RCA Review

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

Contents

March 1971 Volume 32 Number 1

Microwave Amplification Using Transferred-Electron Devices in Prototype Filter Equalization Networks

Barry S. Perlman, RCA Electronic Components
Princeton, N. J.

Abstract—A new type of negative-resistance device has evolved from an extensive study of the transferred-electron effect in bulk GaAs. The use of this two-terminal device in cw reflection-type amplifiers offers the possibility of instantaneous bandwidths of several octaves, output power levels above 1 watt, and a linear dynamic range in the order of 100 dB. The characterization and application of several typical devices with filter prototype equalizer networks has permitted the realization of wideband amplifiers operating in C-, X-, and KU-bands with bandwidths near 4.0 GHz, power outputs near 300 mW at a gain compression of -1 dB, a saturated output approaching 1 watt, and a noise figure of 15 dB.

Introduction

Recently, solid-state epitaxial GaAs devices exhibiting the transferred-electron effect have been utilized to achieve reflection-type microwave amplification[1] with both wide bandwidth and large dynamic range.[2] In this paper, these devices are characterized as stable negative real immittances over bandwidths exceeding an octave, and are used in filter type equalizer networks to design circulator-coupled cw amplifiers operating at frequencies from 4 to over 16 GHz.

Several of these stabilized transferred-electron devices have been measured and characterized by an equivalent circuit consisting of a negative real part and a capacitive component. A typical device has a negative-resistance component at frequencies from below 6.0 GHz to above 12.4 GHz. The available gain and bandwidth associated with a given device have been studied in terms of the validity of its equivalent

* This work was supported in part by the Naval Research Laboratory, Washington, D.C., Contract N00014-70-C-0383 and by the U.S. Army Advanced Ballistic Missile Defense Agency, Wash., D.C., in conjunction with the U.S. Army Electronics Command, Fort Monmouth, N.J., Contract DAA 1307-69-0444.

circuit and the choice of a suitable prototype filter impedance-equalizing network. The sensitivity of the amplifier gain to variations in the negative resistance of the active device has been used as a basis for predicting limitations in the design procedure.

In addition to providing the means for very-wide-band amplification, these same devices have been characterized by an unusually large linear dynamic range. A single active device with a typical noise figure of 15 dB has given a −1 dB gain-compression output power above 250 mW and a saturated power output near 1 W. The basic characteristics of these amplifying devices and their utilization with suitable microwave filter circuit techniques to achieve wide-band reflection-type amplification are described.

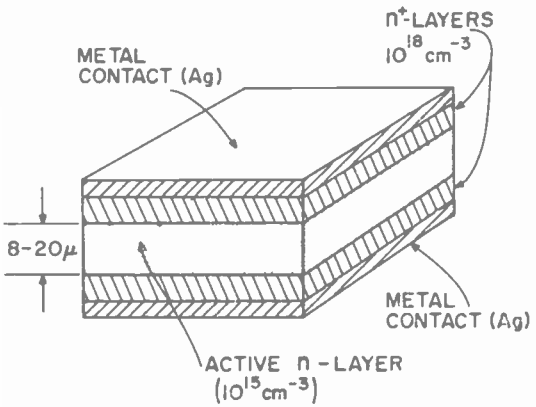


Fig. 1—Cross section of an epitaxial GaAs transferred-electron device. The active n-layer is sandwiched between metallized low-resistivity capping layers.

Material and Device Structure and Properties

A representation of an n+-n-n+ epitaxial GaAs device geometry is shown in Fig. 1. The n+ substrate wafer material is initially sliced from an ingot of heavily doped, low-resistivity (i.e., .001 ohm-cm) single-crystal GaAs. The wafer is chemically polished and cleaned on one side and introduced into a vapor-phase reaction furnace. By means of a controlled chemical reaction between a gaseous mixture of arsenic and gallium compounds, a thin epitaxial layer of high-resistivity n-type (i.e., 1 ohm-cm) GaAs deposits on the substrate surface.[3,4] The resulting structure retains the crystalline properties of the substrate. The thickness and doping characteristics of the active n-layer are

specified in accordance with the achievement of a particular mode of device operation. The remaining low-resistivity n+ capping layer is grown during the same process by introducing an appropriate dopant element into the gaseous mixture.

The low-resistivity n+ layers are extrinsic with a carrier density near $10^{18}/cm^3$ and are utilized to establish an electrical contact to the high-resistivity n layer. The thickness of the epitaxial n+ layer is minimized in order to provide a low thermal impedance path for heat removal and thereby maximize the power handling capability of the device. The n layer is usually doped with a carrier density in the low $10^{15}/cm^3$ range. At this concentration, the diffusion length for carriers from the n+ layers is in the order of 1-2 μm. The effect of a specified doping profile and carrier diffusion in thin samples as they relate to the stability and performance of the active device has been discussed in the literature.[5,6] The doping-density–length product, nl was chosen to be greater than $5 \times 10^{11}/cm^2$. The thickness of the active layer was chosen to cause the frequency associated with the transit time of carriers to lie within the desired amplifier pass band. A typical X-band (8.0-12.4 GHz) device has a layer thickness near 10 μm while at C-band frequencies (4.0-8.0 GHz), a layer thickness near 20 μm is used.

The n+-n-n+ sandwich structure is metalized with an evaporated thin film of silver as shown in Fig. 1 in order to form bondable ohmic contacts. The electrical properties of the metal–semiconductor interface are carefully controlled by sintering the material at an elevated temperature in an inert atmosphere. The contacted material is diced or chemically milled into individual devices, with the particular geometry determined by a power dissipation limitation. Each device is bonded onto the copper pedestal of a coaxial metal–ceramic package. In order to achieve a low thermal resistance bond, the thinner epitaxial side of a square diced chip is placed onto the copper pedestal. A typical epi-side down bonded device in a microwave package is shown in Fig. 2. The total thermal resistance for these devices has been measured to be typically less than 8°C/watt.[7]

A measurement of the dc voltage (I-V) characteristic of the individual devices is used to provide a preliminary estimation of the quality of the process techniques and the low-field material properties. The low-field electron mobility (mobility = velocity/electric field) is calculated from Hall and magnetoresistance measurement data and is typically between 6000-8000 cm²/volt-sec. The negative differential mobility, $dV/dE \simeq -2{,}400$ cm²/volt-sec is found using theoretical data for fields between the threshold value of approximately 3200 V/cm and

10 kV/cm. A typical static I-V characteristic for material with $nl > 5 \times 10^{11}/\text{cm}^2$, is shown in Fig. 3a, where an abrupt current decrease and bias circuit oscillations are observed for voltages above the current saturation threshold. The dc component associated with the formation of a space-charge instability with a characteristic microwave frequency appears as an instantaneous current drop followed by a saturated value. The bias circuit oscillations indicate the presence of a low-frequency negative resistance. This type of supercritical device, which is oscillating in the Gunn mode, exhibits a negative resistance from dc to above the transit-time frequency.[8] In a subcritically doped device, with $nl \leq 5 \times 10^{11}/\text{cm}^2$, the current will saturate above threshold, since there is no low-frequency negative resistance,[9] as shown in Fig. 3b. Subcritically doped devices have a positive resistance at low frequencies and thus exhibit neither a current drop nor bias oscillations.

Device Behavior at Microwave Frequencies

Stable negative-resistance devices have been fabricated from epitaxial GaAs, which has a doping density length product (nl) greater than $5 \times 10^{11}/\text{cm}^2$. These devices have been used to construct linear reflection-type transferred-electron amplifiers.[1,2,10-12] The specific physical

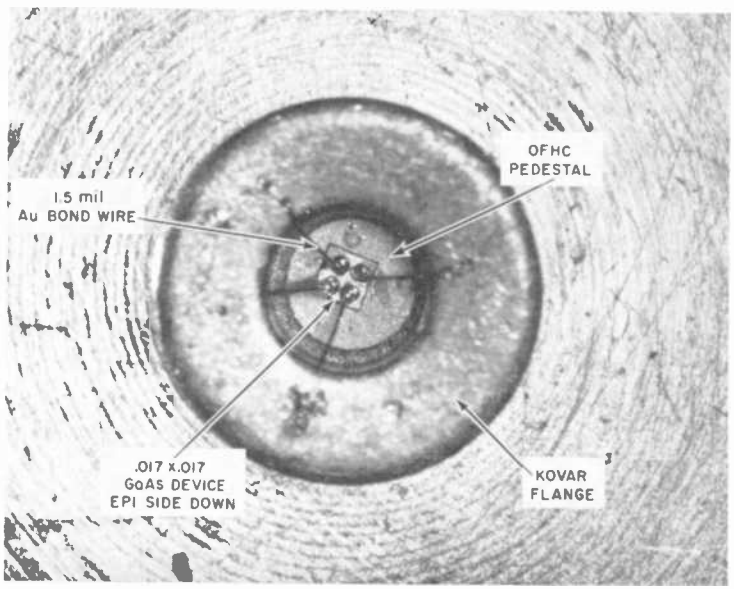
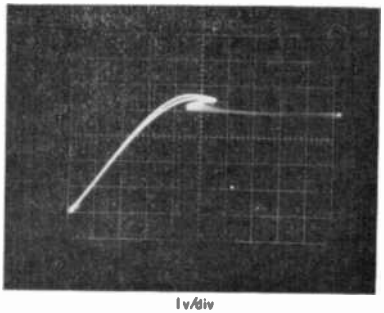


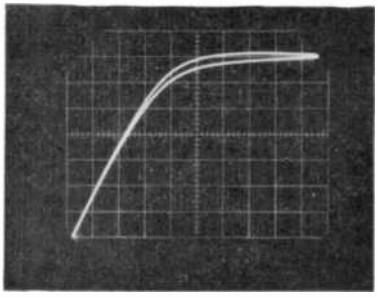
Fig. 2—A single active GaAs device bonded epitaxial side down onto a copper pedestal in a microwave coaxial package.

mechanism associated with the achievement of a nonoscillatory negative-resistance mode of operation with supercritically doped GaAs devices is complex and not fully understood. It has been found that a variation in the internal device parameters (i.e., doping profile and contacts) can dramatically alter the external characteristics of the active device. The additional constraints imposed by the circuit interaction with the active device as a function of both impedance, bias voltage and ambient temperature can also affect the stability.

Amplification has been observed in the presence of growing space-charge waves with devices in which $nl < 2 \times 10^{12}/\text{cm}^2$ when the bias voltage is within a narrow range ($\sim$25%) of the threshold value ($V_{th} \simeq 3.5$ volts for 10 μm material).[13] These results indicated that in this mode of operation, the bias voltage range over which amplification occurred decreased to zero as the nl product increased toward $2 \times 10^{12}/\text{cm}^2$. At higher bias voltages, large-signal instabilities were observed and found to be independent of the circuit impedance. This



(a) SUPERCRITICALLY DOPED AMPLIFIER TYPE DEVICE



(b) SUBCRITICALLY DOPED DEVICE

Fig. 3—*I-V* characteristics for epitaxial GaAs devices.

mode of post threshold amplification has been discussed analytically by Kroemer.[14]

The mode of operation that has been found to result in an unusually broad-band negative resistance occurs when the bias voltage is significantly larger than the threshold value. The usually observed instabilities cease to exist for bias voltages in excess of a new stability threshold when the device is operated in a suitable low-resistance circuit. The level of voltage where stability is achieved is a function of the physical characteristics and temperature of the device as well as the circuit impedance. The maximum net saturated power available from the amplifying device is equivalent to the output that would be realized from the same device operated as an oscillator.

The active device impedance may be measured by operating the device in a 50-ohm transmission line circuit that has been calibrated for operation with a reflection-type network analyzer. By this method, an accurate swept-frequency measurement may be made under variable-bias conditions. The stray and package reactances are initially measured and are eliminated from the device measurement.

An accurate electrical circuit representation of the active device consists of the combination of a negative conductance, G_d, in shunt with a capacitive susceptance, B_d. The measured admittance for a transferred-electron-amplifier device with an active layer length of 12 μm and biased to 2.5 times threshold (10 volts) is shown in Fig. 4a. The negative conductance is greater than -6 mmhos with a maximum of -16 mmhos over the frequency band 7.0-12.4 GHz. The shunt susceptance is essentially due to the parallel-plate capacitance of the bulk device ($C_d \simeq 0.8$ pF). The deviation of the capacitive susceptance from a linear function of frequency above 9.0 GHz is attributed to the variation of the dielectric constant of GaAs ($\epsilon_r \simeq 12.5$).

For the purpose of this paper, it is desirable to represent the active device as a *series* equivalent network. The impedance data shown in Fig. 4b was found by inverting the admittance data [Fig. 4a]. The latter *series RC* result is completely equivalent to the former parallel *GB* data. The primary difference involves a consideration of the circuit impedance criteria for stability applicable to each representation.

The significance of these results is that the device equivalent circuit may be represented as the combination of a negative resistance and a capacitive reactance over a wide range of frequencies and power levels. If the equivalent circuit resistance and capacitance are approximated as constants over their usable frequency range, the use of a filter-type imbedding network with a *series* resonant active device as a load may be used to obtain amplification over very wide bandwidths. The

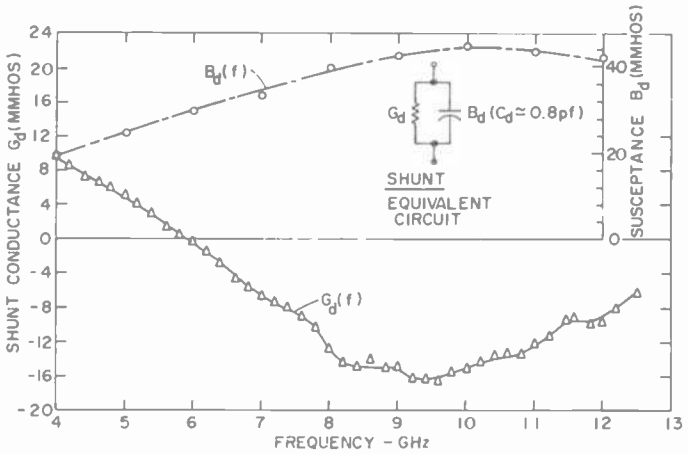


Fig. 4a—Admittance components of a stabilized active transferred-electron device as a function of frequency. The shunt representation is the physically correct circuit for this voltage-controlled device.

resonant condition is readily achieved by properly specifying the package lead inductance. The use of these techniques and the relationship between the circuit and device impedance to achieve a specific amplifier gain and stability are discussed in the following section.

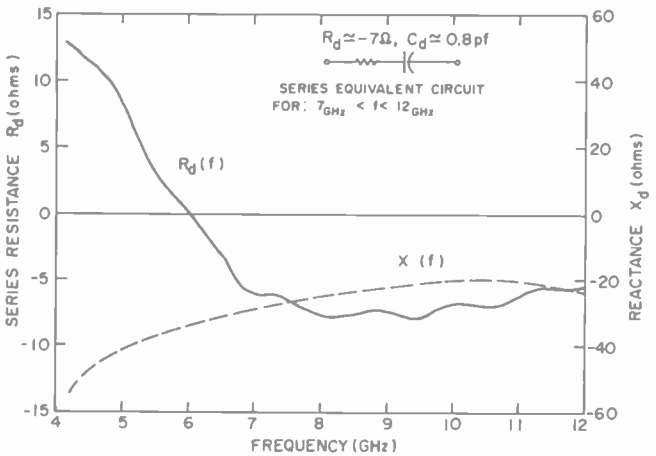


Fig. 4b—Equivalent series impedance components of the stabilized active transferred-electron device as a function of frequency. An equivalent $-RC$ network is also indicated.

Gain and Bandwidth of a Transferred-Electron Amplifier

A preferable method of achieving reflection-type amplification with a negative-resistance device is to utilize a nonreciprocal circulator to separate the input and output signals. Although reciprocal coupling circuits may be used (e.g., a shunt-coupled transmission line), instead of the circulators, the maximum available gain is limited to be 6 dB less than the gain attainable with the circulator.[15] The use of a single 90° hybrid allows a 3 dB increase in power since two devices are used, but the attainable gain is seriously dependent on the relationship of the source and device circuit impedances. The following discussion is concerned with the amplification of a nonreciprocal circulator and the impedance equalizer circuitry necessary to achieve a large gain–bandwidth product with a transferred-electron device.

A circulator-coupled reflection-type amplifier is shown schematically in Fig. 5. The circulator is assumed to be lossless and ideal (input

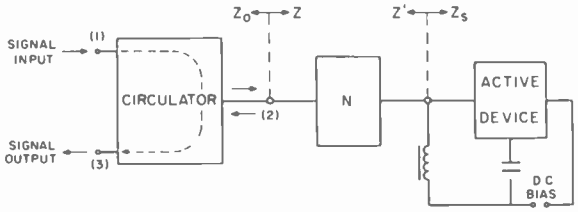


Fig. 5—Schematic representation of a circulator coupled impedance equalized reflection type amplifier. The network N is used to provide both real impedance transformation and reactive equalization.

impedance is real and constant with respect to frequency). In practice, the circulator may be characterized by a finite bandwidth and an in-band reactance function that is a monotonic function of frequency. Since the purpose of this discussion is to study the available performance of the transferred-electron device, the band-limiting property of the circulator will be neglected. The network N represents a lossless reciprocal two-port equalizer used to attain a particular amplifier gain frequency response. This network provides the necessary positive real impedance transformation and reactive filter structure between the negative resistance and the circulator load. If we represent the active device as a constant negative resistance over a finite frequency range, a known filter prototype may be used as the reactive network.[16-18] The reactive component of the device immittance is then included in the filter structure. The maximum attainable bandwidth is then limited

by the quality factor, Q_d, of the device and the range of frequency over which the equivalent circuit for the device and the constant transformed real load impedance are valid assumptions.

The reflection power gain, which is equal to the magnitude squared of the circuit reflection coefficient Γ between the series coupled circulator impedance Z_o and the input impedance Z of the loaded filter circuit, may be shown to be given by the equivalent expression,[19]

$$G = |\Gamma|^2 = \left| \frac{Z' - Z_d^*}{Z' + Z_d} \right|^2, \qquad [1]$$

where the Z_d is the equivalent series impedance of the device and Z' is prescribed in terms of a specific filter design and gain response. The same expression is valid for a shunt representation of the equivalent networks by substitution of admittances for the impedance functions, Z' and Z_d. Because the reactive component of the device impedance will be included in the filter structure, the device impedance Z_d will be specified equal to the assumed constant value for the negative resistance $-R_d$. The function Z' then includes all the reactive elements in an imaginary term $X(\omega)$ in addition to a real load impedance R_o. The power gain may now be expressed as

$$G = \left| \frac{R_o + R_d + j\,X(\omega)}{R_o - R_d + j\,X(\omega)} \right|^2, \qquad [2]$$

where the midband resonant ($X(\omega) = 0$) gain is given by

$$G_o = \left| \frac{R_o + R_d}{R_o - R_d} \right|. \qquad [3]$$

For a midband power gain of 10 dB, the required real circuit impedance is given by

$$R_o \simeq 1.9\,R_d. \qquad [4]$$

Eq. [4], is valid *only* when the device is *series* resonated.

Stability is maintained by avoiding a condition where the total circuit reactance ($X(\omega) = \mathrm{Im}(Z')$) becomes equal to zero at a frequency where the total circuit resistance ($R_d + R_o$) $\leqq 0$. Such a frequency is referred to as a critical frequency, at which the amplifier oscillates. If the device were antiresonant (shunt tuned), stability would

require that the total parallel circuit conductance be positive at any critical frequency. This condition would be met by requiring that the resonant, positive load conductance be larger than the small-signal negative device conductance, G_d. This criteria for the stability of a negative-resistance transferred-electron device has been discussed by Sterzer.[20]

The measured data described in Fig. 4 may be used to provide an equivalent circuit for the active device. The quality factor at the resonant frequency ω_o is given by

$$Q_d = \frac{1}{\omega_o R_d C_d} \text{ (series representation)}, \qquad [5a]$$

$$Q_d = \omega_o R_d C_d = \frac{B_d}{G_d} \text{ (parallel representation)}. \qquad [5b]$$

A representation of the *series* equivalent circuit of the active device and the resonant-filter-type coupling circuit is shown in Fig. 6. The circulator impedance is *assumed* equal to R_o in order to establish the desired midband gain.

If only one resonant network is used in order to obtain a singly tuned gain response, the attainable bandwidth may be readily found from Eq. [2]. By specifying that the gain is decreased to xG_o at the band edge frequencies ω_1 and ω_2, the realizable fractional bandwidth is found to be

$$B_f = \frac{\omega_2 - \omega_1}{\omega_o} = \frac{\sqrt{2}}{Q_d}\left(\frac{g_o}{(g_o - 1)\sqrt{x g_o^2 - 1}}\right), \qquad [6]$$

with

$$\omega_o = \sqrt{\omega_1 \omega_2}, \qquad [7]$$

and

$$g_o = G_o^{1/2}, \qquad [8]$$

and where the expression

$$X(\omega) = Q_d\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right) \qquad [9]$$

has been used for the singly tuned reactance function.

In order to increase and flatten the gain–bandwidth response a more complex network is required. Several possible filter response functions are available (i.e., maximally flat Butterworth, equal-ripple Tchebychev, linear phase, etc.). Regardless of the type of reactive equalizing network, the attainable flat fractional bandwidth for a specific gain with a negative-resistance device is bound and independent of the choice of network, and has been shown to be,[15]

$$B_f = \frac{2\pi}{Q_d \ln G_o}, \qquad [10]$$

where a flat gain of G_o has been assumed in the pass band $\omega_1 \leq \omega \leq \omega_2$. This upper bound for B_f depends only on the choice of gain G_o and on the device quality factor Q_d. It can never be exceeded, regardless of how the network N is constructed.

When the device is incorporated into a specific filter structure as the last element and the real impedance level of the filter network is maintained constant and equal to R_o over the entire filter bandwidth, the available fractional bandwidth becomes

$$B_f = \frac{ng_1}{Q_d}, \qquad [11]$$

where

$$n = \frac{R_o}{R_d} = \frac{g_o + 1}{g_o - 1}, \qquad [12]$$

and g_1 is the normalized filter prototype parameter associated with the end resonator.

The g_k parameters of the normalized filter circuit have been tabulated in several references[21-23] with the number of resonators and the ripple deviation (Tschebychev) from flat transmission as parameters. For the purpose of this analysis, an odd-order Tschebychev prototype will be used as a model for the broad-band amplifier design. An even-order prototype may be used to achieve a real impedance transformation, but with an increased ripple factor. The odd-order response is characterized by a higher insertion loss in the stop band as compared with that of a Butterworth prototype, and the bandwidth is defined at the equal-ripple end points. For the odd-order prototype, the filter network is symmetric with regard to element values. Once a particular

set of g_k values are specified, the corresponding network may be designed as shown in Fig. 6, where the series element values are

$$L_{k_{\text{even}}} = \frac{g_k R_o}{\Delta\omega}, \quad C_{k_{\text{even}}} = \frac{1}{\omega_o 2L_k}, \qquad [13]$$

and the shunt elements are

$$C_{k_{\text{odd}}} = \frac{g_k}{\Delta\omega R_o}, \quad L_{k_{\text{odd}}} = \frac{1}{\omega_o 2C_k}. \qquad [14]$$

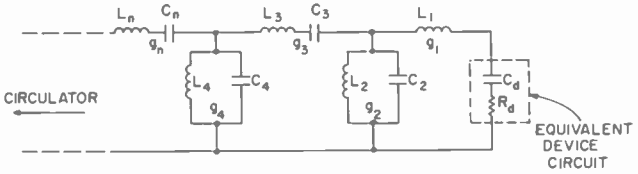


Fig. 6—Band-pass filter prototype for imbedding the active device into the end resonator. The individual element values are given by the normalized g_k parameters that are related to the resonator Q's.

A comparison of the attainable fractional bandwidths described by Eqs. [6], [10], and [11] as a function of midband gain G_o for both single- and multiple-resonator circuitry is indicated in Fig. 7, where the product $Q_d B_f$ has been plotted against G_o(dB). Several types of response characteristics are indicated as parameters. The closest approximation to the shape of the ideal response is the 0.01 dB ripple Tschebychev case, which offers nearly one half the bandwidth capability of the ideal network. However, by increasing the allowable ripple deviation from a flat response (0.25-dB case) the bandwidth capability of the Tschebychev design is almost equal to that of the ideal case. In reality the actual amplifier response will usually not reproduce the filter response associated with the prototype used. Since the basis of this approximation is the validity of the active-device equivalent circuit, this design method is used only to achieve a reasonable network design that may be modified to be more compatible with the specific device used.

In order to obtain both wide-bandwidth and high-gain operation, the use of a passive network that faithfully provides a constant real

load impedance and an equalized reactance with respect to frequency requires that the negative resistance of the device remain essentially constant with respect to frequency. A measure of the frequency sensitivity of the amplifier gain to a variation in the relative negative resistance may be used to determine the limiting average pass-band gain that would be compatible with a specific gain fluctuation for a

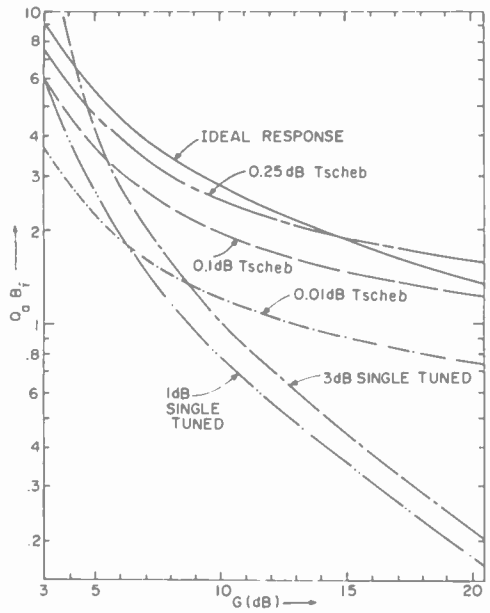


Fig. 7—Normalized fractional bandwidth $Q_d B_f$ as a function of the logarithmic gain G(dB) with the type of equalizer network as a parameter.

given device. Using the gain Eq. [3], assuming reactive equalization across the pass band, the gain sensitivity in terms of the relative device resistance variation becomes

$$S_{G_o} = \frac{dG_o}{G_o} = \frac{(g_o + 1)(g_o - 1)}{g_o} S_{R_d}, \qquad [15]$$

with

$$S_{R_d} = \frac{dR_d}{R_{d_o}}, \qquad [16]$$

where $-R_{d_o}$ represents the value of negative resistance used to achieve the gain $G_o = g_o^2$ and dR_d and dG_n represent the variation in these parameters. The sensitivity ratio S_G/S_{R_d} is plotted in Fig. 8 against the power gain G_o expressed in dB. The effect of a 10% variation in $-R_d$ is seen to produce a 28.5% (~1 dB) variation of gain for a nominal gain of 10 dB.

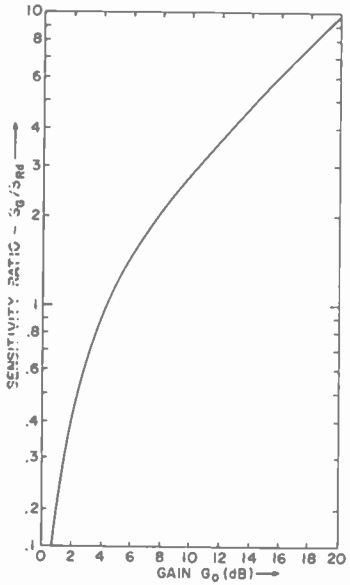


Fig. 8—The relative variation S_G of the amplifier gain to a change in the nominal value of the negative resistance S_{Rd} (i.e. as a function of frequency, etc.) as a function of the average gain G_o is given by a sensitivity ratio S_G/S_{Rd}.

The device data and *series* equivalent circuit shown in Fig. 4b have been used in the amplifier design examples shown in Table 1. The bandwidth is chosen to correspond to a reasonable circulator capability as well as to the device characteristic. The gain is specified at a moderate value in order to minimize the gain sensitivity. Once the gain, bandwidth, and diode Q_d are specified, a reasonable choice for the filter prototype is made by use of Fig. 7. For the 12-dB case, a 0.1-dB Tschebychev should be used; while the 0.01-dB prototype may be used for the lower gains. The corresponding lumped equivalent filter circuit element values are readily found and realized as either distributed or semilumped resonators. A lumped equivalent amplifier filter network

Table 1—Amplifier Design Example

Equiv. Device R_d			$-7\ \left(\begin{smallmatrix}+0.5\\-1.0\end{smallmatrix}\right)$ ohms		
Equiv. Device C			0.8 pF		
Device Q_d			3		
Center Freq. f_o			8.8 GHz		
Bandwidth			7.0-11.0 GHz		
B_f			0.45		
Gain (dB)	6.0	7.0	8.0	10.0	12.0
Gain Sensitivity	0.32 (1.2 dB)	0.38 (1.4 dB)	0.44 (1.6 dB)	0.6 (2 dB)	0.78 (2.5 dB)
Filter Impedance	21	18.3	16.3	13.5	11.7

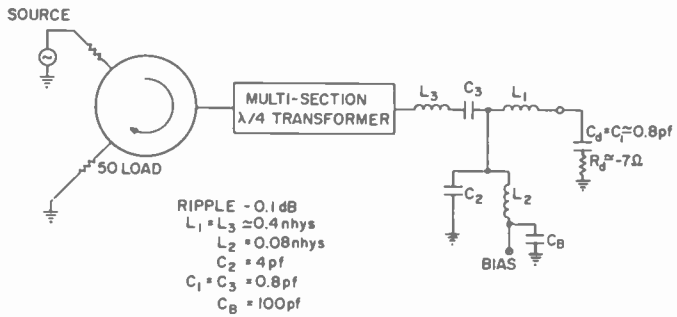


Fig. 9—Circulator-coupled reflection-type amplifier network comprising a multisection impedance transformer and a three-section filter-type equalization network. The active device equivalent circuit is included in the final filter resonator.

is shown schematically for a 3-section 0.1-dB Tschebyshev prototype in Fig. 9. For the lower gain cases an even-order prototype ($n = 4$) could be used to provide the necessary impedance transformation in addition to reactive equalization. A 1-dB-ripple, 4-element prototype would provide an impedance transformation ratio of 2.66.

Several wide-band transferred-electron amplifiers were designed using the filter prototype design technique and realized in a coaxial geometry as shown in Fig. 10. A wide-band quarter-wave transformer is used to adjust the circulator impedance to achieve a specific gain over

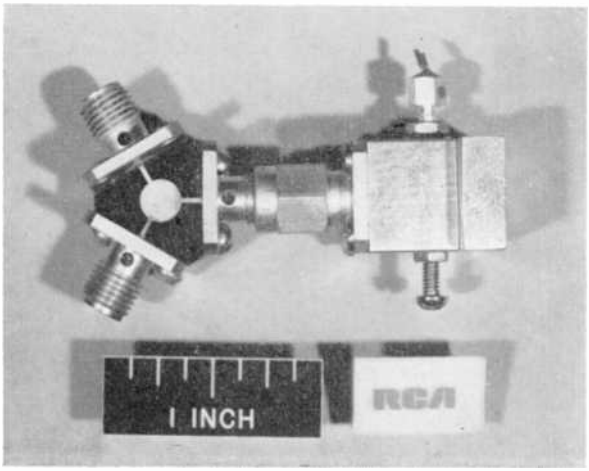


Fig. 10—Coaxial circuit realization of the filter imbedding network given in Fig. 9. A microstrip three-port circulator is used for input–output isolation.

the band 7.0 to 11.0 GHz. Both semilumped resonators and a self-resonant packaged device were used to obtain the necessary filter-type network. The actual performance of this amplifier is described in the next section.

Amplifier Performance

A single-stage transferred-electron amplifier was designed as described in the previous section (Fig. 10). To permit evaluation of amplifier performance, several values of filter impedance were used to adjust the midband gain.

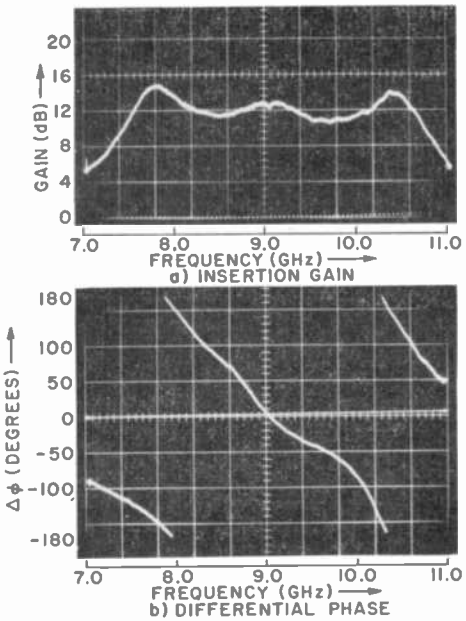


Fig. 11—(a) Gain response for an amplifier that has been designed for a midband gain of 12dB and (b) differential phase response for amplifier. The deviation from a linear phase characteristic is the result of the residual phase error inherent in the dual-channel heterodyne measuring circuit.

The active device was biased at 10 volts, which corresponds to a voltage about 3 times the current-saturation threshold value and is greater than the minimum voltage required for stable amplification. The insertion gain and differential phase shift are measured by comparing the amplified signal with an input reference signal. The gain for an amplifier that has been designed for a midband gain of 12 dB is shown in Fig. 11a. The measured bandwidth has been limited by

the frequency sensitivity of the transformer and circulator networks as well as an alteration in filter element values in order to achieve a reasonably flat response.

The differential phase response for this amplifier is shown in Fig. 11b. The deviation from a linear phase characteristic is the result of the residual phase error inherent in the dual-channel heterodyne measurement circuit.

The primary significance of these results is that a predictable wide bandwidth, substantial gain, and a linear phase response are attainable by use of the negative-resistance property of a single transferred-electron device in conjunction with a filter-type network. An increased amplifier gain is achieved by the use of several cascaded stages. The gain of a cascaded pair of similar 12-dB stages having a multiport circulator network is shown in Fig. 12.

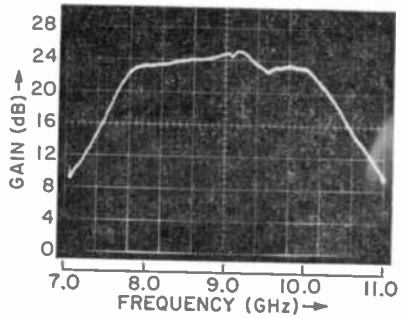


Fig. 12—The gain response obtained by cascading two identical 12-dB stages.

One of the important considerations in determining the gain of the final stage is maintaining linearity. The output power corresponding to a gain in compression of −1 dB is increased as the gain is decreased. The sensitivity of this gain is related to the negative-resistance sensitivity as shown in Fig. 8. An increasing output power causes a decrease in the negative resistance and consequently a decrease in reflection gain. Because this gain sensitivity increases with increasing gain, a low output stage gain is desirable. The power input–output characteristic for a single stage is shown in Fig. 13. The saturated output power for this amplifier occurs with a gain near 3 dB. In general, the large signal gain corresponding to the condition where the device power $(P_o - P_i)$ is equal to its maximum value (i.e., equivalent to the power available from the same device operated as an oscillator) is related to the small signal gain.[24] The maximum available power from the amplifier is related to the sum of the input signal and the saturated device powers.

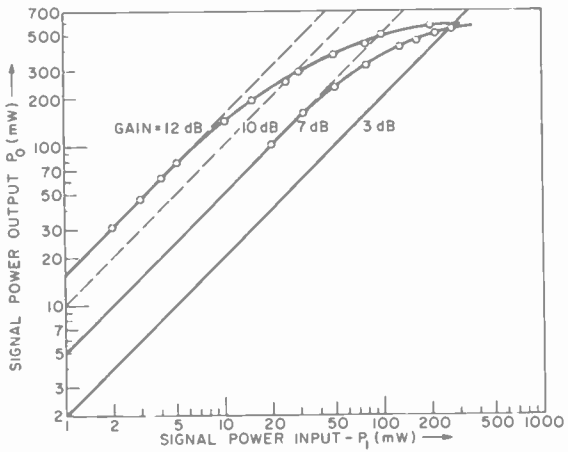


Fig. 13—Power input–output characteristic for a single 12-dB stage.

The saturated output power is observed to approach ½ watt for the amplifiers described here. The output power at −1 dB gain compression is nearly 300 mW for a gain of 7 dB but is only 200 mW when the gain is 12 dB. The lower limit on the dynamic range is related to the noise figure, which has been measured to be 15 dB.

The measured results shown here for an X-band amplifier are representative of the performance that has been obtained for similar amplifiers with center frequencies from 6.0 to 14.0 GHz and bandwidths compatible with present circulator technology. A summary of these results is shown in Table 2.

Table 2—Amplifier Performance

Frequency Band	Small-Signal Gain (dB)	3-dB Bandwidth (GHz)	P_o (watts) (−1 dB) Gain Compression	P_o (watts) Gain = 3 dB
C	8 (22)*	4.5-8.0 (5.0-7.5)	0.250	1
XL	12 (24)*	7.5-10.75 (7.65-10.25)	0.200	0.55
X	6	8.0-12.0	0.250	0.5
KU	6	12.0-16.0	0.250	0.5

* Two Stages.

Conclusion

The feasibility of constructing high-level cw wide-band solid-state microwave amplifiers has been demonstrated using GaAs semiconductor devices that exhibit the transferred-electron effect. These devices are capable of providing negative real impedances over bandwidth exceeding an octave and have been used in circulator-coupled low-impedance filter-type networks to achieve wide-band reflection-type amplifiers operating at frequencies from 4.0 to 16.0 GHz.

The use of filter prototype networks offers a means of realizing amplifier bandwidths and gains compatible with the negative resistance bandwidth and frequency sensitivity and the availability of broad-band transformer and circulator networks.

Single-stage small-signal gains from 6 to 12 dB have been obtained for center frequencies from 6.0 to 14.0 GHz with bandwidths up to 4.0 GHz. The −1 dB gain compression power output has been typically 250 mW with a saturated power output exceeding 500 mW for both C- and X-band devices. The amplitude and phase linearity of these amplifiers has been found to be comparable with traveling-wave-tube amplifiers. With a noise figure of 15 dB, these wide-band amplifiers have demonstrated a dynamic range near 90 dB.

The use of multiple stages has permitted the realization of both high gain and wide bandwidths with the transferred-electron amplifiers. These linear solid-state amplifiers may replace tubes in many microwave systems applications because of their small size and weight, reliability, and low power requirements. As the available output power, gain, and bandwidth increase, this new type of negative resistance amplifier should find increased applications in microwave system design.

Acknowledgment

The author would like to express his appreciation to Mrs. A. San Paolo and W. Klatskin for the fabrication and evaluation of the low thermal resistance bonded devices, R. Enstrom, I. Heygi and A. Dreeben for the epitaxial material, and C. Upadhyayula and R. Marx for their technical assistance, and to B. Berson for many stimulating discussions.

References:

[1] B. S. Perlman and T. E. Walsh, "Stabilized Supercritical Transferred Electron Amplifiers," Digest of Technical Papers, ISSC, Philadelphia, Pa., Feb. 1969 and **IEEE Trans. on Solid State Circuits**, Vol. SC-4, No. 6, Dec. 1969.
[2] B. S. Perlman, L. C. Upadhyayula and R. E. Marx, "Wideband Reflection Type Transferred Electron Amplification," **IEEE Trans. Microwave Theory and Techniques**, Vol. MTT-18, No. 11, Nov. 1970.

[3] R. E. Enstrom and C. C. Peterson, "Vapor Phase Growth and Properties of GaAs Gunn Devices," **Trans. Met. Soc. AIME,** Vol. 239, p. 413, 1967.
[4] J. J. Tietjen and J. A. Amick, "Preparation and Properties of Vapor Deposited Epitaxial $GaAs_{(1-x)}P_x$ Using Arsine and Phosphine," **J. Electrochem. Soc.,** Vol. 113, p. 724, 1966.
[5] J. Magarshack and A. Mircea, "Stabilization and Wideband Amplification Using Over-Critically Doped Transferred Electron Diodes," MOGA 70, Eighth Int'l. Conference on Microwave and Optical Generation and Amplification, Amsterdam, Sept. 1970.
[6] M. Suga and K. Sekido, "Effects of Doping Profile Upon Electrical Characteristics of Gunn Diodes," **IEEE Trans. Electron Devices,** Vol. ED-17, No. 4, p. 275, April 1970.
[7] S. Y. Narayan, R. E. Enstrom and A. Gobat, "High Power CW Transferred Electron Oscillators," **Elec. Ltrs.,** Vol. 6, No. 1, p. 17, Jan. 1970.
[8] H. W. Thim, "Linear Negative Conductance Amplification with Gunn Oscillators," **Proc. IEEE** (Letters), Vol. 55, p. 446, March 1967.
[9] J. A. Copeland, "Characterization of Bulk Negative-Resistance Diode Behavior." **IEEE Trans. Elec. Dev.,** Vol. ED-14, No. 9, p. 461, Sept. 1967.
[10] B. S. Perlman, "CW Microwave Amplification from Circuit Stabilized Epitaxial GaAs Transferred Electron Devices," Digest of Technical Papers, ISSC Conference, Philadelphia, Pa., p. 134, Feb. 1970 and **IEEE Trans. Solid State Circuits,** Vol. SC-5, No. 6, p. 331, Dec. 1970.
[11] J. Magarshack and A. Mircea, "Wideband CW Amplification in X-Band with Gunn Diodes," Digest of Technical Papers, ISSC Conference, Philadelphia, Pa., p. 132, Feb. 1970.
[12] B. S. Perlman and R. E. Marx, "Linear Microwave Solid State Transferred Electron Power Amplifier with a Large Gain Bandwidth Product," Digest of Technical Papers, PTGMTT Symposium, p. 227, May 1970.
[13] B. W. Hakki and S. Knight, "Microwave Phenomena in Bulk GaAs," **IEEE Trans. Electron Devices,** Vol. ED-13, p. 94, Jan. 1966.
[14] H. Kroemer, "The Gunn Effect Under Imperfect Cathode Boundary Conditions," **IEEE Trans. Electron Devices,** Vol. ED-15, No. 11, p. 819, Nov. 1968.
[15] D. C. Youla and L. I. Smilen, "Optimum Negative Resistance Amplifiers," **Proc. Symposium on Active Networks and Feedback Systems,** Polytechnic Institute of Brooklyn, April 1960.
[16] W. J. Getsinger, "Prototypes for Use in Broadbanding Reflection Amplifiers," **IEEE Trans. MTT,** Vol. MTT-11, No. 6, p. 486, Nov. 1963.
[17] W. J. Getsinger and G. L. Matthaei, "Some Aspects of the Design of Wideband Upconverters and Nondegenerate Parametric Amplifiers," **IEEE Trans. MTT,** Vol. 12, p. 77, Jan. 1964.
[18] G. L. Matthaei, "A Study of the Optimum Design of Wideband Parametric Amplifiers and Upconverters," **IRE Trans. MTT,** Vol. MTT-9, p. 23, Jan. 1961.
[19] H. W. Bode, **Network Analysis and Feedback Amplifier Design,** D. Van Nostrand, Inc., Princeton, New Jersey, 1945.
[20] F. Sterzer, "Stabilization of Supercritical Transferred Electron Amplifiers," **Proc. IEEE,** Vol. 57, No. 10, p. 1781, Oct. 1969.
[21] L. Weinberg, **Network Analysis and Synthesis,** McGraw-Hill, 1962.
[22] J. L. Altman, **Microwave Networks,** D. Van Nostrand, Inc., Princeton, New Jersey, 1964.
[23] G. L. Matthaei, et. al., **Design of Microwave Filters, Impedance-Matching Networks, and Coupling Structures,** McGraw-Hill, 1964.
[24] B. S. Perlman, L. C. Upadhyayula, and W. W. Siekanowitz, "Microwave Properties and Applications of Negative Conductance Transferred Electron Devices," **Proc. IEEE,** Aug. 1971 (to be published).

A New Type of Gallium Arsenide Field-Effect Phototransistor

G. A. Swartz, A. Gonzalez, and A. Dreeben
RCA Laboratories, Princeton, N. J.

Abstract—A gallium arsenide field-effect phototransistor, which responds to infrared radiation at a wavelength of 1.5 μm, has been developed. The device is fabricated from chromium-doped semi-insulating GaAs with a thin epitaxial n-type surface layer. The carrier depletion field exists in the p^+-n-type junction formed between the substrate and the epitaxial layer. The radiation serves to increase the depletion field by decreasing the substrate resistance. The phototransistor has a small-signal power gain of 30 dB. The noise equivalent power of the device at 1.5 μm is 10^{-6} watt/cm².

Introduction

A gallium arsenide field-effect phototransistor, which is sensitive to light with energy less than the bandgap energy, has been developed. In the past, several types of phototransistors have been proposed and developed.[1-5] Field-effect phototransistors developed by H. Nelson[3] and R. R. Bockemuehl[4] have similarities to the device described in this paper. However, important differences in structure, material, and operation result in properties that are unique to each type of phototransistor. Nelson's phototransistor consists of a thin semiconducting n-layer with source and drain contacts and a p-layer underneath the n-layer to form a p-n junction. The current through the n-layer is controlled by back-biasing the junction with a constant-current source. Light striking the n or p layer creates hole–electron pairs, which reduce the junction resistance and thus lower the voltage drop across the junction. The reduced junction voltage results in increased current flow from source to drain. The phototransistor developed by Bockemuehl consists of a thin rectangular slab of cadmium sulfide with source and drain contacts at opposite ends and a gate contact on one surface between the source and drain. The slab conductance is a function of gate potential and the light intensity and wavelength. The high output impedance (~50

megohms) and the low transconductance (20 micromhos) limits the usefulness of the device.

The device reported in this paper consists of a thin n-type epitaxial layer of GaAs, with source and drain contacts, grown on an insulating chromium-doped GaAs substrate. The charge carriers in the n-layers are depleted by simultaneously biasing the substrate negative and reducing the substrate resistivity with infrared-excited carriers from the chrome traps. The depletion field is a consequence of the junction formed by the substrate and the n-layer.[6] Different substrate dopants in a series of host materials could be used to produce phototransistors sensitive to light in a wide range of frequencies.

Device Construction ond Operation

Gallium arsenide with a chromium density of about 10^{17} atoms/cm^3 is used as a semi-insulating substrate for the growth, by vapor-phase techniques, of an n-type layer of GaAs about 2 μm thick. The surface resistivity of the n-layer, as determined by measurement of the partial absorption of a 3 GHz wave reflected off the surface, is 600 ohms/sq to 7,000 ohms/sq depending on the particular wafer. The average layer thickness, d, is determined from the weight gain of the substrate during the growth process. Assuming an electron mobility of 6000 cm^2/V-sec, the carrier density in the n-layer varied from 10^{15} to 7×10^{15} cm^{-3}. Although the device operates with n-layers of high or low surface resistivity, most of the tests were performed on n-layers with a 600 ohms/sq surface resistivity.

Au-Ge-Ni (90%-5%-5%) contact strips, 0.03 inch wide and 0.05 inch apart, are deposited on the n-layer in vacuum and sintered into the wafer at 450°C. The wafer is diced into sections. For a typical section, the length l is 0.08 inch, the width W is 0.02 inch, and the substrate height h is 0.014 inch. The substrate is attached to a base plate with silver paint, and one-mil gold wires are bonded to each contact. A schematic drawing of a test device together with the bias circuits is shown in Fig. 1. The source and drain contacts are biased to V_A by a voltage source in series with a load or current-measuring resistor. The base contact is biased by the voltage source, V_B, at a negative potential with respect to the source contact. Light can penetrate the test sample through the n-layer or the sides of the substrate.

Typical current–voltage characteristics of the n-layer for various values of V_B are shown in Fig. 2 where the device is in the dark and in Fig. 3 where the device is illuminated. The illumination is supplied by the filament in a microscope lamp heated to a dull red and the fila-

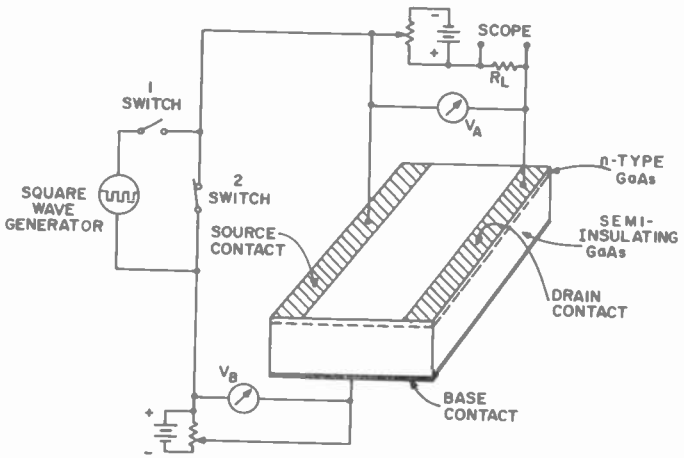


Fig. 1—Schematic drawing of test device with associated circuitry.

ment is focused on to the device. A plot of the source to drain current, I_A, as a function of the base voltage, V_B, is used to determine the device transconductance $\Delta I_A/\Delta V_B$. A plot of I_A versus V_B is shown in Fig. 4 with and without the device illuminated. Figs. 2, 3 and 4 show that the device operates as a transistor with or without illumination; however, illumination increases the transconductance of the device.

The device is sensitive to light in two energy categories; (1) the energy required to excite electron–hole pairs and (2) the energy required to excite electrons from the chromium impurities into the con-

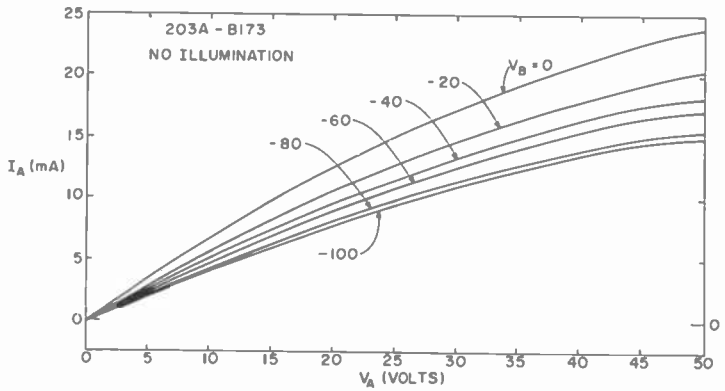


Fig. 2—Source-to-drain current, I_A, as a function of voltage, V_A, for various values of base voltage, V_B, with the device 203A-B173 not illuminated.

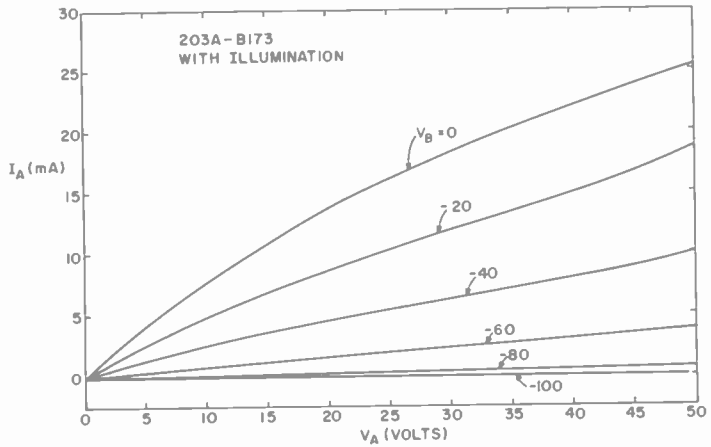


Fig. 3—Source-to-drain current, I_A, as a function of voltage, V_A, for various values of base voltage, V_B, with the device 203A-B173 illuminated.

duction band, which is about 0.8 eV.[7-10] This is illustrated by the two curves in Fig. 5 which show the current I_A as a function of the bulk filament temperature. The data for curve A are taken without a filter between the light source and the device, the data for curve B with a silicon filter in front of the device. The silicon with a bandgap of 1.1 eV absorbs the radiation at 1.4 eV which could excite electron–hole pairs in GaAs. Both curves in Fig. 5 show a sharp decrease in I_A at a low bulb-filament temperature (~800°C). At a high filament temperature (>1200°C), I_A increases when the filter is absent and bandgap radiation

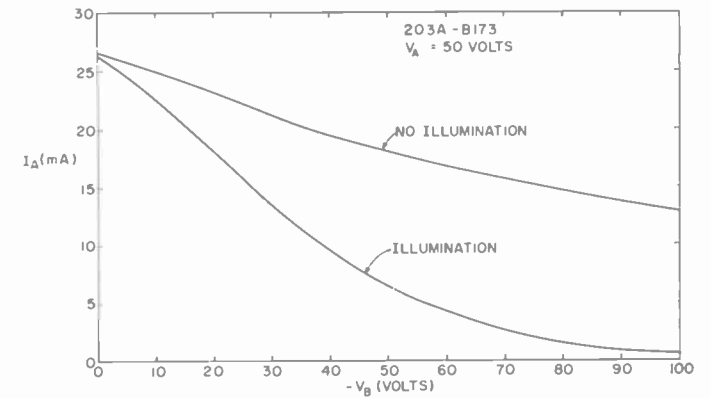


Fig. 4—I_A as a function of V_B for illuminated and unilluminated device 203A-B173.

generated by the high-temperature filament is allowed to reach the device (curve A). I_A does not increase when the silicon filter prevents the bandgap radiation from striking the device (curve B). A germanium filter in front of the device, absorbs all radiation above an energy of 0.7 eV, including the chrome excitation energy of 0.8 eV, and eliminates all control of I_A by the microscope light. The experiment with the germanium filter together with the work by D. R. Heath, et al, indicates that the light does not excite an appreciable density of holes from the chrome levels into the valence band.[6] The excitation energy of the holes is 0.6 eV. Light with this energy is passed by the germanium and does not affect the device operation. The experiments

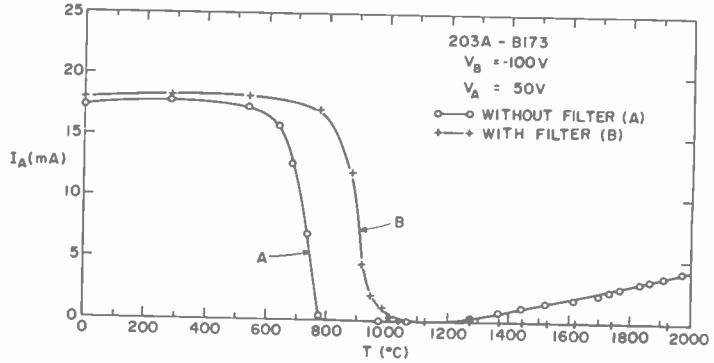


FIG. 5

Fig. 5—I_A as a function of bulb-filament temperature for filtered and unfiltered illumination. The filter is a silicon wafer 0.007 inch thick.

indicate that the sharp decrease in I_A with increasing bulb voltage at low filament temperature is caused by the radiation at the electron excitation energy. The increase in I_A with increasing filament temperature is caused by radiation at the bandgap energy.

With a silicon filter in front of the device to eliminate the bandgap radiation, ΔI_A is plotted in Fig. 6 as a function of relative light intensity. The light intensity is varied by changing the distance between a point light source and the device.

Figs. 5 and 6 graphically illustrate the switching capability of the device. The illumination causes I_A in several of the test devices to be decreased by a factor of 5×10^3.

To estimate the light sensitivity of the device, a black body was placed about 6 mm from the test sample and a light chopper positioned between the black body and the test device. The radiation from a 210°C

black body chopped at 100 Hz causes a periodic variation in I_A that is equal to the noise level of the device. Using Plancks radiation law, the noise equivalent power is 0.5 microwatt/cm^2 at a light energy of 0.8 eV.[11]

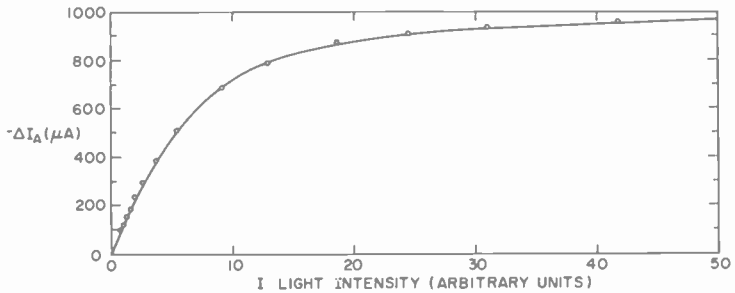


Fig. 6—The decrease in source-to-drain current, $-\Delta I_A$, as a function of relative light intensity. The light is filtered with a silicon wafer.

Various device properties such as power gain, saturation power output, and frequency response were measured. A square-wave generator or a pulse generator is connected in series with the base-bias voltage and a load resistor equal in value to the resistance between source and drain contacts is placed in series with the bias voltage, V_A, as shown in Fig. 1 when S1 is open and S2 is closed. The potential, V_L, across the load resistor, R_L, is monitored, as a 1.0-kHz square wave is applied to the base contact. The rise time of V_L indicates the frequency response and the amplitude of V_L indicates the power output. The value of the rise time, τ_r, is decreased by illumination or heating of the device. Examples of the input and output waveforms for illuminated, unilluminated, and heated devices are shown in Figs. 7(a), 7(b), and 7(c). With sufficient infrared illumination to produce a rise time of 2×10^{-4} sec, the output power from device 203A-B173 is measured as a function of input power as shown in Fig. 8. Device operating parameters are:

Device width $W = 0.02$ inch
Input resistance $R_{in} = 12$ megohms
Output resistance $R_{out} = 1000$ ohms
n-layer bias voltage $V_A = 25$ volts
Cutoff frequency $f_c = 1.5$ kHz
Transconductance $g_m = 0.5$ millimho
Small-signal gain $G = 30$ dB
Saturated power output $P_{out} > 64$ milliwatts

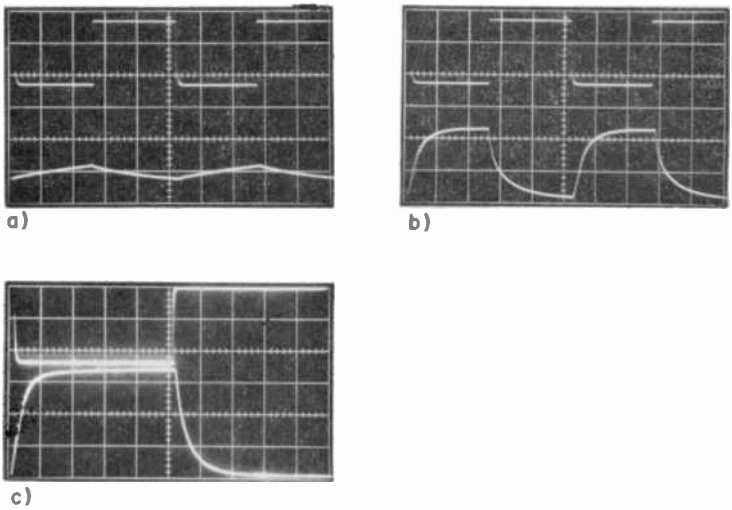


Fig. 7—Input (top traces) and output (bottom traces) waveforms for device 203A-B173. Vertical scale for top trace in a, b, and c – 20 volts/div., dc value of base potential, $V_{B_0} = 55$ volts, $V_A = 25$ volts. (a) Horizontal scale—0.2 millisec/div, bottom trace; vertical scale—10 mA/div. Device is unilluminated. (b) Horizontal scale—0.2 millisec/div, bottom trace; vertical scale—10 mA/div. Device illuminated with filtered light. (c) Horizontal scale— 0.1 millisec/div., bottom trace; vertical scale—2.0 mA/div. Device at temperature of 140°C.

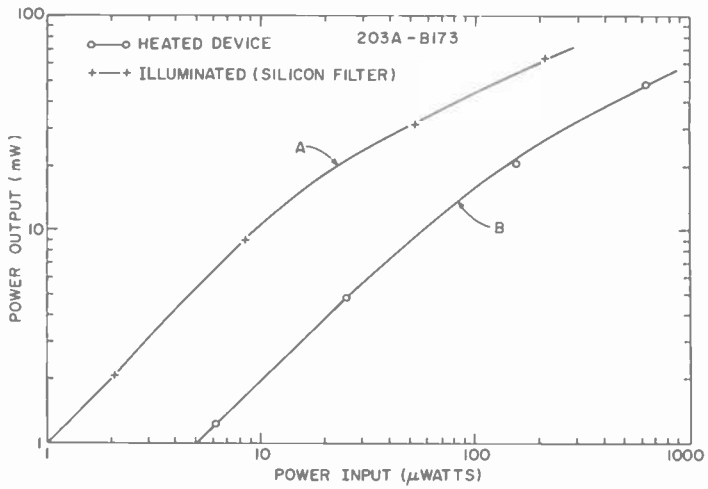


Fig. 8—Power output as a function of power input for device 203A-B173. Device is illuminated with infrared light ($\lambda > 1.1$ μm).

Heating the device to 140°C so that the rise time is decreased to 5×10^{-5} sec ($f_c = 6.4$ kHz) reduces the small-signal gain to 20 dB. A cutoff frequency of 100 kHz has been observed in a heated device with a gain of 9 dB.

The electric field required to deplete carriers from the n-layer is related to the carrier density, n_-, and the ionized donor density, N_+ by the Poisson equation in one dimension;

$$\frac{\partial E}{\partial X} = \frac{(N_+ - n_-)e}{\epsilon}. \qquad [1]$$

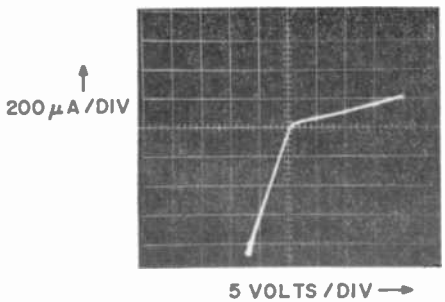


Fig. 9—Current from base contact to n-layer as a function of base potential for sample 208-B173. Horizontal scale—5 volts/div, vertical scale—200 μa/div. Sample temperature = 140°C and is illuminated.

For a spatially uniform donor density and a fully depleted layer

$$E = \frac{N_+ de}{\epsilon}. \qquad [2]$$

With $N_+ d = 10^{12}$ cm^{-2}, $E = 1.45 \times 10^5$ volts/cm. To achieve this high electric field with a typical base potential of 100 volts, a junction is assumed to be present under the n-layer.[6] The presence of this junction is not readily verified by measuring the I-V characteristic between the base contact and the n-layer because of the large series resistance presented by the semi-insulating substrate. To minimize the large series resistance, the substrate of sample 208-B173 with $W \times l = 55 \times 65$ mils^2 was polished to a 0.002 inch thickness and then heated to 140°C while being illuminated. The I-V characteristics between the base contact and the n-layer, as shown in Fig. 9, is that of a junction.

and unilluminated sample. The measured value is within 10% of the calculated value. R_s is determined directly from the high-frequency (500 kHz) resistance measurement on the heated and illuminated sample. C_j and R_j are determined from the low-frequency (5 kHz) measurements using Eqs. [3] and [4]. With $C_s = 5.47$ picofarads, $R_s = 15.5$ kilohms, $C_j = 470$ picofarads and $R_j = 155$ kilohms, curves of effective capacitance and resistance are calculated as a function of frequency using Eqs. [3] and [4]. In Fig. 12 plots of the calculated values together with the experimental points show that the agreement is very good.

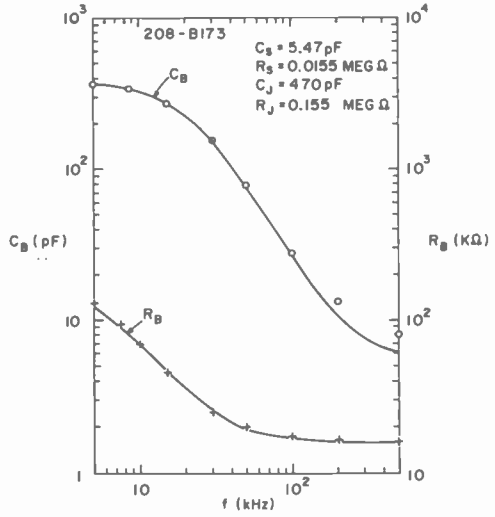


Fig. 12—Effective capacitance and resistance of substrate in series with junction as a function of measurement frequency for sample 208-B173. The data marks are measured values and the lines are calculated using $C_s = 5.47$ picofarads, $R_s = 15.5$ kilohms, $C_j = 470$ picofarads and $R_j = 155$ kilohms.

At a bridge frequency of 5 kHz, C_B was measured as a function of bias voltage. In the low frequency limit

$$C_B = C_j\left(1 - 2\frac{R_s}{R_j}\right). \qquad [5]$$

The dc bias voltage, V_B, is related to the junction voltage by

$$V_j = V_B \frac{1}{1 + (R_s/R_j)}. \qquad [6]$$

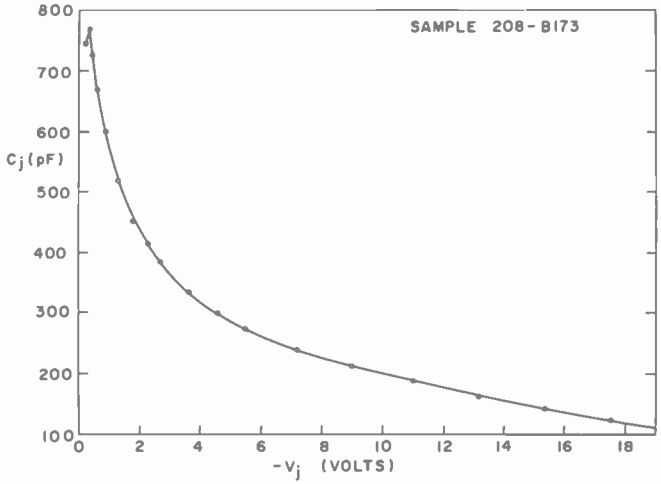


Fig. 13—Junction capacitance, C_j, as a function of base potential, V_B, for sample 208-B173. Sample temperature is 140°C and is illuminated.

Using Eq. [5], which is applicable at 5 kHz for sample 208-B173, C_j is shown in Fig. 13 plotted as a function of V_j. Given the condition $P \gg N$, a density profile of the n-layer is computed[12] from a plot of C_j vs V_j and is shown in Fig. 14. The density profile of the n-layer shows a density decrease at a distance of 1.0 μm from the n-layer–substrate interface. The surface resistivity calculated from the density profile is 1000 ohms/sq and the measured value is 600 ohms/sq.

The equivalent circuit shown in Fig. 11 is used to understand the

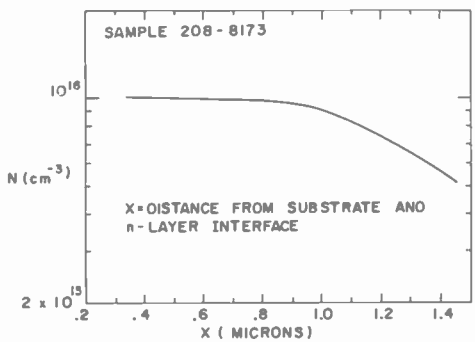


Fig. 14—Donor density as a function of distance from n-layer–substrate interface. Density profile calculated from C-V measurement in Fig. 13.

time constant that is observed when a step-function or square-wave base potential is applied. Application of a step voltage V_B to the two end terminals results in a junction voltage, V_j, with the form

$$V_j = V_B \left[\left(\frac{C_s}{C_s + C_j} - \frac{R_j}{R_s + R_j} \right) e^{-t/\tau} + \frac{R_j}{R_s + R_j} \right], \qquad [7]$$

where

$$\frac{1}{\tau} = \frac{1}{C_s + C_j} \left(\frac{1}{R_s} + \frac{1}{R_j} \right). \qquad [8]$$

At $t = 0$,

$$V_j = V_B \frac{C_s}{C_s + C_j}, \qquad [9]$$

and at $t = \infty$, Eq. [6] is applicable. Eq. [7] assumes C_j is independent of V_j, which is not valid. This assumption is made to linearize the circuit equations for the purpose of understanding which parameters determine the device response time. In the Appendix, the rise time of I_A together with the circuit equations are used to estimate the value of R_s and R_j in device 203A-B173. The resistances are $R_j = 0.5$ megohm and $R_s = 3.5$ megohms. Since $C_s \ll C_j$ and $R_s \gg R_j$

$$\frac{1}{\tau} = \frac{1}{R_j C_j} = \frac{\sigma_j}{\epsilon}, \qquad [10]$$

where σ_j is the effective conductivity of the depletion layer. The response time to an input step potential, which is derived in the Appendix, is

$$\frac{I_A - I_{A_o}}{I_{A_\infty} - I_{A_o}} = (1 - e^{-t/\tau})^{1/2}, \qquad [11]$$

where $I_{A_o} = I_A$ at $t = 0$ and $I_{A_\infty} = I_A$ at $t = \infty$.

The I_A vs V_B curve in Fig. 4 together with the resistances calculated in the Appendix applied to Eq. [6] indicate a layer depletion at $V_B \cong -80$ volts and $V_j \cong -10$ volts.

The transconductance, $\Delta I_A/\Delta V_B$, which is derived in the Appendix, is

$$\frac{\Delta I_A}{\Delta V_B} = \frac{(\epsilon/2Ne)^{1/2}}{d} \frac{\dfrac{1}{1+(R_s/R_j)}}{\left[V_B\left(\dfrac{1}{1+(R_s/R_j)}\right)+V_c\right]^{1/2}}. \qquad [12]$$

With no illumination $R_s \gg R_j$ and the value of the transconductance is small. An increase in device illumination decreases the ratio R_s/R_j by generating carriers in the substrate, where chromium is present, at a faster rate than in the junction where chromium is absent. This decrease of R_s/R_j causes an increased junction voltage, depletion field, and transconductance. Application of heat to the device increases the thermally generated hole density, which also reduces R_s/R_j and improves the device performance.

The device has a potential use in millimeter-wave switching. With an initial epitaxial-layer surface resistance (377 ohms/sq) equal to the impedance of free space, the device will absorb all, part, or none of an impinging millimeter-wave beam depending on the surface conductivity, which is controlled by the base potential and light intensity. The conductivity of a number of small sections on a large area wafer can be sequentially controlled. The switching time, which is $\sim(R_jC_j)$, is sufficient to scan 10^2 points in 0.01 to 0.1 sec. Thus, scanning a millimeter wave image focused on a GaAs wafer seems feasible.

Summary

Field-effect phototransistors have been fabricated from chromium-doped semi-insulating GaAs with a thin epitaxial n-type surface layer. The transistor has a source and drain contact on the n-layer with a gate contact on the semi-insulating substrate. A p_+-n type junction forms between the substrate and the n-layer. Application of a negative potential to the gate contact back-biases the junction. For a given gate or base potential, the junction voltage is determined by the ratio of the substrate resistance to the junction resistance. Excitation of electrons from the chrome impurities in the substrate with infrared radiation increases the junction voltage and n-layer depletion field by reducing the substrate resistance with respect to the junction resistance. Heating the device has the same effect by thermally generating holes from

the chrome impurities. Thus, control of the n-layer conductivity is achieved with gate potential and either infrared radiation at an energy less than the bandgap energy or device heating.

The field-effect phototransistor is capable of 30 dB small-signal gain and a saturated power output greater than 64 milliwatts. The cutoff frequency of the device is a function of the RC time constant of the junction. For an illuminated device this frequency is typically 2.0 kHz. The noise equivalent power at 1.5 μm is 10^{-6} watt/cm^2.

This type of GaAs phototransistor could utilize other substrate impurities with deep energy levels such as iron or nickel[9] to produce sensitivity to light at wavelengths greater than 1.5 μm. Other host materials such as germanium or silicon could be used to fabricate the phototransistor when the substrate is doped with p-type, deep-level impurities to create a semi-insulating material. Each impurity would respond to a different light frequency. By utilizing various host and impurity materials devices sensitive to light in a wide range of wavelengths could be fabricated.

Acknowledgments

We wish to thank B. Hershenov and R. H. Dean for many helpful suggestions and L. S. Napoli for the use of his computer program in the analysis of the CV measurements, as well as his useful advice. We also wish to thank A. Triano, W. M. Anderson, and N. Klein for their technical assistance.

Appendix

Application of a step-function voltage to the base contact causes a decrease in the current I_A. The time constant associated with the decrease in I_A is used to estimate the value of R_s and R_j in device 203A-B173. From the Poisson equation,

$$X = \left(\frac{2\epsilon}{Ne} V_j\right)^{1/2}, \qquad [13]$$

where X is the carrier depletion distance and N is the donor density, which is assumed to be uniform with X. The current is dependent on the depletion distance

$$I_A = I_{A_o}\left(\frac{d - X}{d}\right), \qquad [14]$$

where I_{A_o} is the total current before the base potential is applied. Combining Eqs. [13] and [14],

$$I_A = I_{A_o}\left(1 - \frac{(2\epsilon/Ne)^{1/2}}{d} V_j^{1/2}\right). \qquad [15]$$

With the assumption that C_j is independent of V_j and using the circuit in Fig. 11, the junction voltage is related to the base potential by

$$V_j = V_B\left[\left(\frac{C_s}{C_s + C_j} - \frac{R_j}{R_s + R_j}\right) e^{-t/\tau} + \frac{R_j}{R_s + R_j}\right], \qquad [16]$$

where

$$\frac{1}{\tau} = \frac{1}{C_s + C_j}\left(\frac{1}{R_s} + \frac{1}{R_j}\right). \qquad [17]$$

The junction barrier potential, V_c, which is the potential difference between the Fermi level in the substrate and the n-layer, is about 0.8 volt and is negligible for the purpose of this analysis.

In device 203A-B173 the current I_A shown in Fig. 7 does not noticeably decrease the instant that the base potential V_B is applied. Thus it is assumed $C_s/(C_s + C_j) < R_j/(R_s + R_j)$. With $C_s/(C_s + C_j)$ negligible,

$$I_A = I_{A_o}\left\{1 - \frac{(2\epsilon/Ne)^{1/2}}{d}\left[V_B \frac{R_j}{R_s + R_j}\left(1 - e^{-t/\tau}\right)\right]^{1/2}\right\} \qquad [18]$$

and

$$\frac{I_A - I_{A_o}}{I_{A_\infty} - I_{A_o}} = (1 - e^{-t/\tau})^{1/2}, \qquad [19]$$

where $I_{A_\infty} = I_A$ at $t = \infty$. When $t = \tau$

$$(1 - e^{-1})^{1/2} = \left(\frac{I_A - I_{A_o}}{I_{A_\infty} - I_{A_o}}\right)_{t=\tau} = 0.795. \qquad [20]$$

From geometric consideration $C_s \ll C_j$ and Eq. [17] becomes

$$\frac{1}{\tau} = \frac{1}{C_j}\left(\frac{1}{R_s} + \frac{1}{R_j}\right)$$

$$= \frac{1}{C_j}\frac{R_j + R_s}{R_s R_j}. \qquad [21]$$

The total resistance R_T is known, where

$$R_T = R_s + R_j. \qquad [22]$$

Thus

$$\frac{1}{\tau} = \frac{R_T}{C_j R_s R_j}. \qquad [23]$$

Using Eqs. [22] and [23],

$$R_j^2 - R_j R_T + \frac{R_T \tau}{C_j} = 0, \qquad [24]$$

and

$$R_j = \frac{R_T}{2}\left(1 \pm \sqrt{1 - \frac{4\tau}{R_T C_s}}\right). \qquad [25]$$

Application of Eq. [19] to the output current of device 203A-B173 shown in Fig. 7(c) gives a value of τ. The value of R_T is the input resistance. The value of C_j is derived from the measured value of C_B on sample 208-B173 at an 8-volt bias on the junction. The surface resistance on wafer B173 is uniform to within 30% and therefore the junction capacitance per unit area for various samples from the same wafer is assumed to be uniform to the same degree. The values for Eq. [25] are

$$\tau = 5 \times 10^{-5} \text{ sec},$$
$$R_T = 4 \times 10^6 \text{ ohms},$$
$$C_j = 105 \text{ picofarads}.$$

The calculated resistances are $R_j = 0.5$ megohm and $R_s = 3.5$ megohms. The sign uncertainty in Eq. [25] is resolved by observation of the I-V characteristic of the substrate resistance in series with the junction resistance. A typical junction characteristic is not observed and thus the series resistance R_s must be greater than R_j. To estimate the error introduced by the assumption that C_j is independent of V_j, the value of C_j for a 4-volt junction bias is used in Eq. [25]. The calculated value of R_j is reduced by 30% and R_s is increased by 5%.

To evaluate the transconductance, g_m, the barrier potential, V_c must be included to have g_m remain finite as $V_B \to 0$. In the dc situation, Eq. [18] with V_c included is

$$I_A = I_{A_0}\left\{1 - \frac{(2\epsilon/Ne)^{1/2}}{d}\left[V_B\left(\frac{1}{1+(R_s/R_j)}\right) + V_c\right]^{1/2}\right\}, \qquad [26]$$

and

$$g_m = \frac{dI_A}{dV_B} = \frac{(\epsilon/2Ne)^{1/2}\dfrac{1}{1+(R_s/R_j)}}{d\left[V_B\left(\dfrac{1}{1+R_s/R_j}\right)+V_c\right]^{1/2}}. \qquad [27]$$

References:

[1] J. N. Shive, "A New Germanium Photo-Resistance Cell," **Phys. Rev.**, Vol. 76, p. 575 (1949).

[2] G. H. Royer, "Circuits Using Special Semiconductor Devices," Ch. 6 in **Handbook of Semiconductor Electronics**, ed. L. P. Hunter, McGraw-Hill Book Co., N.Y., N.Y. (1956).

[3] H. Nelson, "The Preparation of Semiconductor Devices by Lapping and Diffusion Techniques," **Proc. IRE**, Vol. 46, p. 1062, June 1958.

[4] H. R. Bockemuehl, "Cadmium Sulfide Field-Effect phototransistor," **Proc. IRE**, Vol. 48, p. 875 (1960).

[5] T. O. Poehler and D. Abraham, "Electric Field Excitation of Electrons from Shallow Traps in CdSe Thin-Film Triodes," **J. Appl. Phys.**, Vol. 35, p. 2452 (1964).

[6] R. H. Dean et al, "Small-Signal Amplification," Interim Tech. Report, Contract No. F33615-69-C-1788, Air Force Avionics Lab., Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, March 1970.

[7] D. R. Heath, P. R. Selway, and C. C. Tooke, "Photoconductivity and Infrared Quenching in Chromium-Doped Semi-Insulating Gallium Arsenide," **Brit. J. App. Phys.**, Vol. 1, p. 29 (1968).

[8] P. R. Selway and J. Franks, "A Gallium Arsenide Infrared-Sensitive Diode Made by the Diffusion of Chromium," **Brit. J. App. Phys.**, Vol. 1, p. 25 (1968).

[9] G. A. Allen, "The Activation Energies of Chromium, Iron, and Nickel in Gallium Arsenide," **Brit. J. Appl. Phys.**, Vol. 1, p. 593, (1968).

[10] F. Hasegawa, "Deep Energy Levels in the High Resistance Region at GaAs Vapor Epitaxial Film-Substrate Interface," **Jap. J. Appl. Phys.**, Vol. 9, p. 638 (1970).

[11] G. Joos, **Theoretical Physics**, pp. 613-616, Hafner Publishsing Co., N.Y., N.Y., (1950).

[12] S. M. Sze, **Physics of Semiconductor Devices**, pp. 370-372, Wiley-Interscience, New York, N.Y. (1969).

MOS Models and Circuit Simulation

John E. Meyer
RCA Laboratories, Princeton, N. J.

Abstract—A four-terminal large-signal model for the MOS transistor is derived in terms of physical device parameters. The theoretical predictions of the model are shown to be in excellent agreement with experimentally measured dc characteristics. A discussion of the effect of the nonlinear device capacitance on computer simulated transient results is included. A program package that performs on-line dc and transient analysis of R, C, and MOS circuits is described. Various output characteristics and waveforms are automatically plotted. A comparison of a computed MOS transient and an experimental measurement demonstrate the excellent results possible using computer-aided-design techniques.

Introduction

The characteristics of metal–oxide–semiconductor (MOS) field-effect transistors, namely, high circuit packing density; low power dissipation; simplified processing, and, therefore, lower cost have made large scale integration (LSI) a reality. MOS integrated circuits are extensively used in minicomputers and remote computer terminal equipment. The 1970 market value of these MOS integrated circuits was over fifty million dollars. Despite the obvious importance of such a market the MOS business has yet to realize its greatest potential—as LSI random-access memories for core replacement. Because of both the complexity and long-term reliability requirements of such LSI chips, it is necessary that their design be very nearly optimum in relation to both circuit and device processing. Another large area of the MOS business involves the design of large custom logic arrays. These custom designs require both fast turn around and accuracy in order that the vendor can earn a profit in a low-volume environment. The best method of achieving these design goals is through the use of computer-aided design (CAD) techniques.

Fortunately for MOS circuit designers the electrical characteristics of the MOS transistor are much more easily described and modeled in

terms of physical device parameters than are the characteristics of bipolar transistors. This fact is increasingly important because the advent of integrated circuits makes inseparable the tasks of circuit design and device design. Until now the overwhelming majority of MOS circuits have been based on the all p-channel MOS (PMOS) approach. Since the PMOS very nearly is an ideal or square-law insulated-gate field-effect transistor (IGFET) simple models[1,2] have generally been adequate for computer-aided design purposes. However, the demand for improved performance in terms of both speed and current-drive capability requires the use of an all n-channel MOS (NMOS) approach or a complementary-symmetry MOS (CMOS) approach requiring both NMOS and PMOS transistors. Both of these approaches require NMOS devices that because of a necessarily more heavily doped substrate are decidedly not square law. Fortunately, again, the additional terms for this bulk effect can be included in an accurate four-terminal MOS device model.[3,4] The model as presented here is in a form most suitable for computer-aided circuit design and has been incorporated into a circuit-analysis package at RCA Laboratories for several years.[5]

In addition, MOS devices do not exhibit any charge-storage effects comparable to bipolar transistors. Therefore, although the distributed gate-channel capacitance is nonlinear it can be adequately represented in a model by a lumped-constant average value. A further simplification results from the high-impedance nature of MOS devices, because of which it is usually possible to ignore the parasitic bulk series resistance associated with both the source and drain terminals.

Derivation of Large-Signal Device Model

1. Four-Terminal I-V Relationships

The structure of an n-channel MOS field-effect transistor is shown in Fig. 1. Let the potentials of the gate, drain, source, and substrate be V_g, V_d, V_s, and V_b, respectively. Further, let $V(z)$ be the potential at the surface of the channel and W, L, and T_{ox} be the channel width, source–drain spacing, and oxide thickness, respectively. At a distance z from the source the total charge per unit area Q_s induced in the silicon substrate will consist of a mobile negative charge Q_m in the surface inversion layer and a fixed negative charge Q_d in the underlying depletion region due to ionized impurity atoms. Thus,

$$Q_m(z) = Q_s(z) - Q_d(z). \qquad [1]$$

By Gauss' law,

$$Q_s(z) = -\frac{\epsilon_{ox}}{T_{ox}}(V_g - V_{tx} - V(z)) \qquad [2]$$

and

$$Q_d(z) = -\sqrt{2\epsilon_{si}qN[V(z) - V_b + 2\Phi_f]}. \qquad [3]$$

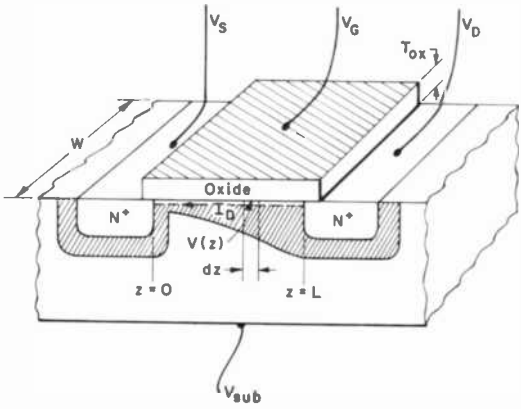


Fig. 1—NMOS structure.

Eq. [2] relates the total charge in the silicon to the impressed field $V_g - V(z)$, and an effective threshold voltage V_{tx}. This effective threshold voltage includes the contributions of the metal–semiconductor work function difference Φ_{ms}, the Fermi potential[6] of the substrate Φ_f, and the equivalent insulator (or interface) charge Q_{ss};

$$V_{tx} = 2\Phi_f - \Phi_{ms} - \frac{Q_{ss}}{C_{ox}}. \qquad [4]$$

Eq. [3] relates the charge of ionized impurities to the channel-substrate potential. The voltage drop across an elemental section of channel is given by

$$dV = IdR = -\frac{I_d dz}{W\mu Q_m(z)}. \qquad [5]$$

Substituting our previous equations into [5] and integrating,

$$-\int_0 I_d dz = \int_{V_s}^{V_d} \frac{W\mu\epsilon_{ox}}{T_{ox}} \left[-(V_g - V_{tx} - V(z) + \frac{T_{ox}}{\epsilon_{ox}} \sqrt{2\epsilon_{si} q N (V(z) - V_b + 2\Phi_f)} \right] dv, \quad [6]$$

or, upon performing the integration,

$$I_d = K \left[-(V_g - V_{tx} - V(z))^2 - \frac{4}{3} k (V(z) - V_b + 2\Phi_f)^{3/2} \right]_{V_s}^{V_d}, \quad [7]$$

where $$K = \frac{\mu\epsilon_{ox} W}{2LT_{ox}}, \quad [8]$$

and

$$k = \frac{T_{ox}}{\epsilon_{ox}} \sqrt{2q\epsilon_{si} N}. \quad [9]$$

Therefore, we note that the nonlinear relationship between channel (drain–source) current and the terminal voltages may be written as the difference of two nonlinear current components, each of which is dependent on only one pair of terminal voltages;

$$I_d = K[f(V_g, V_s) - f(V_g, V_d)], \quad [10]$$

where

$$f(V_g, V) = (V_g - V_{tx} - V)^2 + \frac{4}{3} k (V - V_b + 2\Phi_f)^{3/2}. \quad [11]$$

We note that for the entire channel to be inverted, the term between the brackets in Eq. [6] (the mobile channel charge) must be negative (minority carriers only) for all z, thus giving the following condition;

if

$$(V_g - V) \leqslant V_{tx} + k\sqrt{V - V_b + 2\Phi_f}, \qquad [12]$$

then

$$f(V) = (V_g - V_{tx} - V_{sat})^2 + \frac{4}{3} k (V_{sat} - V_b + 2\Phi_f)^{3/2}, \qquad [13]$$

where V_{sat} is found by solving Eq. [12] for V when the equality applies. Thus

$$V_{sat} = V_b - 2\Phi_f + \left(-\frac{k}{2} + \sqrt{\frac{k}{4} + V_g - V_{tx} - V_b + 2\Phi_f} \right)^2. \qquad [14]$$

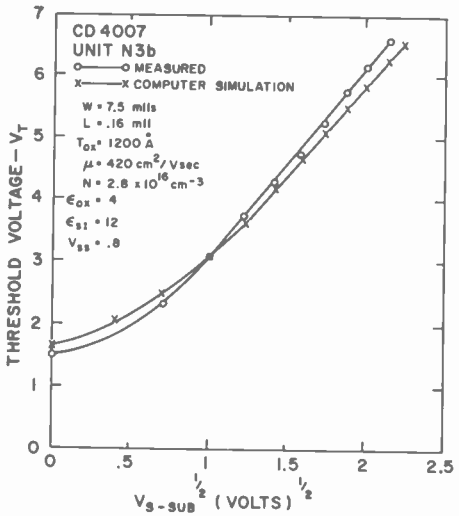


Fig. 2—Comparison of MOS Models with measured threshold variation.

It is important at this point to note that the condition of Eq. [12] when applied to the source end of the channel ($z = 0$, $V = V_s$) yields the familiar threshold voltage relationship, which is plotted in Fig. 2. When the same criterion is applied at the drain end ($z = L$) the equality yields the drain voltage for onset of drain current saturation, i.e., the pinch-off voltage.

As noted previously, an interesting feature of this model is that it is completely bidirectional and can be put in the form of two current sources, one dependent only on $V_g - V_s$ and the other only on $V_g - V_d$. This form is emphasized in Eq. [1] and is illustrated by the simple two-current source model of Fig. 3a. Proceeding one step further, we have the Ebers–Moll "like" model of Fig. 3b.[5,7,8] Both forward and

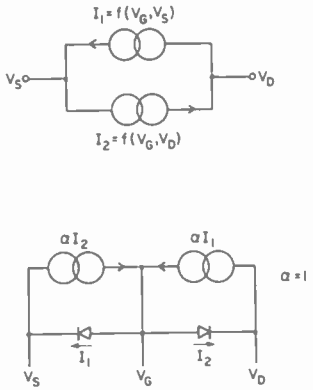


Fig. 3—(a) Two-current-source MOS model and (b) Ebers-Moll "like" MOS model.

reverse α's must equal unity in order to keep $I_g = 0$, as it must for an IGFET. The diode elements in the model are *field-effect diodes* as described by the *I-V* relationships of Eq. [11] through [14]. Thus we see that only one nonlinear characteristic is required to completely specify all the voltage–current characteristics of the device. Such characteristics are shown in Fig. 4.

The reader concerned about the physical meaning of the active characteristic, i.e., $I \neq 0$ at $V = 0$, of such field-effect diodes should realize that these diodes are merely mathematical fabrications that come about by separating the equations in this manner and cannot be measured experimentally. It is also interesting to note that the Ebers–Moll form of the MOS model is a form most easily programmed for SCEPTRE[9] users. This is so because SCEPTRE dc analysis can only handle diode-like functions that are functions of the voltage across the diode (V_b is treated as a constant parameter).

If the substrate impurity density N is small, the depletion charge term Q_d in Eq. [3] is negligible. Under these conditions Eq. [7] reduces to

$$I_d \cong K[(V_{gs} - V_{tx})^2 - (V_{gd} - V_{tx})^2], \qquad [15]$$

i.e., the familiar square-law characteristics.

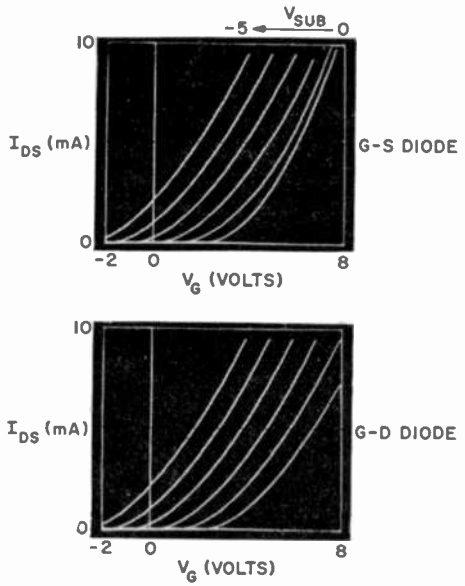


Fig. 4—Characteristics of FET diodes for $V_s = 0$ and $V_a = V_D$.

2. Drain Conductance in Saturation

Eq. [14] defines the drain voltage V_d at which $Q_m(z)$ at the drain becomes zero. Since in practice $Q_m(z)$ cannot become zero, thus implying zero drain current, we have made the approximation in Eq. [13] that beyond pinch-off, the drain conductance becomes zero while the drain current remains constant at a value equal to that achieved at pinch-off. However, as will be noted from the typical device characteristics in Fig. 5, the drain conductance is by no means zero in the saturation region. This conductance is generally attributed to two feedback mechanisms by which a change in drain voltage can affect the drain current. The first[10,11] mechanism is due to a spreading of the depletion region near the drain, which results in a reduction of the active channel length. It has been shown that the saturation conductance can be predicted in good agreement with experimental measurements[12] if the fringing fields introduced by the gate electrode

are accounted for in the model. The second mechanism by which a change in drain voltage can affect the drain current is by electrostatic coupling of the drain contact region to the channel. If the substrate is very lightly doped (or essentially does not exist as is the case for silicon-on-sapphire MOS[13]) some of the field lines emanating from the drain can penetrate the substrate and induce charge in the channel, thus acting as a somewhat inefficient gate.[14]

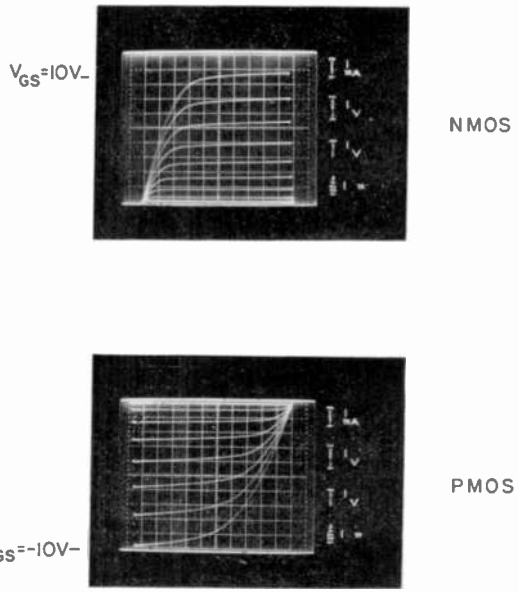


Fig. 5—I-V characteristics of CD4007 COS/MOS transistors.

For the purpose of large-signal computer-aided MOS analysis, it is usually sufficiently accurate to account for the finite drain conductance in an empirical manner[1] by defining the drain current as

$$I_{ds} = I_{ds}'(1 + \gamma |V_{ds}|), \qquad [16]$$

where I_{ds}' was derived previously. This equation in effect linearizes the results of Ref. [12]. Including this in our model, the theoretical device characteristics of an NMOS and PMOS device are shown in Fig. 6. Excellent agreement with Fig. 5 is achieved. A comparison of MOS gate characteristics is shown in Fig. 7.

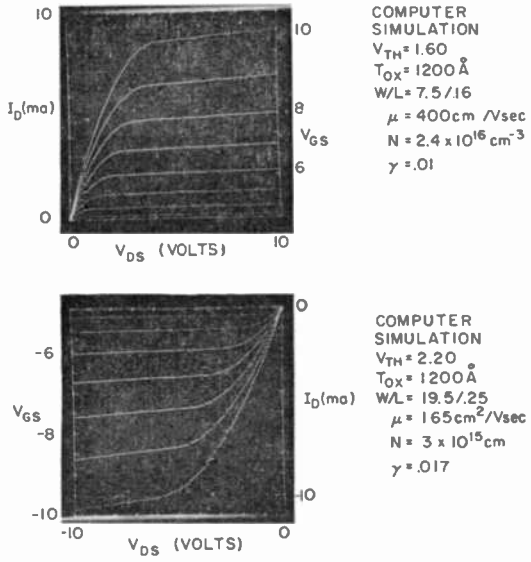


Fig. 6—Computer simulation of CD4007 COS/MOS transistors.

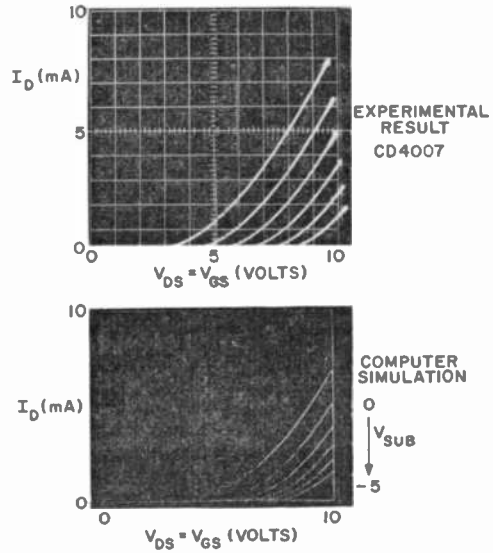


Fig. 7—Gate characteristics of NMOS.

3. Saturation Effects

There are two other separate saturation mechanisms that may or may not have an appreciable effect on the MOS device characteristics. The first mechanism is that of G_m saturation due to parasitic drain and source contact resistance R_s. It is easily shown that for an ideal MOS, reduction in the transconductance results, i.e.

$$G_m' = \frac{G_m}{1 + G_m R_s}, \tag{17}$$

where

$$G_m = \frac{\mu \epsilon_{ox} W (V_{gs} - V_{tx})}{L T_{ox}}. \tag{18}$$

The parasitic resistance becomes important only when

$$\frac{1}{R_s} \approx G_m, \tag{19}$$

or

$$V_{gs} - V_{tx} = \frac{L T_{ox}}{\mu \epsilon_{ox} W R_s}. \tag{20}$$

For typical MOS devices in integrated circuits, this voltage is of the order of 50 volts. Therefore, except in cases where long diffused interconnections are in series with the source electrode (resulting in very large series resistance), G_m saturation can be neglected.

A second saturation effect involves the decrease in inversion layer mobility due to diffuse surface scattering.[15] The result of this scattering is very similar to that of parasitic source and drain resistances, i.e., the flattening out of the G_m versus V_{gs} curve. Since no satisfactory theoretical expression is available to account for this dependence, the present model uses the empirical results of Ref. (16) leading to the following expression for the average effective mobility:

$$\mu_{eff} = \mu_{eff}(0) \left(\frac{E_{s0}}{E_s} \right)^r \tag{21}$$

where

$$E_{s0} = 6 \times 10^4 \text{ V/cm},$$

$$\mu_{eff}\,(0) = \text{constant (low-field) mobility for } E_s < E_{s0},$$

$$r = \text{an empirical constant},$$

$$\bar{E}_s = -(V_{gs} - V_{tx} - 0.5\,V_{ds})/T_{ox}; \qquad [22]$$

$\bar{E}_s$ is the average surface field. Typical curves of mobility variation in complementary MOS devices are given in Ref. (17). Values of $r = 0.12$ and $r = 0.20$ where necessary to match these curves for the n- and p-channel devices, respectively. The comparison between theoretical and measured curves for mobility variation is shown in Fig. 8.

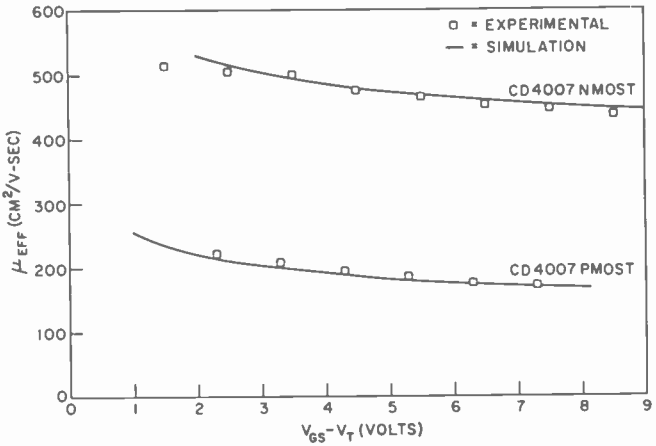


Fig. 8—Variation of effective mobility of CD4007 COS/MOS transistors.

This completes our discussion of the large-signal device model of an MOS device. Further evidence of the excellent correlation between theory and experiment is shown by the comparison of voltage and current transfer characteristics of a complementary MOS inverter in Fig. 9.

Voltage-Dependent Capacitances

In the absence of all stray capacitance, the fundamental speed limitation of an MOS device, the source–drain transit time, is well below

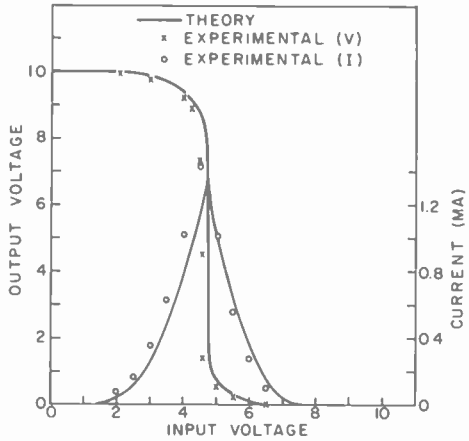


Fig. 9—Transfer characteristics of C4007 COS/MOS inverter.

1 ns and in fact is comparable to that of the fastest bipolar devices.[18] This transit-time calculation does account for the voltage dependent gate–channel capacitance. In practice, however, the speed of an integrated MOS device on bulk silicon is limited by the time to charge the many stray capacitances. These include wiring capacitance between interconnections and the bulk silicon substrate, gate-to-source and drain-overlap capacitance and, most importantly, the voltage-dependent junction capacitance of the source and drain electrodes. All these capacitances are indicated schematically in Fig. 10.

To point out the importance of these harmful stray capacitances, one need only compare the performance of bulk silicon MOS circuits with similar circuits in silicon-on-sapphire. Circuit delays approxi-

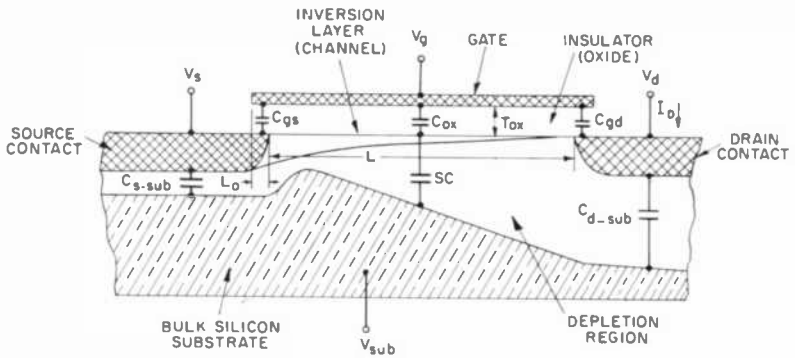


Fig. 10—Cross-section of bulk-silicon insulated-gate FET showing capacitances.

mately an order of magnitude smaller have been demonstrated[19,20] in complex silicon-on-sapphire MOS integrated circuits because nearly all stray capacitance is eliminated through the use of an insulating substrate and through diffusing the source and drain electrodes. The structure of complementary MOS devices in silicon-on-sapphire is shown in Fig. 11.

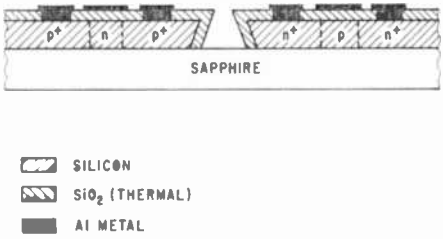


Fig. 11—Structure of CMOS devices in silicon on sapphire.

For the purpose of computer-aided transient circuit analysis, it would be desirable that all of the above capacitances be constant. This approximation greatly simplifies the numerical analysis involved by eliminating the need to invert the circuit capacitance matrix at each time step. This can be done successfully for MOS circuits[1] provided a suitable average value over the appropriate expected voltage variation is substituted for each voltage-dependent capacitor.

The expression for the junction capacitance depends on the doping profile between the source (drain) and the bulk material. If it is assumed that a step-junction exists, C_{sb} and C_{db} are of the form

$$C_{sb} = A\left(\frac{qN\epsilon_{si}}{2}\right)^{1/2}(V_{sb} + 2\Phi_f)^{-1/2}, \qquad [23]$$

$$C_{db} = A\left(\frac{qN\epsilon_{si}}{2}\right)^{1/2}(V_{db} + 2\Phi_f)^{-1/2}, \qquad [24]$$

where $2\Phi_f$ is approximately equal to the contact potential at the source (drain) contacts and A is the junction area. For a transient analysis, this expression should be averaged over the voltage variation, yielding

$$\overline{C} = \frac{1}{V_f - V_i}\int_{V_i}^{V_f} \frac{K_1}{V + 2\Phi_f}\,dV, \qquad [25]$$

or

$$\overline{C} = \frac{2K_1}{V_f - V_i}\left[(V_f + 2\Phi_f)^{1/2} - (V_i + 2\Phi_f)^{1/2}\right], \qquad [26]$$

where V_i and V_f are the initial and final potentials, respectively, of the contact-substrate junction and K_1 includes the remaining constant terms.

The active gate–channel capacitance of an MOS device is voltage dependent since it is the series combination of the geometrical oxide capacitance C_{ox} and the space-charge capacitance C_{sc} of the semiconductor channel. This distributed capacitance must be split up into two lumped capacitances, C_{gd} and C_{gs}. This can easily be done for an *ideal* MOS by first finding the total gate charge[18,21];

$$Q_g = \int_0^L Q(z)\,dz = \int_0^L C_{ox}V_{ox}(z)\,dz, \qquad [27]$$

and, therefore,

$$Q_g = \frac{2}{3}C_{ox}\left[\frac{(V_{gd} - V_t)^3}{(V_{gd} - V_t)^2 - (V_{gs} - V_t)^2} - \frac{(V_{gs} - V_t)^3}{(V_{gd} - V_t)^2 - (V_{gs} - V_t)^2}\right]. \qquad [28]$$

We may then define the two lumped capacitances

$$C_{gs} = \frac{\partial Q_g}{\partial V_{gs}}, \quad C_{gd} = \frac{\partial Q_g}{\partial V_{gd}}. \qquad [29]$$

The resulting relationships are

$$C_{gs} = \frac{2}{3}C_{ox}\left[1 - \frac{(V_{gd} - V_t)^2}{(V_{gs} - V_t + V_{gd} - V_t)^2}\right],$$

$$C_{gd} = \frac{2}{3}C_{ox}\left[1 - \frac{(V_{gs} - V_t)^2}{(V_{gs} - V_t + V_{gd} - V_t)^2}\right] \qquad [30]$$

A straight-line approximation for these capacitance variations is presented in Fig. 12. The actual constant average value to be utilized in any one transient analysis computation depends heavily upon how the particular device in question is expected to perform. Some engineering judgement is required for each individual application.

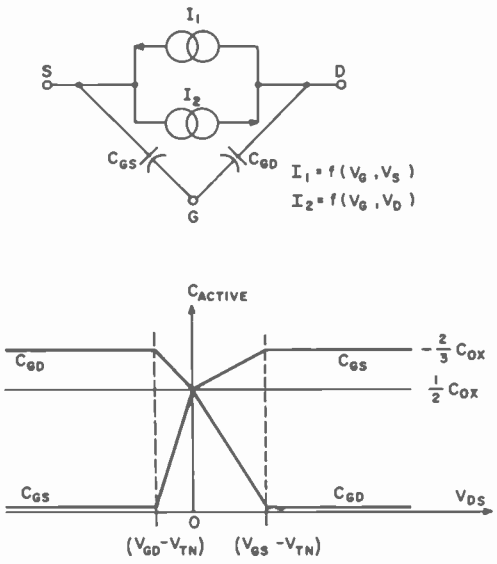


Fig. 12—Approximate gate capacitance variation of n-type FET.

Application to Transient Circuit Analysis

The flow diagram of an MOS computer simulation program called FETSIM[5] is shown in Fig. 13. The system is modular and makes extensive use of the RCA basic time-sharing system (BTSS). To begin, an interactive program called INPUT is used to enter the circuit description. This description includes topology, device and chip physical parameters, and node voltages, including any time-dependent input waveforms. The program allows the user to perform (as many times as necessary) dc analysis on-line. Convergence is found by using a pseudo Newton–Raphson iteration technique that assumes a diagonal nodal admittance matrix. This technique simplifies the numerical analysis involved considerably, but does depend on the user to make a good initial guess for the solution. Fortunately, this is usually possible for high-impedance MOS circuits and, in addition, the square-law

behavior of MOS devices (vs. exponential for bipolars) varies slowly enough to avoid overflow problems and to usually converge. The INPUT program can also plot various dc transfer characteristics on a graphic data plotter. For small enough increments in the abscissa, the above iteration technique is very adequate for on-line design usage. Various data can be altered and any of these analyses rerun to perform a worst-case analysis. When the user is satisfied with the circuit dc performance, a circuit description file (or pseudo tape) is created and stored.

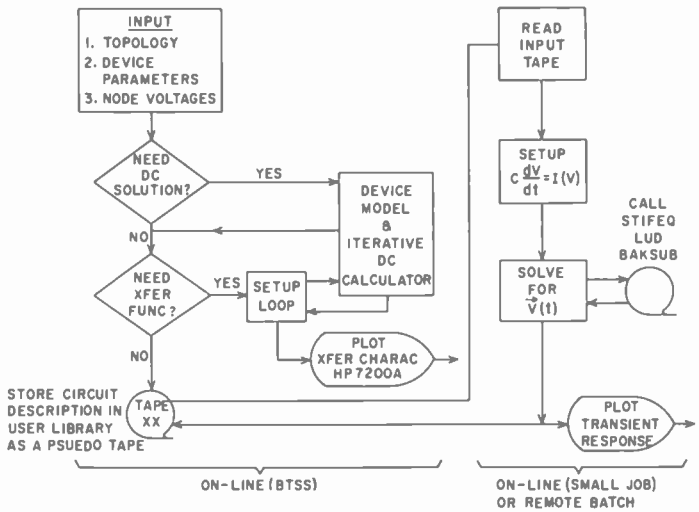
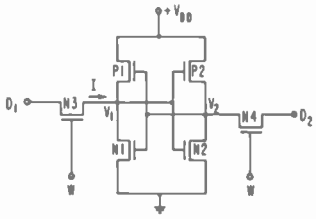


Fig. 13—Flow diagram of MOS circuit-simulation programs.

Transient analysis can be accomplished either on-line or in a remote-batch mode depending on the circuit complexity. In either case, the FETSIM program transforms the data into a set of n nodal equations

$$\frac{d\overline{V}}{dt} = \overline{C}^{-1}\overline{I}(V) = \overline{f}(v), \qquad [31]$$

where $\overline{C}^{-1}$ is the inverse of the nodal capacitance matrix, $\overline{V}$ is the vector of node voltages, and $\overline{I}$ is the vector of node currents (i.e., the sum of branch currents). These n ordinary differential equations are then integrated numerically. This system is said to be "stiff" when



IF W = + V_{DD}
D_1 = + V_{DD}, D_2 = 0, WRITE "0"
D_1 = 0, D_2 = + V_{DD}, WRITE "1"
D_1 = D_2 = + V_{DD}, READ
IF W = 0, CELL UNSELECTED

Fig. 14—Word-organized CMOS cell.

the eigenvalues of the Jacobian matrix

$$\overline{J} = \left(\frac{\partial f_i}{\partial V_j}\right) \qquad [32]$$

are widely separated in magnitude. In such a case, classical numerical integration methods have their step size limited by the largest such eigenvalue (smallest time constant) and thus require large amounts of computer execution time. This problem has attracted much attention recently, since the Jacobian matrix of large nonlinear electronic circuits is frequently stiff.[22] In the FETSIM system, matrix operations are accomplished by decomposition and back-substitution.[23] Numerical integration makes use of a class of PECE algorithms consisting of an application of a predictor, followed by application of one iteration of a pseudo Newton–Raphson method to a corrector. Such algorithms re-

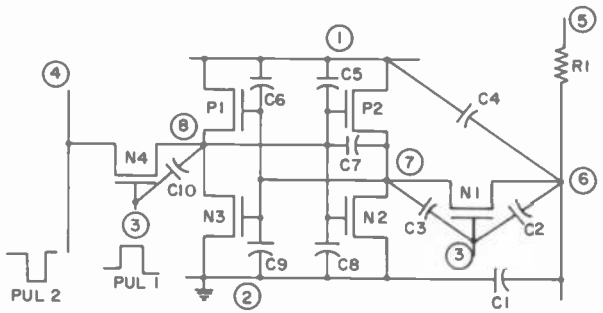


Fig. 15—Labeled schematic for circuit analysis.

quire precisely two evaluations of the derivative function for each forward step. The stability properties of such algorithms compare favorably with those obtained with application of the Newton–Raphson method to the corrector iterated to convergence.[24,25]

The output data from a transient simulation may be printed, plotted on-line, or simply stored on the circuit-description tape. The computed transient response of integrated MOS circuits compares favorably with experimental results.[1,4,5] As an example, consider the silicon-on-sapphire MOS memory cell in Fig. 14, which was designed and simulated using the FETSIM programs. To begin a design, the user must first make a proper drawing and label all components and nodes as shown in Fig. 15. This description is then entered via INPUT as shown in the example in Table 1.

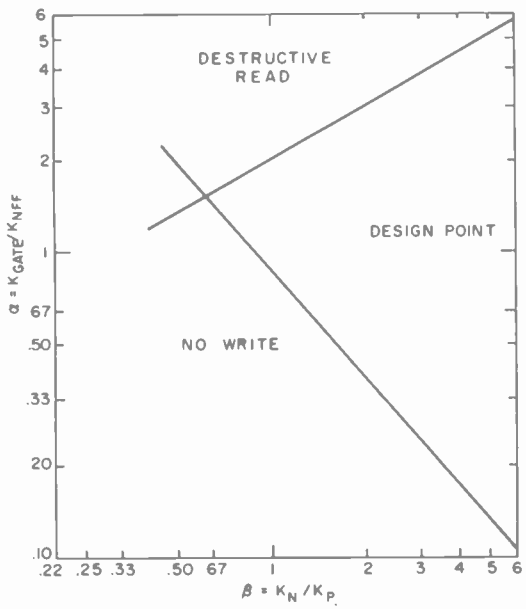


Fig. 16.—Memory-cell design plane.

Since all physical parameters are fixed by the standard fabrication process, the only design parameters for such a memory cell are the device geometries. To determine the optimum such design, numerous passes were made through the dc calculator to determine the worst-case limits of stable operation, i.e., destructive readout and no write. This information can best be presented as a dc design plane, as shown in Fig. 16. The two scales are the natural logarithms of relative device

Table 1—Example session with input program.

```
        /CODE INPUT
  2310 -

ARE YOU ENTERING A NEW CKT OR USING STORED TAPE ?   NEW

 TYPE # OF INDEPENDENT NODES,DEPENDENT NODES,RESISTORS,
CAPACITORS,PMOS,NMOS,PULSES

NI=5
ND=3
NR=1
NC=10
NP=2
NN=4
NPUL=2

 PMOS # =VT,CH.WIDTH(MILS),CH.LENGTH(MILS),D NODE,G NODE,S NODE

PMOS 1=-2,1,.3,1,7,8
PMOS 2=-2,1,.3,1,8,7

 NMOS # =VT,CH.WIDTH,CH.LENGTH,D NODE,G NODE,S NODE

NMOS 1=1.5,1,.3,6,3,7
NMOS 2=1.5,1,.3,7,8,2
NMOS 3=1.5,1,.3,8,7,2
NMOS 4=1.5,1,.3,8,3,4

 RESISTOR # =VALUE(KOHMS),NODE,NODE

R 1=.43,5,6

 CAPACITOR # =VALUE(PF),NODE,NODE

C 1=5,6,2
C 2=.135,6,3
C 3=.045,7,3
C 4=.09,6,1
C 5=.15,8,1
C 6=.06,7,1
C 7=.06,7,8
C 8=.06,8,2
C 9=.15,7,2
C10=.045,8,3

 ENTER CHIP PARAMS

FOX=3.8
FSI=11.7
TOX=1000
UEFFN=350
ACCEPT=3.E16
UEFFP=200
DONORS=2.E15
SLN=0
SLP=.01

 ENTER PULSE BASE,HEIGHT,DELAY(NS),WIDTH,TAU RISE,TAU FALL,#NOI

PULSE 1=0,10,6,20,2,2,3
PULSE 2=10,0,20,20,3,3,4
```

Table 1 (continued)

```
 ENTER NODE VOLTAGES
 NODE 1 MUST BE MOST +;NODE 2 MOST -

 V 1-V 8 = 10,0,0,10,10,8.5,1,9


 SOS ?   YES

ANY CORRECTIONS ? USE *--- OR RETURN

  2310

DO YOU NEED DC SOLUTION?   YES
CALC DONE IN   3  ITERATIONS
V(6)= .1000000E+02
V(7)= .3994340E-02
V(8)= .9999322E+01
  2310

DO YOU WANT A XFER FUNC ?   NO

USE RETURN TO WRITE TAPE FOR TRANSIENT RUN

  2310

WHAT TAPE# ?   60

 ENTER CONTROL INFO(NS)

TPRINT=1
TSTOP=50
TINTEG=1

TIME SHARE OR BATCH ?   TS
  2315
```

geometries. The actual design point was chosen to be well within the limits of operation. Once a stable circuit has been chosen the engineer is next interested in a transient analysis. The FETSIM simulation of the read operation of this cell is compared with the measured response of a typical test array in Fig. 17. The agreement is seen to be excellent; the small discrepancy in delay is due to the differences in the input waveforms. Further analyses would then be done to include decoding and sense circuitry before committing to the design of a larger array.

Summary

Despite the lack of large, widely distributed general-purpose program packages specifically for MOS, computer-aided circuit simulation of MOS integrated circuits is of widespread use and importance. The regularity of large MOS arrays and the accuracy of MOS models make design optimization in terms of physical device parameters and other

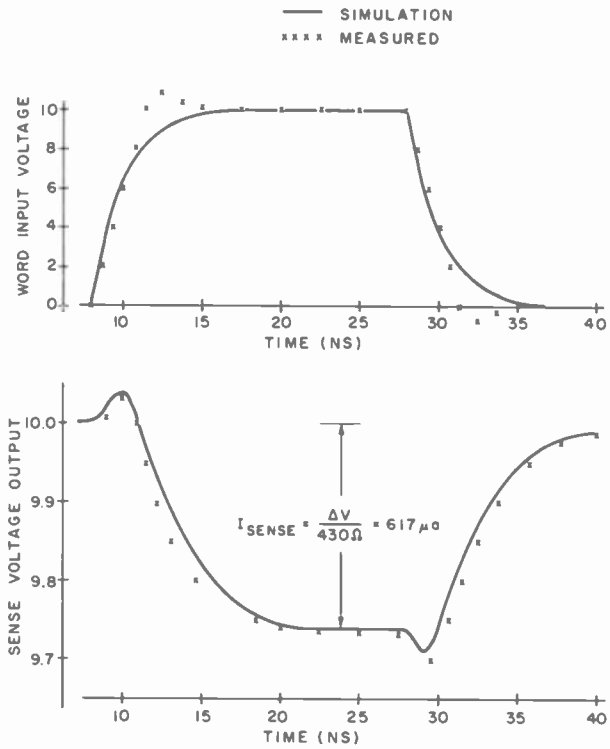


Fig. 17—Comparison of measured and computer simulated read access operation.

processing information a practical matter for the design engineer. Such programs with the inclusion of all parasitic resistance and capacitance give the designer a breadboard simulator. Typical computer execution time for an on-line transient analysis on an RCA Spectra 70/45 is 2-20 minutes. With further improved numerical analysis techniques, larger circuits can be done rapidly with time sharing, thus making such programs truly *design* aids rather than just *analysis* tools.

References:

[1] D. A. Hodges and H. Schichman, "Modeling and Simulation of Insulated-Gate Field-Effect Transistor Switching Circuits," **IEEE J. Solid State Circuits,** Vol. SC-3, p. 285, Sept. 1968.

[2] A. Feller, "Computer Analysis and Simulation of MOS Circuits," Digest of International Solid-State Circuits Conference, 1969.

[3] A. S. Grove, **Physics and Technology of Semiconductor Devices,** John Wiley and Sons, Inc., 1969.

[4] D. Frohman-Bentchkowsky and L. Vadasz, "Computer Aided Design and Characterization of Digital MOS Integrated Circuits," **IEEE J. Solid State Circuits,** Vol. SC-4, p. 57, April 1969.

[5] J. E. Meyer, "Computer-Aided MOS/SOS Design," **Proc. Computer-Aided System Design Seminar,** Cambridge, Mass., April 1969.

[6] C. T. Sah and H. C. Pao, "Effects of Fixed Bulk Charge on the Characteristics of Metal-Oxide-Semiconductor Transistors," **IEEE Trans. Electron Devices,** Vol. ED-13, p. 393, April 1966.

[7] J. J. Gibson, MOS Seminar Notes, Princeton Section IEEE, Nov. 1967.

[8] J. J. Gibson and A. W. Lo, unpublished notes.

[9] H. W. Mathers, S. R. Sedore, and J. R. Sents, "Automated Digital Computer Program for Determining Responses of Electronic Circuits to Transient Nuclear Radiation (SCEPTRE)," Vol. 1, IBM Space Guidance Center, Owego, N. Y., IBM File 66-928-611, Feb. 1967.

[10] H. K. J. Ihantola, "Design Theory of a Surface-Field-Effect Transistor," SEL Technical Report No. 1661-1, Stanford Electronics Laboratories, Stanford University, California, Sept. 1961.

[11] V. K. G. Reddi and C. T. Sah, "Source to Drain Resistance Beyond Pinch-Off in Metal-Oxide-Semiconductor Transistors (MOST)," **IEEE Trans. Electron Devices,** Vol. ED-12, p. 139, March 1965.

[12] D. Frohman-Bentchkowsky and A. S. Grove, "Conductance of MOS Transistors in Saturation," **IEEE Trans. Electron Devices,** Vol. ED-16, p. 108, Jan. 1969.

[13] F. P. Heiman, "Thin-Film Silicon-on-Sapphire Deep-Depletion MOS Transistors," **IEEE Trans. Electron Devices,** Vol. ED-13, p. 855, (1966).

[14] S. R. Hofstein and G. Warfield, "Carrier Mobility and Current Saturation in the MOS Transistor," **IEEE Trans. Electron Devices,** Vol. ED-12, p. 129, March 1965.

[15] J. R. Schrieffer, "Effective Carrier Mobility in Surface Space Charge Layers," **Phys. Rev.,** Vol. 97, p. 611, (1955).

[16] D. Frohman-Bentchkowsky, "On the Effect of Mobility Variation on MOS Device Characteristics," **Proc. IEEE** (Letters), Vol. 56, p. 217, Feb. 1968.

[17] E. J. Boleky, "The Performance of Complementary MOS Transistors on Insulating Substrates," **RCA Review,** Vol. 31, p. 372, June 1970.

[18] J. R. Burns, "High Frequency Characteristics of the Insulated Gate Field-Effect Transistor," **RCA Review,** Vol. 28, p. 390, Sept. 1967.

[19] J. R. Burns and J. H. Scott, "Silicon-on-Sapphire Complementary MOS Circuits for High Speed Associative Memory," **AFIPS Conference Proc.,** Vol. 35, Nov. 1969.

[20] J. E. Meyer, J. R. Burns, and J. H. Scott, "High Speed Silicon-on-Sapphire 50-Stage Shift Register," 1970 International Solid State Circuits Conference, Digest of Technical Papers, p. 200, Feb. 1970.

[21] C. T. Sah, "Characteristics of the Metal-Oxide-Semiconductor Transistor," **IEEE Trans. Electron Devices,** Vol. ED-11, p. 324, July 1964.

[22] F. H. Branin, Jr., "Computer Methods of Network Analysis," **Proc. IEEE,** Vol. 55, p. 1787, Nov. 1967.

[23] N. Sato and W. F. Tinney, "Techniques for Exploiting the Sparsity of the Network Admittance Matrix," **IEEE Trans. Power and Apparatus,** Vol. 82, p. 944, Dec. 1963.

[24] R. W. Klopfenstein and C. B. Davis, "PECE Algorithms for the Solution of Stiff Systems of Ordinary Differential Equations," **Mathematics of Computation** (to be published in Vol. 25, No. 115, July 1971).

[25] W. Liniger and R. Willoughby, "Efficient Integration Method for Stiff Systems of Ordinary Differential Equations," **SIAM Jour. Numerical Analysis,** Vol. 7, No. 1, p. 47, March 1970.

Radiochemical Study of Semiconductor Surface Contamination

III. Deposition of Trace Impurities On Germanium and Gallium Arsenide*

Werner Kern
RCA Laboratories, Princeton, N. J.

Abstract—The contamination of germanium and gallium arsenide surfaces by various trace impurities during device processing has been studied by radioactive-tracer techniques. Processing solutions include NaOH, KOH, HF, HCl, HNO_3, H_2SO_4, H_2O_2, $(NH_4)_2S_2O_8$, and mixtures of these reagents typically used for semiconductor etching. Radionuclides used to label reagent trace constituents are Au^{198}, Cu^{64}, Fe^{59}, Cr^{51}, Zn^{65}, and Mn^{54}. Desorption data for water at various temperatures are presented, and attempts to inhibit adsorption by adding chelating agents to the reagent solutions are briefly described. To permit comparison of the various impurities, the data in this summary have been normalized to a constant impurity concentration of 4 ppm (w/v) per impurity and are expressed in terms of atoms per cm^2 of geometrical surface area.

Ge wafers etched in mixtures of HF + HNO_3, with or without CH_3CO_2H or I_2, and containing just one specific radiotracer showed surface concentration levels as follows: 10^{13} — 10^{14} Au/cm^2, 10^{13} Cr/cm^2, 10^{12} Cu/cm^2, and less than 10^{11} Mn/cm^2. These quantities represent small fractions of one monoatomic layer coverage of the silicon. Gold gives the highest surface concentration because of the very large difference between the oxidation potential of Au ions and Ge and the corrosion resistance of metallic gold, which prevents redissolution of the metal deposit. For deposition from HF the situation is considerably more severe if the protective oxide layer is allowed to dissolve, resulting in uninhibited electrochemical displacement plating on the exposed Ge surface. Surface concentrations of 1×10^{16} Cu/cm^2, 1×10^{15} Au/cm^2, and 3×10^{14} Cr/cm^2 are obtained with this reagent labeled with the corresponding ions. Deposition from 5% NaOH under etching conditions (100°C) was 1×10^{16} Cu/cm^2, 4×10^{14} Fe/cm^2, and 3×10^{15} Au/cm^2. The structure and finish of the Ge surface (other than oxide layers), the doping type and concentration,

* Ed. Note: Part I, Adsorption of Reagent Components, and Part II, Deposition of Trace Impurities on Silicon and Silica, appeared in the June 1970 issue of **RCA Review.**

and the bath immersion time had negligible effects on the extent of the impurity deposition. Most of the deposits were uniformly distributed.

Attempts to inhibit impurity deposition on Ge by adding chelates to the reagent solutions were particularly successful for H_2O_2; EDTA and similar chelates resulted in a suppression of up to 2700 fold. Several chelating agents were able to remove deposited metals from the Ge surface. The most consistently excellent results were achieved by complexing dissolution in acidic H_2O_2. HCl solution was outstanding in its efficiency for desorbing Fe deposits.

Special etch mixtures used for GaAs wafers led to high levels of metal contamination for those solutions that contained alkali. H_2O_2 + NaOH etch caused the deposition of 2×10^{15} Cu/cm^2. Anodic electroetching in KOH solution led to 10^{15} Cu/cm^2 located predominantly on the n-type wafer surface. H_2O_2 + H_2SO_4 polishing etch led to a uniform surface concentration of 3×10^{13} Au/cm^2; H_2O_2 + HF led to similarly low levels. However, the Au and Cu deposits were strongly chemisorbed and could not be desorbed with boiling water.

1. Introduction

The results of radiochemical studies of *silicon* and *germanium* contamination by specific *components* of processing reagents were presented in Part I of this series.[1] Contamination of *silicon* and *silica* surfaces by inorganic *trace* contaminants in reagents commonly employed in the processing of semiconductor devices was described in Part II.[2] The present paper presents results on the contamination of *germanium* and *gallium arsenide* surfaces by inorganic *trace* contaminants in solutions. A survey of the literature covering the subject matter of all parts was presented in the introduction of Part I, which included a brief discussion of sources of contaminants and their effects on the electrical performance of devices.

In the present paper, the radionuclides used as tracers in the work described include Au^{198}, Cu^{64}, Fe^{59}, Cr^{51}, Zn^{65}, and Mn^{54}. Substrate materials include wafers of germanium and gallium arsenide. The concentration of specific impurities on these substrates was determined by conventional radiotracer techniques, and desorption tests were made to explore the nature of the bonding and the chemical form of the impurity on the surface and also to indicate means for its removal. Selected results of these investigations have been published previously.[3,4]

2. Experimental Methods

Important nuclear data for the radionuclides used as tracers, the preparation of the radioactive reagents, and a brief discussion of the

chemistry of the impurity ions were presented in Part II.[2] The preparation of the semiconductor substrates, the immersion treatment in the various etchants and reagent solutions, and the various techniques for the subsequent initial rinsing and the desorption treatments were all described in detail in Part I.[1] The radioactivity measurements and calculations, gamma spectrometry, autoradiographic analysis, and lifetime determination of the nuclides were also carried out by the techniques described there.

As before, the resulting surface concentrations of impurities discussed in this paper are expressed in terms of the number of atoms (indicated by the symbol for the element) per cm^2 of geometric surface area, without distinguishing between alternating chemical states of the deposited species.

3. Contamination by Gold

As in the case of silicon, gold in solutions, particularly in the presence of hydrofluoric acid, tends to deposit readily on germanium and III-V compound semiconductors, primarily by an electrochemical replacement mechanism leading to metallic gold deposits. Gold contamination of semiconductor surfaces gives rise to lifetime degradation of the semiconductor on subsequent heating. Contamination of germanium by gold is particularly critical because of its high diffusivity,[5] although the solid solubility is only 10^{15} atoms/cm^3. Similar considerations hold for gallium arsenide.

Adsorption data for germanium under both etching and nonetching conditions at gold ion concentrations of 3 to 16 ppm are presented in Table 1. Tests made with gallium arsenide as adsorbent are summarized in Table 2. As previously,[1] semiconductor etchants containing HF were quenched with deionized water to stop the etching before withdrawing the wafer. All samples were initially rinsed in 5 portions of 20 ml deionized water for 6 seconds per rinse, followed by an additional 30-second rinse sequence in acetone. The samples were then air dried, and the gamma radioactivity was determined by scintillation counting.

3.1 Deposition of Gold on Germanium Using Au^{198} as Tracer

Surface concentrations on germanium immersed in concentrated HF containing 6 ppm* of gold closely paralleled the surface concentrations

* All concentrations stated as "ppm" in this paper are on a weight/volume basis.

Table 1—Deposition of Gold on Germanium (Au^{198} Used as Radioactive Tracer)

Reagent	Gold Solution		Adsorbent Wafer			Immersion	Adsorbate	
Solution Composition	Au ppm	Au ions/cm³	Doping Type	Resistivity ohm-cm	Initial Surface	(Minutes at 23°C)	Au atoms/cm²	Distribution
49% HF	6.2	1.9×10^{16}	n	0.05	lapped	30	2.1×10^{15}	
49% HF + 10% CM*	5.6	1.7×10^{16}	n	0.05	lapped	30	1.1×10^{15}	Uniform
95 vol HF 49% + 5 vol HNO_3 70%	15.5	4.7×10^{16}	p	0.009	lapped/ polished	0.5	1.0×10^{15}	Uniform
95 vol HF 49% + 5 vol HNO_3 70% + 1% CM*	15.5	4.7×10^{16}	p	0.009	lapped/ polished	0.5	9.2×10^{14}	
95 vol HNO_3 70% + 5 vol HF 49%	15.5	4.7×10^{16}	p	0.009	lapped/ polished	0.5	8.0×10^{13}	
95 vol HNO_3 70% + 5 vol HF 49% + 1% CM*	15.5	4.7×10^{16}	p	0.009	lapped/ polished	0.5	8.0×10^{13}	Uniform
93 vol HNO_3 70% + 7 vol HF 49%	6.2	1.9×10^{16}	p	1.4	lapped	15	1.7×10^{14}	
93 vol HNO_3 70% + 7 vol HF 49% + 10% CM*	7.8	2.4×10^{16}	p	1.4	lapped	15	3.3×10^{13}	
Iodine Etch*	3.1	9.5×10^{15}	p	1.4	lapped	1.0	2.3×10^{13}	
Iodine Etch*	3.1	9.5×10^{15}	p	1.4	lapped	1.0	1.5×10^{13}	
Iodine Etch* + 10% CM*	2.8	8.6×10^{15}	p	1.4	lapped	1.0	2.2×10^{13}	
Iodine Etch*	6.2	1.9×10^{16}	n	0.04	lapped	1.0	7.8×10^{13}	
Iodine Etch* + 10% CM*	5.6	1.7×10^{16}	n	0.04	lapped	1.0	7.5×10^{13}	
Iodine Etch*	6.2	1.9×10^{16}	p	1.4	lapped	1.0	2.8×10^{13}	
15% H_2O_2	15.5	4.7×10^{16}	p	0.009	lapped/ polished	10**	1.3×10^{13}	
5% NaOH	10	3.0×10^{16}	n	0.05	lapped	1.0**	5.7×10^{15}	Uniform

* See Appendix for Compositions
** 95-100°C

found for silicon under the same conditions: 2×10^{15} Au/cm^2. Similar concentrations resulted from etch mixtures consisting of 95 vol HF and 5 vol HNO_3.

From germanium etch mixtures containing excess concentrated HNO_3, in which a fresh semiconductor surface is continuously being exposed to the etchant, considerably less adsorption occurred than from HF solutions of comparable gold concentration. For example, an etching mixture consisting of 7 vol HF and 93 vol HNO_3 and containing 6 ppm of gold ions led to a germanium surface concentration of 2×10^{14} Au/cm^2. An etch mixture of 5 vol HF and 95 vol HNO_3 resulted in slightly less adsorption (8×10^{13} Au/cm^2) than observed for the 7:93 mixture, even though the gold concentration in the solution had been increased to 16 ppm. Etching of germanium is therefore preferably done with mixtures containing a minimum of HF to minimize gold contamination.

Germanium wafers etched in iodine etch containing 3 to 6 ppm of gold ions led to surface concentrations similar to those observed for silicon under the same conditions:[2] (2 to 8) $\times 10^{13}$ Au/cm^2. This represents a 0.01 to 0.05 monoatomic layer of gold. No significant differences in contamination levels were detected for various resistivity types and different resistivity levels, which ranged from 0.009 to 1.5 ohm-cm.

Hot NaOH solution containing 10 ppm of gold deposited uniformly the large quantity of 6×10^{15} Au/cm^2 on the germanium surface, similar to the level previously observed for silicon.[2]

Addition of chelating agents to the reagent solutions did not substantially decrease the contamination. In fact, EDTA added to hot 15% H_2O_2 increased the surface concentration. Autoradiography of this wafer revealed the presence of many sharply localized "pin-points" of high radioactivity.

Desorption of gold from germanium surfaces was found to be extremely difficult, since gold ions usually deposit on germanium in metallic form that requires removal by strong oxidizing agents followed by complexing to prevent its re-adsorption, but the germanium surface may etch under these conditions. The hot HCl-H_2O_2-H_2O that was found to be very effective for desorbing gold from silicon[2,6] may serve as an example. However, some degree of purification can be attained with this agent even at 23°C. Gold deposited from HF-containing reagents could be removed to the extent of 70 to 80% in 15 minutes treatment. As would be expected, deionized water at 95°C is virtually ineffective.

3.2 Deposition of Gold on Gallium Arsenide Using Au^{198} as Tracer

Gold deposited on gallium arsenide wafers during polish-etching in H_2O-H_2O_2-H_2SO_4 mixture by strong and irreversible adsorption. Gold solution concentrations of 0.11, 1.1, and 11 ppm were tested with three wafers for each concentration. The wafers were (100)-oriented gallium arsenide crystals that had been etched in non-radioactive solution of the same composition. Etching in the radioactive-gold-containing solutions was conducted with rotation at 23°C for 5 minutes using 20 ml etchant per wafer. The etching was terminated by rapidly diluting the etch

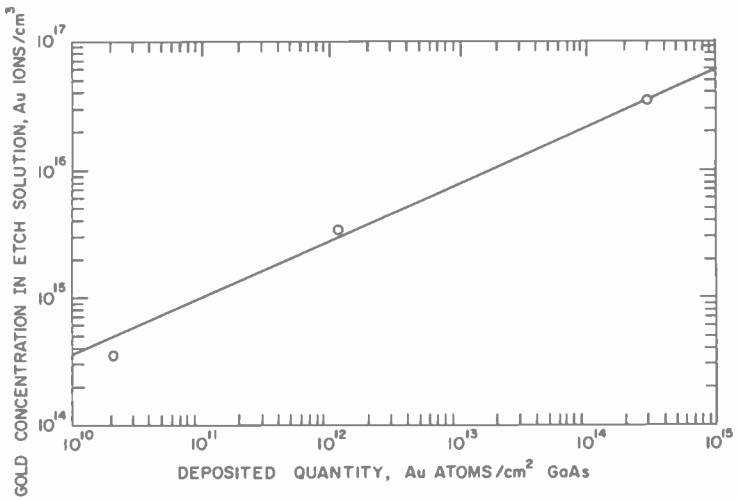


Fig. 1—Quantity of gold deposited on GaAs from H_2SO_4-H_2O_2-H_2O etchant as a function of gold solution concentration. Polishing etch (5 vol 98% H_2SO_4 + 1 vol 30% H_2O_2 + 1 vol H_2O) containing Au^{198} as radioactive tracer; (100)-oriented GaAs as wafers used as substrate. Each data point is the average of measurements from 3 wafers.

solution with deionized water followed by rinsing of the wafers for 1 minute in flowing deionized water. The resulting gold surface concentrations are shown in Table 2 and are presented graphically in Fig. 1 as a function of gold solution concentration. As observed previously for other metal ions absorbed by silicon[1,2] and III-V compounds,[7] the logarithm of the quantity of atoms deposited on the semiconductor surface is linearly proportional to the logarithm of the metal ion concentration in the solution. As in the case of silicon and germanium, the mechanism of gold deposition from acid solutions on

Table 2—Distribution of Gold on Gallium Arsenide* (Au^{198} Used as Radioactive Tracer)

Reagent	Gold Solution		Immersion	Adsorbate	
Solution	Au ppm (added)	Au ions/cm^3 (added)	(Min. at 23°C)	Au atoms/cm^2	Distribution
1 vol H_2O + 1 vol H_2O_2 30% + 5 vol H_2O_4 98%	0.11	3.4×10^{14}	5.0	1.91×10^{10} 2.19×10^{10} 2.19×10^{10} Average 2.1×10^{10}	uniform uniform uniform
1 vol H_2O + 1 vol H_2O_2 30% + 5 vol H_2O_4 98%	1.1	3.4×10^{15}	5.0	1.24×10^{12} 1.24×10^{12} 1.16×10^{12} Average 1.2×10^{12}	uniform uniform uniform
1 vol H_2O + 1 vol H_2O_2 30% + 5 vol H_2O_4 98%	11.0	3.4×10^{16}	5.0	2.85×10^{14} 3.04×10^{14} 3.21×10^{14} Average 3.0×10^{14}	uniform uniform uniform

* (100)-oriented GaAs wafers chemically pre-etched in nonradioactive solution of the same composition.

III-V compounds is an electrochemical process resulting in the deposition of elemental gold.

Adsorption of gold on gallium arsenide from hydrogen peroxide and from ammonium fluoride buffered hydrofluoric acid (oxide etch) is also very strong. It is possible to specifically purify these non-etching reagents by exposing them to a large surface of gallium arsenide using the techniques previously demonstrated with silicon.[3] Solutions containing 11 ppm of Au^{198} ions were used with crushed gallium arsenide crystal pieces as adsorbent to measure the purification effectiveness. Stirring of 50 ml 30% H_2O_2 with 10 g of GaAs reduced to gold concentration by 93% in 3 minutes, by 97% in 15 minutes, and by 99.3% in 30 minutes. Passing a 50-ml batch once through a column (1 cm wide, 8 cm high) of the GaAs pieces resulted in elimination of 99.7% of the Au^{198}. Buffered HF etch processed through a similar column resulted in 99.93% removal of the gold.

4. Contamination by Copper

The degrading influence of copper on the lifetime of semiconductor devices and the thermal conversion effect of this fast-diffusing metal in semiconductor crystals are well known. Electronic effects of trace quantities of copper on germanium surfaces have been studied in some detail.[8-10] Contamination tests with Cu^{64} have been carried out for germanium[8,11] and III-V compound semiconductors.[7] The work presented in this section summarizes results on the deposition of copper from many typical reagent solutions, inhibition of adsorption by chelating additives, and desorption of copper deposits with complexing and chelating agents.

The etching and initial rinsing treatments of the germanium wafers were conducted as described for gold. The results for germanium have been summarized in Tables 3 to 5, and for gallium arsenide in Table 6.

4.1 Deposition of Copper on Germanium Using Cu^{64} as Tracer

Germanium wafers etched in typical mixtures containing HF, HNO_3, CH_3CO_2H, iodine, and 2 ppm of radioactive copper retained only 4×10^{11} copper atoms per cm^2 (Table 3). This is the equivalent to a 0.0002 monoatomic layer of copper.

A 4:8:9 etch mixture of HF-HNO_3-CH_3CO_2H containing 2 ppm of Cu^{64} led to 3×10^{12} Cu/cm^2. The copper was uniformly distributed on the germanium surface except for a few scattered dots of high concentration.

Table 3—Deposition of Copper on Germanium (Cu^{64} Used as Radioactive Tracer)

Reagent	Copper Solution		Immersion		Adsorbate	
Solution	Cu ppm (added)	Cu ions/cm^3 (added)	Duration (Min.)	Temp. (°C)	Cu atoms/cm^2	Distribution
49% HF	2.1	2.0×10^{16}	50	23	3.5×10^{15}*	uniform
49% HF	2.1	2.0×10^{16}	50	23	2.9×10^{15}*	but patchy
49% HF	59	5.6×10^{17}	30	23	1.3×10^{17}	
4 vol HF 49% + 8 vol HNO_3 70% + 9 vol CH_3CO_2H	2.1	2.0×10^{16}	1.0	23	3.0×10^{12}	uniform plus a few dots
4 vol HF 49% + 8 vol HNO_3 70% + 9 vol CH_3CO_2H	2.1	2.0×10^{16}	1.0	23	3.4×10^{12}	uniform plus a few dots
7 vol HF 49% + 93 vol HNO_3 70%	29	2.8×10^{17}	0.50	23	1.0×10^{13}	
7 vol HF 49% + 93 vol HNO_3 70%	29	2.8×10^{17}	1.0	23	3.0×10^{13}	
7 vol HF 49% + 93 vol HNO_3 70%	29	2.8×10^{17}	5.0	23	1.6×10^{13}	
95 vol HF 49% + 5 vol HNO_3 70%	59	5.6×10^{17}	1.7	23	1.3×10^{14}	
Iodine Etch**	2.1	2.0×10^{16}	1.0	23	3.9×10^{11}	
Iodine Etch**	2.1	2.0×10^{16}	1.0	23	4.2×10^{11}	
5% NaOH	2.1	2.0×10^{16}	1.0	100	4.6×10^{15}	uniform
15% H_2O_2	109	1.1×10^{18}	10	95	2.8×10^{16}	

* Solution/surface ratio only 0.5 cm^3/cm^2 Ge which accounts for the relatively low values.

** See Appendix for composition.

Lapped germanium wafers used as adsorbent.

Deposition of copper from a 7:93 HF-HNO_3 mixture to which 29 ppm of labeled copper ions was added was studied as a function of 0.5 to 5 minutes etch time showed etch time independence, the values ranging from (1 to 3) $\times 10^{13}$ Cu/cm^2; this is within the normally observed fluctuation range for etching conditions. As predicted, an HF-rich etch, 95:5 HF-HNO_3, led to a larger quantity of deposited copper. Compared with the deposition of gold under similar conditions, copper deposits to a lesser extent; this is probably due to the fact that copper, in contrast to gold, is being redissolved by the extremely corrosive etch components.

Germanium immersed in 49% HF that had been contaminated with 2 ppm of radioactive Cu(II) ions caused the deposition of at least 4×10^{15} Cu/cm^2.* Autoradiographic analysis showed that the copper deposits were somewhat patchy, but were uniform in the sense that they did not show any accumulated spots of localized high concentration. As in the case of silicon,[2] copper contamination from HF solutions proceeds by electrochemical deposition leading to metallic copper, even at low solution concentrations. The relationship of the copper concentration in solution and the resulting quantity on the germanium surface appears to be similar to that presented for silicon.

Contamination of germanium by copper during etching with hot NaOH solution was studied using 5% NaOH to which radioactive cupric ions had been added. Because of the trace concentrations involved, the slightly soluble copper(II) hydroxide, $Cu(OH)_2$, appeared to be completely dissolved in the etchant. Solutions containing 2 ppm of Cu^{64} led to 5×10^{15} Cu/cm^2.

Table 4 summarizes results of inhibited copper deposition on germanium from various reagents. Large quantities (up to 109 ppm) of radioactive Cu(II) ions were added to these solutions, which contained various concentrations of chelating and complexing additives, to study the effects of inhibition under severe contamination conditions. The addition of 1.4% EDTA-Na_2 to 15% H_2O_2 containing Cu^{64} decreased copper deposition by a factor of 200. A formulation developed by the author (see Appendix), a special chelating–surfactant mixture at a concentration of 1%, suppressed deposition by a factor of 250. Inhibi-

* As indicated in Table 3 the ratios of the HF solution volume to germanium surface area in a few samples were such that the germanium depleted the copper in solution leading to relatively low quantities of adsorbed copper. Because of this stripping effect, accurate determinations of copper deposition require a large ratio of copper-in-solution to copper-being-deposited. This ideal condition could not be attained in the present study; the stated Cu/cm^2 values for HF should therefore be considered minimum values.

Table 4—Inhibited Deposition of Copper on Germanium (Cu^{64} Used as Radioactive Tracer)

Reagent		Copper Solution		Immersion		Adsorbate
Solution Composition	Inhibitor Additive	Cu ppm (added)	Cu ions/cm^3 (added)	Duration (Min.)	Temp. (°C)	Cu atoms/cm^2
15% H_2O_2	none	109	1.1×10^{18}	10	95	2.8×10^{16}
15% H_2O_2	1% CM*	108	1.1×10^{18}	10	95	1.1×10^{13}
15% H_2O_2	0.05M EDTA-Na_2**	109	1.1×10^{18}	10	95	1.4×10^{14}
15% H_2O_2	0.01M EDTA-Na_2**	11	1.1×10^{17}	10	95	8.8×10^{12}
49% HF	none	59	5.6×10^{17}	30	23	1.3×10^{17}
49% HF	1% CM*	58	5.5×10^{17}	30	23	3.0×10^{16}
49% HF	10% CM*	53	4.9×10^{17}	30	23	4.9×10^{13}
49% HF	1% EDTA-Na_2**	58	5.5×10^{17}	30	23	1.1×10^{17}
7 vol HF 49% + 93 vol HNO_3 70%	none	29	2.8×10^{17}	1.0	23	3.0×10^{13}
7 vol HF 49% + 93 vol HNO_3 70%	1% HCO_2H	29	2.8×10^{17}	1.0	23	2.6×10^{13}
7 vol HF 49% + 93 vol HNO_3 70%	2% HCO_2H	29	2.8×10^{17}	1.0	23	1.5×10^{13}
7 vol HF 49% + 93 vol HNO_3 70%	4% HCO_2H	28	2.7×10^{17}	1.0	23	1.4×10^{13}

* See Appendix for composition.
** Ethylene diamine tetra acetic acid disodium salt.
Lapped germanium wafers used as adsorbent.

tion of deposition by a factor of 2700 from concentrated HF was achieved by use of this chelating mixture, but a large amount (10%) of additive was required. EDTA-Na_2 in HF had no effect, and neither did the addition of formic acid to HF-HNO_3 etchant from which surface concentrations in the range from $(1.4\text{-}3.0) \times 10^{13}$ Cu/cm^2 were measured.

4.2 Desorption of Copper Deposited on Germanium from Cu^{64}-Containing HF Solutions

Germanium wafers that had been immersed in the radioactive 49% HF as described in the preceding sections were used for systematic desorption studies. Two initial concentration levels of deposited copper were chosen, a high concentration of 10^{17} and a low one of 10^{13} Cu/cm^2. The samples were radiocounted, immersed in the desorbing solutions for the time and at the temperature specified in Table 5, dipped in acetone to facilitate drying, and recounted. The residual Cu^{64} surface concentrations were calculated as a percent of the initial value and are listed in Table 5. The most efficient formulations are those based on metal complexing with hot acidic hydrogen peroxide mixtures with which copper desorption of better than 99% in 5 minutes was achieved. A more detailed discussion of these systems primarily for use in silicon device technology has been presented in a separate publication.[6] There is evidence from the germanium data that these solutions are even effective at room temperature so that possible etching of the germanium surface can be avoided. This is especially interesting in the light of the fact that the removal of copper from germanium surfaces with other desorbing agents was generally much less effective than that from silicon.[2]

4.3 Deposition of Copper on Gallium Arsenide Using Cu^{64} Tracer

The extent of contamination of gallium arsenide wafers by copper from three frequently used etchants was determined in the following manner: 5 ppm of copper labeled with Cu^{64} was added to each of the etch solutions. Following etching, each sample was initially given 5 water rinses and then one acetone rinse, each of 6 seconds duration. After this, gamma scintillation counting and autoradiographic analyses on Kodak-Type AA x-ray film were conducted. A summary of the data is presented in Table 6.

The first etchant tested was the polishing etch consisting of HF and H_2O_2. Two lapped n-type GaAs wafers with a carrier concentration

Table 5—Desorption of Copper from Germanium With Various Complexing and Chelating Solutions

Desorbing Solution	Temp. (°C)	Ge Sample Type	Cu^{64} Residue (%) After Desorbing		
			1 min.	5 min.	10 min.
3% Hydrogen peroxide + 10% Formic acid	90	A	0.10	<0.1	
	23	A	0.10	<0.1	
		D		11	2.2
3% Hydrogen peroxide + 10% Hydrochloric acid (1N)	95	A	0.03		
3% Hydrogen peroxide + 10% Acetic acid	95	A	0.15	<0.1	
10% Formic acid	90	D		12	2.8

Copper Concentration Levels:

Copper deposited from Cu^{64}-containing HF etchants resulting in the following initial (100%) surface concentrations:

Samples A: $1\text{-}2 \times 10^{17}$ Cu atoms/cm^2

D: $1\text{-}3 \times 10^{13}$ Cu atoms/cm^2

Reagent Solutions:

All solutions are in deionized and distilled water unless stated otherwise.

Ineffective Agents:

The following reagents (excluded from the table) were found ineffective or poorly effective for desorbing copper:

Acetone (23°)
D.I. and distilled water (23°, 100°)
3% Hydrogen peroxide (23°, 90°)
Sat. Ethylene diamine tetra acetic acid (90°)
0.1M Ethylene diamine tetra acetic acid disodium salt (23°, 90°)
0.1M Oxalic acid (90°)

of about $10^{16}/cm^3$ were used. Etching was performed with 50 ml per wafer at 23°C for 5 minutes, followed by quenching with 300 ml deionized water. The resulting copper concentration on the surfaces amounted to only $(2\text{-}3) \times 10^{13}$ Cu/cm^2. Consequently, the autoradiograms were extremely weak but indicated an apparently uniform distribution of the copper. Rinsing in deionized water at 95°C had negligible effects, decreasing the concentration to 92% of the initial value in 15 minutes.

The second etchant was an aqueous solution of 5% H_2O_2 and 0.83% NaOH, used as a crystallographic orientation etch. Etching was carried out with 50 ml solution per wafer at 70°C for 30 seconds, followed by quenching with deionized water. The same type of wafers were used as above. The initial surface concentration measured $(1\text{-}2) \times 10^{15}$ Cu/cm^2 and again was distributed uniformly on both sides of the wafers. Rinsing in deionized water at 95°C lowered these levels only slightly to 93% in 5 minutes, and to 92% in 15 minutes, indicating a strongly chemisorbed copper adsorbate.

The third etch system was an electrolytic etch employing agitated 16% KOH as the electrolyte. Four GaAs wafers were tested. The substrate concentration of these wafers was $1.4 \times 10^{16}/cm^3$ n-type. One side of the wafer was polished and doped p-type by an Mn diffusion. The n-type reverse side was lapped. The samples were made the anode and were illuminated on one side during etching with a beam of light from a projector lamp. Five or ten 2-ampere pulses were applied, as noted in Table 6. A constant bias current of a few milliamperes was maintained during the etching and first water rinse.

The concentrations measured on these anodically etched wafers were all close to 1×10^{15} Cu/cm^2, if averaged for the total geometric wafer area. Removal and analysis of GaAs surface layers employing abrasion techniques coupled with autoradiographic analysis and radioactivity counting showed that the predominant quantity of the copper was deposited fairly uniformly on the lapped n-type back surface of the wafers. The polished p-type wafer surfaces which had been illuminated during the electro-etching exhibited various degrees of nonuniformity in the distribution of the much lower concentration Cu^{64} deposits. It is remarkable how much copper actually deposits during this type of anodic etching, considering that the positively charged copper ions tend to be repelled from the anode.

Contamination of GaAs in 1N HNO_3 containing 5.6 ppm copper(II) ions is reported[7] to result in 1.4×10^{14} Cu/cm^2. This value is intermediate between the present results for the HF-H_2O_2 and the NaOH-H_2O_2 etches.

Table 6—Deposition of Copper on Gallium Arsenide (Cu^{64} Used as Radioactive Tracer)

Reagent	Copper Solution		Immersion		Adsorbate	
Solution Composition	Cu ppm (added)	Cu ions/cm^3 (added)	Duration (Min.)	Temp. (°C)	Cu atoms/cm^2	Distribution
2 vol HF 49% + 11 vol H_2O_2 30%	5.0	4.7×10^{16}	5.0	23	2.4×10^{13}	nonuniform
2 vol HF 49% + 11 vol H_2O_2 30%	5.0	4.7×10^{16}	5.0	23	3.1×10^{13}	nonuniform
1 vol H_2O_2 30% + 5 vol NaOH 5%	5.0	4.7×10^{16}	0.50	70	1.3×10^{15}	both wafer sides uniform
1 vol H_2O_2 30% + 5 vol NaOH 5%	5.0	4.7×10^{16}	0.50	70	2.2×10^{15}	both wafer sides uniform
16% KOH, Electrolytic; 5 pulses of 2 A for 3 sec per pulse	5.0	4.7×10^{16}	0.25	25	8.7×10^{14}	p-side nonuniform uniform n-side
	5.0	4.7×10^{16}	0.25	25	1.2×10^{15}	nonuniform p-side
	5.0	4.7×10^{16}	0.25	25	1.3×10^{15}	nonuniform p-side
16% KOH, Electrolytic; 10 pulses of 2 A for 3 sec per pulse	5.0	4.7×10^{16}	0.50	25	1.3×10^{15}	uniform n-side slightly nonuniform p-side

It can be concluded that copper from various reagent solution deposits on GaAs by the same type of electrochemical mechanism as has been observed for silicon and germanium, leading to irreversibly chemisorbed copper that is difficult to remove without attacking the GaAs surfaces.

5. Contamination by Iron

The occurrence of iron in "analytically pure" etch reagents is widespread. Typical iron concentrations in various reagents were listed in Part II[2] where it was pointed out that sodium and potassium hydroxide may contain up to 10 ppm of iron. Since sodium hydroxide solution is frequently used for etching of germanium and gallium arsenide wafers and devices, the present investigation is concerned primarily with iron contamination from this source.

Small aliquots of ferric chloride-HCl solution containing $Fe^{59}Cl_3$ were added to either HCl or NaOH solutions to yield total iron ion concentrations of 0.01 to 4 ppm. The trace quantities of the acido-complex ion $(FeCl_4)^-$ added were essentially soluble in the NaOH, forming sodium ferrate(III), $NaFeO_2$. Nevertheless, the preparations were centrifuged before use to remove any possibly present hydroxide precipitates.

The germanium samples used as adsorbent were (111) oriented lapped p-type wafers of 1 to 10 ohm-cm resistivity. Following the etching treatments, the wafers were given an initial rinse sequence in 3 beakers of 20 ml deionized water each, followed by a methyl alcohol rinse; each rinse duration was 5 seconds. The results are presented in Table 7.

5.1 Deposition of Iron on Germanium Using Fe^{59} as Tracer

Wafers immersed at 23°C in 19% HCl containing 4 ppm of radio-iron adsorbed only 6×10^{10} Fe/cm^2, 5 times less than silicon adsorbed under the same conditions.[2] Germanium powder equilibrated hot and at room temperature with acid of the same composition followed by one acid rinse and three water rinses was completely free of iron.

Germanium wafers etched for 60 seconds in 5% NaOH solutions at 100°C were contaminated with iron to a lesser extent (approximately 10 times) than silicon. Uniform surface distributions were found by autoradiography, but the data relating iron surface concentration with iron solution concentration fluctuated considerably, similar to those for silicon.[2]

Table 7—Deposition of Iron on Germanium* (Fe^{59} Used as Radioactive Tracer)

Reagent		Iron Solution		Immersion		Adsorbate	
Solution		Fe ppm (added)	Fe ions/cm^3 (added)	Duration, (Min.)	Temperature, (°C)	Fe atoms/cm^2	Distribution
6N	HCl	3.7	4.0×10^{15}	30	23	5.6×10^{10}	
0.0125N	NaOH	0.010	1.1×10^{14}	1.0	100	1.3×10^{13}	
0.125N	NaOH	0.15	1.6×10^{15}	1.0	100	7.8×10^{13}	
1.25N	NaOH	0.088	9.5×10^{14}	1.0	100	6.6×10^{12}	uniform
1.25N	NaOH	0.10	1.1×10^{15}	1.0	100	3.2×10^{12}	
1.25N	NaOH	1.0	1.1×10^{16}	1.0	100	8.6×10^{13}	uniform

* All absorbents are lapped p-type wafers of approx. 1-10 ohm-cm resistivity.

Water and HCl desorption rates, and the residual iron surface concentrations on germanium, were similar to those observed for silicon. Comparative desorption curves for both semiconductors were given in Fig. 7, Part II,[2] where it is seen that deionized water at 23°C has very little effect, whereas 19% HCl is capable of desorbing the iron adsorbates within one minute of rinsing to 2×10^{12} Fe ions per cm^2 germanium (1×10^{12} Fe/cm^2 for silicon). The mechanism of iron deposition from hot NaOH solution on germanium appears to proceed by chemisorption of the $(FeO_2)^-$ or by electrochemical deposition as metallic iron, as discussed for silicon.[2]

Adsorption of iron by germanium during etching in HF-HNO_3 mixtures was not tested but would be expected to be similar to the value found for silicon,[2] which is less than 2×10^{10} Fe/cm^2. A literature reference[11] for an etch mixture consisting of 5 vol HF, 8 vol HNO_3, and 15 vol CH_3CO_2H with the same iron concentration (4 ppm) cites a value of $<7 \times 10^{10}$ Fe/cm^2, in good agreement with this conjecture.

6. Contamination by Chromium Ions

The adsorption of commonly occurring chromium ions by germanium wafers from solutions of HF and etch mixtures of HF-HNO_3 with and without iodine was examined for a chromium ion concentration of 2.5 ppm, with Cr^{51} serving as the labeling species. For the acid etch mixtures this concentration represents a 12- to 25-fold excess over the chromium contamination normally present.[2] All wafers were (111) oriented, p-type, and of 3 to 9 ohm-cm resistivity. Following removal from the radioactive baths all samples were rinsed initially by the usual 5 rinses of deionized water of 20 ml per wafer, 6 seconds per rinse. This was followed by a 6-second dip in acetone to facilitate drying before commencing the radioactivity measurements and autoradiographic analysis. A summary of the experimental data is presented in Table 8.

6.1 Deposition of Chromium on Germanium Using Cr^{51} as Tracer

The chromium in HF solutions is likely to occur as a fluoride hydrate complex that may be adsorbed by the germanium as the molecular species rather than as metallic chromium.

Germanium wafers immersed in concentrated HF for 30 minutes attained an average concentration of 2×10^{14} Cr/cm^2. The adsorbates were uniformly distributed on the surface. Treatment with a solution of the disodium salt of ethylene diamine tetra acetic acid at 23°C

Table 8—Deposition of Chromium on Silicon and Germanium (Cr^{51} Used as Radioactive Tracer)

Reagent	Adsorbent Wafer			Immersion		Adsorbate	
Solution Composition*	Material**	Initial Surface	Wafers Used	Duration (Min.)	Temp. (°C)	Cr atoms/cm²	Distribution
49% HF + Cr^{51}	Ge	lapped/ polished	4	30	23	$1.59 \pm 0.70 \times 10^{14}$	uniform
5 vol HF 49% + 95 vol HNO_3 70%	Ge	lapped	1	1.7	23	2.4×10^{13}	uniform
Iodine Etch† + Cr^{51}	Ge	lapped	1	1.0	23	4.5×10^{12}	uniform

* All solutions contain an added 2.3 ppm Cr; total Cr concentration is 3×10^{16} Cr ions/cm³.
** p-type Ge, 3-5 ohm-cm resistivity.
† See Appendix for Compositions.

desorbed 75% of the layer in 30 minutes; deionized and distilled water at 95°C without additives removed 77% of the chromium in the same period. A mixture of 4 vol 6% H_2O_2 and 1 vol 90% formic acid at 23°C was much more effective, desorbing 85% in 30 minutes (elevated temperatures were not used since the germanium might then be etched). These data indicate that whereas the quantities of chromium adsorbed by germanium are similar to those for silicon,[2] the bonding appears to be weaker.

A germanium wafer etched for 100 seconds in $HF\text{-}HNO_3$ followed by rapid dilution with water had acquired a surface concentration corresponding to 2×10^{13} Cr/cm^2. Again, this deposit was very effectively desorbed by acidic hydrogen peroxide at 23°C; over 95% was removed in only 5 minutes.

Etching for 60 seconds in iodine etch followed by rapid dilution with water led to weakly adsorbed layers of 5×10^{12} Cr/cm^2, of which half could be removed by rinsing with cold water.

Adsorption of chromium during etching of germanium in NaOH or KOH solutions was not tested experimentally, but a high degree of contamination can be predicted since addition of OH^- ions to a solution of Cr^{+3} ions tends to precipitate chromic oxide hydrate, $Cr_2O_3 \cdot H_2O$. Aqueous HCl should effectively desorb the deposits since it forms soluble salts, similar to the reaction with iron adsorbates discussed in the preceding section.

7. Contamination by Zinc

The adsorption by zinc ions from selected etch solutions was investigated for gallium arsenide as adsorbent. This work was done in connection with a study on zinc diffusion where the degree of adsorption of zinc during the etch removal of doped layers had to be ascertained.[12] The mechanism and extent of zinc deposition may serve as an example of a divalent metallic ion that exists in solution only with +2 oxidation state, and as will be shown, exhibits weak and reversible adsorption. This type of physical adsorption behavior might be expected from the great tendency of zinc to form relatively stable complex ions. The results are summarized as follows.

7.1 Deposition of Zinc on Gallium Arsenide Using Zn^{65} as Tracer

Aqua regia solutions containing additions of 0.001, 0.01 and 0.1 ppm of zinc tagged with Zn^{65} were used as etchants. Polycrystalline gallium arsenide wafers were immersed in each solution for 60 seconds and

then rinsed for 5 seconds each in three beakers of 20 ml deionized water per wafer. Each test was run in triplicate. The radioactivity measurements indicated that the zinc adsorption was below 1×10^{10} Zn/cm^2. Several zinc complexes with chloride ions are known[13] that could account for the extremely weak adsorption behavior.

Similar experiments were made with 49% HF containing Zn^{65}. Again, adsorption on gallium arsenide, similar to silicon, was very weak and readily reversible by water rinsing. The ZnF^+ complex that forms with F^- ions[13] may be responsible for preventing zinc deposition.

Buffered HF (oxide etch) tagged with Zn^{65} led to surface concentrations below 1×10^{10} Zn/cm^2 on the polished $(\bar{1}\bar{1}\bar{1})$ face of gallium arsenide wafers. However, the backside of the wafer with a roughly etched (111) surface adsorbed 6×10^{11} Zn per cm^2 of geometrical apparent surface area.

Aqueous solutions of 20% ammonium peroxydisulfate, $(NH_4)_2S_2O_8$, used at 50°C, were also tagged with radioactive zinc at sub-ppm concentrations. Fewer than 1×10^{10} Zn/cm^2 were adsorbed on the $(\bar{1}\bar{1}\bar{1})$ face of gallium arsenide crystals.

The adsorption of zinc ions from HNO_3 on gallium arsenide is also of the physical type. For example, crushed gallium arsenide crystals equilibrated with 1N HNO_3 solution containing 1.7 ppm of zinc ions followed by two washes with 1N HNO_3 retained only 8.9×10^{11} adsorbed zinc ions per cm^2.[7]

8. Contamination by Manganese

A few experiments were carried out to examine the contamination levels that can be expected from manganese ions during typical etching conditions.

8.1 Deposition of Manganese on Germanium Using Mn^{54} as Tracer

The difference between the standard oxidation potentials[14] for the Mn(0)-Mn(II) and the Ge(0)-Ge(IV) couples in acidic solution is such that electrochemical deposition of manganese on germanium is not possible. One can therefore expect to encounter physical and hence easily reversible adsorption. This hypothesis is born out by the experimental data.

$Mn^{54}Cl_2$ was added both in carrier-free form and with 0.1 ppm $MnCl_2$ as isotopic carrier to standard iodine etch solutions. Germanium wafers were etched in these solutions for 100 seconds, followed by a 30-second water rinse. A residual manganese concentration of less than

4×10^9 manganese ions per cm^2 of germanium was detected. In terms of monolayer quantities, this extremely low concentration is equivalent to less than 10^{-6} monoionic layer of manganese.

9. Conclusions

[1] Adsorption of metal ions from mineral acids is low as long as the semiconductor is covered by a protective layer of oxide. Contamination of both Ge and GaAs surfaces by metallic trace impurities from reagent solutions is severest for etchants that dissolve the protective oxide layer, such as dilute HF, without etching the semiconductor. The large difference between the oxidation potentials of the semiconductor and the metal ions in solution therefore leads to large quantities of electrochemically deposited Cu, Au, and Cr. The adsorption of Fe is much lower, and that of Zn is negligible.

[2] Contamination of Ge during etching in mixtures containing HF and HNO_3, with or without CH_3CO_2H and I_2, leads to 3-30 times less Au, 10-50 times less Cr, and 10^3-10^4 times less Cu than contamination from HF because of redissolution of the metal deposits. Mn ions become adsorbed to a very low extent.

[3] During etching of Ge in hot NaOH solution where the semiconductor surface is being dissolved, the contamination by Cu, Fe, and Au is very severe since the deposits accumulate.

[4] Inhibition of metal deposition by the addition of EDTA-containing chelating agents to the solutions is particularly effective for H_2O_2 used for etching germanium. Complexing by acidic H_2O_2 is the most effective method for removing metal contaminants from Ge surfaces.

[5] Special etchants used in GaAs technology can lead to severe surface contamination by heavy metals, especially those etchants that contain alkali such as KOH solutions for electroetching or $NaOH + H_2O_2$ for chemical etching.

[6] Deposition on GaAs of Au from $H_2O_2 + H_2SO_4$ polishing etch, or of Cu from $H_2O_2 + HF$ solution, was considerably less severe, but the adsorbates were also strongly chemisorbed. In contrast, the adsorption of zinc ions from HF and other etchants was very weak and reversible in nature.

Acknowledgments

The author wishes to thank G. M. Loiacono for his proficient technical assistance, A. E. Widmer for contributing to the electroetching experi-

ments, A. Mayer for reviewing the results, and J. A. Amick for critically reading the manuscript.

Portions of this work were performed under Contract AF33(657)-11615, sponsored by the Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Ohio.

Appendix—Composition of Special Reagents

Iodine Etch

1.00 vol HF 49.5%
1.00 vol HNO_3 70.5%
1.40 vol CH_3CO_2H 100%
0.22 vol H_2O (introduced with tracer)
0.24% Triton—X100 nonionic surfactant (Rohm and Haas Co.)
0.15% Iodine Crystals

The ppm concentrations of radioactive impurities added to this (and all other solutions) is expressed on a weight/volume basis.

Chelating Mixture (CM)

This additive with chelating agents plus a surface active agent was designed primarily for use with HNO_3-containing etchants to inhibit the deposition of contaminants.

To prepare this formulation the following materials are emulsified by heating and mixing:

a) 20 ml Versenex 80, Dow Chemical Co.—(aqueous solution of pentasodium diethlylenetriamine penta acetate).
b) 20 ml Versenol 120, Dow Chemical Co.—(aqueous solution of trisodium N-hydroxyethylene diamine tri acetate).
c) 10 g EDTA—(ethylenediamine tetra acetic acid).
d) 20 g Tergitol Nonionic NP 40, Union Carbide Co.—(nonyl phenyl polyethylene glycol ether).

The emulsion is cooled to room temperature and cleared by adding some HNO_3. The solution is then adjusted to 100 ml with deionized water.

References:

[1] W. Kern, "Radiochemical Studies of Semiconductor Surface Contamination—I. Adsorption of Reagent Components," **RCA Review**, Vol. 31, No. 2, p. 207, June 1970.

[2] W. Kern, "Radiochemical Studies of Semiconductor Surface Contamination—II. Deposition of Trace Impurities on Silicon and Silica," **RCA Review**, Vol. 31, No. 2, p. 234 June 1970.

[3] W. Kern, "Radioisotopes in Semiconductor Science and Technology," **Semiconductor Products and Solid State Technology,** Vol. 6, No. 10, p. 22, and Vol. 6, No. 11, p. 23 (1963)

[4] A. Mayer and N. Goldsmith, "Surface Passivation Techniques for Compound Solid State Devices," Techn. Rep. No. AFAL-TR-65-213, August 1965, Contract No. AF 33 (657)-11615.

[5] W. C. Dunlap, Jr., "Gold as an Acceptor in Germanium," **Phys. Rev.,** Vol. 97, p. 614, Feb. 1, 1955.

[6] W. Kern and D. Puotinen, "Cleaning Solution Based on Hydrogen Peroxide for Use in Silicon Semiconductor Technology," **RCA Review,** Vol. 31, No. 2, p. 187, June 1970.

[7] G. B. Larrabee, "The Contamination of Semiconductor Surfaces," **J. Electrochem. Soc.,** Vol. 108, p. 1130 (1961).

[8] S. R. Morrison, "The Evaluation of Germanium Surface Treatments," **Phys. Chem. Solids,** Vol. 14, p. 214 (1960).

[9] D. R. Frankl, "Effects of Copper on Fast Surface States of Etched Germanium," **J. Electrochem. Soc.,** Vol. 109, p. 238, March 1962.

[10] P. J. Boddy and W. H. Brattain, "Effect of Cupric Ion on the Electrical Properties of the Germanium-Aqueous Electrolyte Interface," **J. Electrochem. Soc.,** Vol. 109, p. 812, Sept. 1962.

[11] G. M. Krembs and M. M. Schlacter, "Effect of Solution Trace Impurities on I-V Characteristics of Electro-deposited Surface Barrier Diode," **J. Electrochem. Soc.,** Vol. 111, p. 417 (1964).

[12] H. Becke, D. Flatley, W. Kern, and D. Stolnitz, "The Diffusion of Zinc into Gallium Arsenide to Achieve Low Surface Concentration," **Metal. Soc. AIME,** Vol. 230, p. 307, March 1964.

[13] H. J. Hicks, **The Radiochemistry of Zinc,** U. S. Atomic Energy Commission, NAS-NS-3015. (1960)

[14] W. M. Latimer, **Oxidation Potentials,** Prentice-Hall, New York, 1952 (Second Ed.).

A Comparative Analytical Study of the Performance of Argon Laser Amplifiers and Oscillators*

I. Gorog
RCA Laboratories, Princeton, N.J.

Abstract—The gas laser master-oscillator power-amplifier problem is considered, and the performance of a 1 kW cw ionized argon laser amplifier is compared with that of an oscillator under identical excitation conditions. Efficient amplifier operation requires that the gain be strongly saturated. For 50% inversion utilization efficiency in the two principal argon lines, the following approximate minimum input flux values are required: 50 watts/cm^2 at 4880 Å and 500 watts/cm^2 at 5145 Å. With realistic discharge-tube geometries the oscillator/amplifier configuration is inferior to the simple oscillator configuration. Rough estimates indicate that 1 kW multiline output can be obtained with approximately 0.2% overall efficiency from an oscillator. The efficiency of an amplifier with 100 watts total multiline input (40 W each in the two principal lines) and 1 kW total output is approximately a factor of two lower.

1. Introduction

The purpose of this paper is (1) to provide a general formulation of the gas laser amplifier problem; the formulation should yield the power output and saturated gain as a function of input power (i.e., detailed phase, frequency, hysteresis, mode competition, and pulling effects are of no concern here); (2) to compare the performance of a high-power cw ionized argon laser amplifier with that of an oscillator under identical excitation conditions; and (3) to obtain numerical estimates for an argon laser amplifier capable of ~1000 watts cw output.

It is assumed that the small-signal unsaturated gain is known from measurements. The laser amplifier gain saturation as a function of incident flux is computed first for single-frequency operation in the

* The research reported here was sponsored by the Advanced Research Projects Agency under USA ECOM Contract DAAB 07-69-C-0246 (ARPA order No. 1365).

absence of magnetic fields. The problem is set up in terms of the rate equations developed by Zory.[1] The solution is obtained in a normalized form similar to the Gorog and Spong[2] formalism that was originally derived to describe the saturation characteristics of single-frequency oscillators. The main difference between the oscillator and the amplifier problem is as follows. In an oscillator the saturated gain must be equal to the losses, and the inversion utilization efficiency can be calculated from the loss to small signal ratio. In an amplifier, however, both the gain saturation and the inversion utilization efficiency are determined by the incident flux.

The solution for the ionized gas laser amplifier operating in an axial magnetic field and amplifying circularly polarized light is given by the zero-field solution. If a linearly polarized light wave is to be amplified, then minimum depolarization and maximum inversion utilization efficiency (in the limit of high power levels) occur when the frequency of the wave to be amplified coincides with the frequency of the unsplit line center. The saturation curve is computed also for this case using typical argon-laser parameters.

Typical argon lasers are operated in axial magnetic fields high enough to resolve right- and left-hand circularly polarized gain components. A measurement method is suggested that can yield the circularly polarized gain from threshold measurements that employ linearly polarized attenuators. Expressions that relate the linearly polarized insertion loss to the circularly polarized unsaturated small-signal gain are derived using Sinclair's method.[3]

All results that relate to emitted power are obtained on a per unit volume basis. In the absence of detailed knowledge of the cross-sectional distributions, one can obtain first-order estimates of macroscopic quantities by assuming uniform cross-sectional distribution of all variables and parameters. If detailed information of the cross-sectional variations is available, the formulation and the results obtained here are well suited for computerized numerical calculations.

As stated above, the solutions are obtained for single-frequency operation. It is well known from single-frequency experiments that the single-frequency and multifrequency power outputs from a given laser are about the same, provided the small-signal round-trip gain is much larger than the losses. Similar behavior should be exhibited by an amplifier, except that for the amplifier the requirement is that the light flux should be much greater than the saturation flux parameter. Furthermore, it is expected that at flux levels much lower than the saturation flux, the single-frequency and multifrequency amplifier performance should be similar as long as only a small fraction of the

incident flux is in the frequency regime corresponding to the tails of the gain curve. One important difference between single-frequency and multifrequency amplification is that if the amplifier is in an axial magnetic field of magnitude typical for ionized gas lasers, a multifrequency linearly polarized input flux will always be depolarized after amplification. It may, however, be possible to devise a single-frequency master-oscillator–power-amplifier combination such that the amplifier output is linearly polarized even in the presence of an axial magnetic field.

Some Important Assumptions

The following list summarizes the approach taken in performing these calculations and indicates the most important limitations on the results:

(1) Small-signal gain is known from threshold measurements.

(2) Pumping rates are not influenced by the induced emission rates.

(3) The effect of "cross-talk" resulting from the simultaneous operation of more than one line (e.g., 5145 Å and 4880 Å in argon ion lasers) is neglected.

(4) The cross relaxation between degenerate substates is much faster than all other transition rates.

(5) The line shape is well approximated by a single Lorentzian–Gaussian convolution (i.e., velocity anisotropies are neglected).

(6) A single-frequency model is employed.

(7) Beam divergence is neglected.

(8) For the analysis of operation in an axial magnetic field, (a) splitting shows perfect symmetry, (b) the split line structure can be approximated by a normal Zeeman pattern, and (c) the cross relaxation between substates is not influenced by the removal of the degeneracy.

2. Gas Laser Amplifier, Zero Magnetic Field

Inversion is established between degenerate upper state "a" and degenerate lower state "b". Amplification takes place via the stimulated emission process from state "a" to state "b". A monochromatic (single-frequency) linearly polarized light flux is traveling in the z direction.

The symbols used are those of Zory[1]: λ_a and λ_b are the excitation rates of the upper and lower states, respectively; γ_a and γ_b are the spontaneous decay rates; ρ_a and ρ_b are the population densities; σ_{ab}

and σ_{ba} are the induced emission and absorption cross sections; I is the photon flux (cm^{-2}-sec^{-1}); A_{ab} is the spontaneous decay rate from state "a" to state "b"; δ is the ratio of the upper state degeneracy to that of the lower state; ω is the angular frequency of the incident light flux; v denotes the atomic (ionic) velocity of the active medium; r, z are the position coordinates, and no angular (ϕ) variation is assumed.

The basic rate equations under steady-state conditions can be written as[1]

$$\lambda_a(v,r,z) + I\sigma_{ba}(\omega,v)\ \rho_b(v,r,z) = [\gamma_a + I\sigma_{ab}(\omega,v)]\ \rho_a(v,r,z), \quad [1]$$

$$\lambda_b(v,r,z) + [I\sigma_{ab}(\omega,v) + A_{ab}]\ \rho_a(v,r,z) = [\gamma_b + I\sigma_{ba}(\omega,v)]\ \rho_b(v,r,z). \quad [2]$$

The ratio of the absorption cross section to the cross section of induced emission is given by the degeneracy ratio: $\sigma_{ba}(\omega,v)/\sigma_{ab}(\omega,v) = \delta$ and $\delta = (2J_a + 1)/(2J_b + 1)$, where J_a and J_b are the angular momenta of states "a" and "b", respectively. The effective inversion density, ρ, is given by

$$\rho = \rho_a - \delta\rho_b. \quad [3]$$

From the above rate equations and definitions

$$\rho(v,r,z) = \frac{\overline{N}W(v)}{1 + T_1\sigma_{ab}(\omega,v)I(r,z)}, \quad [4]$$

where

$$T_1 = \frac{\delta(\gamma_a - A_{ab}) + \gamma_b}{\gamma_a\gamma_b}, \quad [4a]$$

and

$$\overline{N}W(v) = \left(1 - \delta\frac{A_{ab}}{\gamma_b}\right)\frac{\lambda_a(v)}{\gamma_a} - \delta\frac{\lambda_b(v)}{\gamma_b}. \quad [4b]$$

Here T_1 is the effective spontaneous decay time constant of the inversion, and $\overline{N}W(v)$ is the excitation or small-signal inversion density per unit velocity range. For most gas lasers the velocity distribution

$W(v)$ is well approximated by a thermal equilibrium Maxwellian distribution; i.e., instead of Eq. [4b] we can explicitly write

$$W(v) = \frac{1}{\sqrt{\pi}u} \exp\left[-\left(\frac{v - v'}{u}\right)^2\right] \qquad [4c]$$

where $mu^2/2 = kT$, with T the kinetic temperature and v' the drift velocity.

The light is assumed to propagate in the z direction. The fractional change in flux is

$$\frac{dI}{I} = dz \int_{-\infty}^{\infty} [\sigma_{ab}(\omega,v)\ \rho_a(r,z,v) - \sigma_{ba}(\omega,v)\ \rho_b(r,z,v)]\ dv$$

$$= dz \int_{-\infty}^{\infty} \sigma_{ab}(\omega,v)\ \rho(r,z,v)\ dv. \qquad [5]$$

The interaction cross section can be obtained from Lamb's theory of gas lasers.[4] The explicit expression is[1]

$$\sigma_{ab}(\omega,v) = \frac{\gamma^2}{\gamma^2 + (\omega_o + Kv - \omega)^2}\,\sigma_o. \qquad [6]$$

where

$$\sigma_o = \frac{\lambda_o^2 A_{ab}}{4\pi\gamma} = \frac{\pi c^2 A_{ab}}{\omega_o^2 \gamma} \qquad [6a]$$

and $\gamma = (\gamma_a + \gamma_b + \gamma_e)/2$ is one-half of the full width of the homogeneous Lorentzian interaction (response) function (γ_e is the environmental contribution to the homogeneous broadening and is primarily the result of long-range phase-perturbing Coulomb collisions). K is the magnitude of the wave vector and λ_o is the wavelength corresponding to the transition.

In typical gas lasers the small-signal inversion can be assumed to be independent of z (no axial variation). In general, the r variation

must be retained. In the following, the saturation problem is treated on a per unit volume basis. Macroscopic quantities (e.g., power flow) can be obtained by integration over the cross-sectional area of the gain medium. Therefore, in the following expressions, the r dependence is not explicitly indicated.

The saturated gain can be computed from Eq. [5]. The line constants required to carry out the computation are available in the literature for most of the important gas laser transitions. The small-signal inversion density depends on the excitation conditions, and can be obtained from small-signal unsaturated gain measurements. In the limit of $I \to 0$ the unsaturated gain is

$$g_o = \lim_{I \to 0} \frac{1}{I} \frac{dI}{dz} = \overline{N} \int_{-\infty}^{\infty} \sigma(\omega, v) W(v) \, dv. \qquad [7a]$$

Then, at line center $(\omega = \omega_o + Kv')$,

$$g_o^{\,o} = \frac{\overline{N}\sigma_o}{\sqrt{\pi}u} \int_{-\infty}^{\infty} \frac{\gamma^2}{\gamma^2 + (Kv_1)^2} \exp\left(-\frac{v_1^{\,2}}{u^2}\right) dv_1, \qquad [7b]$$

where $v_1 = v - v'$. Now define a normalized line-width parameter ϵ

$$\epsilon = \frac{\gamma}{Ku}. \qquad [8a]$$

Eq. [7b] can be written as

$$g_o^{\,o} = \frac{\overline{N}\sigma_o \epsilon^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{\exp(-x^2)}{\epsilon^2 + x^2} \, dx = \overline{N}\sigma_o \epsilon Z_i(0, i\epsilon) \qquad [8b]$$

where $Z_i(0,i\epsilon)$ is the imaginary part of the Fried–Conte plasma dispersion function[5] for 0 real and ϵ imaginary argument. From Eq. [8b] the small-signal inversion density is

$$\overline{N} = \frac{g_o^o}{\sigma_o \epsilon Z_i(0, i\epsilon)} . \qquad [9]$$

From Eq. [5] the saturated gain at an arbitrary frequency ω is, (using Eqs. [4], [6] and [9]),

$$g = \frac{1}{I}\frac{dI}{dz} = \frac{g_o^o}{\epsilon Z_i(0, i\epsilon)} \frac{\gamma^2}{\sqrt{\pi} u} \int_{-\infty}^{\infty} \frac{\exp\left[-(v - v')^2/u^2\right] dv}{(\omega_o + Kv - \omega)^2 + \gamma^2(1 + IT_1\sigma_o)} . \qquad [10]$$

By defining the normalized operating frequency (or rather the normalized detuning from line center) as

$$x_i = \frac{\omega - \omega_o - Kv'}{Ku} , \qquad [11a]$$

we obtain, from Eq. [10],

$$g = \frac{g_o^o}{\epsilon Z_i(0, i\epsilon)} \frac{\epsilon^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{\exp(-x^2)}{(x - x_i)^2 + \epsilon^2(1 + IT_1\sigma_o)} dx. \qquad [11b]$$

If the saturation flux parameter is defined as $I_s = (T_1\sigma_o)^{-1}$ and the normalized flux as $F = I/I_s$, the above gain equation can be written

$$g = \frac{g_o^o}{Z_i(0, i\epsilon)} \frac{Z_i(x_i, i\epsilon\sqrt{1 + F})}{\sqrt{1 + F}} . \qquad [11c]$$

For computational purposes it is convenient to express the above as the ratio of the saturated gain to the small signal unsaturated gain at line center:

$$\frac{g}{g_o^o} = \frac{Z_i(x_i, i\epsilon\sqrt{1 + F})}{Z_i(0, i\epsilon)\sqrt{1 + F}} \qquad [12]$$

From Eq. [11b], in the high-intensity limit when $F \to \infty$, the saturated

gain approaches zero as

$$\lim_{F \to \infty} g = \frac{g_o^o}{Z_i(0, i\epsilon)} \frac{1}{\epsilon F} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp(-x^2)\, dx = \frac{g_o^o}{\epsilon F Z_i(0, i\epsilon)} = \frac{\overline{N}\sigma_o}{F},$$

and from the definition of the fractional gain, the differential normalized flux increase is, in this limit,

$$\frac{dF}{dz} = gF = \overline{N}\sigma_o, \qquad [13a]$$

or

$$\frac{dI}{dz} = N\sigma_o I_s = \frac{\overline{N}}{T_1}. \qquad [13b]$$

But in the high-intensity limit,

$$\lim_{I \to \infty} \hbar\omega \frac{dI}{dz} = P_\infty$$

is the theoretical limit of the power emission per unit volume of active medium. Therefore,

$$P_\infty = \frac{g_o^o}{\epsilon Z_i(0, i\epsilon)} \hbar\omega I_s \qquad [14]$$

is the total power available per unit volume as a function of the small-signal gain at line center. For finite flux, the power emitted per unit volume is $P = \hbar\omega I_s dF/dz$. Therefore, the inversion utilization efficiency of a gas laser amplifier can be expressed as

$$\frac{P}{P_\infty} = \epsilon F \frac{Z_i(x_i, i\epsilon\sqrt{1+F})}{\sqrt{1+F}}. \qquad [15]$$

Equations [12], [14] and [15] are the single-frequency cw gas laser amplifier master equations. Quick estimates of amplifier per-

formance can be made if the small-signal gain along the tube center is known. (I_s and ϵ can be computed from spectroscopic data available in the literature.) If the detailed cross-sectional variation of $g_o{}^o$ is known these "per unit volume" expressions can be readily employed in computerized numerical integrations.

3. Gas Laser Oscillator, Zero Magnetic Field

In general, the saturated gain and power emission expressions, as derived in the previous section for a gas laser amplifier, are not directly applicable to an oscillator. For the oscillator one must start with rate equations that include two oppositely traveling waves that saturate the inversion at two velocity classes. As a result of the output coupling and of the non-zero drift velocity, the gains experienced by the two oppositely traveling waves may not be equal. In order to overcome this difficulty one may obtain an approximate solution by defining an effective gain coefficient, $\overline{g}$, that is the average of the gains experienced by the two traveling waves, i.e., $\overline{g} = (g_+ + g_-)/2$.

The rate equations appropriate to the oscillator are[1]

$$\lambda_a + (I_+\sigma_{ba}{}^+ + I_-\sigma_{ba}{}^-)\rho_b = (\gamma_a + I_+\sigma_{ab}{}^+ + I_-\sigma_{ab}{}^-)\rho_a \qquad [16]$$

$$\lambda_b + (I_+\sigma_{ab}{}^+ + I_-\sigma_{ab}{}^- + A_{ab})\rho_a = (\gamma_b + I_+\sigma_{ba}{}^+ + I_-\sigma_{ba}{}^-)\rho_b \qquad [17]$$

where

$$\frac{\sigma_{ba}{}^+}{\sigma_{ab}{}^+} = \frac{\sigma_{ba}{}^-}{\sigma_{ab}{}^-} = \delta,$$

$$\sigma_{ab}{}^\pm = \sigma_o \frac{\gamma^2}{\gamma^2 + (\omega_o - \omega \pm kv)^2}, \qquad [18]$$

and the rest of the quantities are as defined in the previous section. The r, v and z dependence of the various quantities is not written out in Eqs. [16] and [17]. Both the pumping rates and the population densities depend on velocity and position (see Eqs. [1] and [2]).

The rate Eqs. [16] and [17] can be solved for the saturated gain in a fashion identical to the one outlined in the previous section. The effective gain coefficient obtained is then

$$\bar{g} = \frac{\bar{N}\sigma_o\gamma^2}{2\sqrt{\pi}u} \int_{-\infty}^{\infty} \frac{\exp\left[-(v-v')^2/u^2\right]dv}{\left[\dfrac{1}{\gamma^2+(\omega_o-\omega+Kv)^2} + \dfrac{1}{\gamma^2+(\omega_o-\omega-Kv)^2}\right]^{-1} + \gamma^2\bar{F}}, \quad [19]$$

where $\bar{F}$ is the average normalized one-way flux. In terms of the previously defined normalized variables, Eq. [19] can be written

$$\bar{g} = \frac{\bar{N}\sigma_o}{2\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{\exp(-x^2)\,dx}{\left[\dfrac{\epsilon^2}{\epsilon^2+(x+x'+x_i)^2} + \dfrac{\epsilon^2}{\epsilon^2+(x+x'-x_i)^2}\right]^{-1} + F^-} \quad [20]$$

Here $x' = v'/u$ is the normalized drift velocity.

The small-signal unsaturated effective gain is given by Eq. [20] in the limit $\bar{F} \to 0$. In terms of the plasma dispersion function this quantity can be written as

$$\bar{g}_o = \frac{\bar{N}\sigma_o\epsilon}{2}\left[Z_i(x'+x_i, i\epsilon) + Z_i(x'-x_i, i\epsilon)\right]. \quad [21]$$

Threshold operation occurs at the frequency at which $\bar{g}_o$ is maximum. If x' is finite, threshold operation occurs at a frequency different from line center. Other effects of the finite drift velocity are a reduction of the maximum power to central tuning dip (Lamp dip) ratio and a reduced inversion utilization efficiency. However, uncertainties in gain measurements and in the line parameters and errors introduced by neglecting the exact standing-wave nature of the interaction problem and by neglecting the difference in intensity between the two oppositely traveling waves are probably more important than the effect of drift velocities typical for gas lasers. The above is especially true at high intensities. Therefore, in the following the drift-velocity term is neglected.

In the absence of drift, threshold operation occurs at line center and the small-signal unsaturated gain is again given by Eq. [8b]. If $x' = 0$, the integral in Eq. [20] can be broken up into two integrals

of equal magnitude. Therefore, with the aid of Eq. [8], Eq. [20] can be written

$$\bar{g} = \frac{g_o^o}{\epsilon Z_i(0, i\epsilon)} S(\bar{F}), \qquad [22]$$

where

$$S(\bar{F}) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{\dfrac{\epsilon^2}{\epsilon^2 + (x + x_i)^2} \exp(-x^2)}{1 + \bar{F}\left[\dfrac{\epsilon^2}{\epsilon^2 + (x + x_i)^2} + \dfrac{\epsilon^2}{\epsilon^2 + (x - x_i)^2}\right]} dx. \qquad [23]$$

Now recall that by definition $dI/dz = \bar{g}I$, and note that because there are two oppositely traveling waves the power emitted per unit cross-sectional area per unit length is

$$P = 2\bar{g}I\hbar\omega. \qquad [24]$$

With the aid of Eqs. [22] and [14] we obtain

$$\frac{P}{P_\infty} = 2\bar{F}S(\bar{F}). \qquad [25]$$

Now note that Eq. [22] can be considered as an integral equation of $\bar{F}$ as a function of $\bar{g}/g_o^o$. Therefore, Eq. [25] can also be written

$$\frac{P}{P_\infty} = 2\epsilon Z_i(0, i\epsilon) \frac{\bar{g}}{g_o^o} \bar{F}\left(\frac{\bar{g}}{g_o^o}\right). \qquad [26]$$

Equation [26] gives the inversion utilization efficiency as a function of gain saturation of a laser oscillator operating at normalized frequency x_i. The average normalized flux as a function of gain saturation, $\bar{F}(\bar{g}/g_o^o)$, can be obtained by numerically solving the integral equation Eq. [22]. Such calculations have been carried out by Gorog

and Spong, who computed the inversion utilization curves for typical argon laser parameters.[2]*

The saturated gain can be easily related to the round-trip losses. Assume that the round-trip cavity length is d and the round-trip fractional power loss is T. Recall that $\bar{g}$ is an average saturated gain per unit length. Then the self-consistency condition of oscillation can be written as

$$(1 + \bar{g}d)\,(1 - T) = 1,$$

or

$$\bar{g}d = \frac{T}{1 - T}. \qquad [27]$$

The excess small-signal unsaturated gain, g_r, is usually defined as

$$g_r = \exp\,(g_o{}^o d) - 1,$$

or

$$g_o d = \ln(1 + g_r). \qquad [28]$$

From Eqs. [27] and [28] we obtain the normalized saturated gain;

$$\frac{\bar{g}}{g_o{}^o} = \frac{T}{(1 - T)\,\ln\,(1 + g_r)} \simeq \frac{T}{\ln\,(1 + g_r)}. \qquad [29]$$

Equations [26] and [29] are the oscillator master equations. If the cavity losses and the small-signal unsaturated round-trip gain are known from measurements, the inversion utilization efficiency can be computed from these equations.

4. Gas Laser Amplifier Operating in an Axial Magnetic Field

Here we assume that an axial magnetic field of magnitude typical for high-power ionized gas lasers (~1 kilogauss) is applied to the active medium. We also assume that the "g"-values of the upper and of the lower laser levels are equal (i.e., normal Zeeman pattern) and that the

* The function of $\bar{F}$ versus $\bar{g}/g_o{}^o$ is generated very easily by assuming successive values of $\bar{F}$ in the range $0 \leq \bar{F} \leq \infty$, substituting these values into Eq. [22], and integrating this equation numerically.

removal of the degeneracy affects neither the cross relaxation of the substates nor the spontaneous decay rates.

As a result of the selection rules,[6] the gain medium is circularly polarized. The peaks of the left-hand and right-hand components are shifted symmetrically with respect to the unsplit line center by amounts $\pm\omega_g$. As a result of the assumptions of the preceding paragraph, neither the interaction cross section σ_o, nor the effective spontaneous decay rate T_1 is affected by the presence of a magnetic field. Therefore, the saturation parameter I_s is also independent of the magnetic field. Then the saturation characteristics of a laser amplifier operating in a magnetic field and amplifying a circularly polarized flux are governed by Eqs. [12], [14], and [15] derived in Section 2, provided that now $g_o{}^o$ is interpreted as the small-signal gain at the peak of the circularly polarized gain component and x_i as the normalized detuning from that peak.

From a practical point of view it is of interest to estimate the performance of a linearly polarized single-frequency oscillator–amplifier combination. Here we assume that the linearly polarized oscillator is tuned to the unsplit line center (ω_o). The linearly polarized input flux must now be decomposed into two oppositely rotating circular components. These two components interact with atoms that belong to two different velocity classes centered at $v_z = \pm\,\omega_z/K$, respectively. Since the doppler broadening is symmetric, the two components are amplified by equal amounts. As a result of the different propagation velocities of the two circularly polarized components through the active medium, some depolarization of the output may nevertheless occur.

In estimating the Zeeman amplifier performance with a linearly polarized input flux it is convenient to normalize the saturated gain with respect to the unsaturated gain at the unsplit line center. Denote this unsaturated gain by $g_o{}^z$. In analogy with Eq. [8b] we can write

$$g_o{}^z = \bar{N}\sigma_o\epsilon Z_i(x_z, i\epsilon), \qquad [30]$$

where

$$x_z = \frac{\omega_z}{Ku}. \qquad [31]$$

Denote the quantities that pertain to the circularly polarized components by the subscript C and those that pertain to the linearly polarized flux by he subscript L. Then

$$F_C = \frac{F_L}{2}.$$

The gain saturation for a Zeeman amplifier operating on a linearly polarized flux at the unsplit line center frequency can be calculated from the expression derived for an oscillator in Section 3. In the oscillator case the two symmetric "holes" burned into the gain profile were caused by the oppositely traveling wave components of a standing wave, while in the Zeeman amplifier case they are caused by the two oppositely rotating circularly polarized components of the linearly polarized input flux. Therefore, in analogy with Eq. [20],

$$g_L(\omega_o) = \frac{\overline{N}\sigma_o}{2\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{\exp(-x^2)\,dx}{\left[\dfrac{\epsilon^2}{\epsilon^2 + (x + x_z)^2} + \dfrac{\epsilon^2}{\epsilon^2 + (x - x_z)^2}\right]^{-1} + F_t}, \qquad [32]$$

where the drift-velocity term has been neglected. From Eqs. [30], [32], and [23], the normalized saturated gain is

$$\frac{g_L(\omega_o)}{g_o^z} = \frac{S(F_C)}{\epsilon Z_i(x_z, i\epsilon)}. \qquad [33]$$

When calculating the power emission per unit volume one must include the contributions from both of the two oppositely rotating circularly polarized waves, i.e.,

$$P_L = 2g_L(\omega_o) I_C \hbar\omega.$$

But in analogy with Eq. [14],

$$P_\infty = \frac{g_C^o}{\epsilon Z_i(0, i\epsilon)} \hbar\omega I_S, \qquad [34]$$

where g_C^o is the unsaturated gain at the peak of the circularly polarized gain components. In analogy with Eq. [8b],

$$g_C^o = \overline{N}\sigma_o \epsilon Z_i(0, i\epsilon) \qquad [35]$$

$$= g_o^z \frac{Z_i(0, i\epsilon)}{Z_i(x_z, i\epsilon)}. \qquad [36]$$

Therefore

$$P_\infty = \frac{g_o^z}{\epsilon Z_i(x_z, i\epsilon)} \hbar\omega I_S. \qquad [37]$$

The inversion utilization efficiency is then

$$\frac{P_L}{P_\infty} = 2 \frac{g_L(\omega_o)}{g_o^z} Z_i(x_z, i\epsilon)\epsilon \frac{I_C}{I_S}. \qquad [38]$$

With the aid of Eq. [33], the above expression can be rewritten as

$$\frac{P_L}{P_\infty} = 2F_C S(F_C). \qquad [39]$$

The master equations are now Eqs. [33], [37], and [39]. The required experimentally measured quantity is again the small-signal unsaturated gain, which can be obtained either from transmission measurements or from threshold measurements in oscillator configurations. Methods of obtaining g_C^o from threshold measurements on a conventional internal Brewster-window-equipped laser is described in the following section.

5. Laser Oscillator Operating in an Axial Magnetic Field

If a gas laser oscillator is operating in an axial magnetic field high enough to resolve the two circularly polarized gain components and without polarization-selecting components in the laser cavity, then the laser output will be circularly polarized. Then the gain saturation and the inversion utilization efficiency can be calculated from Eqs. [26], [29], and [14]. (Naturally g_o^o must now be replaced by g_C^o. See Section 4.)

Practical high-power ionized gas lasers are usually constructed with internal Brewster windows. Such systems have been the subject of a number of past investigations.[7-9] It has been found that (1) if the laser is operated in a simple two-mirror cavity and in magnetic fields

larger than ~1 kilogauss, threshold operation occurs at two frequencies close to the peaks of the circularly polarized gain components; this observation is independent of whether the discharge tube is terminated in Brewster windows or perpendicular antireflection-coated windows;[10] (2) for a given mirror reflectivity, the required threshold gain is lower with antireflection-coated windows than with Brewster windows;[8] (3) far above threshold, the power outputs with antireflection-coated windows and with Brewster windows are virtually the same.[3,7,8]

The apparent anomaly associated with the efficient operation of lasers operating with internal Brewster windows in magnetic fields high enough to cause complete splitting of the Zeeman components has essentially been resolved by Sinclair.[3] By solving the laser oscillation condition equation with circularly polarized gain and partially linearly polarized cavity, he has found that if the small-signal gain is much larger than the cavity losses, then the Brewster windows introduce a negligible loss.

Based on the above experimental observations and on Sinclair's analysis, we conclude that the inversion utilization efficiency of high-power Brewster-window-terminated Zeeman lasers can be estimated by using the equations derived in Section 3.* In order to estimate P_x, however, one must measure $g_C{}^o$.

The small-signal gain is usually obtained from threshold measurements. As pointed out above, at threshold the Brewster window losses are not negligible. Furthermore, the threshold measurements are usually performed with partially linearly polarized attenuators. In the following section typical threshold measuring arrangements are analyzed and expressions that relate $g_C{}^o$ to the attenuator losses are derived.

6. Measurement of the Small-Signal Gain of a Zeeman Laser

The small-signal unsaturated gain is usually measured by inserting a calibrated attenuator into the laser cavity and thereby reducing the excess gain till threshold operation is reached. The most commonly employed attenuator consists of a set of rotatable dielectric plates. The calculations that follow are carried out with this type of attenuator in mind. Naturally, the results are valid for any attenuator that is describable by a diagonal loss matrix in a linearly polarized (x,y) representation.

The formalism used here is that of the Jones vector. The calcula-

* This statement is further supported by the amplifier P_L/P_x and P/P_o curves given in Section 7.

tions presented are extensions of Sinclair's work on the polarization characteristics of ionized gas lasers in magnetic fields.[3]

Consider a general plane wave electric field propagating in the z direction: $\vec{E} = \vec{E}\exp[i(\omega t - kz)]$, where, in the Jones vector representation, $\vec{E}_L = \{\vec{E}_x, E_y\}$. The electric field propagates through a gain medium that is circularly polarized. (We assume that the magnetic field is high enough that at threshold the effect of the σ_- gain component on a wave centered on the σ_+ component, and vice versa, is negligible.) The single-pass circularly polarized amplitude gain is described by the gain matrix

$$[G_C] = \begin{bmatrix} G & 0 \\ 0 & 1 \end{bmatrix}. \qquad [40]$$

Any vector described in the (x,y) representation can be transformed into the circularly polarized $(+,-)$ representation by the matrix $[C \leftarrow L]$:

$$[C \leftarrow L] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}. \qquad [41]$$

The inverse operation that transforms a matrix from the $(+,-)$ representation into the (x,y) one is

$$[L \leftarrow C] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}. \qquad [42]$$

Therefore, the circularly polarized gain, Eq. [40], in the (x,y) representation is

$$[G_L] = [L \leftarrow C]\,[G_C]\,[C \leftarrow L] \qquad [43]$$

$$= \frac{1}{2} \begin{bmatrix} 1+G & i(1-G) \\ -i(1-G) & 1+G \end{bmatrix}.$$

In the (x,y) representation the effect on the field amplitude of a partially reflecting mirror is given by

$$[R] = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}. \qquad [44]$$

The effect of a Brewster window is given by

$$[M_B] = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}, \qquad [45]$$

where a typical value for two surfaces of glass is $\alpha = 0.853$. Similarly, a partially linearly polarized attenuator is described by

$$[D] = d \begin{bmatrix} \delta & 0 \\ 0 & 1 \end{bmatrix}. \qquad [46]$$

Now consider an *asymmetric attenuator* arrangement, i.e., consider a cavity consisting of two mirrors having amplitude reflectivities r_1 and r_2, respectively, a Brewster-window-terminated laser discharge tube, and an attenuator of the type described by Eq. [46] placed between the discharge tube and mirror No. 2. The oscillator self-consistency equation is

$$\{[G_L][M_B][D][R_2][D][M_B][G_L][M_B][R_1][M_B] - [I]\}\vec{E}_L = 0, \qquad [47]$$

where $[I]$ is the identity matrix.

The characteristic equation is obtained by setting the determinant of the bracketed expression equal to zero. After carrying out the matrix multiplications indicated in Eq. [47], and with the aid of Eqs. [41]-[46], we obtain

$$0 = \det \begin{bmatrix} \rho\alpha^2\{\tau(1+G)^2 + (1-G)^2\} - 1 & i\rho(\tau+1)(1-G^2) \\ -i\rho\alpha^2(\tau+1)(1-G^2) & \rho\{\tau(1-G)^2 + (1+G)^2\} - 1 \end{bmatrix}, \qquad [48]$$

where

$$\rho = \frac{r_1 r_2 d^2}{4}, \qquad [49]$$

and

$$\tau = \alpha^2\delta^2. \qquad [50]$$

Solving Eq. [48] for G, the one-way circularly polarized amplitude gain as a function of losses is

$$G = \frac{B - \sqrt{B^2 - 4AC}}{2A} \qquad [51]$$

where

$$A = \rho[16\rho\alpha^2\tau - (1+\tau)(1+\alpha^2)], \qquad [51a]$$

$$B = 2\rho(1-\tau)(1-\alpha^2), \qquad [51b]$$

$$C = 1 - \rho(1+\tau)(1+\alpha^2), \qquad [51c]$$

and ρ and τ are given in Eqs. [49] and [50]. The minus sign in front of the square root in Eq. [51] was established by letting $\tau \to \alpha^2$ and evaluating Eq. [51]. The limit $\tau \to \alpha^2$ corresponds to the removal of the attenuator from the laser cavity, and in this limit Eq. [51] reduces to Sinclair's result.[3]

A much simpler expression for G is obtained if the measurements are performed with a *symmetric attenuator* arrangement. Consider a cavity consisting of two mirrors of amplitude reflectivities r_1 and r_2, respectively, a Brewster-window-terminated discharge tube, and two identical attenuators, one on each side of the discharge tube between the Brewster windows and the mirrors. The self-consistency equation is then

$$\{[R_1][D][M_B][G_L][M_B][D][R_2][D][M_B][G_L][M_B][D] - [I]\}\vec{E}_L = 0. \qquad [52]$$

Since the mirror matrixes are simply $[R_1] = r_1[I]$ and $[R_2] = r_2[I]$, they commute with all the other matrixes in Eq. [52]. Therefore, Eq. [52] can be written as

$$\{[M]^2 - [I]\}\vec{E}_L = 0, \qquad [53a]$$

where

$$[M] = (r_1 r_2)^{1/2}[D][M_B][G_L][M_B][D]. \qquad [53b]$$

Now recall that[11]

$$[M]^2 - [I] = \{[M] + [I]\}\{[M] - [I]\} \qquad [54]$$

and that for any two matrixes of the same order[12]

$$\det\{[A][B]\} = \{\det[A]\}\{\det[B]\}. \qquad [55]$$

Therefore, the characteristic equation for Eq. [53a] can be written as either

$$\det\{[M] - [I]\} = 0, \qquad [56]$$

or

$$\det\{[M] + [I]\} = 0. \qquad [57]$$

By carrying out the solution for Eq. [56] and comparing our result in the appropriate limit with Sinclair's result,[3] we find that it yields the physically meaningful answer. The explicit form of Eq. [56] is

$$0 = \det \begin{bmatrix} \rho'\alpha^2\delta^2(1+G) - 1 & i\rho'\alpha\delta(1-G) \\ -i\rho'\alpha\delta(1-G) & \rho'(1+G) - 1 \end{bmatrix}, \qquad [58]$$

where

$$\rho' = \frac{r_1 r_2 d^2}{2}. \qquad [59]$$

By solving Eq. [58], we obtain

$$G = \frac{-2 + \sqrt{r_1 r_2}d^2(1+\alpha^2\delta^2)}{2r_1 r_2 d^4\alpha^2\delta^2 - \sqrt{r_1 r_2}d^2(1+\alpha^2\delta^2)}. \qquad [60]$$

Equation [60] gives the single-pass circularly polarized amplitude gain of the laser medium as a function of the amplitude losses for a symmetric attenuator arrangement.

The above results can be used to determine the small-signal unsaturated gain of a Zeeman laser, provided that the magnetic field is high enough that the σ_+ and σ_- gain components are well resolved. The power gain is simply G^2 and the gain coefficient per unit length for an amplifier of length l is obtained from the usual exponential gain expression, $G^2 = \exp(g_C{}^o l)$, as

$$g_C{}^o = \frac{\ln(G^2)}{l}. \qquad [61]$$

From the experimentally measured threshold splitting of the argon ion laser in an axial magnetic field[10] we conclude that Eq. [40]

describes the threshold gain well at magnetic fields higher than approximately 1 kilogauss. Since typical argon lasers are operated in the region 750 to 1500 gauss, the above results can be considered as good approximations for most practical cases of interest.

7. Numerical Results

Table 1 lists representative values for the parameters of the dominant argon laser lines. The most important parameters required for the evaluation of the saturation equations derived in these notes are the line shape parameter ϵ and the saturation parameter I_S. Both ϵ and I_S are dependent on the discharge conditions; the values included in the table should be considered only as typical values obtained from measurements performed on narrow-bore discharge tubes (diameter $\leqslant 3$ mm).

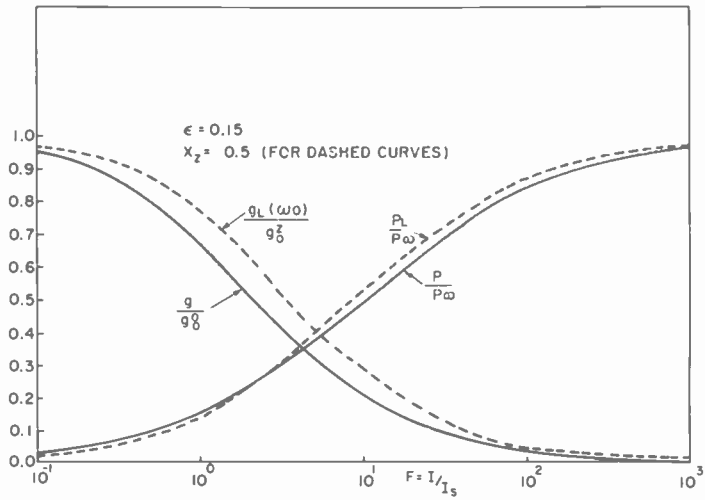


Fig. 1—Argon laser amplifier gain and power emission per unit volume as a function of incident flux.

On Fig. 1 the normalized gain- and power-saturation curves are plotted as functions of the normalized incident flux for the line-shape parameter typical for argon lasers. (See Eqs. [12], [14], [15], [33], [37], and [39].) The solid curves describe an amplifier operating without an axial magnetic field and the dashed curves describe a Zeeman amplifier amplifying a linearly polarized flux tuned to the unsplit line center. (As discussed in Section 4, the behavior of a Zeeman amplifier amplifying a circularly polarized flux is described by the same satura-

Table 1—Dominant Argon Laser Line Parameters

	4880 Å	5145 Å	NOTES
Upper state, a	$3p^4 4p\ ^2D^o_{5/2}$	$3p^4 4p\ D^o_{5/2}$	i
Lower state, b	$3p^4 4s\ ^2P_{3/2}$	$3p^4 4s\ ^2P_{3/2}$	i
Connecting A coefficient, A_{ab}	8.96×10^7 sec^{-1}	0.707×10^7 sec^{-1}	i
Observed "g" value of upper state, g_a	1.241	1.334	i
Observed "g" value of lower state, g_b	1.334	1.334	i
Degeneracy ratio, $\delta = (2J_a + 1)/(2J_b + 1)$	3/2	3/2	
Decay rate of upper state, γ_a	1.03×10^8 sec^{-1}	1.30×10^8 sec^{-1}	
Decay rate of lower state, γ_b	5.52×10^8 sec^{-1}	5.52×10^8 sec^{-1}	i i
Inversion decay time, T_1	0.97×10^{-8} sec	1.01×10^{-8} sec	ii
Doppler temperature of ions	~ 2,500°K	~ 2,500°K	iii
Mass of ions	0.664×10^{-22} grams	0.664×10^{-22} grams	
Thermal speed of ions, u	1.02×10^5 cm/sec	1.02×10^5 cm/sec	
Doppler parameter, u/λ	2.08 GHz	1.98 GHz	
Lorentzian (homogeneous) full width	~ 600 MHz	~ 600 MHz	iv
ϵ	~ 0.15	~ 0.15	
Saturation parameters, $\hbar\omega I_s$	5 Watts/cm^2	50 watts/cm^2	iv
Electron temperature at 250 A/cm^2	~ 5 eV	~ 5 eV	v
Electron density at 250 A/cm^2	~ 6×10^{14} cm^{-3}	~ 6×10^{14} cm^{-3}	v

i. H. Statz et al., *J. Appl. Phys.*, Vol. 36, p. 2278 (July 1965).

ii. P. Zory, *IEEE J. Quant. El.*, Vol. QE-3, p. 390 (Oct. 1967).

iii. The Doppler temperature of the ions is dependent on the discharge conditions. The 2,500 °K value quated in the table is a typical approximate value as obtained from the following sources: A. L. Bloom et al., *Physics of Quantum Electronics*, P. L. Kelley et al., editors, McGraw Hill, New York, 1966, p. 688; E. A. Ballik et al., *Appl. Phys Lett.* Vol. 8, p. 214 (15 April 1966); I. Gorog and F. W. Spong, *IEEE J. Quant. El.*, Vol. QE-3, p. 692 (Dec. 1967). However, according to V. F. Kitaeva et al., *IEEE J. Quant. El.*, Vol. QE-5, p. 72 (Feb. 1969), the above quoted ion temperature value may be too low by more than a factor of two.

iv. Approximate representative values taken from the reference in Note ii above.

v. C. B. Zarowin, *Appl. Phys. Lett.*, Vol. 15, p. 36 (1 July 1969).

tion curves as the no-magnetic-field amplifier.) The crossover of the inversion utilization curves is related to the fact that the Zeeman amplifier amplifying a linearly polarized flux saturates at two velocity groups. However, the difference between the Zeeman and non-Zeeman curves is smaller than either the expected accuracy of the model developed here or the reproducibility of most laser measurements at high power levels.

On Fig. 2 the non-Zeeman oscillator saturation curve is shown (See Eq. [26]). As discussed in Section 5, the same curve can also be used

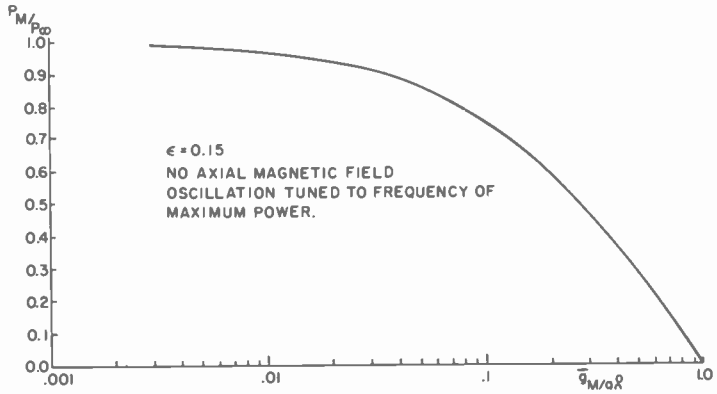


Fig. 2—Argon laser oscillator power emission as a function of gain saturation.

to obtain rough estimates of the Zeeman oscillator performance. (The small difference between the computed P_L/P_∞ and P/P_∞ curves in Fig. 1 provides an additional support for the arguments presented in Section 5.)

From Figs. 1 and 2 we conclude that a typical argon laser oscillator operating at 5145 Å can be easily constructed to work with better than 50% inversion utilization efficiency; the same efficiency will be attained by an amplifier only if it is operating at flux levels in excess of 500 watts/cm^2.

In order to illustrate further the relative performance of oscillators and amplifiers, a specific example will be treated in greater detail.

We are interested in obtaining 1 kW total multiline output. From low-power threshold measurements, the unsaturated gain of the 5145-Å line is approximately four times lower than that of the 4880-Å line. From Table 1 the saturation parameter of the 5145-Å line is approximately ten times larger than that of the 4880-Å line. Therefore,

P_∞ (5145 Å)/P_∞ (4880 Å) $\simeq$ 2.5. In an amplifier arrangement the inversion utilization efficiency of the 4880-Å line is expected to be somewhat better than that of the 5145-Å line. (For a given power input a lower saturation parameter implies higher normalized flux.) Therefore, we estimate that in a high-power multiline power amplifier output, ~60% of the total will be in the 5145-Å line and ~40% in the 4880-Å line. The secondary lines are expected to contribute negligibly to the amplifier output.

The amplifier input power at 5145 Å is assumed to be 40 watts. Preliminary estimates suggested that in order to amplify 40 W to 600 W, the amplifier length will have to be in the vicinity of 20 meters. From the point of inversion utilization efficiency we would like to work with the highest possible flux levels and therefore with the smallest possible bore sizes. On the other hand, from the points of view of diffraction losses, of available power per unit length, and of over-all discharge efficiency, large bore sizes are desirable.[13,14] As a compromise we choose 10-mm-diameter discharge tubes and we assume 250 A discharge current. The available power per unit length and the electrical power dissipation are computed from empirical scaling laws.[15] P_∞ is then computed from the available power per unit length assuming uniform transverse distribution. For purposes of calculation we assume that the input power is linearly polarized single frequency at the unsplit line center and that the amplifier is in approximately 700 gauss axial magnetic field. The small-signal gain $g_0{}^z$ is computed from P_∞ using Eq. [37].

The numerical integration of the emitted power (see Eq. [39]) along the amplifier length is carried out with the aid of Fig. 1 and Table 1. The results are shown in Fig. 3. We find that ~600 watts is obtained from an 18-meter-long amplifier. Table 2 summarizes the computed amplifier characteristics.

For comparison, we now estimate the performance of the same 18-meter-long discharge arranged as an oscillator. At a power level of a few watts, ~40% of the total output power is in the 5145-Å line. We assume that the same relative distribution of powers holds at the 1 kW level. We also assume that the oscillator inversion utilization efficiency will be ~60% and that the output coupling is so high that the internal losses can be neglected. (The internal light flux of the oscillator will have to be kept at the lowest possible level to minimize damage to the optical surfaces; with less than 50% mirror reflectivity, the overall utilization efficiency of the oscillator can be maintained at approximately the same level as the efficiency of the last amplifier section.) With the above assumptions we calculate for the total power output

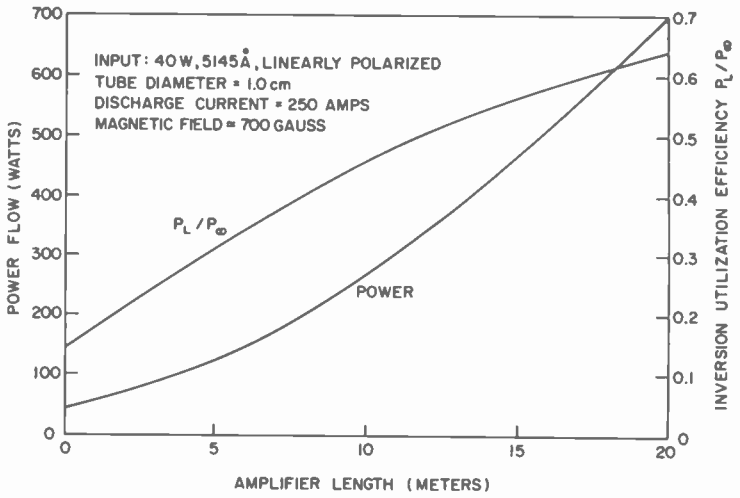


Fig. 3—Performance of an argon laser amplifier.

(all lines) 2.2 kW; and an overall oscillator efficiency of ~0.2%. Therefore, the overall oscillator efficiency is expected to be approximately twice that of the amplifier.

Table 2—Summary of Argon Laser Amplifier Data

Wavelength = 5145 Å
Discharge tube diameter = 10 mm
Discharge current = 250 amperes
Magnetic field = 700 gauss
Power available per unit length = 80 watts/meter
$P_\infty = 1.02$ watts/cm^3
$g_o z = 0.375$ meter^{-1}
Electrical power dissipation = 62 kW/meter (Assuming 1-meter discharge sections and including magnet, electrode and discharge power requirements.)
Optical power input = 40 watts
Normalized input quantities:
$F_L = 1.02$, $g_L(\omega_o)/g_o z = 0.77$, $P_L/P_\infty = 0.14$
Optical power output = 600 watts
Normalized output quantities:
$F_L = 15.3$, $g_L(\omega_o)/g_o z = 0.23$, $P_L/P_\infty = 0.61$
Amplifier length required = 18 meters
Total electrical power dissipation = 1.12 megawatts
Total optical power out (4880 Å and 5145 Å) = 1 kW

8. Conclusion

The calculations presented here suggest that the master-oscillator–power-amplifier configuration is inferior to the simple oscillator configuration for obtaining 1 kW output from an ionized argon laser. In high-power argon lasers the 5145-Å line is the most powerful one. To operate an argon amplifier at 5145 Å with 50% inversion utilization efficiency, the input flux must be approximately 500 watts/cm^2. Therefore, with realistic discharge geometries, 50% inversion utilization efficiency from an oscillator/amplifier configuration is achievable only if the amplifier contributes less then half of the total 1 kW output.

An obvious means of achieving the high power densities required for efficient inversion utilization would appear to be employment of beam-transforming optics between the master oscillator and the power amplifier. This solution, however, is not desirable on two counts:

[1] If the amplifier bore is narrowed, then the available power per unit length is reduced and therefore the amplifier length must be increased in order to obtain a given output. The increased length of the amplifier necessitates the use of periodic beam folding and transforming optics. It has been well established experimentally that at flux levels in excess of a few hundred watts per square centimeter the losses in laser optics rise sharply.[12-14] Therefore, the large number of optical surfaces associated with the beam-folding and beam-transforming optics make the narrow-bore amplifier configuration impractical.

[2] All experimental evidence indicates that the pumping efficiency of argon lasers increases with increasing discharge bore radius.[2,7,9,13-15] Therefore, the increase in inversion utilization efficiency obtained by the reduction of the amplifier bore radius is accompanied by a decrease in pumping (discharge) efficiency.

Based on the above, we conclude that the most practical and efficient way to obtain a 1 kW cw argon laser output is to construct a large-bore oscillator with internal mirrors (no Brewster windows). Our estimates suggest that bore size and output coupling can be so chosen that the power level inside the cavity can be kept within the capabilities of currently available laser optics.

Recent experiments and calculations[13,14] suggest that in the high-current regime both oscillator power output and electrical power dissipation depend on the product of the current density and tube radius (jR). In the jR greater than or approximately equal to 200 A/cm regime the efficiency increases as $(jR)^2$. No experiments performed in this regime have been reported in the literature. The experimental

difficulties associated with the high jR values are considerable but not unsurmountable. With quartz tubes, 150 A/cm has been achieved. The scaling laws suggest that at jR greater than or approximately equal to 200 A/cm the efficiency increases linearly with radius for a constant power dissipation per unit wall area.

Acknowledgment

The author thanks W. Zernik and K. G. Hernqvist for a number of valuable discussions.

References:

[1] P. Zory, "Single Frequency Operation of Argon Ion Laser," **IEEE J. Quant. El.**, Vol. Q.E.-3, No. 10, p. 390, Oct. 1967.

[2] I. Gorog and F. W. Spong: "Single Frequency Argon Laser," **RCA Review**, Vol. 30, No. 2, p. 277, June 1969.

[3] D. C. Sinclair, "Polarization Characteristics of an Ionized-Gas Laser in a Magnetic Field," **J. Opt. Soc. Am.**, Vol. 56, No. 12, p. 1727, Dec. 1966.

[4] W. E. Lamb, Jr., "Theory of an Optical Maser," **Phys. Rev.**, Vol. 134, No. 6A, p. A1429, June 1964.

[5] B. D. Fried and S. D. Conte, **The Plasma Dispersion Function** Academic Press, Inc., New York, 1961.

[6] E. U. Condon and G. H. Shortley, **The Theory of Atomic Spectra**, p. 63, Cambridge At The University Press, 1956.

[7] E. F. Labuda, E. I. Gordon and R. C. Miller, "Continuous-Duty Argon Ion Lasers," **IEEE J. Quant. El.**, Vol. QE-1, No. 6, p. 273, Sept. 1965.

[8] I. Gorog and F. W. Spong, "Experimental Investigation of Multiwatt Argon Lasers," **RCA Review**, Vol. 28, No. 1, p. 38, March 1967.

[9] W. B. Bridges and A. S. Halsted, "Gaseous Ion Laser Research," Technical Report AFAL-TR-67-89, Defense Documentation Center for Scientific and Technical Information, Cameron Station, Alexandria, Virginia, May 1967.

[10] I. Gorog and F. W. Spong, "An Approximate Linewidth Determination Method and the Magnetic Field Tunable Stable Spectrum of the Argon Laser," **IEEE J. Quant. El.**, Vol. Q.E.-3, No. 12, p. 691, Dec. 1967.

[11] R. A. Frazer, W. J. Duncan and A. R. Collar, **Elementary Matrices**, pp. 39-40, Cambridge At The University Press, 1969.

[12] **Ibid.**, p. 16.

[13] G. Herziger and W. Seelig, "Berechnung der Besetzungsdichte und Ausgangsleistung von Ionenlasern," **Z. Physik**, Vol. 215, p. 437, 1968.

[14] G. Herziger and W. Seelig, "Ionenlaser hoher Leistung," **Z. Physik**, Vol. 219, p. 5, 1969.

[15] J. R. Fendley, Jr., **Argon Laser Design**, Master's Thesis, University of Pennsylvania, The Moore School of Electrical Engineering, Philadelphia, Pennsylvania, May 1969 (Unpublished).

Computer Calculation of Electron Trajectories in Television Camera Tubes

Otto H. Schade, Sr.
RCA Electronic Components, Harrison, N. J.

Abstract—A computer program has been designed to trace electron trajectories in camera tubes immersed in an axially symmetric magnetic field in the presence of a transverse magnetic deflection field and radial and axial E-fields with axial symmetry (drift fields, accelerating or decelerating fields with or without field mesh). The axial, radial, and transverse magnetic fields of the focus and deflection coils and the potential fields in the *X,Y,Z* space are specified by six tables read into the computer. Four magnetic-field tables are generated outside the program by measurements of the imaging systems to be analyzed, and potential and E-field tables are obtained from field plots (computer or resistance board).

Dimensions, field intensities, and deflection-field values can be scaled up or down by built-in multipliers, and electron beams can be computer focused for any number of nodal points. The computation error is very small (1.2×10^{-6} cm for a trajectory 28.3 cm in length) and is established for paraxial trajectories in camera tubes by a built-in analytic test for axial chromatic aberration. The aberration is computed from the magnification of the electron optic obtained from one trajectory.

Introduction

This paper describes a computer program capable of tracing electron trajectories in camera tubes immersed in an axially symmetric magnetic field in the presence of a transverse magnetic deflection field and radial and axial *E*-fields with axial symmetry (drift fields, accelerating or decelerating fields with or without field mesh). Although many of the details and safeguards contained in the actual program have been omitted from this description for the sake of brevity and simplicity, it becomes quite obvious from the material that is discussed that a hand calculation of the functions performed by the computer is out of the question. The RCA Spectra 75 computer runs approximately 44 minutes to compile the 500-instruction FORTRAN program, focus the

beam, and calculate 25 trajectories, including a 10-point table of the axial chromatic aberration.

The purpose of the numerical calculations described in this paper is an analysis of electron motion in a camera tube and the imaging characteristics of the electron optic. This analysis explains and predicts the effects of operating parameters and electron-gun characteristics on beam formation, aberrations, and electron landing conditions at the target.

The motion of an electron in the magnetic and electric fields of a television camera tube, such as a vidicon, cannot be expressed in terms of elementary functions. This electron trajectory is generated by progressive addition of a large number of incremental sections generally computed with the assumption that the fields are uniform within an incremental part $\Delta(X,Y,Z)$ of the space. The calculation of aberrations in a high-definition camera requires extreme accuracy—an accuracy obtainable by the use of the tangent angles of flux lines and electron paths at the entrance and exit points of an increment. These angles establish a common radius of curvature for arc approximations and angular subdivision within a data increment.

Electron motion in a camera tube is calculated in two parts. The principal path of the electron is the equivalent deflected magnetic-flux line in the (X,Y,Z) space and includes deflections from radial electric fields; the radial electric fields are replaced by equivalent circular magnetic fields. The exit coordinates for an incremental section are computed by successive approximation under the assumption that a constant radius of curvatures within an increment is determined by the angles $(\delta_n, \delta_{(n-1)})$ and $(\epsilon_n, \epsilon_{(n-1)})$ of the flux-line projections in the (X,Z) and (Y,Z) planes, respectively.

The circular-motion component of the electron in the plane (X',Y') normal to the principal path is computed as a deviation from the principal path and accumulated separately by the sums $\Sigma\Delta X'$ and $\Sigma\Delta Y'$. An electron may gain or lose radial velocity when the principal path (the flux line) has curvature, i.e., the electron tends to move in a tangential direction as a result of its inertia. Thus, when the electron enters an incremental section, the differential angles β_n equal to $\delta_n - \delta_{(n-1)}$ and ρ_n equal to $\epsilon_n - \epsilon_{(n-1)}$ are added vectorially to the tangent angle $\phi_{(n-1)}$ of the electron according to its angle of rotation $\mu_{(n-1)}$. The approximation consisting of lumping the gradual change of ϕ within an increment into a single step causes the (cyclotron) radius of motion to expand or contract by increments at the section boundaries illustrated by the model in Fig. 1 (a model showing the cylindrical flux tube sections enclosed by the circular motion). Ade-

quate precision requires small increments and the use of mean angles $(\overline{\beta},\overline{\rho})$ and mean field values $(\overline{B},\overline{V})$ to calculate the incremental rotation $\Delta\mu$, the cyclotron radius R_n, and the circular path in the increment. Thus, the relatively large data increments (n) are divided into much smaller subincrements (u) in which the arc approximation and angular subdivision of the principal path provide the required accuracy for good interpolations.

In principle, camera-field data can be specified by equations describing the axial magnetic-field function of the focus coil, the axial-potential function in the camera tube, and the transverse-and-axial-field functions of the deflection coils for a given direction of deflection; off-axis

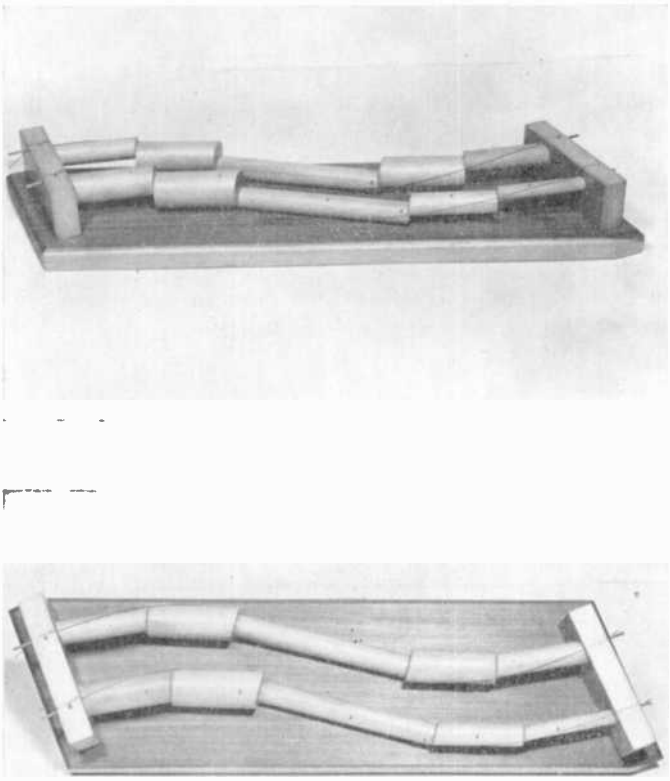


Fig. 1—Model of electron trajectories in the drift space of a camera tube.

and radial fields can be computed from series expansions. The derivations of accurate expressions for the field functions from measured data, however, requires time-consuming computer calculations. On the other hand, measured data generally lack the accuracy required for series expansions with higher-order derivatives. In the techniques described, the axial- and radial-field functions for several radial distances were measured and tabulated. The data tables (smoothed by inspection) were then fed into the computer. The comparatively low accuracy of measured data (perhaps 2% error) does not limit the accuracy of computer solutions.

The Computer Program

Field Data

The long drift space between gun aperture and field mesh of a vidicon, for example, is divided into 28 uniform Z increments specified by ΔZ_1, and the short decelerating space between field mesh and storage surface is divided into 12 smaller increments specified by ΔZ_2 and ΔZ_3. The number of increments between the electron gun and the mesh plane are given by the input data for 4 to 7 values of the parameter p, the general symbol for the distance in an (r,z) plane for which the deflection field is given. The data are read in and stored in a 7×50 matrix for each of the following six functions:

Function No.	Function
1a	B_{Fz} (z,p) axial field of focus coil (gauss)
1b	B_{Dz} (z,p) axial field of deflection coils (gauss)
2a	B_{Fr} (z,p) radial field of focus coil (gauss)
2b	B_D (z,p) transverse field of deflection coils (gauss)
3	V_z (z,p) potential function (volts)
4	V_r (z,p) radial potential gradient (volts per centimeter)

A typical set of functions for a high-resolution return-beam vidicon of 28.31-cm focal distance is given in Tables 1 through 3. Functions 1(a) and (b), and 2(a) and (b) are pairs that have equal distance parameters p to facilitate interpolation of the summation fields as follows:

$$B_z = (B_{Fz} + B_{Dz})\,(z,p) \qquad [1]$$

and

$$B_d = (B_{Fr} + B_D)\,(z,p) \qquad [2]$$

Table 1—Focus-Coil Fields*

	(Axial Component) $B_{Fz}(z,p)$				(Radial Component) $B_{Fr}(z,p)$			
n#	$p=0$‡	1.0	2.0	3.0	$p=0$	1.0	2.0	3.0
0	17.5	17.5	16.0	14.0	0.0	−2.3	−4.6	−6.9
	23.0	23.2	21.8	18.5		−3.25	−6.5	−9.7
	30.5	30.4	29.0	24.5		−4.12	−8.3	−12.3
	39.5	39.3	38.0	37.0		−4.375	−8.75	−13.1
	48.0	48.0	48.5	49.0		−4.5	−9.0	−14.0
5	57.5	57.7	58.2	59.5		−4.27	−8.55	−12.8
	65.1	65.5	67.0	68.5		−3.575	−7.15	−10.7
	71.8	72.0	73.5	76.0		−3.02	−6.05	−9.1
	77.2	77.5	79.0	81.2		−2.25	−4.5	−6.75
	80.8	80.9	82.0	84.1		−1.25	−2.5	−3.75
10	82.5	82.5	83.8	85.5		−0.55	−1.1	−1.65
	83.0	83.2	84.0	85.5		0.0	0.0	0.0
	82.5	82.6	83.0	84.0		0.45	0.9	1.35
	81.2	81.2	81.4	82.1		0.675	1.6	2.5
	79.8	79.8	79.8	79.5		0.78	1.8	3.15
15	78.1	77.9	77.5	76.7		0.675	1.5	2.5
	77.1	76.8	76.0	74.8		0.3	0.6	0.9
	76.9	76.5	75.5	74.0		−0.025	−0.05	−0.075
	77.2	77.1	76.2	74.9		−0.6	−1.3	−2.2
	78.9	78.8	78.2	77.1		−0.9	−2.0	−3.5
20	80.8	80.8	80.8	80.8		−1.0	−2.25	−3.9
	82.8	83.0	83.2	84.0		−0.9	−1.9	−3.2
	84.1	84.2	85.2	86.8		−0.525	−1.05	−1.7
	84.9	85.0	86.0	87.6		0.0	0.0	0.0
	84.1	84.3	85.0	86.5		0.475	1.2	2.1
25	83.0	83.0	83.4	84.0		0.75	1.85	3.25
	81.0	81.0	81.0	80.9		0.975	2.15	3.6
	79.1	79.0	78.5	77.5		0.75	1.7	3.1
28	78.0	77.8	76.6	74.5	0.0	0.0	0.0	0.0
	↓	↓	↓	↓	↓	↓	↓	↓
40	78.0	77.8	76.6	74.5	0.0	0.0	0.0	0.0

* Field measured in gauss.
\# The field mesh is located at $n = 28$.
‡ $p = r$.

Table 2—Deflection-Coil Fields*

	(Axial Component) $B_{Dz}(z,p)$				(Transverse Component) $B_D(z,p)$			
n	$p=0$#	1.0	2.0	3.0	p = 0	1.0	2.0	3.0
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.067	0.0461	0.0575	0.0385	0.0385	0.0193	0.0
	0.0	0.067	0.115	0.123	0.081	0.077	0.0193	−0.008
	0.0	0.115	0.22	0.258	0.231	0.193	0.077	−0.0193
	0.0	0.23	0.41	0.50	0.425	0.366	0.154	−0.058
5	0.0	0.375	0.56	0.92	0.81	0.695	0.385	−0.077
	0.0	0.575	1.23	1.81	1.29	1.15	0.81	0.154
	0.0	0.73	1.56	2.62	2.5	1.96	1.96	1.31
	0.0	0.79	1.71	2.92	2.73	2.73	2.73	2.73
	0.0	0.69	1.42	2.42	3.46	3.5	3.7	4.12
10	0.0	0.50	0.98	1.46	4.12	4.23	4.5	4.96
	0.0	0.31	0.58	0.73	4.44	4.54	4.77	5.1
	0.0	0.185	0.308	0.385	4.65	4.74	4.88	5.1
		0.085	0.154	0.193	4.75	4.81	4.95	5.08
		0.0385	0.058	0.077	4.8	4.85	4.96	5.08
15		0.0192	0.031	0.0385	4.82	4.86	4.96	5.08
		−0.0193	−0.031	−0.0385	4.82	4.86	4.96	5.08
		−0.0385	−0.0575	−0.115	4.71	4.85	4.96	5.08
		−0.115	−0.162	−0.239	4.7	4.77	4.93	5.08
		−0.231	−0.425	−0.54	4.55	4.65	4.82	5.08
20		−0.385	−0.77	−1.16	4.25	4.35	4.62	5.04
		−0.575	−1.21	−2.0	3.73	3.85	4.2	4.62
		−0.73	−1.54	−2.61	3.05	3.05	3.23	3.31
		−0.77	−1.66	−2.7	2.24	2.2	2.16	1.93
		−0.685	−1.38	−2.27	1.54	1.425	1.155	0.54
25		−0.5	−0.96	−1.42	0.925	0.8	0.47	0.00
		−0.315	−0.58	−0.71	0.52	0.447	0.231	−0.06
		−0.185	−0.32	−0.35	0.27	0.231	0.0835	−0.0395
		−0.085	−0.154	−0.189	0.112	0.116	0.0385	−0.0695
		−0.079	−0.138	−0.17	0.141	0.105	0.033	−0.0695
30		−0.072	−0.124	−0.152	0.131	0.095	0.028	−0.0696
		−0.066	−0.11	−0.135	0.120	0.085	0.023	−0.0697
		−0.06	−0.098	−0.129	0.110	0.075	0.019	−0.0698
		−0.054	−0.085	−0.10	0.10	0.066	0.014	−0.070
		−0.048	−0.072	−0.085	0.09	0.057	0.010	−0.070
35		−0.022	−0.06	−0.069	0.08	0.0475	0.005	−0.070
		−0.0385	−0.046	−0.05	0.07	0.0395	0.0	−0.070
		−0.037	−0.043	−0.046	0.07	0.0395	0.0	−0.070
		−0.036	−0.040	−0.042	↓	↓	↓	↓
		−0.035	−0.038	−0.038				
40	0.0	−0.035	−0.035	−0.035	0.07	0.0395	0.0	−0.070
					.09	.06	.02	.01

* The fields are given for a current excitation ($U = 1$) providing a deflection of approximately 1 centimeter. Field measured in gauss.

$p = x$.

Table 3—Potential Fields*

	$V\ (z,p)$#				$V_r(z,p) = \frac{\Delta V_z}{\Delta p}$			
n	$p = 0$‡	1.0	2.0	3.0	$p = 0$	1.0	2.0	3.0
0	500.0	500.0	500.0	500.0	0.0	0.0	0.0	0.0
	500.0	500.0	500.0	500.0				
	510.0	510.0	510.0	510.0				
	520.0	520.0	520.0	520.0				
	530.0	530.0	530.0	530.0				
5	540.0	540.0	540.0	540.0				
10								
15								
	540.0	540.0	540.0	540.0		0.0	0.0	0.0
20	540.18	540.16	540.12	540.05		−0.03	−0.07	−0.07
	540.31	540.29	540.2	540.10		−0.06	−0.12	−0.10
	540.60	540.54	540.38	540.17		−0.12	−0.19	−0.19
	541.10	541.0	540.73	540.33		−0.18	−0.35	−0.32
	541.97	541.84	541.41	540.63		−0.29	−0.59	−0.71
25	543.38	543.24	512.62	541.56		−0.38	−0.71	−1.51
	545.35	545.26	544.0	544.4		−0.19	−0.39	−2.66
	547.63	547.63	547.67	548.16		0.01	0.18	0.28
28	550.0	550.0	550.0	550.0		0.0	0.0	0.0
40	550.0	550.0	550.0	550.0	0.0	0.0	0.0	0.0

* Fields measured in volts and volts per centimeter.

\# The potentials V_z in the decelerating space between mesh ($n = 28$) and target ($n = 40$) are simply duplicate cards of the mesh potential not used by the computer. The computer has instructions to compute the correct values for the target potential V_T at $n = 40$ when tube contains a field mesh.

‡ $p = r$.

Although measured, the values B_{Fz} and B_{Fr} have been smoothed to conform more accurately to the derivative relation; i.e., the tabulated values agree with the relation

$$B_{Frp} = -\left(\frac{\Delta B_{Fzo}}{\Delta Z}\right)\frac{p}{2}.$$

The potential fields given in Table 3 were measured on a resistance board in the field-mesh region.

The axial distance increments ΔZ_1 between electron-gun aperture ($n = 0$) and field mesh ($n = 28$) are 1 cm in length. The short decelerating space (0.31 cm) between field mesh and target ($n = 40$) is divided into 8 increments ($n = 29$ to 36) that are 0.0375 cm in length and specified by ΔZ_2 and 4 increments ($n = 37$ to 40) that are 0.0025 cm in length and specified by ΔZ_3. The magnetic field is substantially constant in the field-to-target region; potentials are not given because they are calculated by the computer. Potentials and B-field data must be given for meshless tubes.

The starting conditions for an electron at the gun aperture depend on the electron-gun characteristics. Typical parameters for a high-resolution gun are:

V_o = total energy spread (0.25 volt)
V_{ro} = radial emission energy (0 to 0.25 volt, gaussian function)
V_{rg} = radial energy acquired in the electron gun (0.0)
X_o, Y_o = starting coordinates at the gun aperture (centimeters)
μ_o = initial rotation (radians)
U = deflection current factor ($U \approx 1$ for a 1 cm deflection)
V_T = target potential (volts)
N_λ = desired number of nodal points
E = focusing instruction (focus, or no change of focus)

The Coordinates (X,Y) of the Deflected Flux Line

The magnetic flux line or principal path of an electron is calculated by piecing together a large number of incremental sections. Each flux line enters a given section of the (X,Y,Z) space at $Z_{(n-1)}$ with an angle $\tau_{(n-1)}$ to the Z-direction and leaves the section at Z_n at a different angle τ_n that becomes the angle $\tau_{(n-1)}$ in the following section. The angle τ in space is the vector sum of the angles δ and ϵ in the projections of the flux-line section onto the (X,Z) and (X,Y) planes, respectively, as illustrated by Fig. 2. The principal deflection occurs in

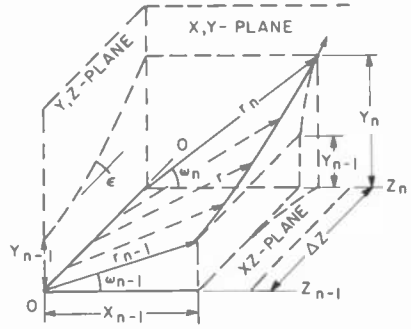


Fig. 2—Section of deflected flux line and projections (broken lines) on the (X,Y), (X,Z), and (Y,Z) planes.

the (X,Z) plane and is caused by the deflection-coil field B_D. Small x and y components are introduced by radial electric fields that are replaced by an equivalent circular magnetic-deflection field B_E that rotates the deflection plane by an angle ω. Entrance and exit angles (δ,ϵ,ω) can be calculated at the data points from the field ratios, whereas the path of the flux line within the increment must be approximated.

High accuracy for interpolation purposes is gained by the constant-curvature approximation illustrated in Figs. 3 and 4(a). The deflection

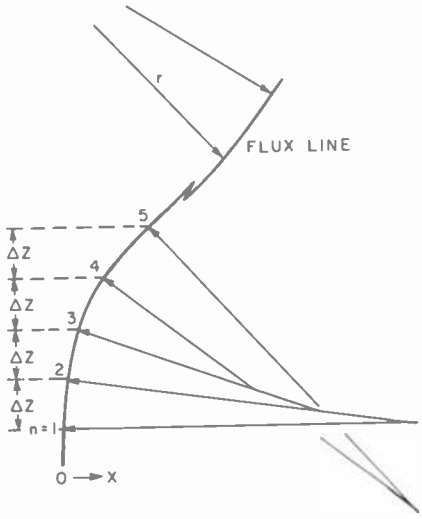


Fig. 3—Arc approximation of flux line between data points.

angle δ_n in an (X,Z) plane projection, for example, is described by

$$\delta_n = \tan^{-1}\left(\frac{B_D}{B_z}\right)_n .$$

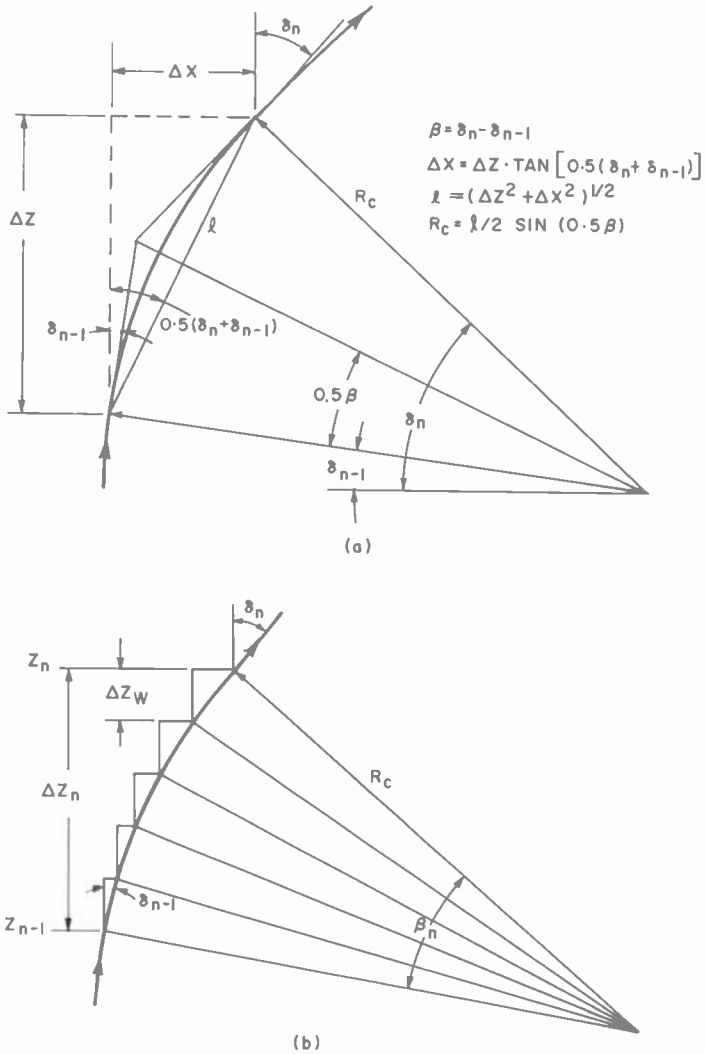


Fig. 4—(a) Detail of arc approximation in ΔZ increment and (b) sub-increments.

The incremental deflection ΔX is given accurately by

$$\Delta X = \Delta Z \tan^{-1} [0.5(\delta_{(n-1)} + \delta_n)], \qquad [3]$$

which determines the coordinate X_n equal to $X_{(n-1)} + \Delta X$ of the exit point. The coordinate Y_n is determined similarly from the deflection angles $\epsilon_{(n-1)}$ and ϵ_n in the (Y,Z) plane projection. Predictions and successive approximations are necessary because the effective fields (B and V) at any one point are vector sums of all components.

The deflection components introduced by radial E-fields are zero at the electron gun and remain zero or negligible until a flux line deflected in the (X,Z) plane approaches the region of the field mesh. The rotation ω by these fields is small and would be only 3° for $r = 3$ cm if the field given in Table 3 for a high-definition return-beam vidicon is replaced by a uniform radial field 10 times stronger ($V_R = -40$ volts) extending over an 8-cm axial distance. It follows that normal V_R fields are essentially y-deflection fields, have a rather small x component, and cause a very small change in the radial distance of the deflected flux line.

The field values and the principal electron-beam path are therefore computed initially without electric fields. The coordinates $(X', 0)$ of the deflected flux line so obtained are then corrected by addition of an equivalent circular B_E field as shown in Fig. 5(a). This field introduces a small rotation ω and changes the coordinates from $(X', 0)$ to (X,Y) with little effect on the values B_{Fz}, B_{Fr}, V_z, and V_r of the fields with axial symmetry. The composite fields B_Z and B_D of Eqs. [1] and [2] must be corrected for incremental distances. This part of the calculations has a feedback loop to permit successive approximation.

Correction of Principal Path for E Fields

In the presence of a strong axial magnetic field, radial electric fields E_r from circular electrodes cause a circular deflection of the principal path of the electron normal to the radius r. The electric field is converted to an equivalent circular magnetic field by equating the respective forces qE_r and $qv_zB_E \times 10^{-8}$ on the electron, yielding B_E equal to $(E_r/v_z) \times 10^8$. Expressed by potentials, B_E, in gauss, is

$$B_{En} = -1.686 \frac{V_{Rn}}{V_{Zn}^{1/2} \cos \tau_n}$$

where τ_n is the angle of the principal path (i.e., the equivalent magnetic flux line in space) to the z-direction.

The equivalent B_E-field is a circular field in the (X,Y) plane, as illustrated by Fig. 5(a) (B_E is zero on-axis and increases with radial distance). The rotation of the deflection plane is counterclockwise for a negative gradient V_r and an axial field directed out of the page. Starting at a point $(X,Y)_{n-1}$, as shown in Fig. 5(b), the incremental

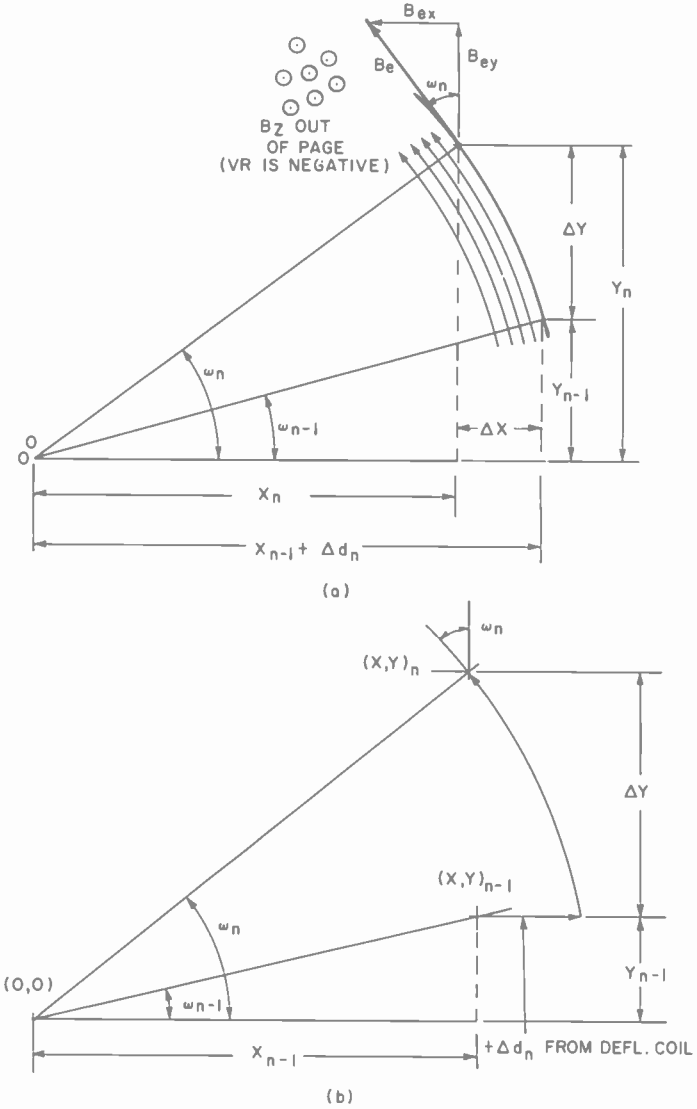


Fig. 5—(a) Circular magnetic-field equivalent of radial E-field from electrode of axial symmetry and (b) deflection components.

deflection $\Delta X'_n$ from the B_D field increases X. The rotation by the equivalent B_E field decreases X and adds a ΔY component. The procedure for calculation of the coordinates $(X,Y)_n$ of the equivalent flux line is given in Appendix 1.

Interpolation of Subincrements

The data-point increments ΔZ are too coarse to allow calculation of the circular-motion component of the electron with adequate accuracy. For this reason, the increments ΔZ are divided into subincrements ΔZ_w equal to $\Delta Z F_w$, where F_w is a fractional multiplier. The value F_{w1} equal to 1/12 is generally satisfactory, but a finer subdivision by a factor F_{w2} less than F_{w1} is needed in the first "loop" of the electron trajectory and in the last increment ΔZ_3 near the target.

The linear subdivision of ΔZ results in nonlinear increments of the incremental deflections ΔX_w and ΔY_w. Fig. 4(b) illustrates that the assumption of constant angular increments $\Delta\beta_n$ is adequate and permits interpolation of the small angles occurring in actual cameras.

The Deviation ($\triangle$X,$\triangle$Y) of the Helical Electron Trajectory

Because of its radial energy, the electron moves away from and back towards the flux line on which it started. In fields of nonuniform flux density, the flux lines in a region adjacent to the principal path are not parallel in space. New flux-line angles (ζ,η) must be computed from the slightly different field values at the electron location when determining the tangential angle (ϕ), the cyclotron radius, the increment of rotation, and the displacement of the electron in an incremental section.

The circular-motion component of the electron trajectory is computed in a separate coordinate system $(\Delta X', \Delta Y')$ normal to the flux-line direction. The origin $(\Delta X' = \Delta Y' = 0)$ coincides with the flux-line coordinates at any value Z. This sliding coordinate system is inclined to the flux-line coordinates by the angle τ. The electron coordinates deviate from the principal path by periodic distance, accumulated as increment sums

$$\sum_{0}^{n,w} \Delta X' \text{ and } \sum_{0}^{n,w} \Delta Y'.$$

An estimate of the incremental change $(\Delta X', \Delta Y')$ in a new w-increment must be made to permit prediction of the electron coordinates

required for interpolation of angles and field values. This prediction, covered in Appendix 2, is made by assuming that the cyclotron radius R_w of the circular motion and the angular increment $\Delta\mu_w$ of the electron rotation do not change appreciably from those computed for the preceding subincrement $(w-1)$.

The flux-line angles $(\zeta,\eta)_w$ at the electron location yield the incremental deflection angles as follows:

$$\beta_w = \zeta_w - \zeta_{(w-1)}$$

and

$$\rho_w = \eta_w - \eta_{(w-1)}.$$

These angles change gradually from zero to their full value at the end of the increment and must be added vectorially to the tangent angle $\phi_{(w-1)}$ of the electron (See Fig. 4(b)). Much accuracy is gained by use of mean values

$$\overline{\beta} = 0.5\ (\beta_w + \beta_{(w-1)})$$

and

$$\overline{\rho} = 0.5\ (\rho_w + \rho_{(w-1)}).$$

These increments change the tangent angle $\phi_{(w-1)}$ of the electron path according to its angle of rotation μ by an inphase component

$$\phi' = \phi_{(w-1)} + \overline{\beta}\cos\mu_{(w-1)} + \overline{\rho}\sin\mu_{(w-1)}$$

and a 90° component

$$\alpha = \overline{\beta}\sin\mu_{(w-1)} - \overline{\rho}\cos\mu_{(w-1)}$$

when it enters the w-section as illustrated by Fig. 6.

Fig. 6(a) shows the cylindrical flux tube in section $(w-1)$ shown enclosed in a box of tangent planes in Fig. 6(b). The electron trajectory leaves plane I at point C with the tangent angle $\phi_{(w-1)}$. Plane II is tilted by the angles β and α to the normal end-face of cylinder $(w-1)$ and is normal to the axis of the cylinder of rotation in section (w). The cyclotron radius R_w of this flux tube is located on plane II as shown, ending at point C. The vector R_w must be advanced or turned back against the exit vector $R_{(w-1)}$ by an angle γ_w to maintain a right angle with the tangent of the electron trajectory at point C, as shown by the shaded rectangular planes in Fig. 6(b) and (c).

It follows from trigonometric relations that the angle γ_w is given by

$$\gamma_w = \tan^{-1}\left(\frac{\sin\alpha}{\tan\phi'}\right). \qquad [4]$$

The solution of this expression has two values. When electrons

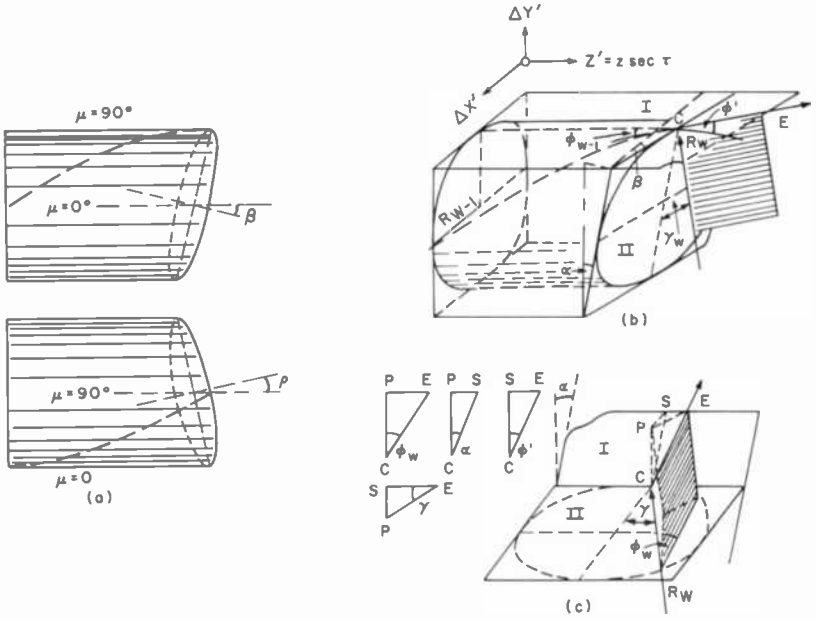


Fig. 6—Section of electron trajectory and flux tube at the junction of two ΔZ increments in a curved field: (a) X,Z and Y,Z projections of flux tube showing the angles β and ρ respectively at end of increment $(n-1)$, (b) flux tube enclosed in tangent planes, and (c) view of end face II, illustrating trigonometric relations of angles (see text).

travel very nearly parallel to a flux line, the angle $\phi_{(w-1)}$ shown in Fig. 6(b) is very nearly zero. At the entrance into a new section, the angle ϕ' may become negative as a result of a negative angle of the flux line. The angle ϕ', however, can never reverse sign unless the axial field is reversed. The correct solution lies on plane *II* extended beyond the top side of plane *I* in Fig. 6(b), and requires that vector R_w turn suddenly around its end point by the angular increment π. Eq. [4] must, therefore, be preceded by the instructions: If ϕ' is positive, use

Eq. [4]; if ϕ' is negative, reverse its sign and compute with the following equation:

$$\gamma_w = \pi + \tan^{-1}\left(\frac{\sin\alpha}{\tan\phi'}\right). \qquad [4a]$$

For equal electron velocities in the preceding and new sections, the tangent angle ϕ_{w1} has the value

$$\phi_{w1} = \cos^{-1}(\cos\phi'\cos\alpha).$$

However, the absolute electron-volt velocity $\overline{V}_w$ may be different in the new section (this effect occurs in accelerating or decelerating fields). The tangent of ϕ_{w1} must therefore be multiplied by the velocity ratio, and ϕ_w is given by

$$\tan\phi_w = \tan\phi_{w1}\left(\frac{\overline{V}_{(w-1)}\cos^2\phi_{(w-1)}}{\overline{V}_w\cos^2\phi_w}\right)^{1/2}$$

which is solved by successive approximation.

The absolute mean volt-velocity of the electron in an incremental section is the mean potential $\overline{V}_w$ given by

$$\overline{V}_w = 0.5\,(V_{zw} + V_{z(w-1)}) + V_o.$$

The mean flux density in the direction of the principal path has the value

$$\overline{B}_w = 0.5\,(B_{zw} + B_{z(w-1)})\sec\tau$$

where

$$\sec\tau = (1 + \tan^2\bar{\xi}_w + \tan^2\bar{\eta}_w)^{1/2}$$

The cyclotron radius can now be computed with

$$R_w = |3.371\sin\phi_w\,(\overline{V}_w)^{1/2}/\overline{B}_w|$$

The vector rotation is computed from the incremental number of "loops" $\Delta N_{\lambda(w)}$ in the principal path length $l = \Delta Z_w \sec\tau$ as follows

$$\Delta N_{\lambda(w)} = \frac{0.0472\ (\Delta Z_w \sec \tau)\overline{B}_w}{\cos \phi_w\ (\overline{V}_w)^{1/2}}$$

and the total vector rotation μ_w is the sum

$$\mu_w = \mu_{(w-1)} + 2\pi\Delta N_{\lambda(w)} - \gamma_w$$

The distance increments $(\Delta X', \Delta Y')_w$ in the coordinate system normal to the flux line are computed from the increment in rotation. These increments are divided by cos ζ and cos η, respectively, to obtain increments $(\Delta X, \Delta Y)_w$ normal to the Z-axis of the camera tube as follows

$$\Delta X_w = -R_w[\sin \mu_w - \sin\ (\mu_{(w-1)} - \gamma_w)]/\cos \zeta$$

and

$$\Delta Y_w = R_w[\cos \mu_w - \cos\ (\mu_{(w-1)} - \gamma_w)]/\cos \eta$$

The change in the flux-line angles from δ,ϵ to ζ,η also requires a correction of the deflection increments ΔX_w and ΔY_w. These corrections move the end position of the electron by the increments

$$\Delta X_c = \Delta Z_w \tan\ [0.5\ (\zeta_w + \zeta_{(w-1)})] - \Delta X_w$$

and

$$\Delta Y_c = \Delta Z_w \tan\ [0.5\ (\eta_w + \eta_{(w-1)})] - \Delta Y_w$$

The total deviation of the electron coordinates from the coordinates (X,Y) of the flux line is finally accumulated by the sums

$$\sum_0^{nw} \Delta X = (\Delta X_w + \Delta X_c) + \sum_0^{n(w-1)} \Delta X \qquad [5]$$

and

$$\sum_0^{nw} \Delta Y = (\Delta Y_w + \Delta Y_c) + \sum_0^{n(w-1)} \Delta Y \qquad [6]$$

These sums complete one w-increment. (The entire calculation could be repeated with computed values for the electron location; however, the gain in accuracy was found to be negligible.) The calculation is repeated for the remaining w-increments and then loops back to compute the flux-line coordinates for the next n-increment.

Focusing Instructions

The program contains 3 multipliers, F_1, F_V, and F_Z. The computer is instructed to multiply the field input tables by these factors when it is desired to change the flux densities or the dimensions of the electron optic. Factor F_Z changes all linear dimensions including parameter values and gradients. The focus factor F_1 is added as a magnetic-field multiplier in essential equations to permit magnetic focusing by consecutive runs without change of data.

The total number of loops N_λ obtained from the input data after the first run with a radial velocity V_r greater than 0 is generally not the desired focus condition. The computer is instructed to first change the field multiplier to F_1 equal to F_1 (N_n/N_λ) where N_n is the desired number of nodal points given as an input constant. During a second run with this factor, the computer is instructed to store the coordinates of deviation Eqs. [5] and [6] existing in the fourth quarter of the final loop in a temporary matrix. The number of nodal points of the second run will be very nearly the desired integral number N_n. This condition, however, is not the optimum focus condition, which is better defined by closest proximity of electron trajectories of specified radial velocity (V_{ro}) having initial rotations μ_o differing by π. Therefore, after the automatically made second run, the computer adds π to the initial angle μ_o and makes a third run. In the last quarter of the final loop, the computer compares the deviation coordinates of the third run with the stored coordinates of the second run until it finds closest proximity. At this point, the ratio of the value $N_{\lambda\pi}$ to the value $N_{\lambda 0}$ computed for the end point of the second run yields a new multiplier for F_1 that is used in a fourth run through the system. The results of the fourth run very closely approximate the fine-focus condition. The fine-focus instruction (second, third, and fourth run) can be repeated. When repetition is made, the initial multiplier F_1 stored in the computer is increased by 1.0005 to assure that a proximity minimum can be found. This "backup" of focus is invoked automatically when the computer cannot find a minimum, i.e., in the case in which the minimum occurs beyond the storage surface.

The program contains instructions that save time by storing the set of field values for the principal path computed for a given focus- and deflection-coil field strength. This method can be used because it is generally desired to run clusters of electron trajectories with different initial conditions (V_{ro}, V_o, μ_o) from a point in the gun aperture for which the main field values along the flux line remain unchanged. Thus, the principal path calculations are bypassed in subsequent runs with identical fields.

Electron Energies, Magnification, and Accuracy of Trajectories

Initial Values for a High-Resolution Electron Gun

The current distribution in the solid angle from a point source in the limiting aperture of the electron gun is well approximated for small apertures by

$$\frac{i}{i_o} \approx \exp - \left\{ \frac{eV_a}{kT} M_g^2 \sin^2 \phi_a \right\} \qquad [7]$$

where $e/kT = 10$ for a cathode temperature T of 1160°K, M_g is the linear magnification in the electron gun, and ϕ_a is the angle of the electron trajectory to the Z axis at the aperture potential V_a. The radial energy V_r of an electron at potential V is, in general,

$$V_r = (V + V_o) \sin^2 \phi \qquad [8]$$

where V_o is the maximum emission energy of the electrons passing through the aperture. With this substitution and $e/kT = 10$, Eq. [7] can be written

$$\frac{i}{i_o} = \exp - \left\{ \frac{10\, M_g^2 V_r}{1 + V_o/V_a} \right\}. \qquad [9]$$

The current drops to 1% for an exponent value of −4.6; a practical cutoff value.

High-resolution electron guns have a magnification M_g of 1.35. At an operating potential of 500 volts, the exponent of Eq. [7] yields the maximum energy spread

$$V_o = V_{ro\,\max} = 0.252 \text{ volts.}$$

A set of typical radial electron energies from a point in the aperture computed with Eq. [8] and [9] is listed in Table 4.

Landing Angles and Magnification of Electron Optic

The theoretical landing angle ϕ_T for paraxial electrons is given by

$$\sin \phi_T = [(V_{ro}/M_o^2)/(V_T + V_o)]^{1/2}$$

where M_o is the linear magnification of the focussing electron optic and V_T is the target potential. The magnification M_o is given by

$$M_o = [V_{ro}/(V_T + V_o)]^{1/2}/\sin \phi_T. \quad [10]$$

Thus, M_o can be computed after a computer calculation of one electron trajectory has solved for the landing angle ϕ_T.

Table 4—Radial Energies and Relative Current for a High-Resolution Electron Gun*

ϕ/ϕ_{max}	Radial Energy (V_{ro})	Relative Current (i/i_o)	ϕ_o (radians)#
1/6	0.0070	0.888	0.00014
2/6	0.028	0.6005	0.00056
3/6	0.063	0.3166	0.00126
4/6	0.112	0.13	0.00224
5/6	0.175	0.0416	0.00350
1.0	0.252	0.01	0.00504

* $V = 500$ volts, $M_g = 1.35$.

The values of $\phi = \tan^{-1} [V_{ro}/(V_a + V_o - V_{ro})]$ are the tangent angles at the aperture.

Radial Aberration in Axial Electron Beams

Given the magnification M_o, the radial landing distance r of electrons from an axial source can be computed analytically by means of the equation[1]

$$r = \Delta r_1 + \Delta r_2 \quad [11]$$

$$\Delta r_1 = \frac{0.5L_1}{(M_o^2 V_D)^{3/2}} V_{ro}^{1/2}(V_{ro,f} - V_{ro}) \quad [12]$$

$$\Delta r_2 = \frac{2L_2}{M_o V_m} V_{ro}^{1/2} \left[\left(V_T + V_o - \frac{V_{ro}}{M_o^2} \right)^{1/2} - \left(V_T + V_o - \frac{V_{ro,f}}{M_o^2} \right)^{1/2} \right]$$

where L_1 is the drift-space length (aperture to mesh) in centimeters, L_2 is the decelerating space (mesh to target) in centimeters, V_D is the drift potential,* V_m is the mesh potential, and $V_{ro,f}$ is the radial energy

* The drift potential is computed from the mean velocity as follows:

$$V_D^{1/2} = \frac{1}{n_1} \sum_{1}^{n_1} (V_n)^{1/2}, V_{n1} = V_m$$

of the electron focused on the target. Eq. [12] gives the chromatic aberration in the drift space L_1 and Eq. [13] gives the chromatic aberration in the decelerating space L_2.

Eq. [12] is exact for constant B-fields in L_1 (i.e., $M_o = 1$) and is a good approximation when the major deviation from the final B fields is limited to a small initial part of the total drift distance L_1. The approximation remains satisfactory for an initial portion equal to $0.36L_1$. Eq. [13] is accurate for a constant B field in L_2 which is true for camera tubes with a field mesh.

These equations have been added to the computer program and serve as a built-in test for accuracy of the radial landing distance $r = (\Sigma\Delta X^2 + \Sigma\Delta Y^2)^{1/2}$ of axial electrons. A focusing instruction specifies $V_{ro,f}$ and initiates calculation of the magnification and radial distance with ϕ_T by means of Eqs. [10], [11], [12], and [13]. The function $r(V_{ro})$ is computed automatically after the first focusing instruction by Eq. [11] for values $V_{ro} = 0.01V_oK^2$, $K = 1,2,3,\ldots 10$ and the preferred focus condition $V_{ro,f} = 0.3V_o$. Good accuracy of trajectory calculations requires "double precision" in the RCA Spectra-75 computer. A subdivision into $F_{w1} = F_{w2} = 20$ yields very high accuracies (see section on Interpolation of Subincrements given previously).

F_w is increased to 40 in the last increment ΔZ_3 in which the potential decreases to the small value V_T of the target. The factors F_{w1} and F_{w2} are read in with other common input data, such as the number of data points and multipliers F_1, F_V, and F_Z.

Use of the Multipliers F_1, F_V, and F_Z

An electron optic (and camera) may be scaled up or down to cover different image formats. The electron optic used for the examples in this paper covers a 50×50 millimeter format. The performance of a camera that covers a 25×25 millimeter format can be computed from the same field data; the dimensions are reduced to one-half by the scale factor F_Z equal to 0.5. When the potentials (and electron velocities) are held constant ($F_V = 1.0$), the magnetic field intensities must be doubled (F_1 is multiplied by 2) to maintain four loops of focus for which the particular electron optic is corrected. A computer printout demonstrates that all distances, including deflections and radial aberrations are then reduced to exactly one half, whereas all angles (δ, ϵ, ϕ) and the magnification M_o remain constant. Thus, the point-spread function of the electron optic decreases to one half, and the total resolution in the format dimension remains constant. Operation with constant magnetic fields requires a reduction of all potentials to one fourth (F_V

equal to 0.25) and the resolution of the electron optic decreases to roughly one sixth; electron landing angles (ϕ) remain the same.

Other combinations can be explored easily (without change of field data) by specifying multipliers F_Z and F_V and letting the computer find the correct B-field multiplier from the focus requested for a given number of nodal points.

Accuracy of Trajectory Calculations

The arc interpolations minimize the cumulative error of the incremental calculation. The following examples are for a high-resolution electron optic of 28.3-cm focal distance with 4 nodal points. The circular motion of 3 electrons starting with μ_o equal to 0 at the aperture of a 4½ inch return-beam vidicon is shown in Fig. 7(a) for the first loop. Landing approaches of six electrons (fourth loop) are shown in Fig. 7(b). The landing point coordinates are compared in Table 5 and Fig. 8 with the analytic solutions computed in the program for the focus conditions $V_{ro,f}$, $0.8V_{ro,f}$, and $V_{ro,f} = 0.3\ V_o$; the latter value yields the highest axial modulation-transfer function. It is evident that the computer focus is slightly in error, focusing for $V_{ro,f}$ equal to approximately 0.057 volt instead of 0.063 volt; the focus error can be decreased by adjustment of the input factor F_1.

The magnification M_o computed from the landing angle and Eq. [10] is 0.5804; the value increases only slightly (to 0.5806) for the highest radial energy, $V_{ro} = 0.252$. M_o can also be computed from the landing-point radius of an electron started from an off-axis point in the gun aperture ($X_o = 5$ micrometers) as shown in the last row of Table 5. The value obtained is M_o equal to 0.57, a value about 1.6% lower than computed by means of Eq. [10], indicating a computational error.

A computational error is also indicated when an axial electron does not return to the axis at a focal point. For a subdivision F_{w1} equal to 1/12, the electrons miss the axis by 0.14 micrometer. The finer subdivision F_{w1} equal to 0.05 used for the sample calculation reduces this error to approximately 0.012 micrometer at the fourth focal point. The small numerical error is proportional to the scale factor of the electron optic, i.e., the relative error is constant.

The total rotation or "No. of loops" ($\mu/2\pi$) is not an integral multiple of 2π and the paths of the electrons are distorted circles because of the non-uniform axial B-field described in Table 1. The electrons rotate in different quadrants according to the starting angle μ_o; the rotation changes only the signs and not the values of the coordinates. The correct sign for a quadrant can be lost in the computer and requires special instructions.

Table 5—Landing Data for Axial Electrons *

X_o, $Y_o = 0.0$		$F_{w1} = 0.05$			$V_T = 1.0$		$V_{ro,f} = 0.063$	Analytic solution for $M_o = 0.5804$ $R_1 + R_2$ for $V_{ro,f}$	
V_{ro}	No. of loops	$N\lambda$	Rn	ϕn	$\Sigma\Delta X$	$\Sigma\Delta Y$	r		
0.007	3.956	3.913	67.6	0.1292	0.3809	−0.09701	0.393	0.3954	0.063
								0.3056	0.0505
0.028	3.956	3.913	135.2	0.2606	0.4504	−0.1132	0.4644	0.4963	0.063
								0.3168	0.0505
0.063	3.957	3.914	202.8	0.3968	−0.07899	0.03577	−0.08671	0.000	0.063
								−0.2693	0.0505
0.112	3.958	3.914	270.4	0.5414	−1.519	0.4363	−1.581	−1.417	0.063
								−1.776	0.0505
0.175	3.959	3.916	388.0	0.6998	−4.234	1.181	−4.396	−4.121	0.063
								−4.570	0.0505
0.252	3.960	3.917	405.4	0.883	−8.714	2.382	−9.034	−8.581	0.063
								−9.119	0.0505
	Off-axis electron; $X_o = 5\ \mu m$, $Y_o = 0$								
0.00001	3.914	3.912	1.639	0.003123	2.745	−0.7666	2.851		
	$M_o = 2.851/5 = 0.57$								

* All lengths in micrometers; ϕ is in radians.

The aberration of deflected electrons depends critically on the shape of the field functions,[1] and accuracy of calculations cannot be tested analytically. Exact agreement of computed and measured characteristics is unlikely because a real camera is not as precise in concentricity, spacing, and flatness of electrodes as is the mathematical model. A slight curvature of the faceplate or field mesh, for example, may be corrected during optimization of the camera fields by a reverse-curvature component; the same component causes aberrations in the more exact computer model.

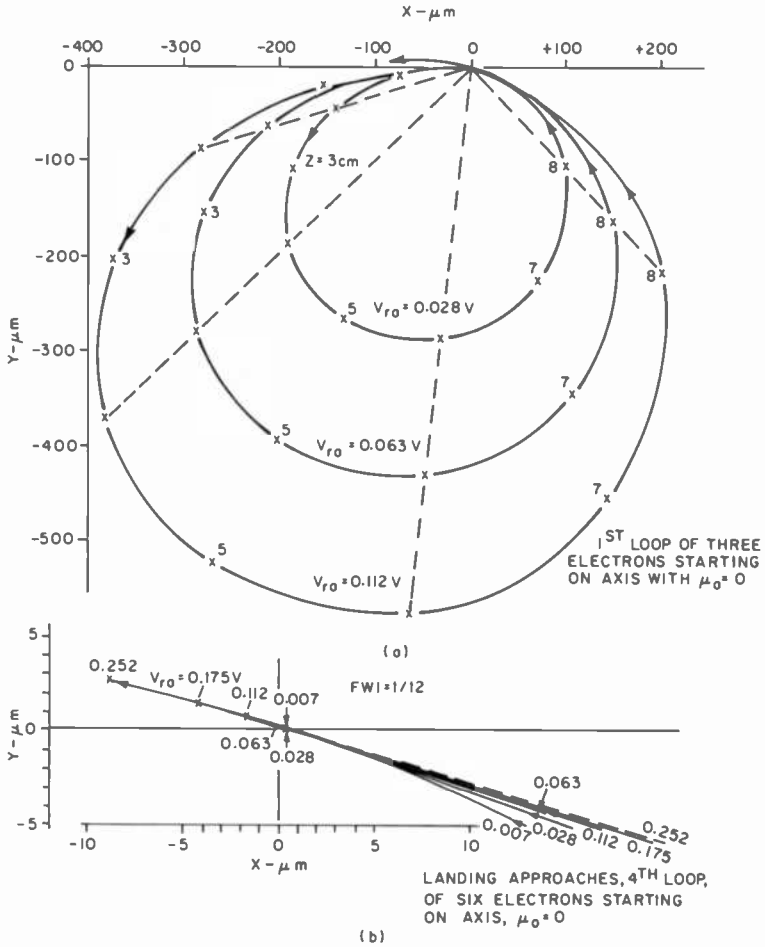


Fig. 7—(a) X,Y projection of first loop of focus computed for three axial electrons in a high-definition camera tube; (b) landing approaches of six axial electrons in 4th loop of focus for the fields given in Tables 1 through 3.

Optimization of fields by observation of camera performance is essential to correct for frequency-dependent second-order effects, such as eddy-current loading of the deflection coils by shields, tube electrodes, and other conductors; these effects are not accounted for in electrostatic or magnetic field calculations or measurements. Improper electron landing caused by zero axial velocity at the target is more pronounced in calculations than in a real camera in which the floating target potential (V_T) increases to meet the beam.

The trajectories of deflected electrons are very complex as illustrated in Fig. 9 by the coordinate deviations $\Sigma(\Delta X, \Delta Y)$ for a 1.913-cm deflection of four point-source electrons. Best focus occurs 25 micrometers in front of the target with a radial spread of 4 micrometers (for V_{ro} equal to 0.063 volt); the spread increases to 10 micrometers at the target. The computational error may be larger than for the undeflected beam and, in addition, the trajectories are sensitive to data errors and axial misalignment of the magnetic and electric fields.

Acknowledgments

The computer program was set up for a numerical analysis of aberrations in high-resolution return-beam vidicon cameras, sponsored by the

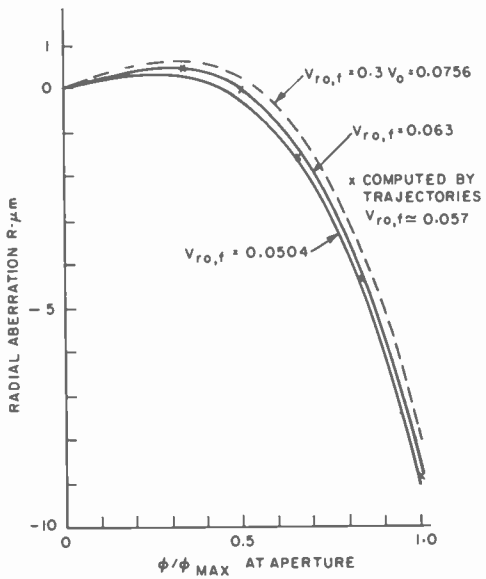


Fig. 8—Analytic solutions (solid lines) of landing point spread in high-definition camera tube and landing points (X) of electron trajectories computed with 820 increments.

Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Ohio under Contract AF33(615)-5347. Much assistance and helpful advice in writing the program was given by G. P. Kirkpatrick of RCA Computer Control Systems, Harrison, N.J.

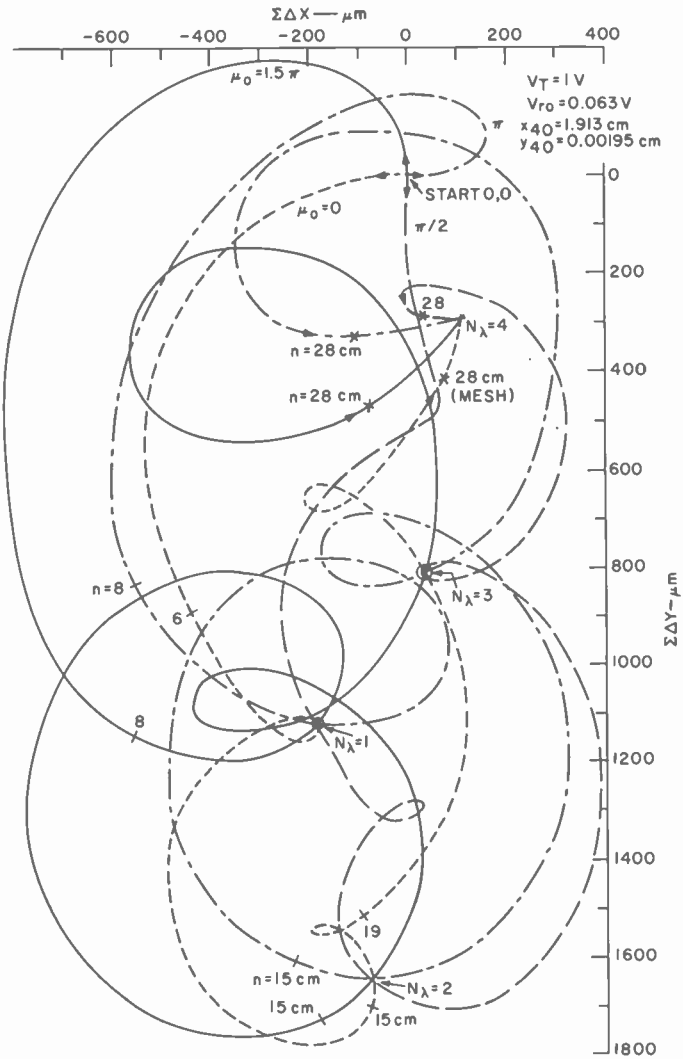


Fig. 9—Circular motion component of four point-source electrons deflected to $X_{40} = 1.913$ centimeters on the target. The principal path is collapsed to the point 0,0.

Appendix 1—Calculation of Coordinates $(X,Y)_n$ of the Equivalent Flux Line

Predict: $X_n = X_{(n-1)}$

$Y_n = Y_{(n-1)}$

Compute: $R_{rn} = (X_n^2 + Y_n^2)^{1/2}$

$$B_{En} = -1{,}686\ V_{rn} \sec \tau_n / [F_1 (V_{Zn} + V_o)^{1/2}] \quad [14]^*$$

$$B_{EX,n} = -B_{En} Y_n / R_{rn} \quad [15]$$

$$B_{EY,n} = B_{En} X_n / R_{rn} \quad [16]$$

$$\epsilon = \tan^{-1} [(g B_{Frn} Y_n + B_{Ey,n}) / B_{Zn}] \quad [17]^{**}$$

$$\Delta y = \Delta Z \tan [0.5(\epsilon_n + \epsilon_{n-1})] \quad [18]$$

$$Y_n = Y_{n-1} + \Delta Y \quad [19]$$

$$\delta_p = \tan^{-1} [B_{dn} + g B_{dn} (X_n - X_n') + B_{Ex,n}] / B_{Zn} \quad [20]$$

$$\Delta X = \Delta Z \tan [0.5(\delta_{p,n} + \delta_{p(n-1)})] \quad [21]$$

$$X_n = X_{n-1} + \Delta X \quad [22]$$

The first term in Eq. [17] is the radial field component of the focus coil. The second term in Eq. [20] corrects the deflection-coil field B_{dn} for the change of X caused by the rotation; B_{dn} was interpolated for the distance X'_n (X'_n being computed without the B_E field), and changes slightly with X as shown in Table 2.

Compute:

$$\omega_n = \tan^{-1} (Y_n / X_n) \quad [23]$$

$$\sec \tau_n = (1 + \tan^2 \delta_p + \tan^2 \epsilon)^{1/2} \quad [24]$$

The calculations required by Eqs. [14] through [24] are repeated until the predicted increment ΔY_p from a preceding calculation agrees within a specified tolerance with the computed value ΔY_n.

Appendix 2—Interpolation of Fields at the Electron Location

The predictions are as follows:

---

* F_1 is the multiplier used to change focus for all B-fields as explained in the section on "Use of the Multipliers F_1, F_r, F_z". F_1 cancels out in ratios of B-fields.

** The prefix g indicates a field gradient.

$$\mu_p = \mu_{(w-1)} + 2\pi\Delta N_\lambda - \gamma \qquad [25]^*$$

$$\sum_0^{n,w} \Delta X_p = \sum_0^{n,(w-1)} \Delta X - [R_n(\sin \mu_p - \sin(\mu_{(w-1)} - \gamma))/\cos\zeta] + \Delta X_c \qquad [$$

$$\sum_0^{n,w} \Delta Y_p = \sum_0^{n,(w-1)} \Delta Y + [R_n(\cos \mu_p - \cos(\mu_{(w-1)} - \gamma))/\cos\eta] + \Delta Y_c \qquad [$$

The second and third terms are the predicted increments. The predicted electron coordinates $(X,Y)_w$, radial distance r_w and rotation ω_w are given by:

$$X_w = S_{Xw} + \sum_0^{n,w} \Delta X_p \qquad [28]\dagger$$

$$Y_w = S_{Yw} + \sum_0^{n,w} \Delta Y_p \qquad [29]\dagger$$

$$r_w = (X_w^2 + Y_w^2)^{1/2} \qquad [30]$$

$$\omega_w = \tan^{-1}(Y_w/X_w) \qquad [31]^{**}$$

Three of the fields at the electron location can be determined from stored values by a correction for radial distance

$$B_{Zw} = B_{Z(n-1)} + F_w w(B_{Zn} - B_{Z(n-1)}) + gB_{Zw} \cdot \text{Dist} \qquad [32]$$

$$V_{Zw} = V_{Z(n-1)} + F_w w(V_{Zn} - V_{Z(n-1)}) + gV_{Zw} \cdot \text{Dist} \qquad [33]$$

$$V_{rw} = V_{r(n-1)} + F_w w(V_{rn} - V_{r(n-1)}) + gV_{rw} \cdot \text{Dist} \qquad [34]$$

where $\text{Dist} = (r_w - X'_w)$ [35]

The equivalent B_E fields are computed with the predictions $\tau_w = \tau_{w-1}$, $\phi_w = \phi_{w-1}$ as follows,

---

* The rotation is counterclockwise for the direction of B_z specified in Fig. 5(a) for the equivalent B_E field.

** For printout only, needs subroutine to establish quadrant.

† The first terms are the coordinates of the principal path accumulated as sums.

$$B_{Ew} = -1.686\ V_{rw} \sec \tau_w / [F_1 (V_{Zw} + V_o)^{1/2} \cos \phi_w] \quad [36]$$

$$B_{XEw} = -B_{Ew} (Y_w / r_w) \quad [37]$$

$$B_{YEw} = B_{Ew} (X_w / r_w) \quad [38]$$

The flux line angles ζ and η at the predicted point $(x,y)_w$ are computed from the total interpolated deflection fields in the X and Y direction. The X deflection field of the principal path is

$$B_{dw} = B_{Zw} \tan \delta_w \quad [39]$$

This field must be corrected similarly to Eq. [20] in Appendix 1 for the difference $(X_w - X_w')$. The correction is made by using Eq. [40] for (X_w) positive and Eq. [41] for (X_w) negative.

$$\zeta_n = \tan^{-1}\{[B_{dw} + gB_{dw}(X_w - X_w') + gB_{Fr}(X_w - X_w') + B_{XEw}]/B_{Zw}\} \quad [40]$$

$$\zeta_w = \tan^{-1}\{[B_{dw} - gB_{dw}(X_w + X_w') + gB_{Fr}(X_w - X_w') + B_{XEw}]/B_{Zw}\} \quad [41]$$

The Y-deflection field is the sum of a radial-focus coil-field component and the B_E field component, which determines the angle

$$\eta_w = \tan^{-1}[(gB_{Fr} Y_w + B_{YEw})/B_{Zw}] \quad [42]$$

The flux-line angle in space has the value

$$\sec \tau_w = (1 + \tan^2 \zeta + \tan^2 \eta_w)^{1/2}. \quad [43]$$

References:

[1] O. H. Schade, Sr., "Electron Optics and Signal Readout of High-Definition Return-Beam Vidicon Cameras," **RCA Review**, Vol. 31, pp 60-119, March 1970.

Analysis and Optimization of A Field-Emitter Array

Jules D. Levine
RCA Laboratories, Princeton, N. J.

Abstract—A simplified electrostatic analysis is made of the electric field and macroscopic current density associated with a dense array of tungsten pins. In practice, the array is used as a cold-cathode field emitter. This analysis is used to determine the optimum geometry for maximizing the current density of the entire array. It is shown that at the optimum condition, the current is the current per pin, multiplied by 0.37 times the number of pins in the array.

Although the field-emission characteristic of an isolated sharp-pointed tungsten pin is well known from field-emission microscopy,[1,2] the combined effect of a dense array of these pins is, at present, unknown. Cline[3] has measured field emission from an array of tungsten points in a nickel matrix. One might suppose that the electrostatic fields of each pin in the array may interfere in some way with the others so as to substantially reduce the total current, but no calculations of pin interference have been made, to the author's knowledge.

The purpose of this note is to perform a theoretical analysis of the electric field at each pin tip in terms of the pin array geometry, and of the optimum geometry for maximizing the current density of the entire array. It will be shown that, under certain conditions, the decrease in current per point due the interference effect noted above can be less than a factor of 3. Certain simplifying approximations will be made, so that one must only consider a few controlling dimensionless variables.

Consider the pin array geometry shown in Fig. 1. The pins are presumed arranged on a square array of side a, and are mounted in, or otherwise fastened to, a substrate matrix that can be a metal or semiconductor (insulators give charging problems). Each pin is taken to be a right circular cylinder of height h and radius R, and each is capped with a hemisphere of the same radius. Reproducible fabrication of tungsten pins in such an array seems to be possible.[3-5] The coordinate origin $(x,y,z) = (0,0,0)$ is located at a typical hemisphere

base, as shown. At $z = L$ there is a metallic plate that serves as the grounded collector (anode). It collects field-emitted current from the array of pin tips, all at a negative potential V_o with respect to ground. (Voltage drops along the pin are normally negligible and are ignored here.)

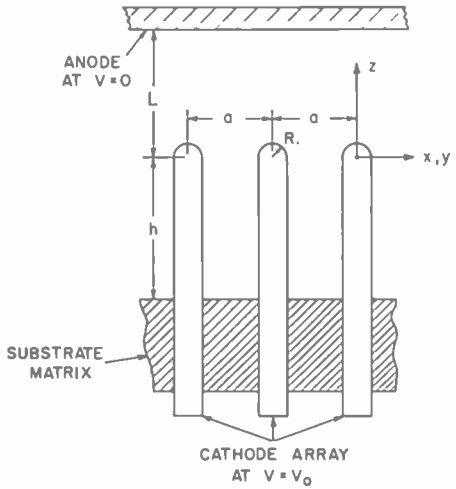


Fig. 1—Geometry of field-emitter array. Materials are available[3,4] wherein $R \sim 0.5\ \mu$, $a \sim 5\ \mu$, and $h \sim 50\ \mu$.

Assume that near the pin tip, the potential $V(x,y,z)$ may be adequately simulated by placing an "effective charge" q at (0,0,0). Then the present problem of the electrostatic potential about a field-emitter array becomes *identical* with the problem of the electrostatic potential about a square array of point charges. The latter problem has already been solved.[6] It was shown that the voltage $V(x,y,z)$ can be expanded in a complete Fourier series of harmonic modes with indices l, m;

$$V = V_{oo}(z) + \sum_{l,m=0}^{\infty'} V_{lm}(x,y,z), \qquad [1]$$

where the primed sum excludes $l = m = 0$. For field-emission applications, the voltage will be conveniently evaluated along each pin axis (0,0,z), for $z \geqslant R$. Then the analysis shows that

$$V_{oo}(z) = 4\pi q(L - z)/a^2, \qquad [2]$$

and

$$V_{lm}(0,0,z) = [q/a(l^2 + m^2)^{1/2}] \cdot \{\exp[-(2\pi/a)(l^2 + m^2)^{1/2}z] - \exp[-(2\pi/a)(l^2 + m^2)^{1/2}(2L - z)]\}. \quad [3]$$

The second exponential term is the contribution of the electrostatic image in the anode of the pin tips; it is negligible in most practical electrode calculations, and will be henceforth dropped. The infinite sum in Eq. [1] can be approximated by an integral as follows:

$$(l^2 + m^2)^{1/2} \rightarrow \xi$$

$$\sum_{l,m=0}^{\infty'} \rightarrow \int_0^{\infty} 2\pi\xi d\xi. \quad [4]$$

The approximation is best when $2\pi z/a \ll 1$; that is, for small tip radii compared to interpin spacing. After the integration, it can be shown that Eq. [1] reduces to

$$V = [4\pi q(L - z)/a^2] + [q/z]. \quad [5]$$

Thus, these approximations give a value of V that appears to be a simple *superposition* of two voltages: that of a one-dimensional capacitor and that of a Coulomb potential, each with an effective charge. To determine the heretofore unknown effective charge q in a self-consistent way, a boundary condition is imposed such that $V = V_o$ at $z = R$, where V_o is the applied (measured) voltage. Substituting this boundary condition in Eq. [5] gives

$$q = V_o\{[4\pi(L - R)/a^2] + [1/R]\}^{-1}. \quad [6]$$

If desired, the effective (measured) capacitance can be computed as $C = q/V_o$. From examination of Eq. [6], the equivalent capacitance circuit is simply a one-dimensional capacitor plus a point capacitor, arranged in series.

The electric field $F(x,y,z)$ on the line $(0,0,z)$ can be computed from Eq. [5] via $F = -\nabla V$. The value of $F(x,y,z)$ at $(0,0,R)$ will be called the effective field F_o; it is

$$F_o = q[(4\pi/a^2) + (1/R^2)]. \quad [7]$$

Elimination of q from Eqs. [7] and [6] gives

$$F_o = V_o[(4\pi/a^2) + (1/R^2)]/\{[4\pi(L - R)/a^2] + [1/R]\}. \quad [8]$$

After multiplying numerator and denominator by R^2, the above reduces to

$$F_o = (V_o/R)[1 + (4\pi R^2/a^2)]/\{1 + [4\pi(L - R)R/a^2]\}. \quad [9]$$

For all practical field-emitted array geometries, $4\pi R^2/a^2 \ll 1$ and $R \ll L$, so Eq. [9] reduces to

$$F_o = F_1/(1 + \rho). \quad [10]$$

Here

$$F_1 = V_o/R \quad [11]$$

is the ideal field at one pin tip,[7] and

$$\rho = 4\pi LR/a^2 \quad [12]$$

is a dimensionless parameter that is a measure of the interference of the other pins. For ρ very large, the effective field of the array is reduced, via [Eq. [10]. Clearly, it is most desirable when $\rho \leqslant 1$. This places a constraint on the geometrical parameters of pin spacing a, pin radius R, and electrode spacing L, such that

$$a^2/(4\pi LR) \geqslant 1. \quad [13]$$

This is, however, a necessary, but not sufficient criterion for the lack of pin interference. The more severe criterion is found below, after examining the optimization of current density.

For one isolated pin, the *microscopic* current density j(A/cm^2 of emitting pin tip area) is given in the simplest form as

$$j = BF_1^2 \exp(-F_r/F_1), \quad [14]$$

where, for tungsten tips, $B = 0.34 \times 10^{-6}$ (if F_1 is expressed in V/cm), and the "reference" field is $F_r = 6.5 \times 10^8$ V/cm. Simplifications used in Eq. [14] are that the field-effect factor k for one isolated pin is unity,* that the Schottky image-force-barrier lowering is ignored,[1,2]

* For a discussion of the field-effect factor, which is dependent on the geometrical details of the emitter contour, see Ref. (1), pp. 45ff and Ref. (2) pp. 98ff.

and that space-charge limitations are ignored.[1,2] To the same order of simplicity, the macroscopic current density J (A/cm^2 of flat array area) is

$$J = (\pi R^2)(a^{-2})BF_o^2 \exp(-F_r/F_o). \quad [15]$$

In Eq. [15], the nominal emitting area per pin πR^2 (tpically 10^{-9} cm^2) and the number of pins per macroscopic square centimeter a^{-2} (typically[3,4] 10^6/cm^2) are explicitly included. Also the effective field F_o is used in Eq. [15], instead of the ideal field per pin F_1 as used in Eq.

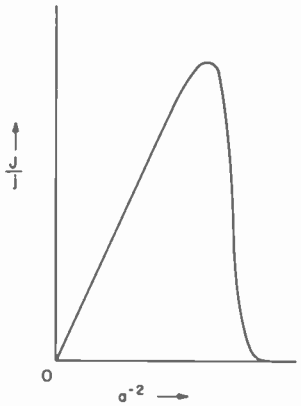


Fig. 2—Ratio of macroscopic current density to microscopic current density (J/j) for a field-emitter array as a function of pin density (a^{-2} pins/cm^2). The plot is schematic, and the optimum values of a^{-2} and J/j are given in the text.

[14]. Taking the ratio of Eqs. [15] and [14] and using $\rho = 4\pi LR/a^2$ from Eq. [12], one has

$$(J/j) = (\pi R^2/a^2)(1+\rho)^{-2} \exp[-(4\pi LRF_r/a^2F_1)]. \quad [16]$$

This is schematically plotted in Fig. 2, where (J/j) is the ordinate and the pin density a^{-2} is the abcissa. Note that the abcissa is proportional to ρ.

To optimize Eq. [16] with respect to the pin density (a^{-2}), set $\partial(J/j)/\partial(a^{-2}) = 0$. The term $(1+\rho)^{-2}$ is not significant in this regard since the optimum (J/j) is at very small ρ (see below). Carrying out the above optimization yields

$$4\pi LRF_r/(a^2F_1) = 1, \quad [17]$$

or equivalently,

$$\rho F_r / F_1 = 1. \qquad [18]$$

Using the formula for F_1 in Eq. [11], Eq. [17] becomes

$$a/R = (4\pi F_r L/V_o)^{1/2}. \qquad [19]$$

Recognizing that $F_r \gg F_1$ for practical field-emission geometries, one sees from Eq. [18] that $\rho \ll 1$, as claimed above. Thus the condition for optimum pin density to give maximum macroscopic current density J creates a sufficient constraint in Eq. [19] between the pin radius R, the electrode spacing L, and the applied voltage V_o. At the maximum (J/j), the exponent in Eq. [16] is exactly unity, and the multiplication of current over that of one pin is the number of pins present in the array divided by $e = 2.72$. Thus, the optimum condition can be a condition of negligible pin interference.

Suppose, for example, that operating conditions of a particular field-emitter-array vacuum diode are desired to be optimized at $V_o = 1000$ V, and $L = 10^{-2}$ cm. Then Eq. [19] yields $a/R \sim 300$ as the design criterion for pin geometry.

In conclusion, a very simple analysis has been made of optimum pin geometry for a field-emitter array. Effects that have been neglected, or glossed over, should be considered in more detail. For example, the equipotentials represented by the three-dimensional version of Eq. [5] have not been adequately treated. Preliminary plots show that the equipotential pin contours can be fairly well matched to the geometrical pin contours of Fig. 1, except that pin optimization carried out here does not explicitly include the pin height, h.

Lacking other theoretical criteria at the present time, the optimum geometry as expressed in Eq. [19] may be found useful, at least qualitatively, for designing realistic field-emitter array geometries.

References:

[1] R. Gomer, **Field Emission and Field Ionization**, Harvard Univ. Press, Cambridge, Mass. (1961).

[2] W. P. Dyke and W. W. Dolan in **Adv. In Electronics and Electron Physics**, Vol. VIII, ed. by L. Marton, Academic Press, N.Y. (1956), pp. 89-179.

[3] H. E. Cline, "Multineedle Field Emission from the Ni-W Eutectic," **J. Appl. Phys.**, Vol. 41, p. 76, Jan. 1970.

[4] R. J. Gerdes, A. T. Chapman, and G. W. Clark, "Refractory Oxide-Metal Composites: Scanning Electron Microscopy and X-ray Diffraction of Uranium Dioxide-Tungsten," **Science**, Vol. 167, p. 979, Feb. 13, 1970.

[5] M. D. Watson, D. N. Hill, and A. T. Chapman, **J. American Ceramic Soc.**, Vol. 48, p. 2 (1970).

[6] J. D. Levine, "The 3-D Electrostatic Potential Surrounding a Square Array of Surface Charges," **Surface Sci.**, Vol. 10, p. 313, 1968.

Contrast Characteristics of X-Ray Images

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

Abstract—The relative importance of various x-ray image parameters, such as the physical properties of the x-ray converting screen, the spectral emission characteristics of the x-ray source, and the x-ray attenuation of the specimen under observation, is considered. It is shown that very smooth thin screens (25-μm thick), which have substantially higher resolution and thus better contrast gradient than conventional x-ray screens, are capable of displaying a 2% thickness variation of a 0.6-cm-thick aluminum layer. No distinct difference in contrast has been found among the three most widely used x-ray screens—ZnS:Ag, ZnS•CdS:Ag, and $CaWO_4$:W.

Introduction

In addition to their application in science and medicine, x-rays are very useful in the field of material inspection. In this application, an x-ray source can be used as an illuminator, often in connection with an x-ray image-intensifier tube.[1] The quality of such a system depends on its ability to display contrast between two elements under observation and to resolve small details. The contrast and resolution of an x-ray image are largely determined by the electrical and physical properties of the x-ray converting screen, the spectral-emission characteristics of the x-ray source, and the x-ray attenuation of the specimen under observation. This paper shows the relative importance of these three parameters in the formation of the x-ray image and evaluates the merits of several phosphors for industrial requirements.

Absorption of X-Rays

When x-rays are traversing a homogeneous substance, their intensity is reduced by a constant fraction per unit length. Thus, the beam intensity of the radiation, I, may be given by

$$I = I_o \exp\left\{-\left(\frac{\mu}{\rho}\right)\rho t\right\}, \qquad [1]$$

where I_o is the x-ray intensity incident on the material, and ρ, t, and μ are the density, thickness, and absorption coefficient of the material, respectively.

The linear absorption coefficient μ of a substance is a function of the x-ray wavelength, the atomic numbers of the constituent elements, and the physical state of the substance. The mass absorption coefficient (μ/ρ), is a more commonly used term and has the useful property of being independent of the physical state of the material. (Tables of μ/ρ values are available in the literature for pure elements.[2])

Eq. [1] indicates that a test object that has a varying thickness or contains nonhomogeneity in its structure translates a homogeneous x-ray beam into local intensity variations. The beam-intensity pattern produced by these variations is made visible by conversion of the x-rays into visible light by use of a fluorescent screen. In this manner, the structural details of the object under test are displayed in terms of brightness variation.

X-Ray Phosphor Screens

The conversion of x-ray energy to light occurs in two stages. In the first stage, the absorbed x-ray energy is converted to high-energy electrons. In the second stage, the kinetic energy of the released electrons is converted to light. The fraction of x-ray absorption, α_p, of a screen at a particular wavelength is given by the following expression:

$$\alpha_p = 1 - \exp\left\{-\left(\frac{\mu}{\rho}\right)\rho t\right\}, \qquad [2]$$

where μ/ρ and t are the mass absorption coefficient and thickness of the phosphor screen, respectively.

The mass absorption coefficient of a material consisting of several elements is the sum of the mass absorption coefficients of the individual elements multiplied by their appropriate weight fractions.[3,4] This rule follows directly from the fact that the mass-absorption coefficient is independent of the physical state of the absorbing material (i.e., whether the element exists in a pure state or as a compound constituent).

The mass absorption coefficient (μ/ρ) of a compound consisting of

n elements is given by

$$\left(\frac{\mu}{\rho}\right) = W_1\left(\frac{\mu}{\rho}\right)_1 + W_2\left(\frac{\mu}{\rho}\right)_2 + \cdots W_n\left(\frac{\mu}{\rho}\right)_n, \qquad [3]$$

where W_n and $(\mu/\rho)_n$ are the weight fraction and the mass absorption coefficient, respectively, of the nth element of the compound.

Absorption coefficients have been calculated for the three most widely used x-ray phosphors: zinc sulfide (ZnS), zinc cadmium sulfide (ZnS•CdS), and calcium tungstate ($CaWO_4$) by the procedure outlined below. The mass compositions and densities of the three phosphors are as follows:

Phosphor	Mass Composition	Density grams/cm³
ZnS	0.67Zn + 0.33S	4.1
ZnS • CdS	0.341Zn + 0.382Cd + 0.277S	4.46
$CaWO_4$	0.142Ca + 0.634W + 0.224O	6.06

The linear absorption coefficients μ of these three phosphors are calculated as follows:

$$\mu_{ZnS} = 4.1\left[0.67\left(\frac{\mu}{\rho}\right)_{Zn} + 0.33\left(\frac{\mu}{\rho}\right)_{S}\right]$$

$$\mu_{ZnS\cdot CdS} = 4.46\left[0.341\left(\frac{\mu}{\rho}\right)_{Zn} + 0.382\left(\frac{\mu}{\rho}\right)_{Cd} + 0.277\left(\frac{\mu}{\rho}\right)_{S}\right]$$

$$\mu_{CaWO_4} = 6.06\left[0.142\left(\frac{\mu}{\rho}\right)_{Ca} + 0.634\left(\frac{\mu}{\rho}\right)_{W} + 0.224\left(\frac{\mu}{\rho}\right)_{O}\right]$$

Fig. 1 shows the absorption coefficients for these three phosphors as a function of wavelength, as calculated by use of published data[2] for the value of (μ/ρ). With μ known, the fraction of x-ray absorption α_p of a solid layer screen can be determined by Eq. [2]. Fig. 2 shows the results of these calculations.

Because the density of the microsettled screens approximately equals the density of a solid-layer screen, it is justifiable to carry the calculations for a solid-layer without considerable error.

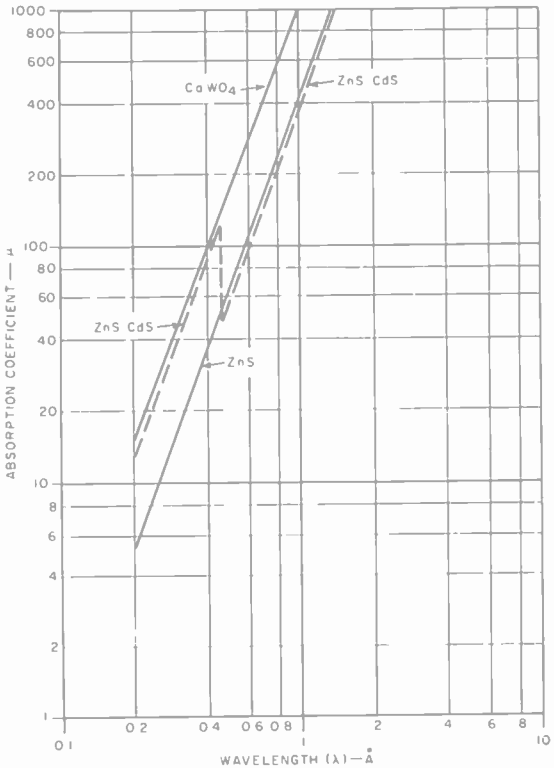


Fig. 1—Adsorption coefficient as a function of wavelength for three phosphor compounds.

As mentioned earlier, the energy of an absorbed x-ray quantum is partially converted into the kinetic energy of the released electron. The kinetic energy E_k of the released electron is given by

$$E_k = E_p - E_b, \qquad [4]$$

where E_p is the energy of the x-ray photon and E_b is the energy required for the complete removal of the electron from a particular shell of the converting atom.

In general, when the x-ray photon is energetic enough to liberate an electron from the K-shell, the probability is high that the K-electron is released; otherwise, electrons are freed from subsequent shells.

It follows from Eq. [4] that, for identical x-ray photon energies, the electrons released by different atoms of a compound possess a

different amount of kinetic energy. Therefore, it is important to know the partial absorption of an element in the compound. From Eq. [3], the partial absorption α_j of an element at a particular wavelength expressed as a fraction of total absorption of the compound is given by

$$\alpha_j = \frac{W_j \left(\frac{\mu}{\rho}\right)_j}{W_1 \left(\frac{\mu}{\rho}\right)_1 + W_2 \left(\frac{\mu}{\rho}\right)_2 + \cdots W_j \left(\frac{\mu}{\rho}\right)_j + \cdots W_n \left(\frac{\mu}{\rho}\right)_n}. \quad [5]$$

For the first electron, the conversion efficiency of x-ray energy η into electron energy is given by

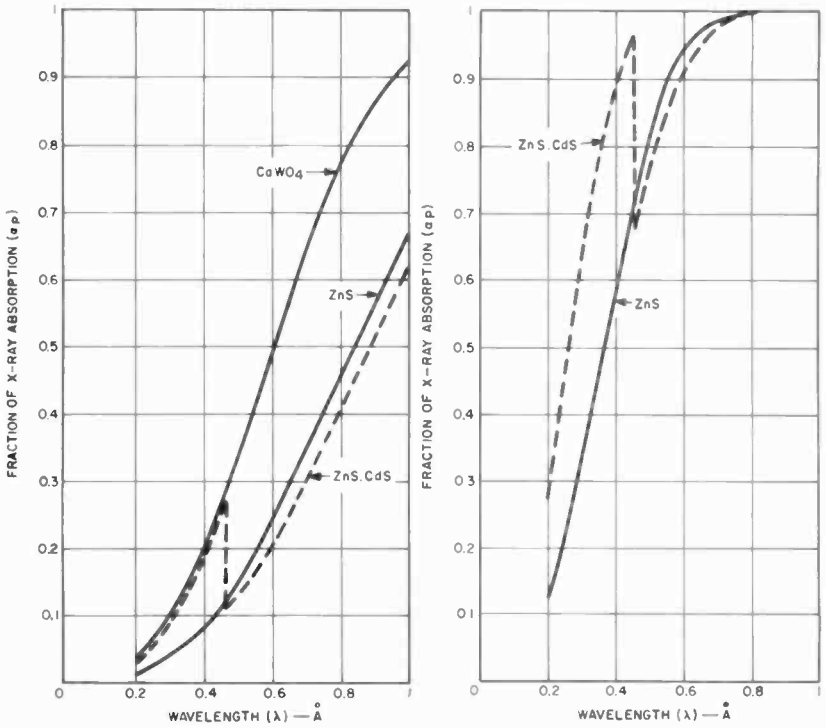


Fig. 2—Fraction of x-ray adsorption of the phosphor screens as a function of wavelength: (a) 2.5×10^{-3} cm-thick phosphor screen; (b) 2.5×10^{-2} cm-thick phosphor screen.

$$\eta = \frac{\alpha_1 E_{k1} + \alpha_2 E_{k2} + \cdots \alpha_j E_{kj} + \alpha_n E_{kn}}{E_p}, \qquad [6]$$

where E_{kj} is the kinetic energy of electrons released by the jth element of the compound.

Values of η for the three x-ray phosphors can be expressed as follows:

$$\eta_{ZnS} = \frac{0.67\left(\frac{\mu}{\rho}\right)_{Zn} E_{k,Zn} + 0.33\left(\frac{\mu}{\rho}\right)_{S} E_{k,S}}{E_p\left[0.67\left(\frac{\mu}{\rho}\right)_{Zn} + 0.33\left(\frac{\mu}{\rho}\right)_{S}\right]}$$

$$\eta_{ZnS\cdot CdS} = \frac{0.341\left(\frac{\mu}{\rho}\right)_{Zn} E_{k,Zn} + 0.382\left(\frac{\mu}{\rho}\right)_{Cd} E_{k,Cd} + 0{,}277\left(\frac{\mu}{\rho}\right)_{S} E_{k,S}}{E_p\left[0.341\left(\frac{\mu}{\rho}\right)_{Zn} + 0.382\left(\frac{\mu}{\rho}\right)_{Cd} + 0.277\left(\frac{\mu}{\rho}\right)_{S}\right]}$$

$$\eta_{CaWO_4} = \frac{0.142\left(\frac{\mu}{\rho}\right)_{Ca} E_{k,Ca} + 0.63\left(\frac{\mu}{\rho}\right)_{W} E_{kW} + 0.224\left(\frac{\mu}{\rho}\right)_{O} E_{KO}}{E_p\left[0.142\left(\frac{\mu}{\rho}\right)_{Ca} + 0.63\left(\frac{\mu}{\rho}\right)_{W} + 0.224\left(\frac{\mu}{\rho}\right)_{O}\right]}$$

where:[5]

$$E_{k,Zn} = (E_p - 9.662) \text{ keV for } \lambda < 1.29 \text{ Å},$$
$$E_{k,S} = (E_p - 2.469) \text{ keV for } \lambda < 5 \text{ Å},$$
$$E_{k,Cd} = (E_p - 26.712) \text{ keV for } \lambda < 0.465 \text{ Å},$$
$$E_{k,Cd} = (E_p - 4.019) \text{ keV for } \lambda = (0.465 - 2.97) \text{ Å},$$
$$E_{k,Ca} = (E_p - 4.038) \text{ keV for } \lambda < 2.83 \text{ Å},$$
$$E_{k,W} = (E_p - 12.094) \text{ keV for } \lambda = (0.18 - 1.025) \text{ Å},$$
$$E_{k,O} = (E_p - 0.532) \text{ keV for } \lambda < 23.3 \text{ Å}.$$

Numerical values calculated from these equations are shown in Fig. 3.

For purposes of energy conversion, only the energy of the first electron has been taken into account. When the excited atom containing a surplus of E_b energy returns to the ground state, it either radiates an x-ray quantum or emits an Auger electron.[6] The first process has a high probability in heavy elements, the second process in light elements. However, in relatively thin screens, both of these

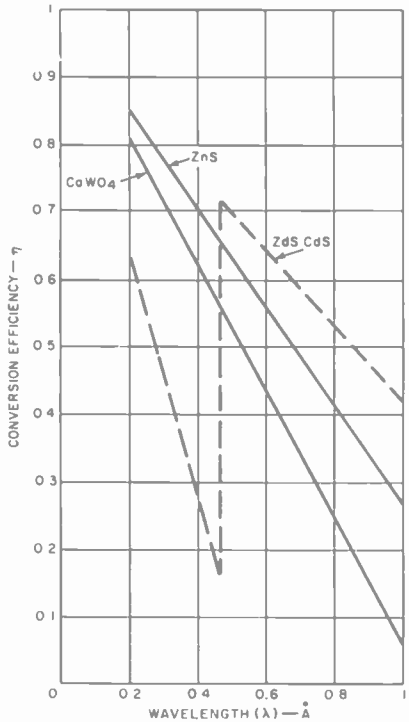


Fig. 3—Conversion efficiency of x-ray energy into electron energy for thin screens.

processes are so inefficient that they can be neglected. The reason for this inefficiency is the low velocity of the Auger or second-generation electrons (electrons produced by the x-ray radiation of the screen) and the low self-absorption of the screens. The curves of Fig. 3, therefore, are representative of the energy conversion of a thin screen. When the conversion efficiency of x-ray energy into electron energy is known, the conversion efficiency for light is obtained as the product of η and the absolute electron-energy-to-light conversion efficiency of the screen.

X-Ray Source

One of the commonly used x-ray sources is an x-ray tube having a tungsten target. Fig. 4 shows the continuous x-ray spectrum[7] of this target for various tube voltages. As the tube voltage decreases, the energy radiation shifts towards longer wavelengths. Fig. 5 shows the x-ray spectrum change when the radiation passes through sheets of aluminum. Aluminum attenuates the radiation and shifts it toward shorter wavelengths. Both of these shifts are important for contrast.

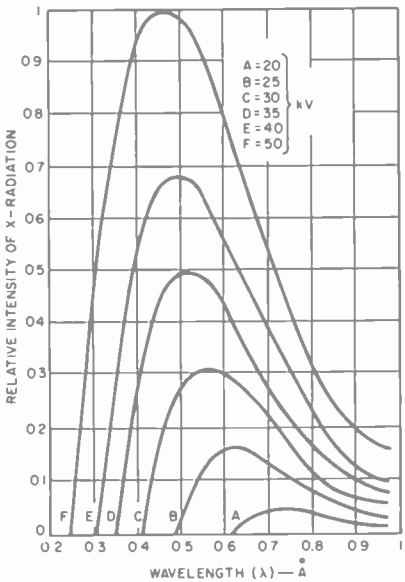


Fig. 4—The continuous x-ray spectrum of a tungsten target for various tube voltages.

When the x-ray tube is operated at low voltages, a greater portion of x-ray energy is absorbed by the phosphor screen, as shown in Fig. 2. On the other hand, the spectrum change shown in Fig. 5 provides contrast percentage that exceeds that of the thickness variation of the specimen, as Fig. 2 indicates.

X-Ray Image Contrast

It is customary to express the x-ray image contrast, C, by the following equation:

$$C = \frac{B_1 - B_2}{B_1}, \qquad [7]$$

where B_1 and B_2 are the luminance of two information elements under observation. When the contrast represent the x-ray intensity variation

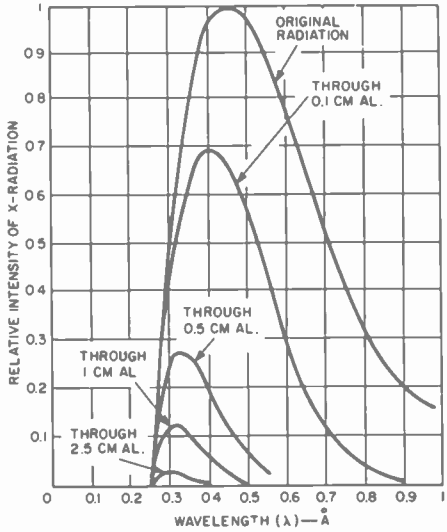


Fig. 5—The change of x-ray spectrum after the radiation passed through sheets of aluminum.

produced by the thickness variation of a specimen, it may be expressed as

$$C = \frac{\int I_o(\lambda)\alpha_p(\lambda)\eta(\lambda)\exp\{-\mu t_1\}d\lambda - \int I_o(\lambda)\alpha_p(\lambda)\eta(\lambda)\exp\{-\mu(t_1+\Delta t}}{\int I_o(\lambda)\alpha_p(\lambda)\eta(\lambda)\exp\{-\mu t_1\}d\lambda}$$

$$= \frac{\int I_o(\lambda)\alpha_p(\lambda)\eta(\lambda)\exp\{-\mu t_1\}(1-\exp\{-\mu\Delta t\})d\lambda}{\int I_o(\lambda)\alpha_p(\lambda)\eta(\lambda)\exp\{-\mu t_1\}d\lambda}, \qquad [8]$$

where $I_o(\lambda)$ is the intensity of the x-ray radiation, μ is the absorption coefficient of the specimen, and t_1 and $(t_1 + \Delta t)$ are specimen thicknesses corresponding to B_1 and B_2. Fig. 4 shows the relative values of $I_o(\lambda)$ for several values of x-ray tube voltages.

One of the requirements in industrial application of x-ray fluoroscopy is to display a thickness variation of 2% of 0.6-cm-thick aluminum layer with an observable contrast. The objective in the study involves selection of the best possible approach that will satisfy this requirement. For this reason, contrast was calculated by alteration of Eq. [8] parameters. The results of this calculation are shown in

Table 1—X-ray Image Contrast Produced by the Thickness Variation of Aluminum Sheets for Several X-Ray Screens at Various X-Ray Tube Voltages

Group No.	Screen	t_s	V_s	t_1	$(+\Delta t)$	C
I	ZnS • CdS	2.5×10^{-3}	50	0.600	0.012	0.034
	ZnS	2.5×10^{-3}	50	0.600	0.012	0.036
	$CaWO_4$	2.5×10^{-3}	50	0.600	0.012	0.035
II	ZnS • CdS	2.5×10^{-3}	30	0.600	0.012	0.06
	ZnS	2.5×10^{-3}	30	0.600	0.012	0.058
III	ZnS • CdS	2.5×10^{-2}	50	0.600	0.012	0.032
	ZnS	2.5×10^{-2}	50	0.600	0.012	0.034
IV	ZnS • CdS	2.5×10^{-3}	50	0.300	0.006	0.025
	ZnS • CdS	2.5×10^{-3}	50	1.00	0.020	0.044
V	ZnS • CdS	2.5×10^{-3}	50	0.600	0.200	0.42
	ZnS	2.5×10^{-3}	50	0.600	0.200	0.44
	ZnS • CdS	2.5×10^{-3}	30	0.600	0.200	0.62
	ZnS	2.5×10^{-3}	30	0.600	0.200	0.61

V_s = x-ray tube voltage in kilovolts.
t_s = thickness of phosphor screen in centimeters.
t_1 = thickness of aluminum sheet in centimeters.
Δt = thickness variation of the aluminum in cm (only in positive direction).

Table 1. Group No. 1 of Table 1 shows the comparison of the three most widely used x-ray screens. In this group, no appreciable contrast difference results from the use of different screens. Group No. II shows that a reduction of the x-ray tube voltage increases the value of the contrast by a significant amount. The reason for this change is the shift in the spectral-emission characteristics of the x-ray source, as shown in Fig. 4. Group No. III indicates that a tenfold increase in screen thickness hardly changes the contrast. Group No. IV shows how the contrast is changed by alteration of the specimen thickness within the thickness variation at the 2% level.

Another important application of the x-ray fluoroscopy is in the field of inspection of small components. For example, in the case of transistors, it is desirable to observe details comparable to a 0.0125-cm-thick wire enclosed in a 0.03-cm-thick envelope. If both the wire and the envelope are made of iron, calculated contrast values are shown approximately in Table 2. Table 2 also contains the contrast for a one-mil wire (2.54×10^{-3} cm).

Table 2—X-Ray Image Contrast C of an Iron Sheet Having a Nominal Thickness of 0.06 cm Overlaid with an Iron Pattern of Thickness Δt (X-Ray Tube Voltage is 50 kV).

Δt	C	Screen
0.0125	0.386	ZnS
0.0125	0.364	ZnS•Cds
0.0125	0.38	$CaWO_4$
0.0025	0.097	ZnS
0.0025	0.09	ZnS•Cds
0.0025	0.095	$CaWO_4$

It can be concluded that an adequate contrast is present that resolves the 2% thickness variation of a 0.6-cm-thick aluminum sheet by application of any one of three screens even at 50-kilovolt x-ray-tube voltage. Moreover, results show that a relatively thin screen produces practically as much contrast as a considerably thicker screen. This effect is very favorable from a slightly different point of view. For example, it is known that if two regions differ in luminance, it is difficult, if not impossible, for an observer to recognize the difference unless the luminance gradient or contrast gradient is sufficiently great. However, high-contrast gradients can be obtained only by the application of relatively thin screens, which have high resolution.

X-Ray Absorption of Phosphor Screens in the Presence of a Specimen

As mentioned previously, the three phosphors under consideration display no distinct difference in contrast. However, inspection of a thick specimen indicates absorption of a large portion of x-ray energy in the substance. Because of this absorption, the number of x-ray photons incident on the phosphor screen is greatly reduced, and the number of x-ray photons absorbed in the screen has a still lower value. Consequently, resolution may be limited by quantum noise. Therefore, it is important to consider the absorption of the phosphor screens that provides the greatest number of absorbed x-ray photons. The fraction of

x-ray absorption, a, of a phosphor screen in the presence of a specimen with a thickness t is given by

$$a = \frac{\int I_o(\lambda)\alpha_p(\lambda) \exp\{-\mu t\} d\lambda}{\int I_o(\lambda) \exp\{-\mu t\} d\lambda} \qquad [9]$$

Values of a calculated for several phosphors are shown in Table 3. This table shows that $CaWO_4$ is the most efficient absorber. However, ZnS•CdS absorbs only 17% less than the $CaWO_4$ screen; the difference is insignificant with respect to the signal-to-noise ratio.

Table 3—X-Ray Adsorption a of Phosphor Screens In the Presence of an Aluminum Sheet. The x-Ray Tube Voltage is 50 kV. The Thickness of Screens is 2.5×10^{-3} cm. The Thickness of Aluminum Sheet t Is Given in cm.

t	a	Screen
0	0.214	ZnS
0	0.228	ZnS•CdS
0	0.413	$CaWO_4$
0.600	0.024	ZnS
0.600	0.049	ZnS•CdS
0.600	0.059	$CaWO_4$

Quantum Noise Limitation of Contrast

The expression for the signal-to-noise ratio produced by the statistical variation of x-ray photon flux absorbed by two information elements of the phosphor screen may be expressed as follows:[8,9]

$$k = \frac{S}{N} = \frac{N_a - N_b}{\sqrt{N_a + N_b}}, \qquad [10]$$

where N_a and N_b are the average numbers of x-ray quantum per storage time per information element from the highlight and lowlight areas, respectively. For reliable detection[8-10] k is about 5; the storage time of the human eye is 0.2 second.

For example, if the observed contrast is 2.5%, $N_b = 0.975\ N_a$. Substitution of this N_b value and $k = 5$ in Eq. [10] gives

$N_a \approx 80{,}000$ quantum per 0.2 second $= 400{,}000$ quantum per second

When the resolution is 5 line-pairs per millimeter, the area of a white bar of a standard "Air-Force" pattern is given by

$$A = 5W^2 = \frac{5}{4R_l^2} = 5 \times 10^{-4} \text{ cm}^2.$$

The absorbed photon flux per cm^2 then becomes

$$400{,}000/(5 \times 10^{-4}) = 8 \times 10^8 \text{ photons per cm}^2 \text{ per sec.}$$

In addition to this information, if the data of Table 3 are used, the necessary minimum output power of the x-ray tube can be calculated as required for the inspection of a 0.6-cm-thick aluminum specimen.

Table 4—X-Ray Image Contrast of Two Adjacent Aluminum Elements Having Thicknesses of 0.635 cm and 0.835 cm ($V_e = 50$ kV).

Screen	C
ZnS	0.46
ZnS•Cds	0.47
$CaWO_4$	0.48

Experiments

The usefulness of the equations developed for the contrast calculations was tested with the following parameters:

$t_1 = 0.635$-cm aluminum sheet

$\Delta t = 0.2$-cm aluminum sheet

$V_e = 50$ kilovolts.

First, the phosphor screen was excited through an aluminum sheet having a thickness $t_1 = 0.635$ cm, and the corresponding luminance B_1 was measured. Second, the luminance B_2 was measured by excitation of the phosphor screen through an aluminum layer with $t_1 + \Delta t = 0.635 + 0.2 = 0.835$-cm. The constrast was obtained by use of Eq. [7]. Table 4 shows the results of these measurements. These results are in reasonable agreement with the theoretical values of contrast given in Table 1, Group No. V.

Additional experiments conducted to resolve a 2% thickness variation of a 0.6-cm aluminum sheet have demonstrated that any of the three phosphor screens displays a contrast large enough to be observable by the eye.

Conclusion

Results of this study have shown that relatively thin screens ($t_s = 25$ μm) have displayed thickness variations as low as 2% of a 0.6-cm-thick aluminum layer with adequate contrast. The advantages of a thin screen are high resolution and improved contrast gradient. These advantages however, can only be utilized with very smooth screens, which poses a very strict requirement on screens. For example, a luminance variation less than 1% requires a screen-thickness variation lower than 0.25 μm for $t_s = 25$ μm, which, in turn, requires the use of a screen particle size smaller than 0.25 μm. In cases where the contrast and resolution are limited by the quantum noise, use of phosphors made of heavy elements is recommended because these phosphors provide high absorption of x-ray radiation. In some cases, contrast may be enhanced by reduction of x-ray tube voltage. However, a decreased x-ray tube voltage also limits the maximum available x-ray energy, and may cause contrast limitation by quantum noise.

References:

[1] Clark, G. L., **The Encyclopedia of X-Rays and Gamma-Rays,** Reinhold Publishing Co., New York, 1963, pp. 503-527.

[2] **Handbook of Chemistry and Physics,** 45th Ed., pp. E 69-70, Chemical Rubber Publishing Co., Cleveland, Ohio, 1964-1965.

[3] **American Institute of Physics Handbook,** 2nd Ed., Chap. 8, p. 96.

[4] Richardson, D. A., "The X-ray and γ-ray Absorption Coefficients of a Number of Glasses," **British J. of App. Phys.,** Vol. 8, Jan. 1957.

[5] **Handbook of Chemistry and Physics,** 45th Ed., p. E73, 1964-1965.

[6] **American Institute of Physics Handbook,** 2nd Ed., Chap. 8, p. 92.

[7] Ulrey, C. T., "An Experimental Investigation of the Energy in the Continuous X-ray Spectra of Certain Elements," **Phys. Rev.,** Vol. 11, p. 401, 1918.

[8] Csorba, I. P., "Resolution Limitations of Electromagnetically Focused Image-Intensifier Tubes", **RCA Review,** Vol. 30, No. 3, p. 36 (1969).

[9] "The Low-Light-Level Performance of the Image Intensifier Orthicon," **Advances in Electronics and Electron Physics,** Vol. 13, p. 183 (1960).

[10] Tol, T., Oosterkamp, W. J., and Proper, J., "Limits of Detail Perceptibility in Radiology Particularly When Using the Image Intensifier," Philips Res. Repts., Vol. 10, p. 141, (1955).

[11] Coltman, J. W., "Scintillation Limitations to Resolving Power in Imaging," **Opt. Soc. Am.,** Vol. 44, no. 3, p. 234 (1954).

Accurate Measurement of Isolation in Three-Port Circulators

Robert L. Ernst
RCA Advanced Communications Laboratory, Somerville, N. J.

Abstract—A method has been developed to accurately determine isolation of a three-port circulator. This method may be implemented with equipment that measures only amplitude of transmission through a terminated circulator, although having a system that also measures phase shift simplifies the measurement considerably. Isolation determined by this method is independent of the VSWR of the sliding load used to terminate the nonisolated port. Hence, the reflection coefficient of the termination may be either greater or less than the true isolation of the circulator without affecting measurements. It is possible to apply a similar technique to correct the VSWR measurement of the circulator, but the correction would be very small.

Introduction

Many three-port circulators have previously been measured to have extremely high isolation over very narrow bandwidths. Isolation in excess of 50 dB has been reported.[1] It has been realized that while these results represent an accurate measurement of the power ratios at the two ports in question, these ratios result from contributions other than the circulator isolation. The other most significant contribution results from the power reflected by the termination used at the nonisolated port. When these two contributions are of the same approximate magnitude and opposite phase, a cancellation occurs resulting in a very high apparent isolation. Therefore, this measured isolation can be very different from the true isolation, and can also be very dependent upon the VSWR of the termination used at the nonisolated port. This paper presents a measurement technique that gives a true value for the circulator isolation independent of load VSWR.

Mathematical Discussion of Circulator Properties

A three-port circulator as shown in Fig. 1 may be described by a set of three scattering parameter equations in the form:

$$b_i = \sum_{j=1}^{3} S_{ij} a_j, \text{ where } i = 1, 2, 3 \qquad [1]$$

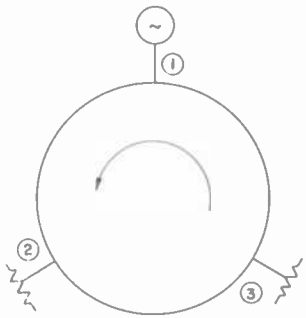


Fig. 1—Circulator model.

where a_j and b_i are, respectively, the incident and reflected power waves at port i. In this discussion. a signal source will be connected to port 1, a termination having a reflection coefficient ρ_2 connected to port 2, and a termination with reflection coefficient ρ_3 connected to port 3. The voltage waves at ports 2 and 3 are then

$$a_2 = \rho_2 b_2 \qquad [2a]$$

$$a_3 = \rho_3 b_3 \qquad [2b]$$

Circulators are usually characterized by assuming that the input reflection coefficient is given by the ratio b_1/a_1; the insertion loss by the ratio b_2/a_1; and the isolation by the ratio b_3/a_1. This assumption is accurate only when perfectly matched terminations are used, i.e., when both ρ_2 and ρ_3 are zero.

Expressions for ratios b_1/a_1, b_2/a_1, and b_3/a_1 can be examined as a function of ρ_2 and ρ_3 to determine how accurately they represent the true values of the circulator properties and how strongly they are influenced by the reflection coefficients ρ_2 and ρ_3.

Evaluation of the complete expressions for these ratios can be

simplified by using the following approximations:

1. The circulator has low insertion loss so that

$$|S_{21}| \approx |S_{32}| \approx |S_{13}| \approx 1. \qquad [3]$$

2. The circulator has high isolation so that

$$|S_{12}| \approx |S_{23}| \approx |S_{31}| \ll 1. \qquad [4]$$

3. The VSWR of the circulator and terminations are low, resulting in

$$|S_{11}| \approx |S_{22}| \approx |S_{33}| \approx |\rho_2| \approx |\rho_3| \ll 1. \qquad [5]$$

4. The isolation coefficients in Eq. [4] are approximately of the same order of magnitude as the reflection coefficients of Eq. [5].

Invoking these approximations results in the following equations

$$\frac{b_1}{a_1} \approx S_{11} + S_{21}\, S_{32}\, S_{13}\, \rho_2\, \rho_3 + S_{21}\, S_{12}\, \rho_2 + S_{31}\, S_{13}\, \rho_3 \qquad [6]$$

$$\frac{b_2}{a_1} \approx S_{21}\,(1 + S_{22}\, \rho_2) \approx S_{21} \qquad [7]$$

$$\frac{b_3}{a_1} \approx S_{31} + S_{21}\, S_{32}\, \rho_2. \qquad [8]$$

Eq. [6] shows that the ratio b_1/a_1 is a close approximation to the input reflection coefficient S_{11}. A dependence upon the terminations does exist, but this dependence is small.

Eq. [7] shows that the ratio b_2/a_1 is essentially equal to the insertion loss of the circulator and that virtually no dependence on the terminations exists.

Eq. [8] shows that the ratio b_3/a_1 is a combination of the true circulator isolation and a reflection from the termination at the non-isolated port. These two terms can easily be of the same order of magnitude so that the ratio b_3/a_1 can in fact be a very poor indication of circulator isolation. When these two components are of opposite phase, a very high value of isolation when expressed in dB would appear to exist. If, on the other hand, the reflection component is significantly

greater than the isolation component, the true isolation is masked. A technique for measuring the true isolation independent of the reflection coefficient is discussed next.

Circulator Measurement Techniques

More accurate information about circulators can be determined if the contributions of the reflections from the terminations can be eliminated. It is not possible to obtain loads that are perfectly matched, so a technique is needed that permits the value of the circulator property to be extracted from the measurement. This may be implemented by using sliding terminations rather than fixed ones.

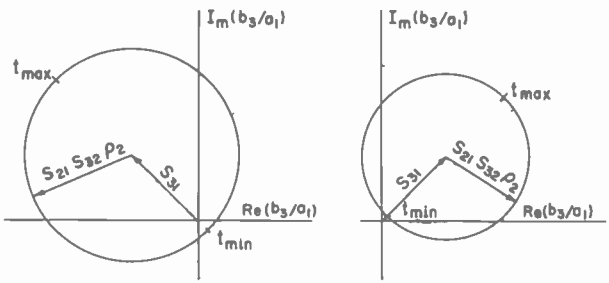


Fig. 2—Circular locus generated by sliding termination.

The remainder of this paper concentrates on isolation measurements because the need for correction is greatest in this case. If a sliding termination is placed on the nonisolated port and the ratio b_3/a_1 measured, the true isolation can be easily and accurately determined. As the termination is moved, it is found that the locus of b_3/a_1 is a circle. The vector representation of Eq. [8] showing the generated circle is given in Fig. 2. The vector from the origin to the center of this circle is the true isolation of the circulator while the radius of the circle is essentially equal to the reflection coefficient of the sliding load. In principle, this is very similar to methods used to find the true directivity of directional couplers.[2,3]

Several techniques may be used to measure these components. The selection of technique is determined principally by the availability of test equipment.

If the measuring equipment can only determine magnitude of the

ratio b_3/a_1, the sliding termination is adjusted to give minimum and maximum readings t_{min} and t_{max}, respectively. These readings occur at points on the circle that intersect a straight line through the center of the circle and the origin as shown in Fig. 2. If the circle does not enclose the origin then

$$|S_{31}| = \frac{|t_{max}| + |t_{min}|}{2} \qquad [9]$$

and

$$|S_{21}\, S_{32}\, \rho_2| = \frac{|t_{max}| - |t_{min}|}{2}. \qquad [10]$$

If the circle does enclose the origin, then

$$|S_{31}| = \frac{|t_{max}| - |t_{min}|}{2} \qquad [11]$$

and

$$|S_{21}\, S_{32}\, \rho_2| = \frac{|t_{max}| + |t_{min}|}{2}. \qquad [12]$$

However, if amplitude measuring equipment is used, it is not possible to determine with one sliding load which set of equations is applicable. By using a second sliding load having a reflection coefficient different than the first, a new pair of values can be calculated. The value that is the same as in the first pair represents $|S_{31}|$, the true circulator isolation.

If the test equipment has phase measuring capabilities, the selection of either Eq. [9] or [11] becomes considerably simplified. When the phase of t_{min} and t_{max} is the same, Eq. [9] applies; while Eq. [11] is used if the phase difference between t_{min} and t_{max} is 180°.

Very rapid isolation measurements are possible using computer-controlled equipment. At each frequency, readings can be made with at least three different positions of the sliding load. These readings may then be used to determine the location of the center of the circle by the computer, and the value of true isolation displayed on the output device. A similar technique may be used to accurately determine input VSWR. Sliding terminations would be required at both port 2 and port 3. Using a somewhat tedious procedure, data can be obtained at each

frequency to permit the determination of $|S_{11}|$ through the solution of four simultaneous equations. As indicated earlier, the correction would be small.

Experimental Results

True isolation in three-port circulators has been measured by using software generated for use with the computer-controlled Hewlett–Packard 8540 Automatic Network Analyzer.[4] This software causes the analyzer to measure reverse transmission through a terminated circulator at six different positions of the termination for each specified frequency. These six measurements are then used to calculate the center of a circle which passes through the data points. Although three data points should be adequate to define a circle, six points are used to improve accuracy. The value of the center is then displayed expressed as isolation in dB.

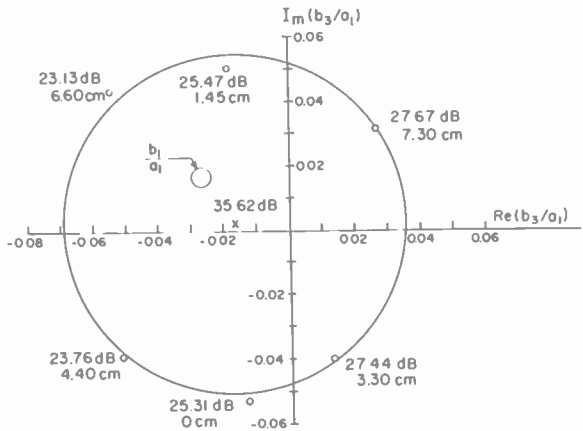


Fig. 3—Isolation data at 3000 MHz.

To verify the proper operation of the software, measurements have been made of the circulator of Reference [1] to obtain points used to calculate the center of the circle. This is done with the standard software supplied with the analyzer system. These six data points as a function of sliding load position, together with the isolation determined by the new software, are shown at 3000 MHz and 3200 MHz in Figs. 3 and 4, respectively. At 3000 MHz, the circle encloses the origin; while at 3200 MHz, it does not. In both cases, the circles drawn by using the measured value of true isolation as centers are in excellent agreement with the data points used to find the centers.

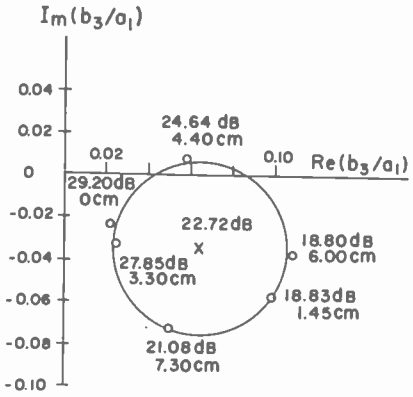


Fig. 4—Isolation data at 3200 MHz.

The significance of the corrected isolation technique is apparent from Figs. 3 and 4. At 3000 MHz, the value of true isolation is substantially greater than any of the data points. At 3200 MHz, it is substantially less than the greatest value of the data points.

The variation of the measured input reflection coefficient (b_1/a_1) as the termination on port 2 is moved is also shown in Fig. 3. It can be seen by comparison that the error introduced in this measurement by the termination reflection is small.

The corrected values of isolation should be independent of the VSWR of the sliding load. The circles shown in Figs. 3 and 4 are generated by a sliding load having a VSWR of about 1.10. Measurements have also been done with a sliding load having a VSWR of about 1.05. The isolation measurements for both cases are compared in Table 1. It can be seen that agreement is excellent.

Table 1—Comparison of Isolation Measurements Taken With Loads of Different VSWR

Frequency (MHz)	Isolation (dB) (VSWR = 1.05)	Isolation (dB) (VSWR = 1.10)
2600	10.02	10.09
2800	15.50	15.55
3000	34.07	33.87
3200	22.58	22.03
3400	20.45	20.07
3600	25.17	25.28
3800	16.68	16.75
4000	11.83	11.84

The true isolation of the circulator reported in Reference [1] is shown in Fig. 5 together with the previously reported results. Measurements have been made at 100-MHz intervals from 2600 to 4000 MHz and at 5-MHz intervals from 2900 to 3100 MHz to ensure that the peak isolation frequency would not be missed. It can readily be seen that the true peak value is much less than the earlier reported value.

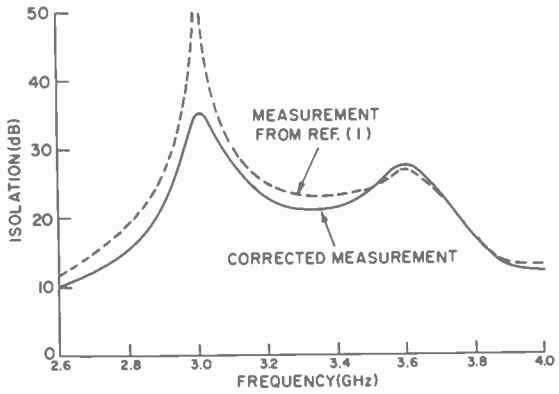


Fig. 5—Isolation of an S-band circulator.

Acknowledgment

The author gratefully acknowledges the assistance of Mr. William A. Schroeder in measuring and verifying the data reported in this paper.

References:

[1] B. Hershenov and R. L. Ernst, "Miniature Microstrip Circulators Using High-Dielectric-Constant Substrates," **RCA Review**, Vol. 30, No. 3, p. 541, Sept. 1969.

[2] M. Sucher and J. Fox (Editors), **Handbook of Microwave Measurements, Volume II**, Polytechnic Press of the Polytechnic Institute of Brooklyn, New York, 1963, pp. 646-670.

[3] R. Kellejian and C. L. Jones, **Microwave Measurements Manual**, McGraw-Hill, New York, 1965, pp. 27-34.

[4] R. A. Hackborn, "An Automatic Network Analyzer System," **Microwave J.**, Vol. 11, No. 5, p. 45, May 1968.

RCA Technical Papers Fourth Quarter, 1970

October

"Coherent Optical Sources for Communications," J. I. Pankove and Coauthors, **Proc. IEEE**
"Cross-Bred Components," B. Shore, **Signal**
"Impurity Distribution in High-Efficiency Silicon Avalanche Diode Oscillators," J. M. Assour, **IEEE Trans. GED**
"Luminescence in Indirect Bandgap $Al_xGa_{1-x}As$," H. Kressel, F. H. Nicoll, F. Z. Hawrylo, and H. F. Lockwood, **Jour. Appl. Phys.**
"Measurement of AM Noise in Pulsed Oscillators and Amplifiers," P. A. Levine, **IEEE Trans. GMTT**
"A Packaged System of a Solid-State Microwave-Biased Photoconductive Detector for 10.6 μm," C. Sun and T. E. Walsh, **Proc. IEEE**
"Simplified Theory for Stable Gunn Domains Including Diffusion," R. A. Sunshine and Coauthor, **Jour. Appl. Phys.**
"Technology and Design of GaAs Laser and Non-Coherent IR-Emitting Diodes," R. Glicksman, **Solid State Tech.**
"Relationship of Surface-State Band Structure to Surface Atomic Configuration of Zinc Blende (110)," J. D. Levine and S. Freeman, **Phys. Rev.** B (15 October)
"Surface States for a Molecular Orbital Model Hamiltonian: Resolvent Methods and Application to Zinc Blende," S. Freeman, **Phys. Rev.** B (15 October)

November

"Anthracene Electroluminescent Cells with Tunnel-Injection Cathodes," J. Dresner and A. M. Goodman, **Proc. IEEE (Letters)**
"A Compact Microstrip High-Power High-Efficiency Avalanche-Diode Oscillator," S. G. Liu, **IEEE Trans. GMTT** (Correspondence)
"Computer Solution of Nonlinear Simultaneous Equations," G. P. Kirkpatrick, **IEEE Trans. GBTR**
"Future of Ham Radio," L. Steller, **Electronics World** (Letters)
"High-Performance Integrated Circuits for High-Gain FM-IF Systems," R. T. Peterson, **IEEE Trans. GBTR**
"Industrial Video Tape Applications to Continuing Engineering Studies Programs," J. M. Biedenbach, **IEEE Trans. GE**
"A method of Matrix Polarization Rotating or Retarding Light-Valve Arrays," G. W. Taylor, **Proc. IEEE**
"Microwave Circuits for High-Efficiency Operation of Transferred Electron Oscillators," J. F. Reynolds, B. E. Berson, and R. E. Enstrom, **IEEE Trans. GMTT**
"One-Year Labor Warranty," T. J. Bradshaw, **Electronics World** (Letters)
"Optical Absorption Spectra of Cuprous Telluride (Cu_2Te)," R. Dalven, **Jour. Appl. Phys.** (Communications)
"Optical Storage Utilizing a Silicon Vidicon," E. Luedicke and R. S. Silver, **Proc. IEEE** (Letters)
"Power Amplification with Anomalous Avalanche Diodes," H. J. Prager, K. K. N. Chang, and S. Weisbrod, **IEEE Trans. GMTT**
"Radio-Frequency-Sputtered Films of β-Tungsten Structure Compounds," J. J. Hanak, J. I. Gittleman, J. P. Pellicane, S. Bozowski, **Jour. Appl. Phys.**
"RCA's Continuing Role in Pacific Data Communications," H. O. Haynie, **Signal**
"Silicon Avalanche-Diode Microstrip L-Band Oscillator," A. Rosen and J. Assour, **IEEE Trans. GMTT** (Correspondence)
"Transferred Electron Amplifiers and Oscillators," S. Y. Narayan and F. Sterzer, **IEEE Trans. GMTT**
"The Use of Computers in Editing and Composition," A. H. Coleman, **IEEE Trans. GAES**
"Wide-Band Reflection-Type Transferred Electron Amplifiers," B. S. Perlman, C. L. Upadhyayula, and R. E. Marx, **IEEE Trans. GMTT**
"Transition Metals and the Fröhlich Mechanism for Superconductivity," A. Rothwarf, **Phys. Rev. B** (1 November)

"An Exact Solution to the Rate Equation for Reversible Photoisomerization," J. Blanc. **Jour. Phys. Chem.** (12 November)
"Li-Defect Interactions in Electron-Irradiated n-Type Silicon," B. Goldstein, **Phys. Rev. B** (15 November)
"Flux Monitoring Boosts Accuracy of Phased Array Radar Systems," H. C. Goodrich and R. C. Tomsic, **Electronics** (23 November)
"IC Limiter Preserves Phase Over 59-dB Dynamic Range," R. J. Turner, **Electronics** (Designer's Casebook) (23 November)
"Strain-Induced Electric Fields in Superconducting Aluminum," B. Abeles and Co-authors, **Phys. Rev. Letters** (30 November)

December

"Annealing of Electron Bombardment Damage in Lithium-Containing Silicon," G. J. Brucker, **IEEE Trans. GNS**
"Characteristics of Silicon Avalanche Diodes as Oscillators and Power Amplifiers in S Band," R. V. D'Aiello and J. M. Assour, **IEEE Jour. Solid-State Circuits** (Special Correspondence)
"Chemical Vapor Deposition of Aluminum Oxide Films from Organo-Aluminum Compounds," M. T. Duffy and W. Kern, **RCA Review**
"Compounds and Alloys for Superconducting Applications," R. E. Enstrom, J. J. Hanak, and G. W. Cullen, **RCA Review**
"Coupled TEM Bar Circuit for L-Band Silicon Avalanche Oscillators," J. F. Reynolds, A. Rosen, B. E. Berson, H.-C. Huang, J. M. Assour, H. W. Becke, and R. Amantea, **IEEE Jour. Solid-State Circuits**
"CW Microwave Amplification from Circuit-Stabilized Epitaxial GaAs Transferred Electron Devices," B. S. Perlman, **IEEE Jour. Solid-State Circuits**
"A Demountable Electrolyte Contact Cell," A. M. Goodman and J. D. Morgan, **Rev. Sci. Inst.** (Notes)
"Deposition and Properties of Silicon Dioxide and Silicate Films Prepared by Low-Temperature Oxidation of Hydrides," W. Kern and A. W. Fisher, **RCA Review.**
"Final Report: Subcommittee on Transmission and Transmission Paths," W. C. Morrison, **IEEE Trans. GBC**
"Heteroepitaxial Growth of Germanium and Silicon on Insulating Substrates," D. J. Dumin, P. H. Robinson, G. W. Cullen, and G. E. Gottlieb, **RCA Review**
"Interface Properties of Chemically Deposited Silica Films on Gallium Arsenide," W. Kern and J. P. White, **RCA Review**
"The Long-Term Effects of Radiation on Complementary MOS Logic Networks," W. J. Poch and A. G. Holmes-Siedle, **IEEE Trans. GNS**
"Low-Temperature Vapor Growth of Homoepitaxial Silicon," D. Richman, Y. S. Chiang, and P. H. Robinson, **RCA Review**
"Microwave Integrated-Circuit Technology—A Survey," M. Caulton and H. Sobol, **IEEE Jour. Solid-State Circuits**
"Permanent Radiation Effects in Hardened Al_2O_3 MOS Integrated Circuits," F. B. Micheletti and K. H. Zaininger, **IEEE Trans. GNS**
"A Photoconductive Sensor for Card Readers," F. V. Shallcross, W. S. Pike, P. K. Weimer, and G. M. Fryszman, **IEEE Trans. GED** (Correspondence)
"Preparation, Properties, and Applications of Chemically Vapor Deposited Silicon Nitride Films," M. T. Duffy and W. Kern, **RCA Review**
"The Preparation of Ternary and Quaternary Compounds by Vapor Phase Growth," B. J. Curtis, F. P. Emmenegger, and R. Nitsche, **RCA Review**
"Radiation Hardened Registers," D. Hampel and K. J. Prost, **IEEE Trans. GNS**
"Standardization in Quadruplex Television Tape Recording," R. N. Hurst, **Jour. SMPTE**
"Stepper Motors: Application and Selection," A. J. Bianculli, **IEEE Spectrum**
"Thin-Film RF Integrated Amplifiers for UHF Broadband Communications Systems," C. Kamnitsis and R. Minton, **Microwave Jour.**
"Vapor Deposition and Characterization of Metal Oxide Thin Films for Electronic Applications," C. C. Wang, K. H. Zaininger, and M. T. Duffy, **RCA Review**
"Vapor Deposition of Semiconducting Mononitrides of Scandium, Yttrium, and the Rare Earth Elements," J. P. Dismukes, W. M. Yim, J. J. Tietjen, and R. E. Novak, **RCA Review**
"Vapor Growth of (II-VI)-(III-V) Quaternary Alloys and Their Properties," W. M. Yim, J. P. Dismukes, and H. Kressel, **RCA Review**
"Vapor Phase Growth of Magnetic Semiconducting Spinels," R. L. Pinch and L. Ekstrom, **RCA Review**
"Vapor-Phase Growth of Several III-V Compound Semiconductors," J. J. Tietjen, R. E. Enstrom, and D. Richman, **RCA Review**
"Magnetic Circular Dichroism Spectra of Divalent Lanthanide Ions in Calcium Fluoride," H. A. Weakliem, C. H. Anderson, and E. S. Sabisky, **Phys. Rev. B** (1 December)

Patents Issued to RCA Inventors Fourth Quarter 1970

October

J. J. Amodei Polarization Controlled Photochromic 'Write-In' System (3,535,021)

L. K. Baker Method of Fabricating Semiconductor Devices (3,535,774)

J. J. Drautman, Jr. Superconductive Magnet Construction (3,534,308)

A. D. Eyster and J. F. Heagy Method of Evacuating and Sealing a Glass Envelope Containing a Photoconductive Device (3,536,462)

A. J. Foster High Efficiency Single Turn Magnetic Head (3,535,466)

A. N. Gardiner Marking Device (3,536,039)

C. J. Hirsch Area Navigation Method and Apparatus for Aircraft with VHF-Omnirange (VOR) and Distance Measuring Equipment (DME) (3,534,399)

L. A. Jones and R. B. Goyer Magnetic 'Pen' for a Graphic Tablet (3,532,817)

K. R. Kaplan Threshold Gate Circuits (3,532,897)

S. Knanishu High Input Impedance Solid State D.C. Amplifier Suitable for Use in Electrical Measurement (3,532,893)

H. Kressel and F. Z. Hawrylo Semiconductor Laser Producing Light at Two Wavelengths Simultaneously (3,537,029)

B. J. Lechner Display Circuit Including Charging Circuit and Fast Reset Circuit (3,532,813)

L. J. Limbaugh and E. C. James Random Access Optical Sound Track Reproducer with Automatic Gain Controlled Amplifiers Responsive to a Reference Zone on Each Track (3,536,858)

A. L. Limberg D. C. Restoration Circuit with Arc-Over Protection (3,535,436)

A. L. Limberg Electrical Circuit for Providing Substantially Constant Current (3,534,245)

A. L. Limberg High Current Transistor Amplifier Stage Operable With Low Current Biasing (3,534,279)

R. M. Matheson and F. A. Helvy Method of Making Photoemissive Electron Tubes (3,535,011)

D. D. Maxson Device for Generation of a Self-Acting Fluid Bearing (3,534,893)

A. Month Method for Preparing Cathodes (3,536,526)

P. L. Nestleroth and W. H. Silvers Method for Making Cathode Assembly for Electron Tube (3,535,757)

J. Ollendorf and F. P. Jones Semiconductor Devices Having Soldered Joints (3,532,944)

J. I. Pankove Confocal Semiconductor Diode Injection Laser (3,537,028)

J. Paull Centrifugal Testing Apparatus (3,534,595)

J. N. Pratt and G. E. Anderson Noise Immune Video Circuits (3,535,444)

A. K. Rapp Control Circuit for Memory (3,533,088)

C. F. Rose Magnetic Tape Cassette Player Apparatus (3,532,293)

R. C. Shambelan Semiconductor Wafer Transporting Jig (3,535,437)

J. C. Sondermeyer Amplifier Protection Circuit (3,536,462)

W. Steinberg Digital Logic Circuit for Deriving Synchronizing Signals from a Composite Signal (3,532,810)

J. A. Vallee Decoder for Self Clocking Digital Magnetic Recording (3,537,082)

P. K. Weimer Coupling System for Panel-Type Array (3,537,071)

D. H. Willis Color Television Receiver (3,535,437)

R. O. Winder Shift Circuits Including Threshold or Other Logic Gates and Having Multiple-Phase Shift Pulses Applied to Each Stage (3,552,991)

B. Zuk Memory Employing Transistor Storage Cells (3,533,087)

November

D. R. Andrews Capstan and Flywheel Arrangement for Magnetic Tape Transport (3,537,332)

T. G. Athanas and D. M. Griswold Ion Bombardment of Insulated Gate Semiconductor Devices (3,540,925)

W. M. Austin Video Circuits Employing Cascoded Combinations of Field Effect Transistors with High Voltage, Low Bandwidth Bipolar Transistors (3,541,234)

R. D. Brand Television Tuner Cast Housing With Integrally Cast Transmission Lines (3,538,466)

P. S. Carnt and T. E. Bart Identification Circuit For Phase Alternating Line-System Operation of Color Video Tape Recorders (3,538,244)

L. A. Cochran Circuit for Eliminating Spurious Modulation of the Subcarrier Frequency Oscillator in a Color Television Receiver (3,541,241)

E. W. Curtis Automatic Beam Current Limiting Using Reference Current Sources (3,541,240)

R. A. Gange Memory System with Defective Storage Locations (3,541,525)

R. A. Gange and P. Hsieh Line Terminating Circuits (3,541,475)

H. J. Gerritsen and H. S. Sommers, Jr. Transistory Hologram Apparatus (3,542,452)

J. E. Goldmacher and J. A. Castellano Electro-Optical Compositions and Devices (3,540,796)

R. J. Green Apparatus for Electroplating a Ribbon (3,537,971)

S. Gubin Satellite Communications Systems (3,541,553)

C. J. Hall and R. Peter Color Temperature Correction Controlled by the Color Killer and Color Oscillator (3,541,242)

J. R. Harford Gain Controlled Amplifier (3,538,448)

W. G. Henderson and J. T. Mark Sputter Ion Pump (3,540,812)

W. R. Isom Cartridge Debris Trap (3,537,710)

W. R. Isom Pressure Roller Construction (3,537,661)

W. Kerm Microminiature Electrical Component Having Indexable Relief Pattern (3,543,106)

J. M. Kresock and T. W. Burrus Keying Circuit (3,541,235)

A. C. Luther and R. G. Breed Servo System (3,542,950)

J. C. Meagher Electromechanical Coupling (3,542,284)

J. C. Miller and C. M. Wine Light Pen Operating With Remote Graphic Display (3,543,240)

M. E. Mol Electro-Optical Scanner (3,541,247)

R. G. Olden Apparatus for Conveying Toner Particles (3,538,994)

D. W. Roe Conductive Coatings of Tin Oxides (3,537,890)

E. Sarkisian and N. J. Amdur Apparatus for Measuring Low Voltages and Currents with Amplifier Protective Means (3,541,402)

J. C. Schira Optical Field Correction Devices for an Electronic Photocomposition System (3,540,361)

H. W. Silverman Use in an Automatic Testing System of a Simulator of an Article Being Tested for Testing the Testing System (3,541,440)

C. C. Spagnoli, B. W. Binkley and D. C. Lavallee Timing System for Readout of Stored Data (3,542,286)

J. M .Walter and F. D. Rando Single Wire Crosspoint Switching Circuit with External Signaling (3,541,515)

H. R. Warren Video Recording and Reproducing Apparatus Utilizing a Single Track on a Magnetic Tape for the Luminance and Color Information Components of a Color Television Signal (3,542,946)

R. K. W. Yee Gated Differential Amplifier (3,541,466)

B. Zuk Logic Circuit (3,539,823)

December

R. W. Ahrons Complementary Semiconductor Matrix Arrays for Low Power Dissipation Logic Application (3,550,089)
G. A. Alphonse and A. Akselrad Photocromic Memory in Which Memory Location Is Selectively Heated During Write Cycle (3,550,096)
J. E. Annis Information Storage Decoder System (3,551,900)
D. M. Baugher Transistor High Current Switching and Inverter Circuits (3,546,627)
H. R. Beelitz Input and Output Emitter-Follower CML Circuitry (3,549,899)
J. R. Burns and J. J. Gibson Clock Logic Circuits (3,551,693)
J. Brous Traveling Wave Tube with Evaporated Nickel Attenuator Coating and Method of Manufacture Thereof (3,544,832)
D. J. Carlson UHF-VHF Tuner with U-Shaped Strip Balanced Transmission Line (3,551,819)
E. F. Cave and F. C. Duigon Charge Storage Device with PN Junction Diode Array Target Having Semincoductor Contact Pads (3,548,233)
R. H. Cornely Semiconductor Laser Logic Apparatus (3,546,495)
L. P. Dague High Voltage Wide Band Amplifier (3,546,611)
J. W. Daniel, Jr. Receiver Gain Control System Providing Negative Resistance Stabilization (3,544,902)
R. A. Dischert, W. J. Cosgrove and R. R. Brooks Vertical and Horizontal Aperture Equalization (3,546,372)
G. Y. Eastman Heat Exchanger for High Voltage Electronic Devices (3,543,841)
D. Espinal Drive Circuit for Digit Lines (3,546,487)
M. S. Fisher Capacitor Discharge Ignition Circuit (2,546,528)
K. C. Gaspar Momentary Retaining Translation Means for Multiple Switches (3,544,738)
G. W. Gray Oscillating Voltage Range Indicator (3,551,696)
D. L. Greenaway, H. J. Gerritsen Sequential Information Hologram Record (3,545,834)
F. G. Hammersand and W. A. Novajovsky Grid Support for Electron Tubes (3,544,831)
J. J. Hanak Multi-Head Magnetic Transducer (3,544,982)
G. H. Heilmier Control of Optical Properties of Materials with Liquid Crystals (3,551,026)
P. E. Justus and T. A. Korn Programmed Method for Manipulating Electronic Fonts in Electronic Photocomposition Systems (3,546,495)
Z. J. Kiss Dark Trace Cathode Ray Tube with Photochromic Image Screen (3,548,236)
W. L. Lehmann and J. C. Marsh Compensated VHF-UHF Automatic Gain Control Delay System (3,548,315)
A. Lichowsky Integrated Power Output Circuit (3,554,860)
A. Macovski Color Television Camera Overload Compensating System (3,546,373)
J. A. McDonald and T. J. Christopher Lock-On Prevention in Transistor Deflection Circuits (3,544,811)
J. A. McDonald and T. J. Christopher Spurious Oscillation Suppression in Transistor Deflection Circuits (3,544,810)
E. D. Menkes Low Power, High Stability Digital Frequency Synthesizer (3,546,618)
J. F. Monahan Burst Flag Generator (3,546,602)
H. Nelson Semiconductor Laser Having High Power Output and Reduced Threshold (3,551,842)
R. G. Olden Electrophotographic Method and Apparatus (3,549,251)
A. Prieto Computer Memory Addressing (3,551,898)
A. K. Rapp Non-Destructive Read-Out Memory Cell (3,549,904)
P. R. Sahm and F. D. Rosi Superconductors (3,544,316)
J. H. Scott, Jr. Variable Threshold Level Field Effect Memory Device (3,549,904)
H. J. Sheetz and R. T. Molloy Expandable-Metal-Structure Making by Etching (3,546,075)
J. A. Tourtellot Automatic Record Changer (3,545,767)
W. R. Walters Thyristor Control Circuit Using D. C. Controllable (3,551,696)
J. P. Watson Tape Level Sensor (3,550,828)
C. P. Wen Tunable Microstrip Band Pass Filter Utilizing Cryomagnetic Materials at the Junction of two Conductive Loops (3,546,637)
D. H. Westwood Digital Frequency Synthesizer (3,546,617)
N. E. Wolff Process of Thermoplastic Deformation Imaging (3,547,628)
E. K. C. Yu Current Mode Switching Circuit (3,549,900)
L. A. Zanoni Alternating Current Driven Load Circuit (3,551,689)

AUTHORS

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.

Arthur Dreeben completed his graduate work in Inorganic and Solid State Chemistry and Physics at the Polytechnic Institute of Brooklyn in 1950. He held a Teaching Fellowship, and did research on Infrared Stimulable phosphors. From 1950 to 1953, he was employed as a research chemist at the General Electric Research and Knolls Atomic Power Laboratories on the development of analytical and radiochemical procedures. In 1953, he joined the research department of the Westinghouse Lamp Division as a Research Engineer. Here, he did research in the field of luminescence, including high-temperature phosphors, lamp phosphors, and transparent luminescent films. He joined RCA Laboratories in 1958 as a Member of the Technical Staff and has worked on photoconductors, electroluminescence, problems in crystal growth, dislocations, and impurity precipitation in semiconductors. At present he is working on the growth and properties of epitaxial layers of III-V compounds for various microwave devices.

He is listed in "American Men of Science" and is a member of the American Chemical Society, the Electrochemical Society, Sigma Xi, and Phi Lambda Upsilon.

Robert L. Ernst received the BEE degree with honor from Manhattan College in 1961 and the MSEE degree from the Polytechnic Institute of Brooklyn in 1965. In 1961 Mr. Ernst joined the Western Union Telegraph Company where he was active in the design and development of microwave components. In 1963, he joined the Defense Advanced Communications Laboratory of RCA. Here he has developed many solid-state microwave circuits such as resistive and reactive frequency converters, varactor diode frequency multipliers, transistor oscillator-frequency multipliers, voltage tunable transistor oscillators, and bulk-effect oscillators. He has done extensive theoretical work in the fields of interference reduction techniques and parametric amplifiers. Currently, he is working with microwave transistor amplifiers, both large signal and small signal, and reduced-size microstrip ferrite circulators. Mr. Ernst is a Member of the IEEE, Eta Kappa Nu, and Tau Beta Pi.

Angel Gonzalez attended the Engineering School of Vigo, Spain, where he received his BSEE. From 1964 to 1965, he worked in the Electronics Department of LTIEMA, Madrid, Spain. From 1966 to 1968, he was with the TECA Corp., White Plains, N. Y. He joined RCA Laboratories, Princeton, N. J., in 1969 as a member of the Integrated Microwave Circuits Group. Presently, he is attending the Stevens Institute of Technology, working toward an MS degree in Computer Science.

Istvan Gorog attended the University of California at Berkeley, where he received the B.Sc (1961), M.Sc (1962), and Ph.D (1964) degrees in Electrical Engineering. His dissertation was a study of the near ultraviolet and visible light emission from a "dynamic pinch." Since September of 1964, Dr. Gorog has been a Member of the Technical Staff of the RCA Laboratories, David Sarnoff Research Center, where his main areas of concern have been quantum electronics and electro-optical systems. Currently he is head of the Optical Electronics Research Group in the Solid State Research Laboratory. During 1968 he was on leave of absence from RCA as a National Science Foundation Post-Doctoral Fellow in Frascati, Italy, where he worked on laser-related problems in plasma research. Dr. Gorog is a member of the American Physicial Society and of Eta Kappa Nu.

Werner Kern received a certificate in chemistry in 1944 from the University of Basle, Switzerland, and a diploma in chemical technology in 1946. He published a thesis on the chromatographic isolation and characterization of fluorescing polynuclear hydrocarbons which he discovered in soil. He was analytical research chemist with Hoffmann-LaRoche in Switzerland, and in 1948 transferred to their research division in New Jersey to develop new radiochemical methods. In 1955 he received an AB degree in chemistry from Rutgers University and in 1958 joined Nuclear Croporation of America where he became chief chemist directing research in nuclear and radiation chemistry. He joined RCA Electronic Components in 1959 primarily to investigate semiconductor contamination and surface passivation by radiochemical methods. Since 1964, he has been at RCA Laboratories as a Member of the Technical Staff, where his activity has centered in semiconductor process research in the areas of device passivation, new methods of chemical vapor deposition of dielectric films, and the development of associated analytical methods.

Mr. Kern is a member of the American Chemical Society, the Electrochemical Society, the Society of Sigma Xi, and AAAS, the Geological Society of New Jersey, and is listed in American Men in Science.

Jules D. Levine received his B.S. degree in mechanical engineering from Columbia in 1959, and his S.M. and Ph.D degrees from MIT in 1961 and 1963, majoring in surface physics. Joining RCA in 1963, he has carried out theoretical and experimental research on surface phenomena. From 1963-1966 he investigated metallic adsorption on metals and insulators, with particular application to thermionic energy conversion. From 1966-1969, he studied surface states on semiconductors, with special emphasis on the III-V and II-VI compounds. More recently, he has completed research on photodischarge spectroscopy of surface states and on Schottky barrier interface states.

Dr. Levine is a member of the American Physical Society, the AAAS, the IEEE, and the American Vacuum Society, Thin Film Division.

John E. Meyer, Jr., received his BSEE degree from Rutgers University in 1961, and the MSEE degree from the University of Pennsylvania in 1964. He is presently completing his work for the Ph.D in Electrical Engineering at Rutgers University. From 1961 to 1963, he worked as an engineer in the Missile and Space Vehicle Department at General Electric, Philadelphia, Pennsylvania, where he was engaged in the design of attitude control systems. From 1963 to 1964, he worked in the Physics Department at Princeton University as a member of the technical staff where he designed R. F. and digital instrumentation for the cyclotron group. Mr. Meyer joined RCA Laboratories in 1964 where he has engaged in research on integrated circuits in both the silicon and thin-film technologies. He has done extensive work on solid state image sensors and on complex silicon-on-sapphire integrated circuits.

Barry S. Perlman received the BEE from the City College of New York and the MSEE from Brooklyn Polytechnic Institute in 1961 and 1964, respectively. He is presently pursuing a Ph.D in Electrophysics at Brooklyn Polytechnic Institute, with emphasis on solid-state physics and advanced network theory. During 1961 he was a design and development trainee with RCA Defense Electronic Products. His work included signal processing, radar receivers, and parametric devices. From 1961 to 1968 he was a member of the Advanced Communications Laboratory in New York City, becoming a senior member of the technical staff in 1966. He was primarily concerned with advanced receiver techniques such as the development of all-solid-state microwave troposcatter and relay subsystems, airborne X-band receivers, integrated circuits, high level parametric circuits, and dielectric and superconducting microwave filters. In June of 1968 he joined the Advanced Technology Laboratory at the RCA David Sarnoff Research Center, Princeton, N. J. He is presently working on transferred electron oscillators and amplifiers. Mr. Perlman is a member of the IEEE and NYSSPE.

Otto H. Schade, Sr. was born and educated in Germany; he came to the United States in 1926. He joined RCA Electronic Components and Devices, Harrison, New Jersey, in 1931. Since 1938, he has specialized in television circuits, camera tubes, and picture tubes. From 1944 to 1957 he worked on a unified general method of image analysis and specification, including practical methods for measuring the "aperature" effect (square-wave and sine-wave response function) and fluctuation levels (noise) of optical, photographic, and electronic image-system components and the eye. He has had the responsibility for the thermal and electrical design of nuvistor tubes. More recently, he has developed an accurate method for calculating the resolving power of television and photographic imaging systems to assist in the evaluation of high-definition television systems, and a new electron optic providing minimum aberrations and uniform focus in television camera tubes with larger (50 × 50 mm) image surfaces. Dr. Schade has received numerous honors, including the Modern Pioneers Award of the National Association of Manufacturers (1940), the Morris Liebmann Memorial Prize of the Institute of Radio Engineers (1950) and a Fellowship (1951) from the Institute of Radio Engineers. In 1951 he was made a Fellow, and also was the first recipient of the David Sarnoff Gold Medal Award, of the Society of Motion Picture and Television Engineers. In June 1953, he was invested with the honorary degree of Doctor of Engineering by Rensselaer Polytechnic Institute. In 1960, he received the Progress Medal Award of the Society of Motion Picture and Television Engineers for his outstanding technical contribution in the engineering phases of the motion picture and television industries.

George A. Swartz received his BS degree from Massachusetts Institute of Technology in 1952 and the M.S. and Ph.D degrees in physics from the University of Pennsylvania in 1954 and 1958, respectively. He joined the technical staff of RCA Laboratories in 1957. His work has been in the field of plasma propulsion, plasma stability, microwave phenomena in gaseous and solid-state plasmas, and solid-state physics.

Dr. Swartz is a member of the American Physical Society and Sigma Xi.