in the center of a white line. Δr_m is the maximum value of transverse focusing error.

The image-current density i_b in the center of a black line is expressed as

$$i_b = 4\int_{w/2}^{\Delta r_m} J(\Delta r)\Delta r \cos^{-1}\frac{w}{2\Delta r}\,d(\Delta r) - 4\int_{3/2w}^{\Delta r_m} J(\Delta r)\Delta r \cos^{-1}\frac{3w}{2\Delta r}\,d(\Delta r) \qquad [20]$$

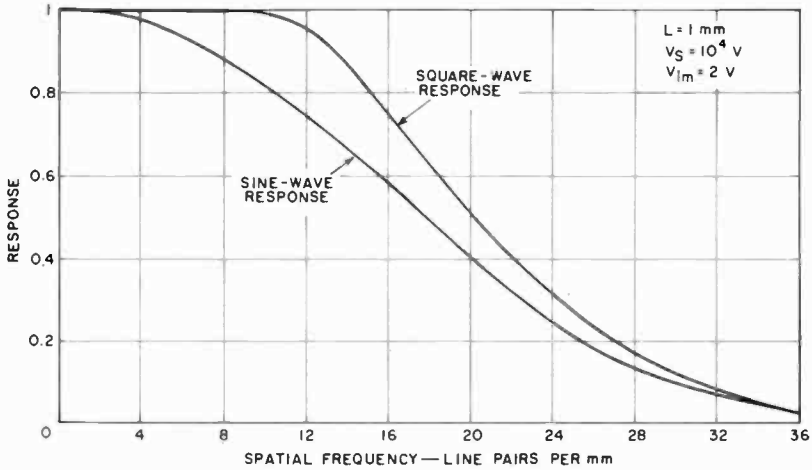


Fig. 14—Response curves of a biplanar electron lens.

The square-wave response C is, then,

$$C = \frac{i_w - i_b}{i_w + i_b} = 1 - 8/I'\int_{w/2}^{\Delta r_m} J(\Delta r)\Delta r \cos^{-1}\frac{w}{2\Delta r}\,d(\Delta r) + 8/I'\int_{\frac{3w}{2}}^{\Delta r_m} J(\Delta r)\Delta r \cos^{-1}\frac{3w}{2\Delta r}\,d(\Delta r) \qquad [21]$$

Fig. 14 shows the calculated square-wave response curve for conditions of screen-to-cathode separation of L at 1 millimeter, screen voltage V_s at 10000 volts, and maximum emission energy V_{im} at 2 volts. For

different values of screen-to-cathode separation, L', screen voltage V'_s, and maximum emission energy V'_{im}, the abscissa of Fig. 3 is multiplied by L/L', $\sqrt{V_s'/V_s}$, and $\sqrt{V_{im}/V_{im}'}$, respectively. With all these considerations, Fig. 14 also reveals that the MTF (sine-wave response) of the biplanar lens may well be approximated by

$$\text{MTF} \simeq \exp\{-12(V_{im}/V_s)(nL)^2\},$$

where L is given in millimeters and n is in line-pairs per millimeter.

As an example, for the case of a cesium-antimony (Cs_3Sb) cathode illuminated by a wavelength of 5000 angstroms, the maximum emission energy eV_{im} of electrons is

$$eV_{im} = \frac{e12395}{\lambda} - e\,(E_g + E_A) = e\left(\frac{12395}{5000} - 2.05\right)$$

$$= 0.43 \text{ electronvolts},$$

where E_g and E_A are the forbidden-band-gap energy and electron affinity,[6] respectively.

The limiting resolution, R_l for L at 1 millimeter and V_s at 10 kilovolts from Fig. 3 (at approximately 3% response) is

$$R_l = 36\sqrt{2/0.43} = 77.5$$

line pairs per millimeter. This resolution figure is in close agreement with the resolution figures observed on developmental biplanar image tubes.

Conclusion

The sine-wave response of electron lenses used in image-intensifier tubes shows that under normal operating conditions the electron-optical resolution ($R(n) \simeq 0.03$) of a magnetic or a two-concentric-sphere-type lens is much higher than the resolution of other optical elements of image tubes (phosphor screens, fiber optics, mica spacer). However, the image-quality loss expressed by the sine-wave response is not negligible in the normal operating spatial-frequency range (0 to 80 line pairs per millimeter). For a biplanar lens, the sine-wave response may be expanded in the spatial-frequency domain by limitation of the maximum emission energy of photoelectrons.

Acknowledgment

The author thanks E. G. Ramberg of the David Sarnoff Research Laboratory, Princeton, New Jersey, for his comprehensive review of this paper.

Appendix—Image-Center Current Density of a Disk as a Function of Axial Emission Energy of Electrons in Sharp Focus at the Phosphor Screen

For very small disks, the $\sin\theta\cos\theta$ product in Eq. [8] is nearly a constant between the emission angles θ_1 and θ_2 that correspond to a transverse displacement r_o at the phosphor screen. Therefore, this product may be brought outside of the integral. In addition, if the second term of Eq. [8] (which provides only a minor contribution to the image-current density at $\theta_f \neq 0$) is omitted, the image-center current density becomes:

$$J(0) = 2I \sin\theta_f \cos\theta_f \int_0^{r_o} d\theta = I \sin 2\theta_f \,(\theta_2 - \theta_1)$$

$$= 2I \sqrt{\frac{V_{fa}}{V_i}} \sqrt{1 - \frac{V_{af}}{V_i}} \,(\theta_2 - \theta_1) \qquad [22]$$

The value of $(\theta_2 - \theta_1)$ may be obtained by calculation of the slope of the transverse-focusing-error curve at the point of θ equal to θ_f from Eq. [1]:

$$\left[\frac{d(\Delta r)}{d\theta}\right]_{\theta=\theta_f} = \frac{2(V_i - V_{af})}{E} = \frac{2r_o}{\theta_2 - \theta_1} \qquad [23]$$

By combination of Eqs. [22] and [23] the image-current density becomes

$$J(0) = \frac{2IEr_o}{V_i} \sqrt{\frac{V_{af}}{V_i - V_{af}}}. \qquad [24]$$

For an emission-energy distribution given by Eq. [9], the image-center current density is:

$$J'(0) = \int_{V_{af}/V_{im}}^{1} \frac{2I'Er_o}{V_i} \sqrt{\frac{V_{af}}{V_i + V_{af}}}\, dN$$

$$= \frac{2\pi^2 I' r_o}{\pi - 2} \frac{E}{V_{im}} \sqrt{\frac{V_{af}}{V_{im}}} \left[\cos\left(\frac{\pi}{2} \frac{V_{af}}{V_{im}}\right) C\left(\sqrt{1 - \frac{V_{af}}{V_{im}}}\right) \right.$$

$$\left. - \sin\left(\frac{\pi}{2} \frac{V_{af}}{V_{im}}\right) S\left(\sqrt{1 - \frac{V_{af}}{V_{im}}}\right) \right], \qquad [25]^*$$

where C and S are the familiar Fresnel integrals[7] and I' is the emission current density of the polyenergetic electron emitter. Fig. 4 shows the plot of Eq. [25].

---

* The method of breaking the integral into two Fresnel integrals was suggested to the author by E. G. Ramberg of the David Sarnoff Research Laboratory, Princeton, New Jersey.

References:

[1] I. P. Csorba, "Resolution Limitations of Electromagnetically Focused Image-Intensifier Tubes," **RCA Review,** Vol. 30, No. 1, p. 36, 1969.
[2] E. Ruska, "Focusing of Cathode-Ray Beams of Large Cross Section," **Zeit für Phys.,** Vol. 83, p. 684, July 1933.
[3] P. Schagen et al, "A Simple Electrostatic Electron-Optical System with Only One Voltage," **Philips, Res. Rev.,** Vol. 7, p. 119, 1952.
[4] V. K. Zworykin et al, **Electron Optics and The Electron Microscope,** pp. 584-585, John Wiley & Sons, Inc., New York, 1945.
[5] J. W. Coltman, "The Specification of Imaging Properties by Response to a Sine Wave Input," **Jour. Opt. Soc. of America,** Vol. 7, No. 2, p. 121, April 1960.
[6] A. H. Sommer, "Semiconducting Photoemitters," **Optica Acta,** Vol. 7, No. 2, p. 121, April 1960.
[7] Jahnke-Emde, **Tables of Functions,** p. 35, Dover Pub., New York, 1945 ed.

Computer-Aided Design of a Fast-Wave Coupler*

M. J. Schindler, RCA Electronic Components, Harrison, N. J.

Abstract—Fast-wave couplers are used in certain crossed-field devices to impress the signal information on the electron beam. In addition to producing the desired increase in the cyclotron radius, this modulation causes trajectory displacements of the electrons and PM-AM conversion of the signal and thereby limits the useful bandwidth of the device. This paper describes the application of computer techniques in the study of this modulation. Gratifying agreement was obtained between the calculated beam behavior and data obtained from a beam analyzer. For a specific application, coupler characteristics can thus be predicted with confidence, and the relevant parameters can be optimized in advance.

Statement of the Problem

The cycloidal motion of electrons drifting in a cross electric and magnetic field can be used to impress information on the beam. The radius of the cycloids is varied by means of the "fast-wave coupler", shown schematically in Fig. 1, which was proposed by Cuccia for the purpose of amplitude modulation at high power levels.[1] Arnett[2] applied this coupler in an M-type device and it was later used in various applications.[3,4,5] Undesirable side effects can, however, seriously curtail the usefulness of the coupler. Prediction of such effect is, therefore, an important aspect of fast-wave coupler design, and represents the main purpose of this paper.

The electron in Fig. 1 is assumed to enter the coupler from the left with a drift velocity u_o and a negligibly small amount of rotation. In Region A only the dc electric field prevails, and the electron moves along an equipotential line V_g, called the "guide line." In Region B, an rf field is superimposed on the dc field, which oscillates at a frequency fairly close to the cyclotron frequency of the electron, thereby super-

* The work described was supported in large part by the Office of Naval Research, Washington, D.C.

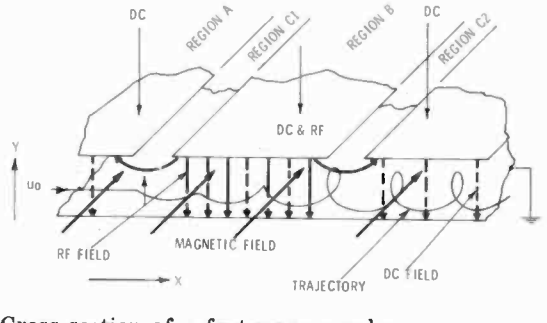


Fig. 1—Cross section of a fast-wave coupler.

imposing a rotational motion of the desired amplitude on the uniform, translational drift motion. As will be shown, the modulation is caused by the Y component of the rf field; the X component, which prevails in the transition Regions (C1, C2) has the detrimental effect of causing the electron to shift to a different "guide line." The amount of this shift depends on the "phase" of the electron; the trajectory illustrated in Fig. 1 is quite undesirable because this particular electron lingers longest in the harmful fringing field present in Region C2. The consequence of this phase sensitivity is a thickening of the entire beam, which can lead to a partial or complete loss of its information content.

The computer-aided design procedure falls naturally into three parts: first, the calculation of the signal voltage V_s across the coupler, which can be obtained with a given input power; second, the calculation of the X and Y components of the resulting rf fields; third, the calculation of the electron trajectories in these fields and the resulting beam modulation and beam thickness, which will be compared with experimental data obtained with a beam analyzer.

The RF Model for the Coupler

The physical configuration of the coupler is shown in Fig. 2. The first transmission line segment T_2 is a short 50-ohm section starting at the

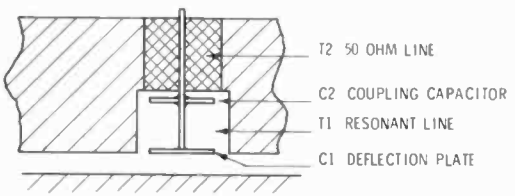


Fig. 2—Electrically the coupler consists of two transmission lines (T1, T2) and two capacitors (C1, C2).

input connector. Capacitor C2 represents the coupling discontinuity (designated "coupling capacitor" in Fig. 2). The "resonator" constitutes a second transmission line segment T1, which is terminated in the capacitance C1 of the "coupler plate". The voltage V_g between this plate and ground determines the magnitude of the rf field in the coupler; its determination for a given input (e.g., 1 watt) is the purpose of the present calculation.

A time-sharing routine (ASPIC) was applied to solve this problem. This routine determines voltages and impedances of microwave networks with discontinuities (see Appendix 1). A voltage of 1 volt is assumed across the load, and the corresponding voltage at the input of the network is calculated back step by step, until the generator power necessary for the appearance of 1 volt across the load is obtained. Finally, this correlation is reversed, and the voltage V_g for a generator power of 1 watt is calculated. Typical curves of V_g as a function of frequency are shown in Fig. 3 for different values of C_2.

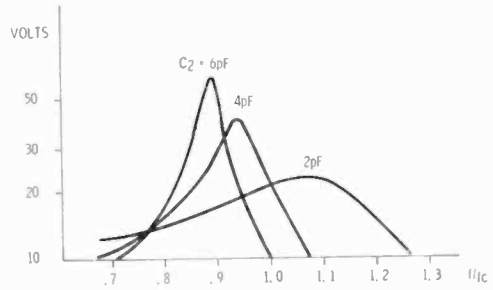


Fig. 3—The voltage across the coupler gap as a function of normalized frequency for three values of the coupling capacitor.

The Electrical Field Components

Both the dc and the rf electric fields in the coupler contain, in addition to useful transverse (Y) components, detrimental components in the axial (X) direction of electron flow. In Fig. 4A, which depicts the field configuration for the rf case, the X components are quite obvious. In the dc case (Fig. 4B), the X components are a much smaller fraction of the total field; on the other hand, the absolute value of the dc field can be 1 or 2 orders of magnitude higher than that of the rf field, and the dc X components can become significant.

The field components were obtained by means of a batch-mode program that solves Laplace's equation in both axial and planar symmetry[6] for both the asymmetrical configuration of Fig. 4A and a symmetrical

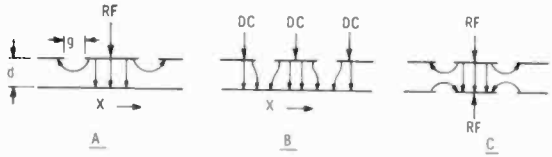


Fig. 4—Fields in an asymmetrical coupler have rf components (A) and dc components (B). The symmetrical coupler (C) differs mainly in its rf fields.

one, as in Fig. 4C. To illustrate their importance, the results are briefly discussed for the rf components of a coupler with a height d of 1 mm (40 mils) and a gap width g of 0.1 mm (4 mils).

The useful modulating field E_m is the Y component, which would ideally jump from 0 to 100% at the edge of the coupler plate. In reality, there is a substantial overshoot in the asymmetrical coupler, to 2.5 E_m as shown in Fig. 5. In the symmetrical case, the effect is much less pronounced (1.5 E_m). (For clarity, Fig. 6 shows only the rear half of the transverse field; it has even symmetry with respect to the X-E_y plane.) The most valuable results of the field calculations, shown in Figs. 7 and 8, are the X components for the asymmetrical and symmetrical case, respectively. In the asymmetrical field, the harmful X component reaches values as high as 3 times the useful Y component, while the symmetrical design cuts this value in two

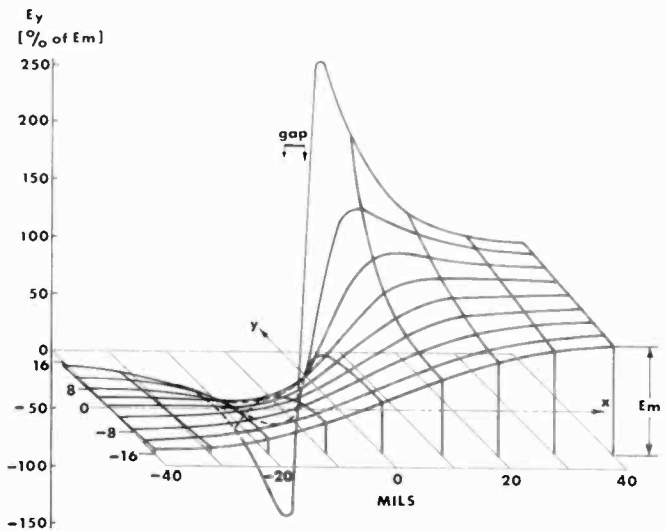


Fig. 5—The transverse (modulating) field in an asymmetrical coupler near the gap.

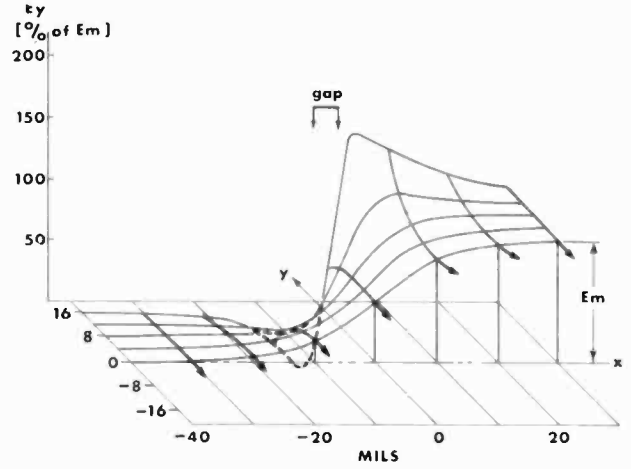


Fig. 6—The transverse (modulating) field in a symmetrical coupler near the gap.

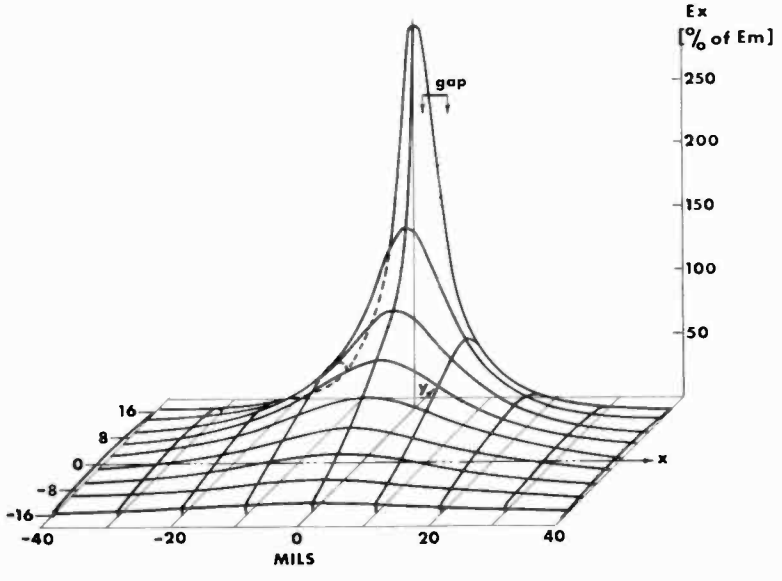


Fig. 7—The axial (harmful) field in an asymmetrical coupler near the gap.

Trajectory Calculation

Having obtained the necessary knowledge of the field components in the coupler, it is now possible to proceed with the calculation of electron trajectories. In uniform crossed electric and magnetic fields, the electron follows a cycloidal path[7] with a translational component u_o that constitutes the drift velocity

$$u_o = E_o/B_o, \qquad [1]$$

where E_o is the uniform electric field and B_o the uniform magnetic field.

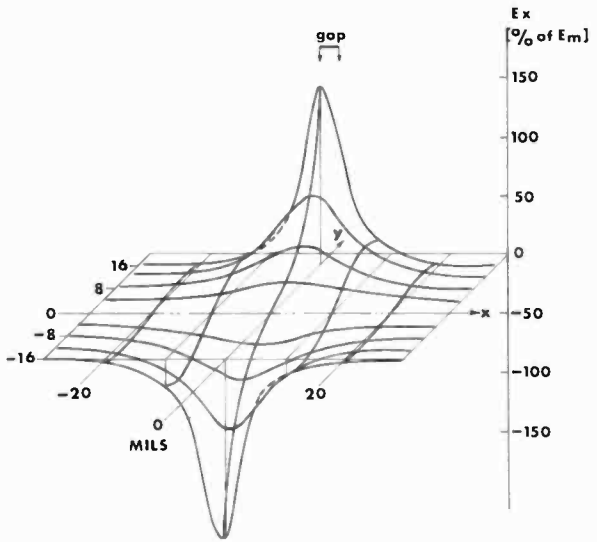


Fig. 8—The axial (harmful) field in a symmetrical coupler near the gap.

The rotational component is described by the frequency of rotation

$$\omega_c = \frac{e}{m_e} B_o = \eta B_o \qquad [2]$$

where e/m_e is the electron charge to mass ratio ($e = 1.6 \times 10^{-19}$ Coul., $m_e = 9.1 \times 10^{-31}$ kg). Finally, the radius of rotation is given by

$$m = \eta \frac{E_o}{B_o^2} \qquad [3]$$

It is helpful to establish some numerical values for these equations. Assuming a coupler height of 1 cm, a voltage of 1 kilovolt and a magnetic field of 1 kilogauss, the drift velocity becomes 10^6 meters per second, the cyclotron frequency is 2799.2 MHz, and the cycloidal radius is 0.056857 mm, or approximately 2 mils.

In an actual coupler the electric fields are, as shown above, neither homogeneous nor constant, and an exact solution of the equations of motion will be cumbersome at best, and certainly not suitable for interactive computer-aided design. A different approach was therefore chosen, based on the impulse concept. The electron is assumed to drift

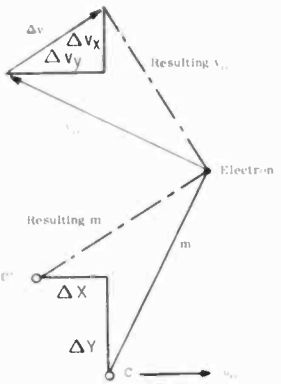


Fig. 9—The orbiting electron before and after an impulse.

with the constant velocity, and a coordinate system that moves at the same velocity is chosen. In this coordinate system, the unperturbed trajectory is a circle with radius m, and the tangential velocity of the electron v_o is given by

$$v_o = 2\pi f_c m = \frac{m}{Q}. \qquad [4]$$

If the electron is exposed to a field impulse, the velocity v_o of the electron will be changed by an amount proportional to, and in the direction of, this impulse. This velocity change Δv can be broken down into a longitudinal component Δv_x in the direction of u_o, and a transverse component Δv_y perpendicular to u_o. The new rotational velocity $v_o + \Delta v$ requires a new center of rotation (Point C′). With the nomenclature of Fig. 9, the coordinates of point C(X_c, Y_c) change by an amount ΔX and ΔY given by

$$\Delta X = -Q\,\Delta v_y \qquad [5]$$

$$\Delta Y = Q\,\Delta v_x \qquad [6]$$

Physically, the changes ΔX represent temporary displacements of the center of rotation, while the changes ΔY represent guideline shifts. These are, as indicated by Eq. [6], caused by axial components of the electric field.

To determine the magnitude of Δv_x and Δv_y, the effect of an electric field on the electron must be considered. In the present analysis, the continuous force on the electrons is replaced by a series of impulses that produce the same effect. This simplified method can not, however, account for the "feedback" effects of the electron on the field. Because the magnetic fields of the moving electrons are negligible anyway, the impulse method basically amounts to a neglect of the space-charge force.

In an electric field of intensity E the force F exerted on an electron is

$$F = eE. \qquad [7]$$

The electron therefore experienced an acceleration a of

$$a = \eta E. \qquad [8]$$

The resulting velocity change in the direction of the field, Δv, is

$$\Delta v = \int_{t_1}^{t_2} a\,dt = \eta \int_{t_1}^{t_2} E\,dt. \qquad [9]$$

In the case of a sinusoidally varying field of frequency f, Δv is given by

$$\Delta v = \eta \int_{t_1}^{t_2} \hat{E} \sin \omega t. \qquad [10]$$

For half a sine wave, the integral is evaluated as follows:

$$\Delta v = \eta \hat{E}\,\frac{1}{2\pi f}\left|\int_0^{t=1/2f} \cos 2\pi f t\right| = 5.6 \times 10^{10}\,\frac{\hat{E}}{f} \qquad [11]$$

The effect of the substituted impulse will equal that of half the sine-wave if the pulse magnitude is also $5.6 \times 10^{10}\hat{E}/f$, which has the dimension of volt-seconds/meter.

If the effect of the fringing fields are to be evaluated, a very close pulse sequence is required, because the electron may pass this region in much less than one cycle. For time intervals short compared with a cycle, Eq. [10] can be approximated by

$$\Delta v \doteq \eta E \sin \omega t \, \Delta t. \qquad [12]$$

To follow an electron through the whole coupler by this method is quite straightforward and ideally suited for application of a time-shared computer.

Some key features of the program (described in Appendix 2) will now be explained. The values of the field components E_x (X component) and E_y (Y component) as obtained from the potential-solving routine are approximated by fairly simple algebraic expressions (surface-matching) and are contained in "Subroutine Field". An electron is introduced into Region A of Fig. 1 at a given guide line position (YC), and the calculations of Eqs. [5] and [6] are performed. The electron is then permitted to drift for a short time (in both translation and rotation) and is again operated upon by Eqs. [5] and [6]. In the transition regions (C1 and C2 in Fig. 1), the time increments are decreased by some factor (usually 4) because of the rapid field changes, and the calculations continue until the electron either leaves the coupler or is intercepted on the electrodes. Additional electrons are then introduced that lag the previous one by a given phase angle (usually 36°). These electrons may linger in the region with high axial field for longer or shorter periods; their trajectories consequently do not coincide. When 360° has been covered, a high and a low value for the guide line positions (YCH and YCL, respectively) are available, as well as a high and low value for the modulation radius (RH and RL), and they are printed out. The calculation is then repeated at a different frequency.

Calculated Results

For the case of an asymmetrical coupler, two typical trajectories in the second transition region (C2 in Fig. 1) are shown in Fig. 10, at the cyclotron frequency f_c (10A) and at 1.1 f_c (10B). The two electrons are 180 degrees out of phase; while one of them spends quite some time in the region of X-fields higher than 0.5, the other is never exposed to

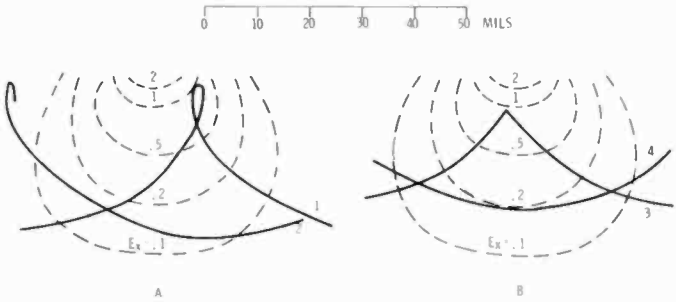


Fig. 10—Two electron trajectories in the gap region at cyclotron frequency f_c (1 and 2), and at 1.1 f_c (3 and 4).

more than 0.2. It should be kept in mind that the X fields (broken lines in Fig. 10) vary with the signal frequency, so that an electron may appear to linger briefly in the high-field region without actually being exposed to the detrimental field.

The results of one complete calculation are summarized in Fig. 11 as a function of the ratio f_s/f_c, where f_s is the signal frequency and f_c the cyclotron frequency. The trajectories numbered 1 through 4 in Fig. 10 correspond to the points 1 through 4 in Fig. 11. Each vertical line represents a 360° phase variation of the signal, approximated by 10 discrete increments; Fig. 11 thus involves 150 trajectory calculations.

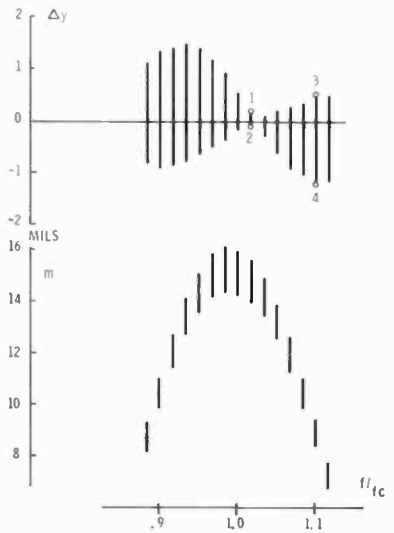


Fig. 11—Guide line shift ΔY and modulation radius m for an asymmetrical coupler.

The bottom part of Fig. 11 shows the modulation radius as a function of frequency, and it, too, is affected by the phase of the modulating signal. The reason for this is most easily understood if one visualizes modulation by just two half waves of sin ωt. If the electron enters this field at $t = 0$, the electron is deflected by the first half-wave and its rebound fully coincides with the second half wave. If it enters at $t = \frac{1}{4}$, however, the field reverses before the electron is ready to swing back at its natural (cyclotron) frequency, and part of the potential modulation is lost.

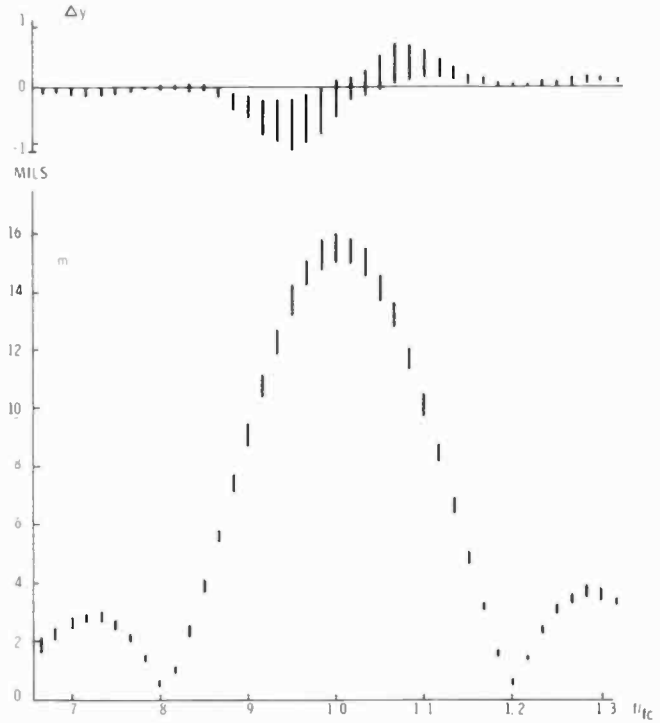


Fig. 12—Guide line shift ΔY and modulation radius m for a symmetrical coupler.

Because, in any coupler with substantial bandwidth, the electron is exposed to only a few half waves (typically 5 to 10), some amplitude modulation due to rf phase (PM-AM conversion) must always be expected.

If the coupler has a symmetrical configuration (Fig. 4C), an electron traveling on the symmetry axis does not experience any harmful axial fields at all, and better modulation characteristics can be expected. This is borne out by the curves of Fig. 12, which otherwise represents

operating conditions nearly identical to those of Fig. 11. Fig. 12 covers a full octave of signal frequency, so that the first side lobes of the $\sin x/x$ modulation envelopes can be seen.

Experimental Results

In order to measure beam position and modulation at the output of a fast wave coupler, a beam analyzer was constructed, as shown schematically in Fig. 13. After leaving the coupler, the beam enters a

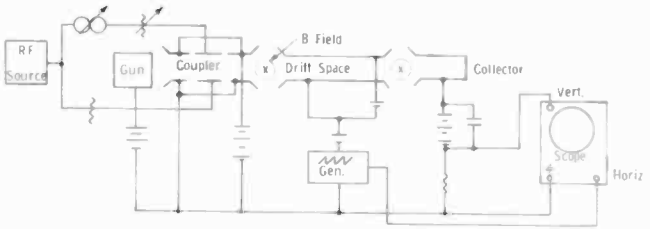


Fig. 13—Experimental apparatus for the analysis of a crossed-field electron beam.

drift space that has a much smaller electric field gradient than the coupler. Consequently, the beam must spread out by the ratio of the two gradients. Thus, relatively small guide-line displacements can easily be measured by varying the potential of the drift space until the beam is intercepted either at the lower or the upper surface. That part of the beam which reaches the collector causes a vertical deflection on the oscilloscope.

Depending on the modulation and thickness of the beam, the oscilloscope display acquires different shapes. Fig. 14A describes the most general beam condition: the unmodulated centerline has the potential p, the beam thickness is t, the modulation radius of the upper boundary layer is M, and that of the lower one m. All dimensions are measured in volts. The ideal unmodulated beam is infinitely thin; as the drift-space potential V_d is increased, the beam will suddenly reach the collector at $V_d + \Delta V = p$, and suddenly disappears again at $V_d = p$ as shown in Fig. 14B. A beam which has the finite thickness t creates the oscilloscope display of Fig. 14C. Modulation manifests itself basically in a narrowing of the display because the whole beam is intercepted if its electrons, at *any* moment during their orbit, reach

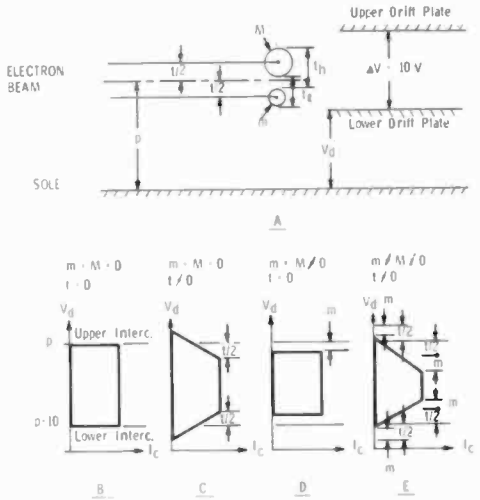


Fig. 14—Definition of terms for a beam in the drift region of the analyzer (A) and resulting oscilloscope displays for the unmodulated ideal beam (A), a beam with finite thickness (C), a modulated ideal beam (D) and the general beam shown above (E).

the potential of the drift electrode (Fig. 14D). In the most general case, where the modulation of a beam is not uniform throughout its thickness, the display becomes asymmetrical, because the beam appears to have a different thickness if approached from the side of higher potentials (t_h), or from the low side (t_l). As illustrated in Fig. 14E, all the beam characteristics can be reconstructed from its oscilloscope display. Fig. 15 shows beam "profiles" for different degrees of modulation; the traces are offset by ½ cm vertically for clarity.

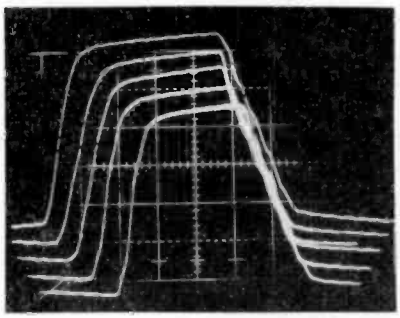


Fig. 15—Photograph of the oscilloscope display for a beam at five different modulation levels. The traces are offset vertically for clarity.

In Fig. 16, the results of such a beam analysis are summarized for the case of a symmetrical coupler. The first column (A) gives the measured modulation radius in mils, as a function of input power, for 7 signal frequencies f_s, in terms of the cyclotron frequency f_c. The second column (B) shows the corresponding beam boundaries, also in mils, again as a function of input power (or coupler voltage). In Columns C and D, the corresponding calculated values are shown. The agreement between measured and calculated modulation is well within the experimental error, an indication that the substitution of discrete

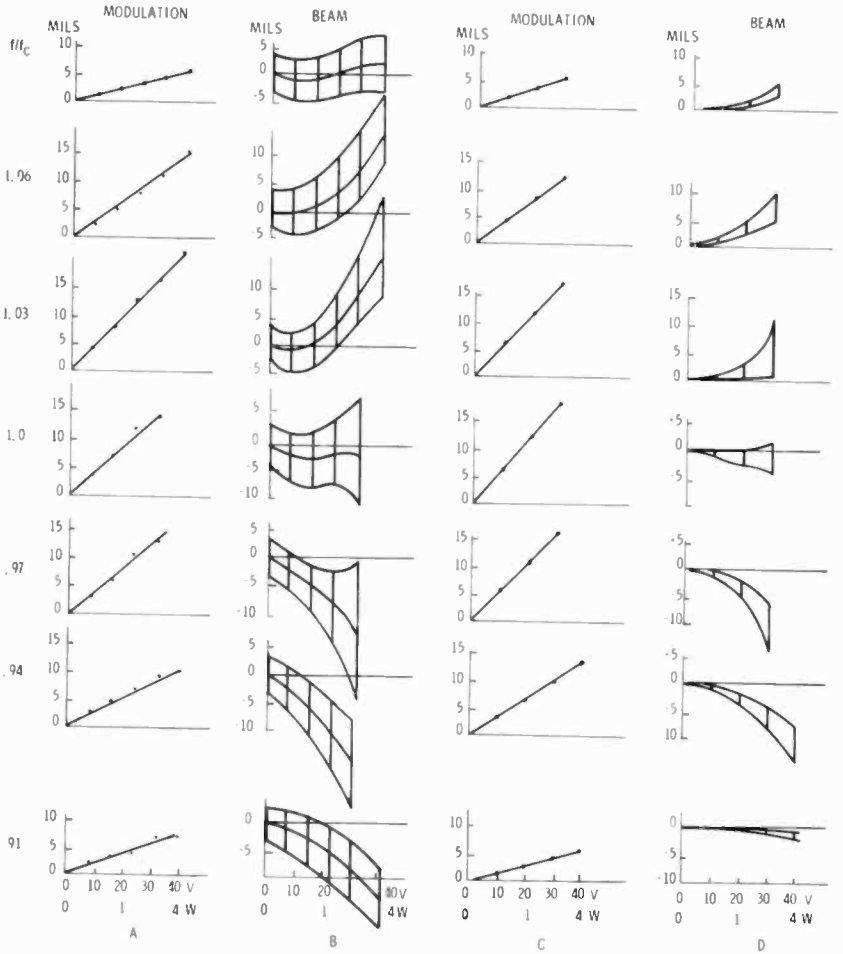


Fig. 16—Comparison of calculated and measured results at seven normalized frequencies, as a function of modulating power: measured modulation radius (A), measured beam configuration (B), calculated modulation radius (C) and calculated beam configuration (D).

impulses for continuous motive forces is adequate. Comparison of columns B and D indicates the major difference between calculated and measured beam behavior very clearly: the actual beam has finite thickness, even before any rf forces are applied. Otherwise the agreement is quite gratifying when it is considered that the calculation is relatively coarse and that the actual beam is affected by the transitions from the injection gun to the coupler and from the coupler to the drift space. There is, furthermore, a small error in the f_s/f_c ratio for the measured data, as evidenced by the asymmetry of the measured against the calculated modulation.

Conclusions

Three computer programs (two of them interactive) were used in the design of a fast-wave coupler for the modulation of a crossed-field

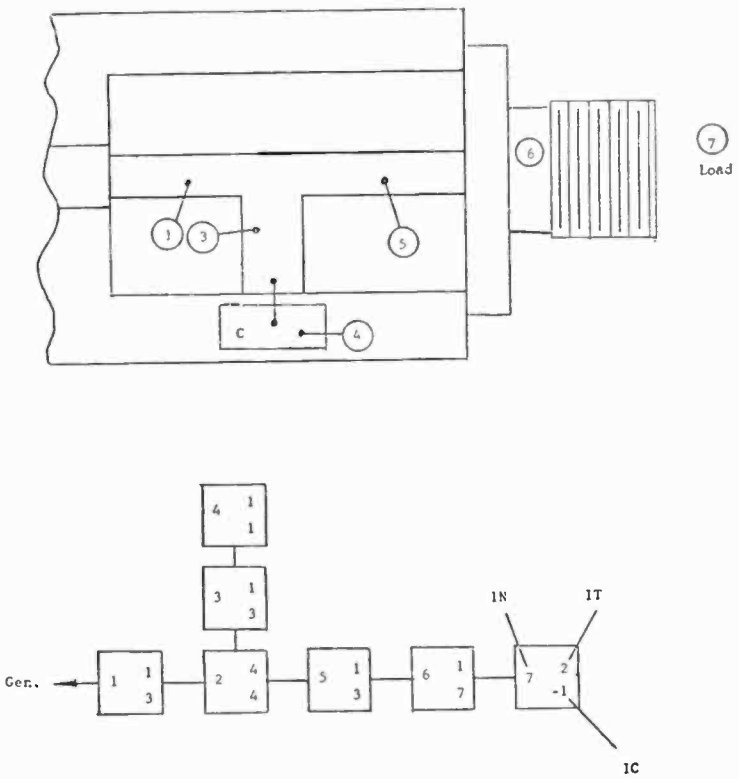


Fig. 17—A simple microwave circuit and its ASPIC representation. Circled numbers correspond to "IN" in schematic.

electron beam. Very good agreement was obtained between the calculated beam behavior and data obtained from a beam analyzer with the exception that the real beam is always somewht thicker than the calculated beam. The calculation can thus be utilized to optimize the coupler design for specific applications and substantial improvements in performance were in fact obtained.

Acknowledgments

The author thanks the many people in RCA who contributed to this work, especially Dr. Erwin Belohoubek, on whose work both the theoretical and experimental aspects of this paper are based, and who provided guidance and stimulating discussions.

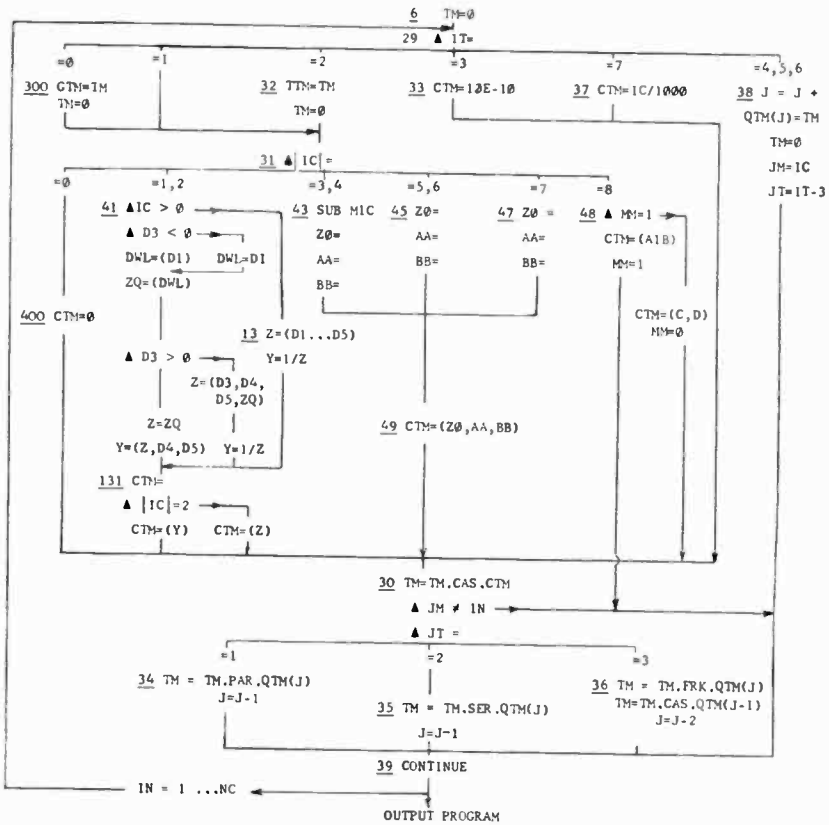


Fig. 18—Flowchart of the circuit analysis section of ASPIC.

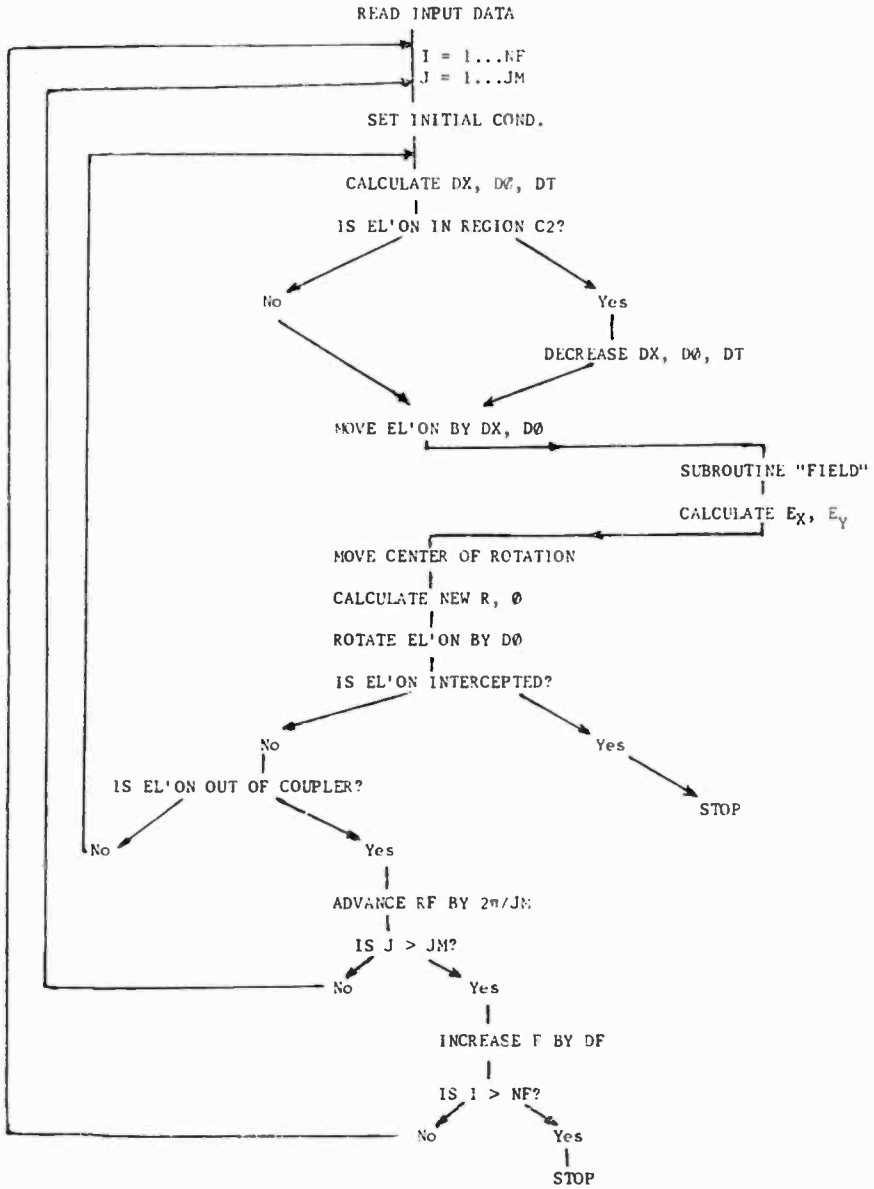


Fig. 19—Flow chart for the trajectory calculations.

Appendix 1—ASPIC, A Time-Sharing Program for the Analysis of Microwave Circuits

ASPIC is a general-purpose network analysis program specifically written to suit the microwave community. The circuit is broken down into its components (lumped elements, transmission lines etc.) as shown in Fig. 17. Each element IN is described as to topology (IT), kind of circuit (IC), and its physical or electrical parameters. The transfer matrix (ABCD matrix) is calculated for each element, and the matrices are manipulated according to the topology, as indicated in the flowchart (Fig. 18). The user can choose from a variety of outputs, such as VSWR, loss, or phase. Complex quantities (e.g. admittances) are available both by real and imaginary part, or as magnitude and angle.

Appendix 2—Trajectory Calculations

The following program was written for a time-shared computer in modified FORTRAN IV. The calculations expressed in Eqs. [5] and [6], and the drift Eqs. [1] and [2] form the heart of the program.

The flow chart of the program is shown in Fig. 19. After the input data are read, the time (DT), translation (DX), and phase ($D\phi$) increments are calculated, and decreased by a factor M if the electron is near the output gap of the coupler. The electron is then translated by DX, and in the appropriate subroutine the X and Y components of the coupler fields are found. The new center of rotation is found, and the electron rotates by $D\phi$. If this brings the electron close to a coupler surface, the calculation stops; otherwise, the above procedure is repeated until the electron leaves the coupler. This calculation is performed for JM values of the phase, and for NF frequencies.

References:

[1] C. L. Cuccia, "The Electron Coupler—A Developmental Tube for Amplitude Modulation and Power Control at Ultra-High Frequencies," **RCA Review,** Vol. 10, p. 270, June 1949.

[2] H. D. Arnett, "Electron Beam Interaction in Cuccia-Coupler and Its Effect on Quadrupole Amplifier Performance," NRL Report No. 5695.

[3] Adler, et al, "Quadrupole Amplifier, A Low Noise Paramp Device," **Proc. IRE,** Vol. 47, p. 1713, Oct. 1959.

[4] Dain, et al, "A Cyclotron Resonance Frequency Multiplier," **IRE Trans. Elec. Devices,** Vol. ED-10, p. 195, May 1963.

[5] J. W. Klüver, "An Electrontically Variable Delay Line," **Proc. IRE,** Vol. 50, p. 2487, Dec. 1962.

[6] H. Kulsrud, "Programming System for Electron Optical Simulation," **RCA Review,** Vol. 28, p. 351, June 1967.

[7] W. J. Kleen, **Electronics of Microwave Tubes,** p. 14, Academic Press, N. Y., 1958.

Synthesis of Optimum Reflection-Type Microwave Equalizers

C. M. Kudsia, RCA Limited, Montreal

Abstract.—This paper presents an algorithmic approach towards optimizing the design of microwave equalizers to achieve the best performance available from an equalizer of a given order under a given set of design constraints. The design constraints chosen are consistent with maximizing the performance of a communication channel. The algorithm yields a trade-off among the order of the equalizer, its group delay response, its amplitude response (or gain–slope), and the unloaded **Q** of the equalizer network. The approach is general and can be adopted to optimize all-pass networks in any frequency range.

1. Introduction

Most of the present microwave communication systems are of the terrestrial type employing i-f-type radio repeaters. These repeaters are spaced so as to ensure line-of-sight propagation. The rf signal is separated into its various channels in these repeaters; each channel is then converted to an intermediate frequency (70 MHz) that is subsequently amplified, equalized for group delay or amplitude variation if necessary, and then reconverted to rf. All such channels are then recombined at the output before transmission to the next repeater or station. This technique of amplification and equalization at i-f offers the advantage of using standard and relatively cheap components available at these frequencies. The weight of the repeater and its associated reliability is of no great consequence as these repeaters are located on ground.

In the recent years, the use of communication satellites as single radio repeaters to achieve global communication has resulted in a somewhat different philosophy as regards amplification and equalization of the signals. First of all, in satellite systems, very wide bandwidth (500 MHz) is employed. The main disadvantages of employing an i-f scheme for interstage amplification and equalization in such a system

are that a high i-f ($\geqslant$500 MHz) has to be used if preselection at 6 GHz (satellite communication band) is to be avoided, the number of local oscillator sources is large, and, most important of all, a high level up-converter is needed to provide sufficient drive for the output traveling-wave tube. This creates a serious spurious and intermodulation problem. On the other hand, an rf scheme utilizing rf amplification and equalization offers the advantages of simplicity, a lower parts count, and the requirement of a single translating local oscillator source. The main disadvantage is the size and weight of the associated branching networks comprising filters and microwave equalizers. This paper presents an approach towards optimizing the design of microwave equalizers so as to achieve the best performance available from an equalizer of a given order under a given set of design constraints and, therefore, a minimum weight for a given performance of the equalizer.

A computer program is developed that determines the optimum location of poles and zeros for a given order of equalizer under the given constraints. These constraints are determined from the overall performance requirement, e.g., the maximum variation in group delay, the maximum gain slope, and the maximum group delay ripple that can be tolerated. These criteria help determine the trade-offs among the various parameters of the equalizer under a given set of conditions. Further, a subroutine in the computer program translates this location of poles and zeros into an equivalent microwave reactance network, similar to that reported by Cristal[1] but consisting of symmetrical inductive irises. The relation between susceptances and mechanical dimensions is based on experimentally determined values of susceptances for different iris dimensions.[5] Theoretical and experimental data of two prototype models are discussed in the paper.

The relevant theory of all-pass commensurate transmission line networks is discussed in Section 2, followed by a discussion on the algorithm for optimization of the equalizer in Section 3. Sections 4 and 5 discuss the physical realization and experimental results.

2. General Theory

The theory of commensurate-transmission-line all-pass microwave networks is well established.[1,2] However, to present the logical development of an appropriate algorithm for optimizing all-pass networks, relevant background on all-pass networks as applied to commensurate transmission line structures is discussed in this section.

In network theory, any two-port electrical network can be described by its transfer function $S_{12}(t)$,

$$S_{12}(t) = \frac{a_0 t^n + a_1 t^{n-1} + \ldots a_n}{b_0 t^m + b_1 t^{m-1} + \ldots b_m}, \qquad [1]$$

where a and b are real coefficients and t is the complex frequency variable. In factorized form, the transfer function can be represented as

$$S_{12}(t) = \frac{a_0 (t - z_1)(t - z_2) \ldots (t - z_n)}{b_0 (t - p_1)(t - p_2) \ldots (t - p_n)}, \qquad [2]$$

where the roots $z_1, z_2, \ldots z_n$ are the zeros of the transfer function (denoted by 0) and $p_1, p_2, \ldots p_m$ are the poles (denoted by x) of the transfer function. In other words, the electrical network can be

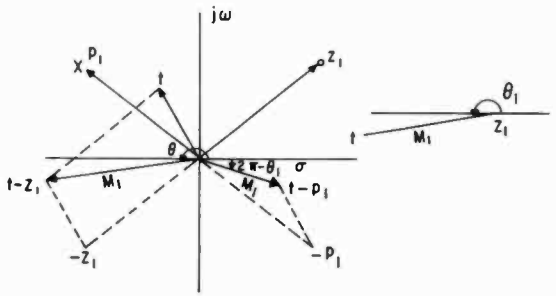


Fig. 1—Complex frequency plane representation of the vectors (t-z_1) and (t-p_1) in polar form.

described in terms of its poles and zeros in the complex frequency plane. In terms of the polar coordinates, the zeros and poles may be represented by

$$t - z_1 = M_1 \exp\{j\theta_1\}, \text{ for the zero}$$
$$t - p_1 = m_1 \exp\{j\phi_1\}, \text{ for the pole}$$

Where $j = \sqrt{-1}$. M_1 and m_1 represent the magnitude while θ_1, ϕ_1 represent the phase of the vectors drawn from t to z_1 and to p_1, respectively, as shown in Fig. 1. The transfer function can, therefore, be represented in terms of magnitude and phase angle by

$$S_{12}(t) = \frac{a_0}{b_0} \frac{M_1 M_2 \ldots M_n \exp\{j(\theta_1 + \theta_2 + \ldots \theta_n)\}}{m_1 m_2 \ldots m_n \exp\{j(\phi_1 + \phi_2 + \ldots \phi_n)\}}. \qquad [3]$$

If the function $S_{12}(t)$ is evaluated at real frequencies, i.e., for $t = j\omega$ and if $M_1 = m_1$, $M_2 = m_2, \ldots$ for all frequencies and $n = m$, then

$$S_{12}(t) = \frac{a_0}{b_0} \exp \{j(\theta_1 + \theta_2 + \ldots + \theta_n - \phi_1 - \phi_2 \ldots - \phi_n)\} \quad [4]$$

The condition $M_1 = m_1$ implies that

$$|\sigma_1 + j(\omega - \omega_1)| = |-\sigma_1' + j(\omega - \omega_1')|$$

where $Z_1 = \sigma_1 + j\omega_1$ and $p_1 = -\sigma_1' + j\omega_1'$.

For the above condition to hold at all frequencies we must have $|\sigma_1| = |-\sigma_1'|$ and $\omega_1 = \omega_1'$. Under these constraints and assuming the arbitrary constants a_0, b_0 to be equal,

$$S_{12}(t) = \exp \{j(\theta_1 + \theta_2 + \ldots \theta_n - \phi_1 - \phi_2 - \ldots \phi_n)\}. \quad [5]$$

Therefore, magnitude $\alpha = |S_{12}(t)| = 1$

$$\text{and phase } \beta = \theta_1 + \theta_2 + \ldots \theta_n - \phi_1 - \phi_2 - \ldots \phi_n$$
$$= -j \ln S_{12}(t)$$

evaluated at real frequencies. Such networks are referred to as all-pass networks. There is no change in the magnitude of a signal that is passed through such networks but the phase varies depending upon the location of poles and zeros.

From the conditions that $n = m$, and $M_1 = m_1$, $M_2 = m_2, \ldots$ at all frequencies, the following conclusions can be drawn for the all-pass networks:

(a) For every pole, there must exist a zero situated symmetrically about the $j\omega$ axis.

(b) For non-real poles and zeros, there must exist quadrantal symmetry about the origin in the complex frequency planes.

This implies that there are two basic types of all-pass networks—one having a real pole and a zero and the other having a pair of conjugate poles and zeros located symmetrically about the origin as shown in Figs. 2a and 2b. Mathematically, an all-pass function may therefore be represented by

$$S_{12}(t) = \frac{H(-t)}{H(t)}. \qquad [7]$$

$H(t)$ is a strict Hurwitz polynomial given by

$$H(t) = \prod_{i=1}^{k} (t + \sigma_i) \prod_{j=1}^{m} (t^2 + 2\sigma_j t + |t_j|^2), \qquad [8]$$

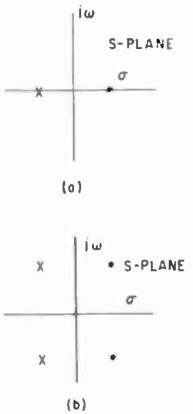


Fig. 2—Pole-zero pattern of all-pass networks: (a) all-pass C-type section and (b) all-pass D-type section.

where σ_i, $\sigma_j > 0$

$$t_j = \sigma_j + j\omega_j.$$

k represents the number of real poles (or zeros), and m represents the number of non-real poles (or zeros). $k + m$ gives the order of the all-pass network. i and j as subscripts of t, σ, and ω refer simply to the ith or jth pole (or zero). In microwave terminology, all-pass sections with a real pole are referred to as C-sections and all-pass sections with nonreal poles (or zeros) are referred to as D-sections.

Thus far, no mention has been made of the nature of the all-pass networks, i.e., if the networks are to be realized as lumped circuits or as distributed circuits. Conclusions drawn thus far are independent of frequency. If, in the physical realization of the network, we must use commensurate transmission line lengths, the complex frequency variable must be replaced by

$$t = \tanh \gamma l$$
$$= \sum + j\Omega \qquad [9]$$

where γ is the complex propagation constant, and l is a commensurate length of the network. Ω refers to the real-frequency axis in the complex frequency plane. This result is owing to Richard's transformation[3] for networks consisting of lumped resistors and commensurate transmission lines. For the lossless case, $\Sigma = 0$ and

$$t = j \tan \frac{2\pi}{\lambda} l, \qquad [10]$$

where λ is the wavelength in the transmission medium.

If n is the effective number of unit elements in the network, the expression for the transfer function becomes[2]

$$S_{12}(t) = \left(\frac{1-t}{1+t}\right)^{n/2} \frac{H(-t)}{H(t)}. \qquad [11]$$

The appearance of factor $[(1-t)/(1+t)]^{n/2}$ in the transfer function is equivalent to adding a delay of a transmission line $n/4$ radians long (each commensurate line length of the network is taken as quarter wavelength at the center frequency).

The phase β of this transfer function would therefore be

$$\beta = -j \ln S_{12}(t), \qquad [12]$$

and the group delay on the real-frequency axis would be

$$\tau = -\left.\frac{d\beta}{d\omega}\right|_{t=j\Omega}$$
$$= -\left.\frac{1}{2\pi} \frac{1}{f_0} \frac{d\beta}{dt} \frac{dt}{dF}\right|_{t=j\Omega} \qquad [13]$$

where $F = f/f_0$, f is the frequency variable, and f_0 is the center frequency.

From Eqs. [8] to [13], it can easily be shown that

$$\tau_N = \tau f_o = \frac{1}{4}\left[n + (1+\Omega^2)\left\{ \sum_{i=1}^{k}\left(\frac{1}{\sigma_i + j\Omega} + \frac{1}{\sigma_i - j\Omega} \right) + \right.\right. \qquad [14]$$

$$\left.\left. \sum_{j=1}^{m}\left(2\frac{\sigma_j + j\Omega}{|t_j|^2 - \Omega^2 + j2\sigma_j\Omega} + 2\frac{\sigma_j - j\Omega}{|t_j|^2 - \Omega^2 - j2\sigma_j\Omega} \right) \right\} \right].$$

If waveguide structure is used, the normalized group delay is (see Appendix 2)

$$\tau_N = S(F, F_o)\frac{1}{4}\left[n + (1+\Omega^2)\left\{ \sum_{i=1}^{k}\left(\frac{1}{\sigma_i + j\Omega} + \frac{1}{\sigma_i - j\Omega} \right) + \right.\right. \qquad [15]$$

$$\left.\left. \sum_{j=1}^{m}\left(2\frac{\sigma_j + j\Omega}{|t_j|^2 - \Omega^2 + j2\sigma_j\Omega} + 2\frac{\sigma_j - j\Omega}{|t_j|^2 - \Omega^2 - j2\sigma_j\Omega} \right) \right\} \right],$$

where $S(F, F_o) = F/(F^2 - F_o^2)^{1/2}(1 - F_o^2)^{1/2}$,

$F_o = f_{co}/f_o$,

f_{co} = cutoff frequency of the waveguide.

This result is the same as Eq. [23] in Cristal's[1] paper and clearly indicates that the group delay of an all-pass structure is simply the summation of group delays of individual all-pass sections with real zeros (C-type) and complex zeros (D-type). This property of all-pass structures for lumped circuits was first postulated by Bode.[4]

This relationship is very useful in evaluating the group delay/frequency characteristics of all-pass structures as we simply have to sum the delays of single C and D-type all-pass sections. From Eqs. [7], [8], [12] and [13], it can be shown that the normalized group delay τ_{NC} for a C-section is given by

$$\tau_{NC} = \frac{1}{2}\frac{\sigma_c(1+\Omega^2)}{\sigma_c^2 + \Omega^2} S(F, F_o); \qquad [16]$$

for the D-section, it is given by

$$\tau_{ND} = \sigma_d(1+\Omega^2)\frac{\Omega^2 + |t_d|^2}{\Omega^4 + 2\Omega^2(\sigma_d^2 - w_d^2) + |t_d|^4} S(F, F_o) \qquad [17]$$

σ_c refers to the real zero of the C-section and $t_d = \sigma_d + jw_d$ refers to the complex zero for the D-section.

These formulae are sufficient to determine an algorithm to optimize the location of poles and zeros of an all-pass structure under a given set of constraints.

3. Algorithm for Optimum All-Pass Networks

The criteria for optimization are chosen so as to be flexible in order to satisfy any or all of the following contraints:

(a) That the total system relative group delay (filter + equalizer) is a minimum.

(b) That the total system relative group delay lies within a prescribed region, or its deviation from this region is a minimum. This allows to determine the trade-off for flatness of group delay (i.e., control of group-delay ripple) in the center frequency region, where most of the energy is transmitted, against flatness at the pass-band edges, towards which the transmitted energy falls off very rapidly.

(c) That the total system relative group delay satisfies either or both of the above conditions with the added constraint that the value of the absolute group delay at center frequency of the equalizer has an assigned value. This added constraint determines the pass-band amplitude flatness of the equalizer and hence that of the system for a given Q_o that can be realized in the equalizer structure.

A typical situation when a communication channel needs equalization is depicted in Fig. 3.

Let the curve C_1 $(f_i, \tau_i;\ i = 1, 2, \ldots N)$ be the set of points that describe the absolute group delay versus frequency characteristics of the filter or any communication channel, f_i being the ith frequency point and τ_i being the corresponding absolute group delay. Let τC_{mi} and τD_{ni} be the group delays generated at a particular frequency f_i by mth C section (for a given σ_{cm}) and nth D section (for a given σ_{dn} and ω_{dn}), respectively. Then

$$\text{curve } C_2 = \left(\sum_m \tau C_{mi} + \sum_n \tau D_{ni}\right), \qquad i = 1, 2, \ldots N$$

represents the group delay of the equalizer. Summation over m and n simply implies the summation of group delays over all the C and D type sections assumed in the equalizer. Let d-level represent an arbitrarily chosen level of the total system group delay. Then we define a system function F by

$$F\ (\sigma_{cm}, \sigma_{dn}, \omega_{dn}) = \sum_{i=1}^{N} \left[d\text{-level} - \left(\tau_i + \sum_m \tau C_{mi} + \sum_n \tau D_{ni} \right) \right]^2 . \quad [18]$$

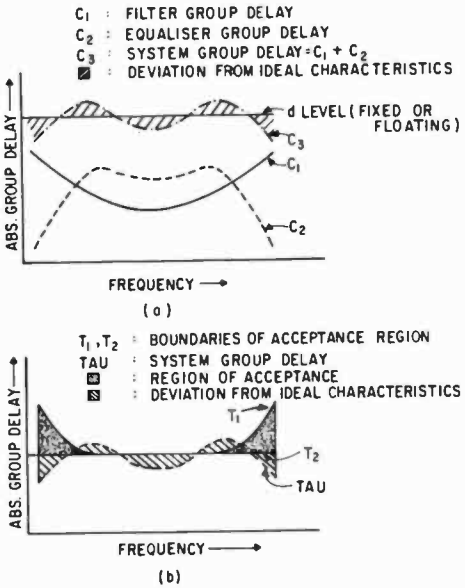


Fig. 3—Optimization criteria for the equalizer. (a) When the system group delay is constant across the desired band, the constant itself having either a preassigned value or allowed to float. (b) When the system group delay lies between the curves T_1 aid T_2 designated as the region of acceptance.

Minimization of this function F will realize an optimum equalizer subject to the constraint that the overall system group delay is fixed at d-level. Making d-level as one of the variables, i.e.,

$$F\ (\sigma_{cm}, \sigma_{dn}, \omega_{dn}, d\text{-level}) = \sum_{i=1}^{N} \left[d\text{-level} - \left(\tau_i + \sum_m \tau C_{mi} + \sum_n \tau D_{ni} \right) \right]^2 , \quad [19]$$

gives F an additional degree of freedom. Effectively it allows d-level to float. Further, let us assume that the overall system group delay (relative) is to lie within a prescribed region, say between curves T_1 and T_2 in Fig. 3b. Also, let

$$B_{1i} = d\text{-level} + T_{1i}$$
$$B_{2i} = d\text{-level} + T_{2i}$$

$$\text{TAU} = \tau_i + \sum_m \tau C_{mi} + \sum_n \tau D_{ni}$$

Where T_{1i} and T_{2i} represent the group delays of curves T_1 and T_2 at the frequency point. Then,

$$F\ (\sigma_{cm}, \sigma_{dn}, \omega_{dn}, d\text{-level}) = \sum_{i=1}^{N} (\text{TAU} - B_{1i})^2 + \sum_{i=1}^{N} (B_{2i} = \text{TAU})^2 \qquad [20]$$

$$\text{all TAU} > B_{1i} \qquad \text{all TAU} < B_{2i}$$

Minimization of F yields the optimum equalizer satisfying the constraint that the relative value of the overall delay lies in the region surrounded by curves T_1 and T_2. The number of C and D sections can be varied until the optimized system delay lies within curves T_1 and T_2. This allows design realization to within exact tolerances. If $T_1 = T_2 = d$-level, F reduces to the previous case (Fig. 3a).

The total number of variables X for F are $X =$ number of C-sections $+ 2 \times$ number of D-sections $+ 1$. As can be seen, the optimization process is achieved by reformulating the problem to a minimization one with X variables. Minimization of the function F, was achieved by modifying an existing scientific subroutine[6] that determines an unconstrained minimum of a function of several variables using the "rapid convergent descent method" proposed by Fletcher and Powell.[7] This assumes that the function F along with its gradient can be computed at any point of interest.

The optimized values of poles and zeros thus obtained are next converted into an equivalent reactance network.

4. Reactance Network from Pole-Zero Configuration

Any reactance network can be represented as[8]

$$Z(t) = \frac{\text{odd } H(t)}{\text{even } H(t)} \text{ or } \frac{\text{even } H(t)}{\text{odd } H(t)}$$

where $Z(t)$ is the impedance function and $H(t)$ is a Hurwitz polynomial. Therefore,

$$Z(t) = \frac{H(t) - H(-t)}{H(t) + H(-t)} \left(\frac{\text{odd } H(t)}{\text{even } H(t)} \right). \qquad [21]$$

Knowing $H(t)$ from Eq. [2] for a given equalizer, $Z(t)$ is obtained in terms of poles and zeros. This impedance function can be expanded in the form of a cascaded commensurate transmission-line network. The impedances involved can be realized in a waveguide structure using impedance inverters (Fig. 9). Evaluation of susceptances in terms of poles and zeros for a three-section (CD) equalizer is presented in Appendix 1.

The computed values of susceptances are then converted to a normalized value defined by

$$B_N = B_o \frac{\lambda_{gN}}{\lambda_{go}}, \qquad [22]$$

where B_N = normalized value of susceptance at f_N,

B_o = computed value of susceptance required at f_o,

λ_{go} = guide wavelength at frequency f_o,

λ_{gN} = guide wavelength at frequency f_N

$f_N = 1.44\ f_c$,

f_c = cutoff frequency of waveguide

This normalized susceptance is then converted to the equivalent microwave reactance network consisting of an obstacle (flat metal slab) located in the center of the waveguide.

This conversion is based on the design curve plotted from the measured data[3], and can be used to determine the required obstacle dimensions. The length of the cavities is determined by the formula

$$l_i = \theta_i \frac{\lambda_{go}}{2\pi}, \qquad [23]$$

where θ_i is the electrical and l_i is the mechanical length of the ith cavity between the centers of the obstacles. All this is achieved by a sub-

routine of the main computer program. It is to be emphasized here that relationship and the interpolation of design curve (Fig. 4) between measured points is subject to same errors as indicated in Ref. (3).

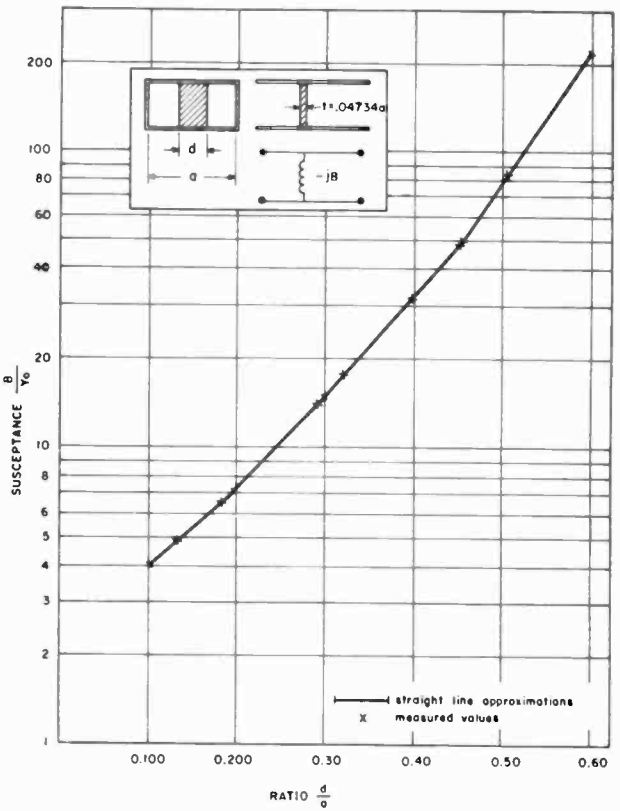


Fig. 4—Normalized susceptance of inductive irises in rectangular waveguide at $f = 1.44 f_c$.

5. Discussion of Theoretical and Experimental Results

A typical output of the computer is shown in Fig. 5. The input parameters are the absolute group delay that is to be equalized, the input gain slope, order of the equalizer, starting values of the poles and zeros for the equalizer, acceptable region if any for the overall system group delay, and the value of the average unloaded Q for the equalizer cavities. The output consists of the optimized values of poles and zeros, group delay/frequency characteristics of the optimum equalizer,

Fig. 5—A typical computer output depicting optimized parameters of an equalized filter with two sections of equalization.

overall system relative delay and gain slope, and the mechanical dimensions of the equivalent microwave reactance network. The overall system group delay at the center frequency can be fixed or kept floating. This provides an additional trade-off between the group-delay response and the overall-gain slope.

The amplitude characteristic of the all-pass microwave network has a shape similar to that of its group delay. Consequently, group delay equalization tends to equalize the amplitude response at the same time; the exact amount of amplitude equalization depends upon the average unloaded Q, Q_o of the equalizer cavities, and the absolute group delay of the all-pass network. The choice of Q_o for the equalizer directly affects the overall gain slope with negligible effect on the group delay. This feature can therefore be used with great advantage in optimizing the overall gain–slope. It is shown in Appendix 3 that if τ_F and τ_E represent the absolute group delays of the filter (or the input group delay) and the all-pass network respectively, then, at any frequency,

$$\alpha = 27.3 \left[\frac{\tau_F}{Q_F} + \frac{\tau_E}{Q_E} \right] f_o$$

$$\text{and} \quad \frac{d\alpha}{df} = 27.3 \left[\frac{1}{Q_F} \frac{d\tau_F}{df} + \frac{1}{Q_E} \frac{d\tau_E}{df} \right] f_o$$

α is the total insertion loss of the input network–equalizer combination; Q_F and Q_E represent the average unloaded Q's of the input network and the equalizer, respectively. It is assumed that Q_F and Q_E remain essentially constant with frequency over the band of interest. τ_E, the absolute group delay of the equalizer, is optimized with respect to the overall group-delay constraints and, consequently, only Q_E is available for gain–slope optimization. This provides one extra degree of freedom to affect the overall gain slope. Since the linear component of the amplitude response causes maximum contribution to intermodulation noise, choice of Q_E is most effective when chosen to minimize this component. The value of Q_E is determined by the design and material used in the physical realization of the equalizer.

The Theoretical and experimental results of C-section and D-section equalizers are depicted in Figs. 6a and 6b. Fig. 7 depicts two microwave reactance networks with different values of Q_E, and Fig. 8 shows a typical equalized filter. The reduced height design of equalized filters (Fig. 8) shows great promise for the design of light-weight multiplexers for communication satellites.

6. Conclusions

An approach towards optimizing the design of microwave equalizers so as to achieve the best performance available from an equalizer of a given order under a given set of design constraints has been presented. This in turn yields a minimum weight for a given performance of the equalizer—a feature that is essential in the design of multiplexing networks for communication satellites. The approach is general, and can be readily adopted to optimize all-pass networks in any frequency range.

The constraints that may be imposed on the design are outlined

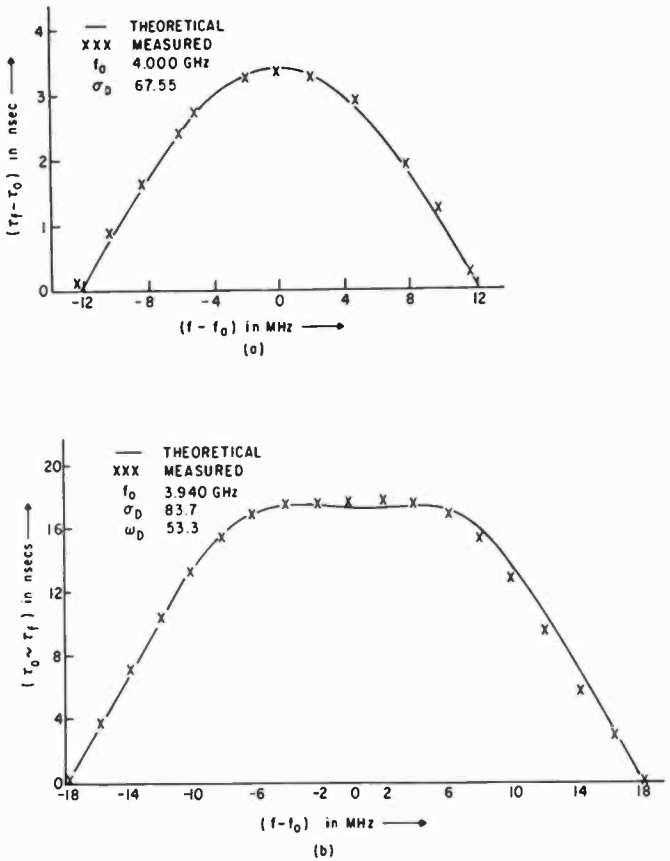


Fig. 6—Group delay response of rf equalizers: (a) computed and measured response of a C-section equalizer and (b) computed and measured response of a D-section equalizer.

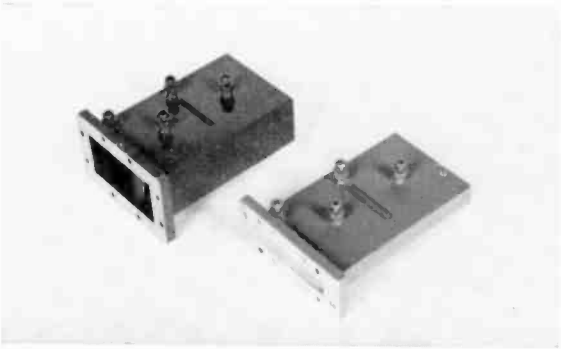


Fig. 7—Typical microwave reactance networks with symmetrical irises. The two networks have identical group delay response but different unloaded Q's.

in Section 3. Using these constraints, it is possible to determine the following trade-offs for all-pass networks in any frequency range:

(a) group-delay response versus the order or complexity of the equalizer;

(b) group delay response versus amplitude response (for an assumed value of average unloaded Q) for a given order of the equalizer;

(c) group delay response (flat) in one part of the desired band versus ripple in the remaining part.

The choice of Q_o (the average unloaded Q) provides an additional degree of freedom. The value of Q_o can be affected by the material and design of the equalizer.

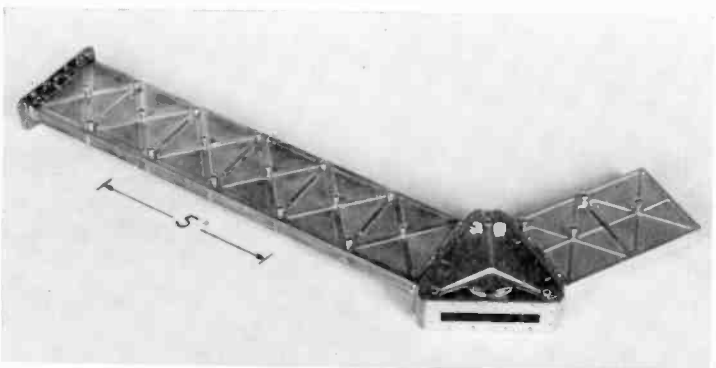


Fig. 8—A reduced height, spark-eroded equalized filter.

This feature can be very useful in practical applications for wideband communication systems. It is possible by choosing an appropriate value of Q_o to equalize the group delay and amplitude variation simultaneously. Further, the cutoff frequency of the waveguide can also be taken as a variable within certain limitations.

The design of the equalizer is obtained in terms of the optimized location of poles and zeros. This is independent of the type of configuration to be used in the realization of the microwave network. A reflection-type arrangement was chosen owing to the simplicity and ease with which a reactance network at microwave frequencies can be realized. It is beyond the scope of this paper to compare the reflection-type arrangement as against transmission-type equalizers. However, it is felt that for second- or higher-order equalization, the reflection-type arrangement yields less weight for the same performance.

Microwave equalizers are currently being incorporated into the design of multiplexing networks for communication satellites. It is felt that the next generation of terrestrial microwave links and communication satellites are likely to use such devices extensively.

Acknowledgments

The author is highly indebted to A. D. R. Walker for his invaluable assistance in the development of the computer program on RCA Spectra 70 Computer. Also, the author wishes to thank S. Singer for testing the equalizers and M. V. O'Donovan for his kind permission to utilize his design curve of normalized susceptance of waveguide obstacles. Thanks are also extended to RCA Ltd. for their generous support of this work and their kind permission to publish this paper.

Appendix 1

For a CD section all-pass structure, the Hurwitz polynomial is given by

$$H(t) = (t + \sigma_c)(t^2 + 2\sigma_d t + |t_d|^2),$$

where $t_d = \sigma_d + j\omega_d$

$$\mathfrak{I}_{IN} = \frac{Z_{IN}}{Z_o} = \frac{H(t) - H(-t)}{H(t) + H(-t)}$$

$$= t\frac{2\sigma_c\sigma_d + |t_d|^2 + t^2}{\sigma_c|t_d|^2 + (\sigma_c + 2\sigma_d)t^2}$$

$$= \frac{at + bt^3}{1 + ct^2},$$

$$a = \frac{(2\sigma_c\sigma_d + |t_d|^2)}{\sigma_c|td|^2};$$

where $$b = \frac{1}{\sigma_c|t_d|^2}, \qquad [24]$$

$$c = \frac{(\sigma_c + 2\sigma_d)}{\sigma_c|t_d|^2},$$

3_{IN} can be expressed as

$$3_{IN} = \frac{aY^2 + b}{Yc}\left(1 + \frac{Y^2}{c}\right)^{-1} \text{ where } Y = \frac{1}{t}.$$

Expanding 3_{IN} in powers of Y we get

$$3_{IN} = \frac{b}{c}Y^{-1} + \frac{1}{c}\left(a - \frac{b}{c}\right)Y + \frac{1}{c}\left(\frac{b}{c^2} - \frac{a}{c}\right)Y^3 + \frac{1}{c}\left(\frac{a}{c^2} - \frac{b}{c^3}\right)Y^5 + \cdots \qquad [25]$$

Next we consider the input impedance of a three-section reactance network consisting of the three impedance inverters K_1, K_2, and K_3 as shown in Fig. 9

$$Z_1 = jZ_0 \tan 2\theta$$

$$Z_2 = K_1^2/jZ_0 \tan 2\theta$$

$$Z_3 = Z_0\frac{Z_2 + jZ_0 \tan 2\theta}{Z_0 + jZ_2 \tan 2\theta}$$

$$Z_4 = K_2^2\frac{Z_0 + jZ_2 \tan 2\theta}{Z_0(Z_2 + jZ_0 \tan 2\theta)}$$

$$3_4 = \frac{Z_4}{Z_0} = a_1\frac{t + t^3}{1 + b_2t^2 + t^4}$$

$$t = j \tan \theta \qquad [26]$$

where $a_1 = 2K_2^2(Z_0^2 + K_1^2)/(Z_0K_1^2)$,

$$b_2 = \frac{2}{K_1^2}(2Z_0^2 + K_1^2),$$

Z_0 = characteristic impedance,

$$Z_5 = Z_0 \frac{Z_4 + jZ_0 \tan 2\theta}{Z_0 + jZ_4 \tan 2\theta},$$

$$Z_{IN} = K_3^2 \frac{Z_0 + jZ_4 \tan 2\theta}{Z_4 + jZ_0 \tan 2\theta},$$

$$\mathfrak{Z}_{IN} = \frac{Z_{IN}}{Z_0} = \frac{K_3^2}{Z_0^2} \frac{1 + 2t\mathfrak{Z}_4 + t^2}{\mathfrak{Z}_4 + 2t + \mathfrak{Z}_4 t^2}.$$

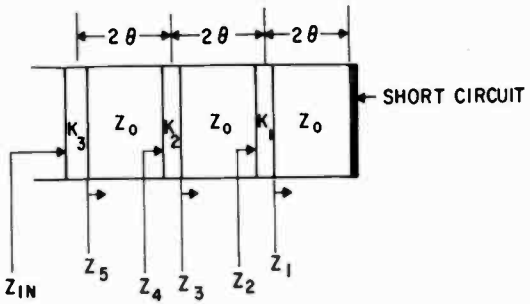


Fig. 9—A three section reactance network with impedance inverters.

Substituting for $\mathfrak{Z}_4$, we get

$$\mathfrak{Z}_{IN} = \frac{K_3^2}{Z_0^2} \frac{1 + (1 + 2a_1 + b_2)\ (t^2 + t^4) + t^6}{(a_1 + 2)\ t + (2a_1 + 2b_2)\ t^3 + (2 + a_1)\ t^5}. \qquad [27]$$

$$= \frac{1 + dt^2 + dt^4 + t^6}{et + ft^3 + et^5}$$

where $d = 1 + 2a_1 + b_2$,

$e = (2 + a_1)\ Z_0^2/K_3^2$, [28]

$f = 2\ (a_1 + b_2) Z_0^2/K_3^2$,

Letting $t = 1/Y$ and expanding 3_{IN} we get

$$3_{IN} = \frac{1}{e} Y^{-1} - \frac{f}{e^2} Y + \left(\frac{f^2}{e^3} - \frac{1}{e}\right) Y^3 + \left(\frac{2f}{e^2} - \frac{f^3}{e^4}\right) Y^5 + \cdots$$

$$+ \frac{d}{e} Y - \frac{df}{e^2} Y^3 + \frac{d}{e}\left(\frac{f^2}{e^2} - 1\right) Y^5 + \cdots$$

$$+ \frac{d}{e} Y^3 + \frac{d}{e}\left(-\frac{f}{e}\right) Y^5 + \cdots$$

$$+ \frac{1}{e} Y^5 + \cdots$$

$$\text{or } 3_{in} = \frac{1}{e} Y^{-1} + \left(\frac{d}{e} - \frac{f}{e^2}\right) Y + \left(\frac{f^2}{e^3} - \frac{1}{e} - \frac{df}{e^2} + \frac{d}{e}\right) Y^3 + \cdots \qquad [29]$$

Equating the coefficient of Y^{-1}, Y, and Y^3, we get

$$\frac{b}{c} = \frac{1}{e}$$

$$\frac{a}{c} - \frac{b}{c^2} = \frac{d}{e} - \frac{f}{e^2}$$

$$\frac{b}{c^3} - \frac{a}{c^2} = \frac{f^2}{e^3} - \frac{1}{e} - \frac{df}{e^2} + \frac{d}{e}.$$

Determining d, e and f in terms of a, b and c, we have

$$e = b/c$$

$$f = \frac{-a\,(c + c^2) + b\,(1 + c + c_2)}{-abc + b^2\,(1 + c)} \qquad [30]$$

$$d = \frac{abc - a^2c^2 - b^2 + b^2c^2 + b^2c}{b^2c^2 - abc^2 + b^2c}$$

a, b and c are related to the poles and zeros of the CD equalizer by Eqs. [24]. Therefore, d, e, and f are determined in terms of poles and zeros.

Next we relate d, e, and f to the susceptance parameters

$$\begin{aligned} d &= 1 + b_2 + 2a_1 \\ &= 1 + \frac{2}{K_1^2}(2Z_0^2 + K_1^2) + 2.2\frac{K_2^2}{K_1^2}\left(1 + \frac{K_1^2}{Z_0^2}\right), \\ e &= \left[2 + 2\frac{K_2^2}{K_1^2}\left(1 + \frac{K_1^2}{Z_0^2}\right)\right]\frac{Z_0^2}{K_3^2}, \\ f &= 4\frac{Z_0^2}{K_3^2}\left[\frac{K_2^2}{K_1^2}\left(1 + \frac{K_1^2}{Z_0^2}\right) + \frac{1}{K_1^2}(2Z_0^2 + K_1^2)\right]. \end{aligned} \qquad [31]$$

Solving for K_1/Z_0, K_2/Z_0, and K_3/Z_0,

$$\begin{aligned} \frac{K_3^2}{Z_0^2} &= \frac{2(d+1)}{f + 2e}, \\ \frac{K_2^2}{Z_0^2} &= \frac{4(eK_3^2/Z_0^2 - 2)}{(K_3^2/Z_0^2)(f - 2e) + 8}, \\ \frac{K_1^2}{Z_0^2} &= \frac{8}{(K_3^2/Z_0^2)(f - 2e)}. \end{aligned} \qquad [32]$$

Following the method of Cohn[8] the normalized susceptance of the obstacles would be given by

$$\frac{X_i}{Z_0} = \frac{K_i/Z_0}{1 - (K_i/Z_0)^2}, \quad i = 1, 2, 3 \qquad [33]$$

and the electrical lengths of the cavities would be

$$\theta = \pi - \frac{1}{2}\left[\tan^{-1}\frac{2X_3}{Z_0} + \tan^{-1}\frac{2X_2}{Z_0}\right], \qquad [34]$$

$$\theta = \pi - \frac{1}{2}\left[\tan^{-1}\frac{2X_2}{Z_0} + \tan^{-1}\frac{2X_1}{Z_0}\right],$$

$$\theta = \pi - \frac{1}{2}\tan^{-1}\frac{2X_1}{Z_0}.$$

Appendix 2

If waveguide is used in the physical realization of the all-pass networks, the complex frequency variable t will be given by

$$t = j\tan\frac{2\pi}{\lambda_g}l,$$

where λ_g is the guide wavelength and l is a commensurate line length of the microwave network. The guide wavelength is given by

$$\lambda_g = \frac{\lambda}{\sqrt{1-\left(\frac{\lambda}{\lambda_{co}}\right)^2}}$$

where λ is the free-space wavelength and λ_{co} is the cutoff wavelength of the waveguide. Assuming l to be a quarter of a wavelength at the center frequency,

$$t = j\tan\frac{2\pi}{\lambda}\sqrt{1-\left(\frac{\lambda}{\lambda_{co}}\right)^2}\ \ \frac{\lambda_{go}}{4}.$$

λ_{go} is the guide wavelength at the center frequency f_o, given by

$$\lambda_{go} = \frac{\lambda_o}{\sqrt{1-\left(\frac{\lambda_o}{\lambda_{co}}\right)^2}}$$

λ_o is the free-space wavelength at the center frequency f_o.

Therefore

$$t = j \tan \frac{\pi}{2} \frac{\lambda_o}{\lambda} \frac{\sqrt{1 - \left(\frac{\lambda}{\lambda_{co}}\right)}}{\sqrt{1 - \left(\frac{\lambda_o}{\lambda_{co}}\right)^2}}.$$

Substituting $F = \frac{f}{f_o} = \frac{\lambda_o}{\lambda}$,

and $F_o = \frac{f_{co}}{f_o} = \frac{\lambda_o}{\lambda_{co}}$,

$$t = j \tan \frac{\pi}{2} F \frac{\sqrt{1 - \left(\frac{F_o}{F}\right)^2}}{\sqrt{1 - F_o^2}},$$

$$\text{or } t = j \tan \frac{\pi}{2} \sqrt{\frac{F^2 - F_o^2}{1 - F_o^2}},$$

$$\text{and } \frac{dt}{dF} = (1 - t^2)\, j \frac{\pi}{2} \frac{1}{2} \frac{(F^2 - F_o^2)^{-1/2}}{(1 - F_o^2)^{1/2}} 2F,$$

$$\text{or } \frac{dt}{dF} = j \frac{\pi}{2} (1 - t^2) \frac{F}{[(F^2 - F_o^2)(1 - F_o^2)]^{1/2}}.$$

For TEM-mode transmission,

$$f_{co} = 0$$

$$t = j \frac{\pi}{2} F = j \frac{\pi}{2} \frac{f}{f_o},$$

and $$\frac{dt}{dF} = j \frac{\pi}{2} (1 - t^2).$$

Appendix 3

The insertion loss α_o at the center frequency f_o for a filter using Cohn's formula is given by

$$\alpha_o = 4.343 \frac{\omega_o}{\Delta\omega} \frac{\sum_{k=1}^{n} g_k}{Q_{uk}},$$

where ω_o is the center frequency in radians, $\Delta\omega$ is the equal-ripple bandwidth, Σg_k is the summation of low-pass prototype parameters, and Q_{uk} is the unloaded Q of the kth cavity. Assuming the same unloaded Q for all cavities,

$$\alpha_o = 4.343 \frac{2\pi f_o}{\Delta\omega} \frac{1}{Q_F} \sum_{k=1}^{n} g_k .$$

The group delay τ_{OF} at the center frequency is given by

$$\tau_{OF} = \frac{1}{\Delta\omega} \sum_{k=1}^{n} g_k .$$

Therefore

$$a_o = 2\pi \times 4.343 \frac{f_o}{Q_F} \tau_{OF}$$

$$= \frac{27.3}{Q_F} f_o \tau_{OF}.$$

In Ref. (10) it is shown that this relationship can be extended to the entire pass-band bandwidth of the filter, i.e.,

$$\alpha_F = \frac{27.3}{Q_F} f_o \tau_F,$$

where α_F and τ_F represent the insertion loss and absolute group delay, respectively, at frequency f in the pass band.

Following similar reasoning, it can be shown that similar relations exist between the insertion loss, and absolute group delays for the all-pass microwave networks in the band of interest, i.e.,

$$\alpha_E = \frac{27.3}{Q_E} f_o \tau_E,$$

and $$\frac{\alpha_E}{\alpha_o} = \frac{\tau_E}{\tau_o}.$$

The subscript E refers to the corresponding quantities for the equalizer.

The total insertion loss, α, through the filter–equalizer combination will therefore be

$$\alpha = \alpha_F + \alpha_E$$

$$= 27.3 \left[\frac{\tau_F}{Q_F} + \frac{\tau_E}{Q_E} \right] f_o,$$

and $$\frac{d\alpha}{df} = 27.3 \left[\frac{1}{Q_F} \frac{d\tau_F}{df} + \frac{1}{Q_E} \frac{d\tau_E}{df} \right] f_o.$$

As an example, if it is desired to have the same insertion loss at the band edge as at the center frequency, then

$$\frac{(\tau_F)_{\text{Edge}} - \tau_{OF}}{Q_F} = \frac{\tau_{OE} - (\tau_E)_{\text{Edge}}}{Q_E}.$$

Knowing the group delays and the insertion loss or unloaded Q of the filter, Q_E—the unloaded Q of the equalizer—can be determined to satisfy the above relationship.

References

[1] E. G. Cristal, "Theory and Design of Transmission Line All-pass Equalizers," **IEEE Trans. Microwave Theory and Techniques,** Vol. MTT-17, p. 28, Jan. 1969.

[2] J. O. Scanlan and J. D. Rhodes, "Microwave All-pass Networks," **IEEE Trans. Microwave Theory and Techniques,** Vol. MTT-16, p. 62, Feb. 1968.

[3] P. I. Richard, "Resistor Transmission-Line Circuits," **Proc. IRE,** Vol. 36, p. 217, Feb. 1948.

[4] H. W. Bode, **Network Analysis and Feedback Amplifier Design,** pp. 239-242, D. Van Nostrand Co., Inc., New York, 1945.

[5] N. K. M. Chitre and M. V. O'Donovan, "Computer Aided Design of Waveguide Filters," RCA Engineer, Vol. 12, No. 1, June, 1966.

[6] 'DFMFP—Unconstrained Minimum of a function of several variables,' System/360, Scientific Sub-routine package (360-A-CM-03 X) Version II, IBM 1966.

[7] R. Fletcher and M. J. D. Powell, "A Rapidly Convergent Descent Method for Minimization," **Computer Jour.,** Vol. 6, No. 2, p. 163-168, 1963.

[8] E. A. Guillemin, **Synthesis of Passive Networks,** New York, Wiley, p. 56, 1957.

[9] S. B. Cohn, "Direct-Coupled Resonator Filters," **Proc. IRE,** Vol. 45, p. 187. Feb. 1957.

[10] L. Young, "Group Delay and Dissipation Loss in Transmission-Line Filters," **IEEE Trans. Microwave Theory and Techniques,** Vol. MTT-11, p. 215, May 1963.

RCA Technical Papers Second Quarter, 1970

Apri

"Comparison of Photomultipliers and Avalanche Photodiodes for Laser Applications," R. J. McIntyre, **IEEE Trans. GED** (April)
"Control of Optical Losses in p-n Junction Lasers by Use of a Heterojunction: Theory and Experiment," H. Kressel, H. Nelson, and F. Z. Hawrylo, **Jour. Appl. Phys.** (April)
"Electrical and Optical Properties of Vapor-Grown GaAs:Si," H. Kressel and H. von Phillipsborn, **Jour. Appl. Phys.** (Communications) (April)
"Epitaxially Deposited Si from a SiH_4 HCl Mixture," A. C. Ipri, **Solid State Tech.** (Tech Brief) (April)
"Field Distribution In a Magnetoplasma-Loaded Waveguide at Room Temperature," R. Hirota and K. Suzuki, **IEEE Trans. GMTT** (April)
"A Latching Ring-and-Post Ferrite Waveguide Circulator," W. W. Siekanowicz, R. W. Paglione, and T. E. Walsh, **IEEE Trans. GMTT** (April)
"Photoemission of Holes and Electrons from Aluminum into Aluminum Oxide," A. M. Goodman, **Jour. Appl. Phys.** (April)
"A Rapid Evaluation Technique for Functional Gunn Diodes," R. D. Larrabee, W. A. Hicinbothem, Jr., and M. C. Steele, **IEEE Trans. GED** (April)
"A Single-Vidicon Television Camera System," L. Briel, **Jour. SMPTE** (April)
"Threshold Logic Asymptotes," R. O. Winder, **IEEE Trans. GC** (Short Notes) (April)
"Tubes and/or Solid-State Devices for Power Generation," L. S. Nergaard, **Microwave Jour.** (April)
"Vacuum-Deposited Thin Film p-Se/n-CdSe Heterojunction Diodes," R. M. Moore and C. J. Busanovich, **IEEE Trans. GED** (April)
"Electroabsorption in Semiconductors: The Excitonic Absorption Edge," J. D. Dow and D. Redfield, **Phys. Rev.** B (15 April)

May

"Complementary MOS Logic and Applications," S. S. Eaton, **Electronic Engineer** (MOS Course, Pt. 4) (May)
"The C-V Technique as an Analytical Tool—Part I," K. H. Zaininger and F. P. Heiman, **Solid State Tech.** (May)
"Design Shortcuts for Microwave Frequency Dividers," W. J. Goldwasser, **Electronic Engineer** (May)
"A Dual Scan Acquisition Technique for a Laser Communication System," A. Waksberg, **IEEE Trans. GAES** (Correspondence) (May)
"A Dynamic Space Division Multiplex Mobile Radio System," H. Staras and L. Schiff, **IEEE Trans. GVT** (May)
"Electronic Composition System," J. S. Greenberg and R. E. Schubert, **IEEE Trans. GAES** (May)
"The Lumped Element Approach to Microwave Integrated Circuits," M. Caulton, **Microwave Jour.** (May)
"Mechanism of Thermal Annihilation of Stacking Faults in GaAs," M. S. Abrahams and C. J. Buiocchi, **Jour. Appl. Phys.** (May)
"Open Loop TV Image Stabilization," J. Dubbury, **IEEE Trans. GAES** (May)
"Simplified Construction and Processing of a Helium-Cadmium Laser," K. G. Hernqvist and D. C. Pultorak, **Rev. Sci. Instr.** (May)
"Triflop—A Three-State Memory Element," A. Turecki, **Electronic Engineer** (IC Ideas) (May)
"Phonon Interference in Thin Films of Liquid Helium," C. H. Anderson and E. S. Sabisky, **Phys. Rev. Letters** (May 11)
"Remote Control for Color TV Goes the All-Electronic Route," W. Evans, C. Moeller, and E. Milbourn, **Electronics** (May 25)

June

"Back Scattering of Material Emitted from RF-Sputtering Targets," J. L. Vossen, J. J. O'Neill, Jr., K. M. Finlayson, and L. J. Royer, **RCA Review** (June)

"Cleaning Solutions Based on Hydrogen Peroxide for Use In Silicon Semiconductor Technology," W. Kern and D. Puotinen, **RCA Review** (June)
"Conversion and Imports," E. W. Herold, **IEEE Spectrum** (Forum) (June)
"Coplaner-Waveguide Directional Couplers," C. P. Wen, **IEEE Trans. GMTT** (June)
"The C-V Technique as an Analytical Tool—Part II," K. H. Zaininger and F. P. Heiman, **Solid State Tech.** (June)
"Detection of Damage on Silicon Surfaces: Origin and Propagation of Defects," A. Mayer, **RCA Review** (June)
"The Epitaxial Growth of Silicon on Sapphire and Spinel Substrates: Suppression of Changes in the Film Properties During Device Processing," G. W. Cullen, G. E. Gottlieb, and C. C. Wang, **RCA Review** (June)
"The Etching of Deep Vertical-Walled Patterns in Silicon," A. I. Stoller, **RCA Review** (June)
"Evaporation of Aluminum with RF-Induced Substrate Bias," J. L. Vossen and J. J. O'Neill, Jr., **RCA Review** (June)
"Fabrication of Al_2O_3 COS/MOS Integrated Circuits," F. B. Micheletti, P. E. Norris, and K. H. Zainlnger, **RCA Review** (June)
"High-Power K-Band Silicon Avalanche-Diode Oscillators," S. G. Liu, J. J. Risko, and K. K. N. Chang, **Proc. IEEE** (Letters) (June)
"An Inexpensive Integrating Photoresist Exposure Control System," G. A. Riley, **RCA Review** (June)
"Low-Threshold $Al_x Ga_{1-x}$ As Visible and IR-Light-Emitting Diode Lasers," H. Kressel, H. L. Lockwood, and H. Nelson, **IEEE Jour. Quantum Electronics** (June)
"Monitoring Silicon Tetrachloride Concentration in Hydrogen Carrier Gas," G. A. Riley and J. A. Amick, **RCA Review** (June)
"New Solid State AM Ampliphase Transmitters," J. L. Preston, **Broadcast News** (June)
"A New Technique for Etch Thinning Silicon Wafers," A. I. Stoller, R. R. Speers, and S. Opresko, **RCA Review** (June)
"New Techniques for Television Tape Programming," A. C. Luther, Jr., **IEEE Trans. GBC** (June)
"A Novel Technique for Forming Glass-to-Metal Seals Using a Silicon Nitride Interface Layer," A. I. Stoller, W. C. Schilp, Jr. and J. Benbenek, **RCA Review** (June)
"Optical Losses in 'Close-Confinement' Epitaxial p-n Junction Lasers—Theory and Experiment," H. Kressel, H. Nelson, and F. Z. Hawrylo, **IEEE Jour. Quantum Electronics** (June)
"Optimization of Charge Storage in the MNOS Memory Device," A. M. Goodman, E. C. Ross, and M. T. Duffy, **RCA Review** (June)
"The Performance of Complementary MOS Transistors on Insulating Substrates," E. J. Boleky, **RCA Review** (June)
"Radiochemical Study of Surface Contamination. I. Adsorption of Reagent Components," W. Kern, **RCA Review** (June)
"Radiochemical Study of Surface Contamination. II. Deposition of Trace Impurities on Silicon and Silica," W. Kern, **RCA Review** (June)
"Recent Developments in GaP(Cs)-Dynode Photomultipliers," H. R. Krall, F. A. Helvy, and D. E. Persyk, **IEEE Trans. GNS** (June)
"Room-Temperature Close-Confinement GaAs Laser with Overall External Quantum Efficiency of 40 Percent," R. G. Gill, **Proc. IEEE** (Letters) (June)
"Selective Electroless Plating by Selective Deactivation," N. Feldstein and T. S. Lancsek, **RCA Review** (June)
"A Sonic Pen: A Digital Stylus System," P. de Bruyne and Coauthor, **IEEE Trans. GC** (Short Notes) (June)
"Spectral Lines: Convention Progress," E. W. Herold, **IEEE Spectrum** (June)
"A Study of Dielectric Defect Detection by Decoration with Copper," W. J. Shannon, **RCA Review** (June)
"Two Room-Temperature Electroless Nickel Plating Baths," N. Feldstein, **RCA Review** (June)
"Vapor Deposited Tungsten as a Metallization and Interconnection Material for Silicon Devices," J. M. Shaw and J. A. Amick, **RCA Review** (June)
"Video Edging," R. J. Butler, **Jour. SMPTE** (June)
"Reassessment of a Determination of e/h, Using Macroscopic Quantum Phase Coherence in Superconductors," B. N. Taylor and Coauthors, **Phys. Rev. B** (1 June)
"Eu^{+3} ^5D Resonance Quenching to the Charge-Transfer States in Y_2O_2S, La_2O_2S, and LaOCl," W. H. Fonger and C. W. Struck, **Jour. Chem. Phys.** (15 June)
"Volume Generation and Hall Mobility of Holes in Anthracene," J. Dresner, **Jour. Chem. Phys.** (15 June)

Patents Issued to RCA Inventors Second Quarter 1970

April

Dallas Roy Andrews, Tape Reeling Search System with Transistor Search Amplifier (3,505,485)

Frederick Orie Bartholomew, Switching Circuit (3,509,362)

Joseph Michael Bielen and Gilbert Victor Morris, Method of Dicing Semiconductor Wafers (3,507,426)

Robert Daniel Browning, Magnetic Tape Reel (3,508,719)

Philip Joseph Donald, Optical Data Selection and Display (3,504,609)

Raymond Louis Giordano, Signal Envelope Discriminator and Gating Circuit (3,505,537)

William Adolph Gottfried, Double Photoresist Processing (3,506,441)

Steven Robert Hofstein, Solid State Clock (3,505,804)

Charles Samuel Jackson, Method of Etching Semiconductive Devices Having Lead-Containing Elements (3,505,132)

Marvin Allan Leedom, Combination of a Container for a Liquid and Means for Dispensing the Liquid (3,507,252)

Robert Levin, Test Tape with Preselected Skew (3,508,231)

Abraham Lichowsky, Centrifugal Force Controlled Transducer (3,505,744)

Joseph Lawrence McLaughlin, Method of Fabricating Photomasks (3,507,592)

Thomas Charles McNulty, Thyristor Controlled Voltage Regulating Circuit (3,509,450)

Adolph Karl Rapp, Multivibrators Employing Transistors of Opposite Conductivity Types (3,509,379)

Robert Fincher Sanford, SYNC Slipper (3,507,986)

Alfredo Soon-Nam Sheng and Emile Hebert, Electronic Timer (3,505,541)

George Van Taylor, Tape Basket (3,508,696)

Anatole Turecki, Counter Circuits (3,508,033)

Sidney Wald and Leroy Henry Werner, Code Converter (3,505,667)

Chandler Wentworth, Method of Making a Laminated Ferrite Memory (3,505,139)

Joseph Wilder and Walter Joseph Gorkiewicz, High Efficiency Light Polarization System (3,508,809)

Robert Owen Winder, Binary Arithmetic Circuits Employing Threshold Gates In which Both the Sum and Carry are Obtained in One Gate Delay Interval (3,506,817)

Robert Owen Winder, Networks of Elements for Implementing Threshold Functions (3,506,845)

Robert Owen Winder, Logic Circuitry (3,508,076)

May

Jack Avins, Transistor Electrical Circuit with Collector Voltage Stabilization (3,512,098)

Lucas John Bazin, Television Camera Power Supply (3,510,378)

Bernard Bruce Bossard and Sui Yuan, Balanced Mixer Circuits (3,513,398)

David John Carlson, Ultra-High Frequency Transistor Oscillator (3,510,802)

David John Carlson, Field-Effect Transistor Amplifier (3,513,405)

Edgar Thomas Casterline and Ronald Rosenzweig, Method of Fabricating Semiconductor Devices (3,513,022)

Kern Ko Nan Chang, Integrated Semiconductor Structure with Frequency Selective Transmission Line Sections (3,513,403)

Michael Cooperman, Current Steering Networks Providing the Exclusive OR Function

Joseph Richard Burns, Signal Translating Circuit Comprising a Plurality of IGFET Ampli-

Edward Wesley Curtis, Protection Circuits for Kinescopes (3,510,722)

James Francis Delany, Photocomposing Apparatus (3,509,803)

Andrew Gordon Francis Dingwall, Integrated Circuit Planar Transistor (3,510,736)

Daniel Joseph Donahue and Austin Eugene Hardy, Photographic Method for Making Tri-Colored Cathode Ray Screen (3,514,287)

Robert Stover Fow, Magnetic Core Memory Plane Assembly and Method (3,513,453)

Hendrik Jurgen Gerritsen and David Leslie Greenaway, Motionless Hologram Imaging (3,511,553)

Walter Joseph Gorklewicz and Joseph Wilder, Apparatus for Forming a Beam of Light (3,512,868)

Charles Glen Hawkins, Tape Basket Stripper Apparatus (3,514,024)

Frederic Paul Heiman, Field Effect Transistor with Improved Insulative Layer Between Gate and Channel (3,513,364)

Douglas Alan Johnson, Robert Joseph Tomsic and Hunter Crary Goodrich, Linear Flux Control Circuit (3,510,675)

Furman Donald Kell, Drum Construction for Helical Scan Tape Recorder (3,510,604)

Jacob Klapper, Varying Bandwidth Frequency Shift Keying Receiver (3,510,779)

Heshmat Khajezadeh and Lewis Alfred Jacobus, Jr., High Voltage Transient Protection for an Insulated Gate Field Effect Transistor (3,512,058)

Richard Joseph Klensch, Deflection Corrector Circuit for Cathode Ray Tube (3,512,039)

Walter Frank Kosonocky, Multi-Layer, Light-Emitting Semiconductor Device (3,514,715)

George Andrew Kupsky, Encapsulated Optical Semiconductor Device (3,512,027)

Charles Brook Leuthauser, RF Power Amplifier (3,513,406)

Charles Joseph Mangiaracina and Charles Boyd Meyer, Optical Multiplexing Apparatus (3,510,657)

James Courtland Marsh, Jr., Transistorized Automatic-Gain-Controlled Amplifier (3,510,579)

Austin Joseph Mortimer, Load Sensing and Compensating Control Circuits (3,510,743)

Lawrence Paul Nahay and Bernard Eugene Patrusky, Digital Demodulator System (3,514,702)

Yoshihiro Okuno, Gain Controlled Transistor Amplifier with Constant Bandwidth Operation over the AGC Control Range (3,510,580)

Roland Norman Rhodes and Albert Macovski, Video Circuits for Color Television Receivers (3,510,573)

William Frederick Schacht and William Francis Fisher, Apparatus for Previewing Slides (3,514,198)

Henry Claude Schindler and James Joseph Drautman, Protective Apparatus for a Superconductive Switch (3,513,421)

Fred Sterzer, Impedance Control Using Transferred Electron Diodes (3,510,805)

Seymour Heywood Winkler and Rudolph Reinhold Laessig, Thermoelectric Generator Suitable for Use at Elevated Temperatures in a Vacuum (3,510,363)

Daniel Aaron Wisner, Speed Control for Automotive Vehicles (3,511,329)

June

Thomas Charles Abrahamsen, Multiple Font Keyboard (3,517,792)

Julius Agin, Dead Track Handling (3,518,625)

Robert Roosevelt Brooks, Preset Sensitivity and Amplification Control System (3,518,371)

Joseph Richard Burns, Signal Translating Circuit Comprising a Plurality of Igfet Amplifiers Cascaded in Direct Coupled Fashion (3,516,004)

David Harry Carpenter, Color Killer and Automatic Chroma Control Circuits (3,517,114)

Francis David Cassidy, Magnetic Memory with Noise-Cancellation Sense Wiring (3,518,640)

John Edwin Croy, Weighted Time Accounting in Time Shared Computer (3,518,633)

Wolfgang Friedrich Wilhelm Dietz, Voltage Regulator (3,517,252)

Walter Gold Gibson and Arthur John Gravel, Portable Color Television Camera System (3,518,360)

Josef Gross and William Henry Barkow, Deflection Yoke and Apparatus for Its Fabrication Utilizing a Magnetic Ramming Technique (3,518,590)

Donald Charles Herrmann and John Lucas Bazin, Push-Pull Current Source (3,518,457)

George Victor Jacoby, Readback Circuit for Information Storage Systems (3,516,066)

Siguard William Johnson and Marvin Allan Leedom, Screening Unit for Half-Tone Color Reproduction (3,517,596)

Richard Joseph Kiensch, Device for Producing Line Halftone Images Similar to the Images Produced by the Woodcut Type Method of Printing (3,517,119)

Hon Chiu Lee and Gregory Hodowanec, Microwave Circuit with Coaxial Package Semiconductor Device (3,517,272)

Marvin Allan Leedom, Electrophotographic Apparatus (3,517,994)

John Evert Meyer, Jr., Coupling System for Elemental Panel Array (3,518,627)

John Joseph O'Toole, Modulation Error Cancelling Apparatus (3,518,370)

Robert Carlton Peyton, Arrangement for Handling Printable Character Bit Codes (3,517,165)

Donald Barry Schilling and Charlotte Dobin, Adaptive Diode Having Mobile Doping Impurities (3,515,953)

Donald Henry Willis, Automatic Chroma Control Circuit (3,517,115)

Paul Carleton Wilmarth, Television Receiver Horizontal Deflection Output Stage Protection Circuit and Direct Voltage Supply (3,518,482)

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 Engineering (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.

I. P. Csorba graduated from the Electrical Engineering Fundamentals at the Electrical Engineering Faculty of the Technical University of Budapest in 1952. In the same year he was admitted to the Communication Engineering Faculty, specializing in communication and electronics. In October, 1954, he received the Diplom Ingenieur Degree in Electrical and Communication Engineering. From 1955 to 1956 he worked as a research engineer at the War Technical Institute, Budapest. In December of 1956 he joined the research group of Rauland Corporation, Chicago, where he worked on electrostatic-type image converter tubes, television picture tubes, scan-converter tubes. From 1959 to 1961 he was with Motorola, Inc., Chicago, working primarily on electrostatic-type scan magnification. In November of 1961, he joined RCA as a member of the Photo and Image Tube Engineering Activity. Mr. Csorba has been active in the design and development of magnetic and electrostatic-type image tubes and photomultiplier tubes.

Michael T. Duffy received the B.Sc. degree from University College, Dublin, in 1958 and completed his Ph.D. studies in Solid State Chemistry in 1963. He was a teaching assistant at this University from 1961 to 1963. In 1964, he joined the University of Toronto as a Research Associate, engaged in molecular beam studies. In July 1966, he joined the staff of RCA Laboratories, Princeton, N. J. Since then, he has been engaged in materials research in the field of integrated electronics and particularly in the area of thin film dielectrics and MIS structures.

Alvin Malcolm Goodman was born in 1930 in Philadelphia, Pa. He attended the Drexel Institute of Technology, Philadelphia, Pa., and received a B.S.E. in 1952. He did graduate work at Princeton University, receiving an M.A. in 1955 and a Ph.D. in 1958. From June 1956 to January 1957 he served as Research Assistant at Princeton University, and then as Assistant Professor of Electrical Engineering at the Case Institute of Technology up to June 1959. He performed research at RCA Laboratories as a summer employee in 1954, 1955, and 1958. He has been a Member of the Technical Staff since June 1959. Dr. Goodman has specialized in solid-state physics. His thesis subject was "Dember Effect and Trap Levels in Silver Chloride". He is a member of the American Physical Society, Sigmi Xi and the Institute of Electrical and Electronics Engineers, and is listed in American Men of Science.

John J. Hughes attended the 7th Army Electronics School in Germany and is presently attending Middlesex County College. He joined ITT Federal Laboratories in 1950, where he worked in the Chemical Department, the Countermeasures Department and the Microwave Tube Department. From 1958 to 1961 he was assigned to Associated Testing Laboratories as a representative of ITT, where he monitored and evaluated the environmental testing of microwave tubes. In 1962 he joined the RCA Electronic Components, Harrison, New Jersey, where he was a production foreman in the Microwave Tube Division. In 1963 he joined the Microwave Research Laboratory at RCA Laboratories, Princeton, New Jersey, where he has worked on plasma-tube studies, klystron design, and integrated solid-state microwave techniques. He has worked on the "Blue Chip" program since its inception and participated in the work on microstrip properties and the development of the integrated receiver.

Mr. Hughes is a member of the Institute of Electrical and Electronics Engineers.

Richard J. Klensch received the BSEE degree from the University of Illinois in 1951. He did graduate work as an incidental student in the Electrical Engineering Department of Princeton from 1952 to 1958. Mr. Klensch joined RCA Laboratories as a research engineer in 1952, associated with high resolution radar and infrared detection devices. From 1954 to 1956 he served in the army as a radar instructor; returning to the RCA Laboratories in 1956. Subsequently, he did research in the areas of microwave scanning anttennas, time division multiplex systems, digital and analog communications, and color television. In 1961, he participated in research leading to an advanced Naval Communications System, continuing this effort until joining Graphic Systems Division in 1966. As part of the Advanced Development Group at GSD, Mr. Klensch has been investigating new electronic halftone generation techniques and CRT display Systems. He is a senior member of the IEEE.

Chandra M. Kudsia received his B. Sc. in physics from Delhi University in 1961, B.E. in Electrical Communication Engineering from Indian Institute of Science, Bangalore in 1964 and M. Eng. in Electrical Engineering from McMaster University, Hamilton, Ontario in 1966. After a year with Amphenol Canada Limited, Toronto, he joined RCA Limited, Montreal, in May 1967. At RCA he has worked on the design of microwave components and sub-systems for high-quality microwave links and ground stations for satellite communications. He was one of the team responsible for the Canadian Domestic Communication Satellite project during its project definition phase. He is presently working on the Communication Technology Satellite which is a joint undertaking by the Government of Canada and NASA.

Mr. Kudsia is a registered Professional Engineer of the Province of Ontario and a member of IEEE.

Dietrich Meyerhofer received his Bachelor of Engineering Physics from Cornell University in 1954 and his Ph.D. degree in physics from Massachusetts Institute of Technology in 1958. Since 1958 he has been a Member of the Technical Staff at RCA Laboratories, Princeton, N. J. His research activities have included the study of electrical properties of semiconductors and insulators, of light emission from semiconductor diodes and lasers, and Q-switching of the CO_2 laser. For three years Dr. Meyerhofer was associated with a laboratory doing applied research in support of the RCA Graphic Systems Division. He investigated electronic applications in the printing industry. This included the use of lasers for exposure of photosensitive printing plates and for direct machining of gravure cylinders; electronic halftone screening and holographic font storage. He is now a member of the Electronic Printing Group at RCA Laboratories.

Dr. Meyerhofer is a member of the American Physical Society, the Optical Society of America, and the IEEE.

John Murr, Jr. attended the Valparaiso Technical Institute in 1955. He received his diploma in electrical engineering in 1957. He joined RCA Laboratories in 1957 where he worked on optical-electronic components. Since 1966 he has been concerned with the fabrication of silicon solid-state devices.

Louis S. Napoli received his B.S. in 1959 and the M.S. In 1961 in Electrical Engineering, both from Rutgers University. He has pursued further studies in electrophysics at the Polytechnic Institute of Brooklyn. He joined the technical staff of RCA Laboratories in June, 1959, and has engaged in research in phase-locked oscillators as a microwave logic device. He has specialized in research relating to microwave phenomena in gaseous plasmas, solid-state microwave devices, and is now engaged in research on solid-state microwave integrated circuits. His work in collaboration with Dr. George Swartz on amplification at 24 Gc by the interaction of an electron beam with a cesium plasma was cited by Industrial Research Magazine as one of the 100 most important achievements in 1963.

Mr. Napoli is a member of Sigma Xi, the Institute of Electrical and Electronics Engineers, Tau Beta Pi, and Eta Kappa Nu.

Edward C. Ross received the BSEE degree from Drexel Institute of Technology in 1964 and the MSE, MA, and Ph.D. degrees from Princeton University in 1966, 1967 and 1969, respectively. He joined RCA Laboratories in 1964, where he has been working in the fields of integrated arrays of MOS field-effect transistors, the electrical and physical characteristics of silicon-on-sapphire films, and the physics and application of field-effect storage transistors.

Dr. Ross is a member of Tau Beta Pi, Eta Kappa Nu, and the Institute of Electrical and Electronics Engineers.

Max J. Schindler received his M.S. degree in Electronic Engineering in 1951, and the degree of Doctor of Technical Sciences in Solid State Physics in 1953, both from the Technische Hochshule in Vienna, Austria. From 1951 to 1954, he served as a research and teaching assistant at the Technische Hochschule in the Department of Metal Physics. From 1954 to 1957, he worked at the Tungsram-Watt tube plant in Vienna. From 1957 to 1958, he was a research scientist at the Wright Air Development Center. Dr. Schindler joined the RCA Microwave activity in 1958, and has since worked on a number of basic technical problems related to the design of traveling-wave tubes, magnetrons, crossed-field devices, and solid-state devices. Among his major contributions are improvements in periodic-permanent-magnet focusing structures and the formulation of an improved design technique for such structures. He has also made unique contributions in a study of thin magnetic films, and in the computer-aided analysis of crossed-field phenomena and other microwave devices. Dr. Schindler is a Senior Member of the IEEE.

Daniel Tarangioli attended the Trenton Technical Institute in 1965 where he completed two years of studies in electrical technology. From 1960 to 1965 he was employed by Bell Telephone Laboratories, working in the area of material technology and solid-state devices. In 1967, he joined RCA Laboratories where he has been concerned with the fabrication of silicon solid-state devices.

John J. Walsh received the BSEE from Oklahoma A&M in 1949 and the MSEE from the University of Pennsylvania in 1956. Mr. Walsh spent 15 years in the Gravure Printing Industry at Triangle Publications, Parade Publications, Publication Corporation, and Detroit Gravure, where he held various engineering and production positions. He was associated with early installations of electronic register control systems and electronic press drives. In 1959 he was responsible for the introduction of the first solid state register control system in the industry. In 1964 he became associated with the U. S. Agency for International Development and served in Istanbul, Turkey as a printing and electric-electronic advisor. In 1966, he joined RCA's Graphic Systems Division in marketing and product planning. He was associated with the Graphic Systems Applied Research group at the Laboratories. For the past five years he has been an independent consultant in the Graphic Arts. He is a registered Professional Engineer in Pennsylvania.