RCA Review

June 1975 Volume 36 No. 2

RCARCI 36(2) 211-380(1975)

JUL 1975
HOUSTON PUBLIC
LIBRARY
Houston, Texas

RCA Review, published quarterly in March, June, September and December by RCA Research and Engineering, RCA Corporation, Princeton, New Jersey 08540. Entered as second class matter July 3, 1950 under the Act of March 3, 1879. Second-class postage paid at Princeton, New Jersey, and at additional mailing offices. Effective Jan. 1, 1971, subscription rates as follows: United States and Canada: one year $6.00, two years $10.50, three years $13.50; in other countries, one year $6.40, two years $11.30, three years $14.70. Single copies (except for special issues) up to five years old $3.00.

RCA Review

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

Contents

213 Time-Dependent Characteristics of Photo-Induced Space-Charge Field and Phase Holograms in Lithium Niobate and Other Photorefractive Media
G. A. Alphonse, R. C. Alig, D. L. Staebler, and W. Phillips

230 Reliability Aspects and Facet Damage in High-Power Emission from (AlGa)As CW Laser Diodes at Room Temperature
H. Kressel and I. Ladany

240 Mobility of Current Carriers in Silicon-on-Sapphire (SOS) Films
S. T. Hsu and J. H. Scott, Jr.

254 Three-Filter Colorimetry of Color-Television Picture Tubes
G. M. Ehemann, Jr.

274 A Study of Second-Harmonic-Extraction Trapatt Amplifiers for X-Band Operation
W. R. Curtice, J. J. Risko, and P. T. Ho

296 Microwave Integrated-Circuit MIS Varactor Phase Shifter
R. L. Camisa, B. F. Hitch, S. Yuan, and M. Ettenberg

316 The Conversion Efficiency of Ideal Shockley P-N Junction Photovoltaic Converters in Concentrated Sunlight
F. Sterzer

324 Silicon Solar Cells for Highly Concentrated Sunlight
R. H. Dean, L. S. Napoli, and S. G. Liu

336 Review on Industrial Applications of High-Power Laser Beams III
I. P. Shkarofsky

369 Technical Papers Published

372 Patents

375 Authors

June 1975 Volume 36 Number 2

RCA Corporation

Robert W. Sarnoff Chairman of the Board and Chief Executive Officer
A. L. Conrad President and Chief Operating Officer

Editorial Advisory Board

Chairman, J. A. Rajchman RCA Laboratories

A. A. Ahmed Solid State Division
E. D. Becken RCA Global Communications
G. D. Cody RCA Laboratories
D. M. Cottler Government and Commercial Systems
N. L. Gordon RCA Laboratories
G. B. Herzog RCA Laboratories
J. Hillier RCA Research and Engineering
E. O. Johnson International Licensing
J. Kurshan RCA Laboratories
C. H. Lane Electronic Components
D. S. McCoy Consumer Electronics
K. H. Powers RCA Laboratories
R. E. Quinn RCA Laboratories
P. Rappaport RCA Laboratories
J. H. Scott, Jr. RCA Laboratories
L. A. Shotliff International Licensing
T. O. Stanley RCA Laboratories
F. Sterzer RCA Laboratories
J. J. Tietjen RCA Laboratories
W. M. Webster RCA Laboratories

Secretary, Charles C. Foster RCA Laboratories

Editor **Ralph F. Ciafone**

Associate Editors

W. A. Chisholm RCA Limited (Canada)
M. G. Gander RCA Service Company
W. O. Hadlock RCA Research and Engineering
D. R. Higgs Missile and Surface Radar Division
W. A. Howard National Broadcasting Company
C. Hoyt Consumer Electronics
E. McElwee Solid-State Division
C. A. Meyer Electronic Components
M. G. Pietz Government and Commercial Systems
C. W. Sall RCA Laboratories
I. M. Seideman Astro-Electronics Division
R. N. Hurst Communications Systems Division

© RCA Corporation 1975 All Rights Reserved Printed in USA

Time-Dependent Characteristics of Photo-Induced Space-Charge Field and Phase Holograms in Lithium Niobate and Other Photorefractive Media

G. A. Alphonse, R. C. Alig, D. L. Staebler, and W. Phillips

RCA Laboratories, Princeton, N. J. 08540

Abstract—The transport equation for the space-charge field in photorefractive media in the presence of arbitrary illumination (within the linear absorption range) and electric fields is solved using Amodei's model for situations where the photo-excited carrier concentration is proportional to the illumination. Rigorous, closed-form solutions are obtained for the complete space–time development of the field in the special cases of holographic recording and erasure, and the results are verified experimentally. One of the important experimental findings is that the model is accurate at low and moderate spatial frequencies in accordance with Amodei's requirement of short electron-migration length compared to the hologram grating wavelength. At higher spatial frequencies, the experiments are in better agreement with the model of Young and coworkers in which the restriction on the migration length is removed.

1. Introduction

The optical inhomogeneity[1,3] induced by light in ferroelectrics such as lithium niobate ($LiNbO_3$), lithium tantalate ($LiTaO_3$), barium titanate ($BaTiO_3$), strontium barium niobate (SBN), potassium tantalate niobate (KTN), and lead zirconate titanate (PLZT) has been used in phase-holography experiments[4–6] with these materials and has raised the possibility of using them as storage media in holographic memory systems. Indeed these "photorefractive" materials

are the most versatile of all storage media—they can be used for read–write applications because they are "erasable";[7] they can also be used for read-only or volume storage applications because they are "fixable."[8] By comparison, photographic emulsions are suitable for read-only applications; magneto-optic materials, which require high optical power density in short times, are suitable for nonholographic storage although they have been the subject of serious holographic developments;[9] and thermoplastics, which require processing, cannot be used in real time. The major drawback of photorefractive media is their lack of sensitivity, but considerable improvement has already been made in that direction.[7,10]

Much progress[3,11,12] has been made toward the understanding of the process that leads to the build-up of index inhomogeneity upon exposure to light of moderate intensity. The prevailing model is that of the generation of a space-charge field by the migration of photoexcited electrons giving rise to electro-optically induced refractive-index changes. Electrons are excited from impurity traps by the incident radiation and, upon migration, are retrapped at other locations, giving rise to a frozen-in electric field. The latter induces refractive index changes via the electro-optic effect. In undoped crystals, the traps are provided by small traces of impurities. In doped crystals, the dopants act as donor–acceptor traps via intervalence exchanges such as $Fe^{2+} \rightleftarrows Fe^{3+}$ in iron-doped lithium niobate.[7,10] In general, the migration of charges occurs under the combined influence of diffusion and an electric field. For the purpose of this paper the source of the field is immaterial. It may be externally applied by means of electrodes, or it may be internal, as in the bulk photovoltaic effect of the type recently reported by A. M. Glass et al.[13] Although the resulting space-charge field can be generated in arbitrary direction, it has been shown[14] that for the class of crystals of interest, electro-optically induced refractive index variations without alterations of the optical properties of the crystal are obtained only when the field is along the optic axis. The major change occurs mainly in the extraordinary refractive index.

When the light intensity is very strong, the space-charge migration model does not account for the correct magnitude of the refractive-index change, as shown by D. von der Linde et al.[15] They attribute the absorption mechanism to a multiphoton process that induces strong macroscopic polarization changes whose field, rather than the space-charge field, contributes to the refractive-index change. For the rest of this paper we are concerned only with moderate light intensities–less than about 200 mW/mm^2.

In holographic practice the ability to predict the diffraction efficiency under a given optical exposure is important for the design of systems and for the prediction of systems performance. Although the subject of the space-charge field resulting in the formation of holograms or other index inhomogeneities has been studied extensively, all analytical work so far has been only approximate. Thus, after Chen's[3] initial calculations confirming the space-charge model, Amodei[16] used the transport equation for the model to derive expressions for the field due to spatially sinusoidal illumination. He assumed that the electron migration length was much shorter than the spatial periodicity and that the concentration of the photoexcited carriers was in equilibrium with the illumination. This implies that the lifetime of the electrons is very short and constant and that the electron concentration is much smaller than the trap density. However, his results were limited to the initial build-up of the field for short times and to its equilibrium value as the time approaches infinity, for the special cases of drift-dominated and diffusion-dominated transport processes. Young et al[17] have removed Amodei's assumption of short migration length and obtained interesting new results, but their time solution is also limited to the linear initial build-up. It is possible to obtain a numerical solution for the complete time development of the field, as was done by King et al[18] for the case of the focused beam, but no prediction can be made from such a solution without the use of a computer.

In this paper we present, based on Amodei's[16] model, a rigorous, closed-form solution for the space–time development of the induced field under arbitrary illumination (within the linear absorption region) and electric fields. The solution for the important cases of spatially sinusoidal illumination (as in holographic recording) and uniform illumination (as in holographic readout) are then given. We use Amodei's model because his assumption of a migration length much smaller (typically smaller than 500 Å) than the hologram's periodicity is applicable over a broad range of hologram recording angles.

2. General Solution to the Transport Equation

The continuity equation for the trapped charge density ρ due to the motion of the photoexcited electrons is

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{J}. \qquad [1]$$

where $\mathbf{J}$ is the current density and t the time. The total electric field

E is related to the charge density by Gauss' law;

$$\nabla \cdot \epsilon \mathbf{E} = \rho, \qquad [2]$$

where ϵ is the dielectric constant. Combining Eqs. [1] and [2], we obtain

$$\nabla \cdot \left(\mathbf{J} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \right) = 0. \qquad [3]$$

The differential equation for the space-charge field $\mathbf{E}_s$ is obtained by substituting in Eq. [3] the proper expression for the current density and separating the total field into its known or applied part $\mathbf{E}_a$ and the unknown $\mathbf{E}_s$, i.e.,

$$\mathbf{E} = \mathbf{E}_a + \mathbf{E}_s. \qquad [4]$$

The current density due to the forces of drift and diffusion is

$$\mathbf{J} = qDn_0\nabla n + q\mu n_0 n\mathbf{E}. \qquad [5]$$

where q is the electronic charge, D the diffusion constant, μ the electron mobility, $n_0 n(\mathbf{r})$ their concentration with equilibrium value n_0. Since the lifetime of the photoelectrons is short (estimated[7] to be about 10^{-11} sec for a trap concentration of 10^{18} cm^{-3}) compared to the dielectric relaxation time

$$T_0 = \frac{\epsilon}{q\mu n_0}, \qquad [6]$$

the electron concentration $n_0 n(\mathbf{r})$ will be assumed in equilibrium with the illumination. Substitution of Eqs. [4], [5], and [6] into Eq. [3] gives

$$\nabla \cdot \left[T_0 \frac{\partial}{\partial t} \mathbf{E}_s + n\mathbf{E}_s + \frac{kT}{q} \nabla n + n\mathbf{E}_a \right] = 0, \qquad [7]$$

where use has been made of the Einstein relation $Dq = kT\mu$, k being Boltzmann's constant and T the absolute temperature.

The integration of Eq. [7] with respect to space gives

$$\frac{\partial}{\partial t} \mathbf{E}_s(\mathbf{r}, t) + \frac{n(\mathbf{r})}{T_0} \mathbf{E}_s(\mathbf{r}, t) = \mathbf{G}(\mathbf{r}, t) - \mathbf{A}(\mathbf{r}) \qquad [8]$$

where

$$\mathbf{A}(\mathbf{r}) = \frac{kT}{qT_0} \nabla n(\mathbf{r}) + \frac{n(\mathbf{r})}{T_0} \mathbf{E}_a, \qquad [9]$$

and where $\mathbf{G}(\mathbf{r},t)$ is an arbitrary function of space and time such that $\nabla \cdot \mathbf{G} = 0$.

The solution of Eq. [8] must be such that

$$\lim_{t \to \infty} \mathbf{E}_s(\mathbf{r}, t) = \mathbf{E}_s(\mathbf{r})$$

in the steady state, with

$$\lim_{r \to \infty} \mathbf{E}_s(\mathbf{r}, t) = 0,$$

since there is no perturbation of the electron density, i.e., no illumination, far from the origin.

With $\mathbf{A}$ a function of space only, so must also be $\mathbf{G}$. To show this, we note that in the steady state $\mathbf{E}_s$ is electrostatic in nature, i.e., $\nabla \times \mathbf{E}_s = 0$. Hence there exists a function $\Phi(\mathbf{r})$ such that $\nabla\Phi = T_0\mathbf{G}/n$. Then from $\nabla \cdot \mathbf{G} = 0$ and from the above boundary condition for $\mathbf{E}_s$ we obtain

$$\nabla \cdot (n \nabla \Phi) = 0 \qquad [10]$$

subject to the boundary condition

$$\lim_{r \to \infty} \nabla \Phi = 0. \qquad [11]$$

Since Eq. [10] is an elliptic differential equation of the second order with a Dirichlet type of boundary condition (Eq. [11]), there exists a unique, stable solution[19] for Φ from which $\mathbf{G}(\mathbf{r})$ can be calculated. Of course the actual form of $\Phi(\mathbf{r})$ will depend on $n(\mathbf{r})$ and, in general, it can only be found numerically.

With $\mathbf{G}$ independent of the time, the solution of Eq. [8] is

$$\mathbf{E}_s(\mathbf{r}, t) = \frac{T_0}{n(\mathbf{r})} [\mathbf{G}(\mathbf{r}) - \mathbf{A}(\mathbf{r})] + \mathbf{B}(\mathbf{r}) \exp\{-n(\mathbf{r})t/T_0\}, \qquad [12]$$

where $\mathbf{B}(\mathbf{r})$ is obtained from the initial conditions.

3. Application to Holography

The form of the expressions for $\mathbf{G}(\mathbf{r})$, $\mathbf{A}(\mathbf{r})$, and $\mathbf{B}(\mathbf{r})$ is determined by the physical situations. We shall now obtain results for writing and erasing holograms.

3.1 Spatially Periodic Illumination—Hologram Formation

The set-up for writing a hologram is shown in Fig. 1. For a medium such as $LiNbO_3$ it is desired to store a spatially periodic space-charge field along the optic or C axis. The crystal is oriented so that the interference between two incident beams I_1 and I_2 at angles $\pm\theta$ respectively with respect to the normal to the sample form a standing-wave pattern along the C axis (z direction). We also make provision for a dc electric field E_a to exist, also along the C axis. The source of E_a is immaterial; it can be external or internal. The polarization of the incident light is not important for "writing", but it must be extraordinary for "reading." The incident illumination is of the form $I = I_0(1 + m\cos Kz)$ where, as shown in Fig. 1, $I_0 = I_1 + I_2$ is the sum of the intensities of the two beams I_1 and I_2, $m = 2(I_1I_2)^{1/2}/I_0$ is the modu-

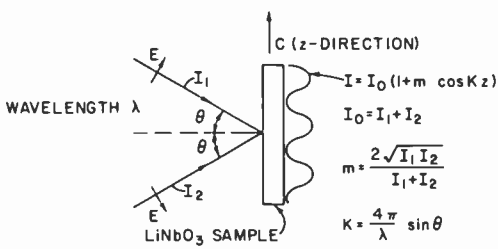


Fig. 1—Setup for writing a hologram (spatially periodic illumination).

lation index, and $K = (4\pi/\lambda)\sin\theta$ the spatial angular frequency. The photocarrier density is in equilibrium with the illumination, i.e., n_0 is proportional to I_0, and

$$n(z) = 1 + m\cos Kz. \qquad [13]$$

The fields **G**, **A**, **B** and $\mathbf{E}_s$ have only a z component, and are functions of the z coordinate only, except for **G** which is a constant in the illuminated region by virtue of the fact that $\nabla\cdot\mathbf{G} = dG/dz = 0$. Then from Eq. [9]

$$A(z) = -\frac{m}{T_0}E_{max}\sin(Kz - \varphi) + \frac{E_a}{T_0}, \qquad [14]$$

where

$$E_{max} = [E_D^2 + E_a^2]^{1/2}, \qquad [15]$$

$$E_D = K\frac{kt}{q}, \quad \varphi = \tan^{-1}\left(\frac{E_a}{E_D}\right). \qquad [16]$$

With the initial condition $E_s(z,0) = 0$, Eq. [12] gives

$$E_s(z,t) = \frac{m}{n(z)} E_{max}[1 - \exp\{-n(z)t / T_0\}] \sin(Kz - \varphi). \quad [17]$$

In Eq. [17] the term $(T_0G - E_a)/n(z)$ has been omitted since the arbitrary constant G can be adjusted to make it vanish.

In this closed-form expression, the periodic factor $n(z)$ in the denominator and in the exponential causes the stored field due to a simple sinusoidal optical exposure to contain harmonics. Thus, the problem is in general nonlinear, especially for $m \simeq 1$. For $m \ll 1$, the solution is essentially linear and the stored field is a phase-shifted sinusoid building up in time as

$$E_s(z,t) \cong mE_{max}[1 - \exp(-t/T_0)] \sin(Kz - \varphi). \quad [18]$$

For $E_A \ll E_D$, we obtain the diffusion-dominated case where $\varphi = 0$; for $E_a \gg E_D$, the drift dominated case is obtained with $\varphi \simeq \pi/2$. It must be noted that the field does not have the singularities for $m = 1$ and Kz equal to some odd multiples of π as in Ref. (16). As the denominator in Eq. [17] approaches zero, so does the argument of the exponential (which is missing in Ref. (16)), and a power expansion of the exponential shows that the singularity is removed.

Several spatial distributions, calculated for different values of m, are plotted in Figs. 2 and 3, and the time development of the Fourier amplitudes of E_s, normalized with respect to mE_D, is shown in Fig. 4. In Fig. 2 the spatial distribution of E_s/mE_D is plotted over one period for $m = 0.9$ and several values of E_a/E_D. The nonlinearity is pronounced; the field contains a dc term and several harmonics. On the other hand, in Fig. 3 where $m = 0.02$, the field is sinusoidal. In Fig. 4 we see how the various Fourier components develop in time for $m = 0.9$. By comparison the fundamental component is also plotted for $m = 0.02$ (the other terms are negligible) as a simple exponential build-up. The nonlinear effect for large m not only creates harmonics, but also causes an increase of the amplitude of the fundamental component.

3.2 Uniform Illumination—Hologram Erasure

The physical situation corresponding to this case is that of reading out a hologram with only one of the beams in Fig. 1. The illumination has the constant value I_0 over the whole hologram and the photoexcited electrons concentration is just n_0. In this case $A(z) = 0$, $G = 0$,

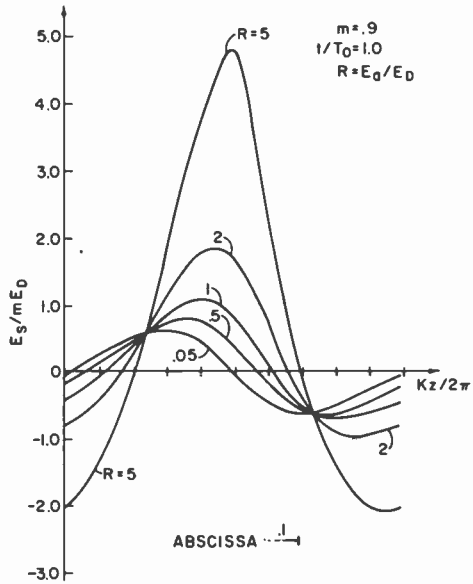


Fig. 2—Spatial distribution of stored field for $m \simeq 1$, with applied field as a parameter, under sinusoidal illumination.

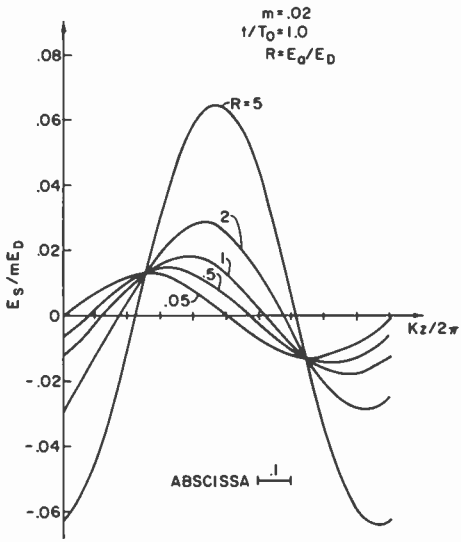


Fig. 3—Spatial distribution of stored field for $m \ll 1$, with applied field as a parameter, under sinusoidal illumination.

and if the initial value of E_s is $E_0(z)$, Eq. [12] gives the exponential decay.

$$E_s(z,t) = E_0(z) \exp(-t/T_0). \qquad [19]$$

Strictly speaking, the net illumination inside the medium is not a constant, even though the incident illumination is. In particular, if $E_0(z)$ is the result of the previous recording, the resulting refractive index modulation constitutes a hologram grating. Then exposure to either beam in Fig. 1 will reconstruct the other beam by diffraction; the interference of the two will result in $A(z) \neq 0$ and may enhance the hologram prior to erasing it (depending upon the direction of the incident illumination) if the diffraction efficiency is high.[20]

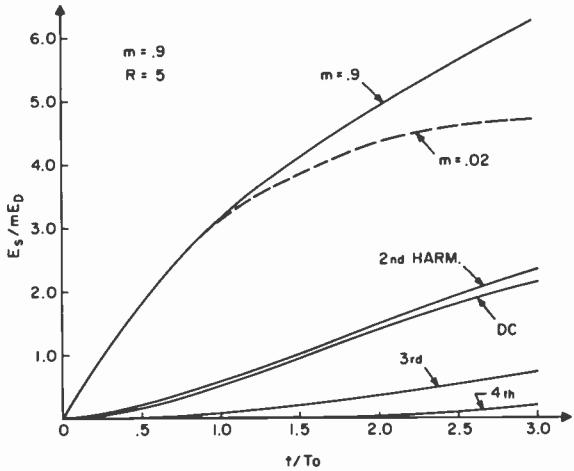


Fig. 4—Harmonic content of stored field for $m \simeq 1$, showing comparison with $m \ll 1$.

4. Experimental Verification by Means of Holography

The temporal dependence of the space-charge field as a function of illumination, its dependence on recording angle and electric field, and the nonlinear effects due to $m \approx 1$ can be observed experimentally. The observations can be made holographically on the fundamental component of the field, which produces a modulation of the extraordinary refractive index via the electro-optic effect. The material behaves as a thick phase grating that diffracts light of a given wavelength if the light is incident at the Bragg angle.[21] The first-order diffraction efficiency η is

$$\eta = \sin^2\left[\frac{\pi l|\Delta n_e|}{\lambda_0 \cos\theta_0}\right], \qquad [20]$$

where l is the thickness of the grating, $|\Delta n_e|$ the magnitude of the fundamental component of the extraordinary refractive index change, λ_0 the incident wavelength, and θ_0 the Bragg angle corresponding to λ_0. When the quantity in the bracket is much less than $\pi/2$, a condition that can always be arranged experimentally, we simply have

$$\eta \cong \left[\frac{\pi l}{\lambda_0 \cos\theta_0}|\Delta n_e|\right]^2 \qquad [21]$$

In terms of the fundamental component E_f of the stored field (along the z-axis), one has[22]

$$\Delta n_e = -\frac{1}{2} n_e^3 r_{33} E_f, \qquad [22]$$

where n_e is the extraordinary refractive index at λ_0 in the absence of field, and r_{33} the component of the electro-optic matrix. Thus, since the diffraction efficiency is proportional to $|E_f|^2$, its measurement at a wavelength for which the material is insensitive would provide a non-interacting measure of $|E_f|^2$.

A series expansion of Eq. [17] gives $E_f(z,t) = |E_f| \sin(Kz - \varphi)$, where

$$|E_f| = mE_{max}\left[f_0(t) + \frac{1}{2}\sum_{r=1}^{\infty}\left(\frac{m^2}{2}\right)^r f_{2r}(t)\right], \qquad [23]$$

$$f_0(t) = 1 - \exp(-t/T_0), \qquad [24]$$

$$f_2(t) = 1 - \{\exp(-t/T_0)\}\sum_{s=0}^{2r}\frac{1}{s!}(t/T_0)^s \qquad [25]$$

Eq. [23] is a power series expansion for $|E_f|$ in terms of $m^2/2$ consisting of a simple exponential build-up $f_0(t)$ and a number of progressively slower growing exponentials $f_{2r}(t)$. For $m \ll 1$ or $t/T_0 < 1$, $f_0(t)$ is the major time contributor to $|E_f|$ and to the total field, which is purely sinusoidal (Fig. 3 and Eq. [18]). For $m \simeq 1$ the other Fourier components of the stored field are non-negligible, and higher-order terms in $m^2/2$ add significantly to $|E_f|$.

The following can thus be verified experimentally:

(a) Since E_{max} is independent of the illumination, the saturation efficiency (the value of η for $t/T_0 \gg 1$) should also be independent of it. The illumination only determines its time development.

(b) E_{max} increases with K, i.e., with the recording-beam angle. Thus the saturation efficiency, or alternatively, the diffraction efficiency for any fixed illumination, should increase as the recording angle θ increases.
(c) E_{max} increases with applied field. Thus, again the saturation efficiency, or alternatively, the diffraction efficiency for any fixed value of the illumination, should be larger for holograms recorded in the presence of an applied field.
(d) The spatial distribution of the stored field for holographic recording is a nonlinear function of the illumination when the modulation index m is near unity. This should reflect in an increase in the diffraction efficiency, not accountable by a simple exponential, for $m \simeq 1$.

The experimental set-up is shown in Fig. 5. The 488-nm line from an argon laser, expanded and apertured to a uniform diameter of 2 mm, is split by the variable beam splitter into two beams I_1 and I_2,

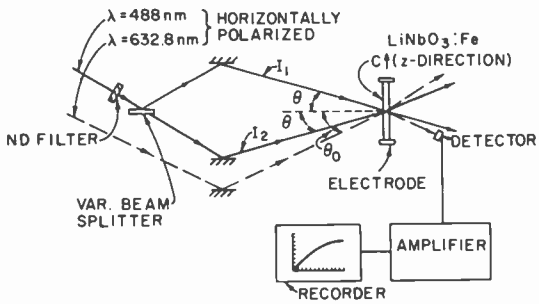


Fig. 5—Experimental setup for making holograms and simultaneously observing diffraction efficiency.

which impinge upon the iron-doped lithium niobate ($LiNbO_3$:Fe) sample at the angle θ. The width of the sample between the +C and −C edges is 3 mm and its thickness is 1.7 mm. The relative intensities of I_1 and I_2 can be controlled by means of the variable beam splitter. The $LiNbO_3$:Fe sample is oriented such that its C or optic axis (z-direction) is parallel to the plane of incidence, and the polarization of the laser is also in the same plane. By means of electrodes placed at the ±C faces of the sample, a voltage of 3 kV could be applied ($E_a \simeq$ 10 kV/cm, neglecting fringes). The total light intensity could be controlled by means of neutral density filters. The 632.8-nm light from a weak (0.2 mW) He-Ne laser was used at its appropriate Bragg angle θ_0 to continuously monitor the events at 488 nm, and the diffracted

632.8-nm light was detected, amplified, and fed to a strip chart recorder. The effect of the monitor light on the experiment was negligible in view of its low intensity and of the low absorption of the sample at its wavelength. It was thus possible to vary at will the total illumination I_0, the recording angle θ, the modulation index m, and the applied electric field E_a.

The $LiNbO_3$:Fe samples studied were prepared so as to maximize their record and erase sensitivity as described in Ref. (7). These samples have low overall Fe concentration and all but a small percentage of the Fe is in the divalent state (the remaining amount being trivalent). Samples were prepared by two methods, high-temperature argon annealing and Li_2CO_3 annealing, both of which are described in Ref. (10). We present here results measured in one argon-reduced sample that is typical of all the others having roughly the same Fe^{2+} and Fe^{3+} concentrations.

In order for any correlation to be possible between calculations and measurements it is necessary to be able to calculate T_0 under arbitrary illumination. Since n_0 is proportional* to I_0, all the material constants in Eq. (6) and in n_0 can be lumped into one single constant C such that

$$T_0 = \frac{C}{I_0} \qquad [26]$$

Also, from Eqs. [19], [21], and [22] the decay of the diffraction efficiency η of any hologram from its initial value η_0 under uniform illumination is of the form

$$\eta/\eta_0 = \exp(-2t/T_0) = \exp(-2I_0t/C). \qquad [27]$$

Thus a single measurement of T_0 for any decaying hologram gives the parameter C, which can subsequently be used in Eq. [26] to calculate T_0 for arbitrary illumination. This parameter C, whose units are that of energy density, truly characterizes a given material. For the sample used in Fig. 6 through 9 we found $C = 2.2$ mJ/mm^2 when I_0 is expressed in mW/mm^2.

Fig. 6 shows the time development curves with I_0 as a parameter. The curves were obtained by continuously monitoring and recording the diffraction efficiency at 632.8 nm while exposing the sample to the hologram interference pattern at 488 nm. For each curve the beams I_1 and I_2 are equal ($m = 1$) and $\theta = 10°$. The total illumination

* Actually[15] $n_0 = \sigma N \tau I_0/\hbar\nu$ where σ is the absorption cross-section, N the trap concentration, τ the electron life time, $\hbar$ Plank's constant and ν the optical frequency. This gives $C = \epsilon\hbar\nu/q\mu\sigma N\tau$.

I_0 was adjusted for each curve by means of neutral density filters. Its values for the experimental (solid) curves were 4.0, 2.0, 1.1, and 0.5 mW/mm^2, respectively. The dashed curves are the theoretical fits. They were obtained in two steps. First, the saturation efficiency η_{max} for one of the curves was calculated. This step was found to be more suitable than the actual measurement η_{max} due to the poor long-time stability of the bench. By successive trials, using Eq. [21] in the form $\eta = \eta_{max}F^2(t)$, where $F(t)$ is the expression in bracket in Eq. [23], and

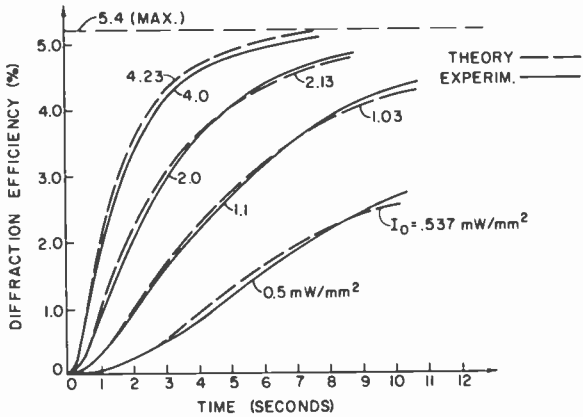


Fig. 6—Time development of diffraction efficiency, with illumination as a parameter. Saturation efficiency is independent of illumination.

using the experimental value of I_0, a value of η_{max} was found that gave the best fit to the experimental curve. By means of this step we were able to avoid the need to measure material parameters such as n_e and r_{33}. We found η_{max} = 5.4%. Next, using that value of η_{max}, the "theoretical" value of I_0 was then found to give a more refined fit to the experiment. Both the theoretical and experimental values of I_0 are indicated in Fig. 6 for each curve. The agreement is excellent.

The angular dependence is shown in Fig. 7, which is a plot of the diffraction efficiency as a function of the spatial frequency $f = K/2\pi$ lines/mm. For each data point the exposure was 2.3 mJ/mm^2 and the modulation index unity. For each new value of K, i.e., each new angle θ, care was taken to avoid errors due to beam misalignments. The results show strong correlation between the theory and the experiments up to a spatial frequency of $f \simeq 1600$ lines/mm, corresponding to θ = 23° or $K \simeq 10^4$ per mm, followed by a striking experimental departure from the theory beyond that turning point. Neither experimental errors nor the angular dependence of the reflection coefficient

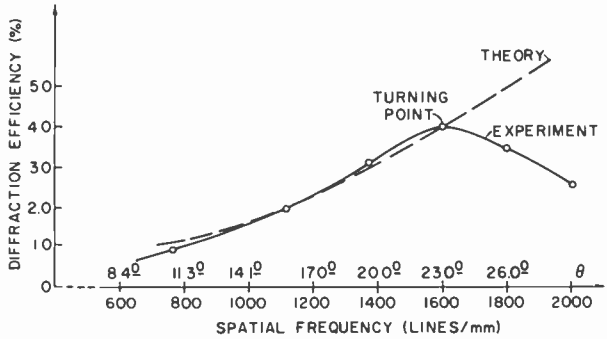


Fig. 7—Dependence of diffraction efficiency on recording-beam angles (grating frequency). At turning point, migration length is comparable to grating spacing.

could account for the deviation. A calculation of the diffusion length[17] (no applied field) $L = (D\tau)^{1/2} = (kT\mu\tau/q)^{1/2}$, using $\mu \simeq 15$ cm^2/V sec and $\tau \simeq 4 \times 10^{-11}$ sec, gives $L \simeq 4 \times 10^{-5}$ mm, i.e., $KL \simeq 0.4$ at the turning point. The field for values of K larger than the turning point decreases in a way consistent with Young's model,[16] which includes situations where $KL > 1$. This model predicts an increase of E_f up to a certain point as KL increases, and an eventual decline in the form of $1/K$ ($1/K^2$ for η) for sufficiently large values of KL. The turning point in Fig. 7 is the point at which Amodei's model breaks down.

The effect of an applied 3 kV voltage is shown in Fig. 8 for $\theta = 20°$ and $m = 1$. From the theory, a curve closely matching the experimen-

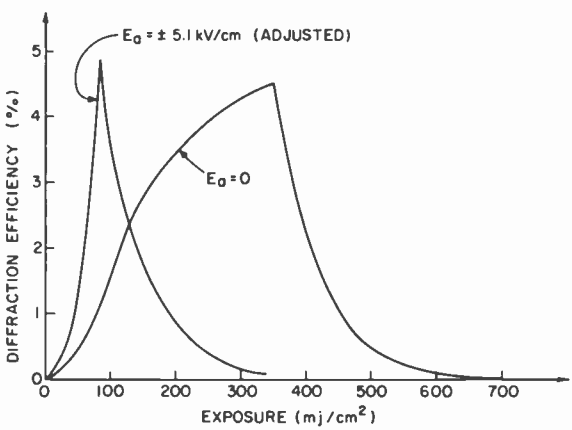


Fig. 8—Effect of applied electric field.

tal data could be obtained by setting $E_a = \pm 5.1$ kV/cm as compared to an expected field of 10 kV/cm. This discrepancy may be attributed to faulty electrodes. Another possible explanation lies in the fact that the migration length in this case is $L = \mu\tau E_a = 6 \times 10^{-5}$ mm for $E_a =$ 10 kV/cm. For $\theta = 20°$ this gives $KL = 0.53$, and since this value is above the turning point, the effectiveness of E_a is reduced. The decay characteristics of the diffraction efficiency are also shown in Fig. 8.

The nonlinear characteristics are shown in Fig. 9 to compare the time development of holograms recorded with the same total illumination of 1.1 mW/mm^2, but with $m = 1.0$ and 0.5. Since T_0 is 2 seconds for both cases, the curves are plotted as a function of t/T_0 in order to compare them. The increase of diffraction efficiency with m

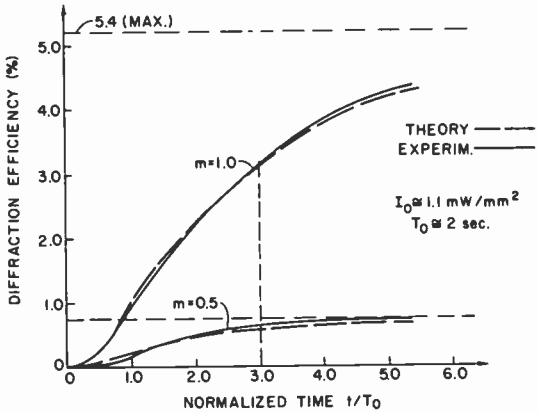


Fig. 9—Effect of modulation index on diffraction efficiency showing the increase for $m \simeq 1$.

is clearly indicated. For example, a simple exponential build-up of the efficiency (in the form of $f_0^2(t)$) would reach 92% of its saturation value in 3 time constants ($t/T_0 = 3$). However, it can be verified that the build-up reaches 87% for $m = 0.5$ and only 55% for $m = 1$ in the same time interval, a behavior that is quite consistent with Eq. [23].

5. Conclusion

The transport equation for the stored space-charge field in photorefractive media in the presence of arbitrary illumination and electric field has been solved rigorously, using the space-charge migration model, and analytical solutions are given for the particular cases dealing with hologram storage and erasure under Amodei's assumption

that the migration distance is much smaller than the hologram grating spacing. The solution, which in general is nonlinear, gives the explicit time development of the stored field as a function of illumination, as well as the field dependence on recording angle and applied field. The theory is verified experimentally for iron-doped lithium niobate by monitoring the diffraction efficiency of the hologram grating resulting from the modulation of the extraordinary refractive index of the material by the stored field. The experimental results are in agreement with those calculated from the model at low to medium spatial frequencies and seem to be in accordance with a model by Young et al at high spatial frequencies.

In calculating the transient behavior of a given material under arbitrary illumination, only one material constant need be known. The constant C, defined in Eq. [26], which combines all the other pertinent material constants and whose units are those of energy density, can be easily obtained by a measurement of the decay time constant of a stored field under uniform illumination. The smaller the constant, the faster the build-up of the stored field in any given material. It is a true measure of the optical sensitivity of the material and, as such, can be used as a means for comparing or optimizing the preparation of materials for holographic storage.

Acknowledgment

The authors are grateful to J. A. Rajchman, G. D. Cody, B. F. Williams, and R. D. Lohman for encouraging this work, to W. C. Saunders for technical assistance, and to several of their colleagues at RCA Laboratories for technical discussions. They are particularly indebted to W. J. Burke and B. F. Williams for their constructive review of the manuscript.

References:

[1] A. Ashkin, G. D. Boyd, J. M. Dziedzie, R. G. Smith, A. A. Ballman, J. J. Levinstein, and K. Nassau, "Optically-Induced Refractive Index Inhomogeneities in $LiNbO_3$," *Appl. Phys. Lett.*, **9**, p. 72 (1966).

[2] F. S. Chen, "A Laser-Induced Inhomogeneity of Refractive Indices in KTN," *J. Appl Phys.*, **38**, p. 3418 (1967).

[3] F. S. Chen, "Optically-Induced Change of Refractive Indices in $LiNbO_3$ and $LiTaO_3$," *J. Applied Phys.*, **40**, p. 3389 (1969).

[4] F. S. Chen, J. T. LaMacchia, and D. B. Fraser, "Holographic Storage in Lithium Niobate," *Appl. Phys. Lett.*, **13**, p. 223 (1968).

[5] J. B. Thaxter, "Electrical Control of Holographic Storage in Strontium Barium Niobate," *Appl. Phys. Lett.*, **15**, p. 210 (1969).

[6] F. Micheron, C. Mayeux, and J. C. Trotier, "Electrical Control in Photoferroelectric Materials for Optical Storage," *Appl. Optics*, **13**, p. 784 (1974).

[7] D. L. Staebler and W. Phillips, "Fe-Doped $LiNbO_3$ for Read-Write Applications," *Appl. Optics*, **13**, p. 788 (1974).

[8] J. J. Amodei, W. Phillips, and D. L. Staebler, "Improved Electro-optic Materials and Fixing Techniques for Holographic Recording," *Appl. Optics*, **11**, p. 390 (1972).

[9] R. S. Mezrich, "Magnetic Holography," *Appl. Optics*, **9**, p. 2275 (1970).

[10] W. Phillips and D. L. Staebler, "Control of the Fe^{2+} Concentration in Iron-doped Lithium Niobate," *J. Electronic Mat.*, **3**, p. 601 (1974).

[11] J. J. Amodei, "Electron Diffusion Effects During Hologram Recordings in Crystals," *Appl. Phys. Lett.*, **18**, p. 22 (1971).

[12] W. D. Johnston, Jr., "Optical Index Damage in $LiNbO_3$ and Other Pyroelectric Insulators," *J. Appl. Phys.*, **41**, p. 3280 (1970).

[13] A. M. Glass, D. von der Linde, and J. T. Negran, "High-Voltage Bulk Photovoltaic Effect and the Photorefractive Process in $LiNbO_3$," *Appl. Phys. Lett.*, **25**, p. 233 (1974).

[14] P. V. Lenzo, E. G. Spencer, and K. Nassau, "Electro-Optic Coefficients in Single-Domain Ferroelectric Lithium Niobate," *J. Opt. Soc. Amer.*, **56**, p. 633 (1966).

[15] D. von der Linde, A. M. Glass, and K. F. Rodgers, "Multiphoton Photorefractive Processes for Optical Storage in $LiNbO_3$," *Appl. Phys. Lett.*, **25**, p. 155 (1974); also "High-Sensitivity Optical Recording in KTN by Two-Photon Absorption," *Appl. Phys. Lett.*, **26**, p. 22 (1975).

[16] J. J. Amodei, "Analysis of Transport Processes During Holographic Recording in Insulators," *RCA Review*, **32**, p. 185 (1971).

[17] L. Young, W. K. Y. Wong, M. S. W. Theriault, and W. D. Cornish, "Theory of Formation of Phase Holograms in Lithium Niobate," *Appl. Phys. Lett.*, **24**, p. 264 (1974).

[18] S. R. King, T. S. Hartwick, and A. B. Chase, "Optical Damage in KTN," *Appl. Phys. Lett.*, **21**, p. 312 (1972).

[19] P. M. Morse and H. Feshback, *Methods of Theoretical Physics*, p. 706, McGraw-Hill Book Co., N.Y., (1953).

[20] D. L. Staebler and J. J. Amodei, "Coupled-Wave Analysis of Holographic Storage in $LiNbO_3$," *J. Appl. Phys.*, **43**, p. 1042 (1972).

[21] H. Kogelnik, "Coupled Wave Theory for Thick Hologram Gratings," *Bell System Tech. J.*, **48**, p. 2909 (1969).

[22] J. F. Nye, *Physical Properties of Crystals*, p. 252, Oxford University Press, London, (1969).

Reliability Aspects and Facet Damage in High-Power Emission from (AlGa)As CW Laser Diodes at Room Temperature*

H. Kressel and I. Ladany

RCA Laboratories, Princeton, N.J. 08540

Abstract—Factors are described that limit the optical power output from (AlGa)As laser diodes (λ = 8100–8300 Å) operating cw at room temperature with uncoated facets. Rapid laser "catastrophic" degradation due to facet damage (in contrast to "bulk" phenomena previously considered) has been found to occur as a result of excessive optical flux density at the facets. The diodes studied are capable of initial cw power emission values of 25 to 100 mW from one facet depending on their dimensions. Data are presented showing long-term constant-current operation at power levels below these maximum values. Preliminary data are also presented on devices utilizing dielectric facet coatings to minimize facet damage.

Introduction

Significant progress has been made in improving the operating lifetime of cw heterojunction laser diodes (and high radiance LED's) at room temperature. This has resulted from improved knowledge concerning the relevant device aspects affecting the formation of nonradiative centers[1–6] and external factors such as facet erosion resulting from ambient conditions.[5] While attention has been paid to internal device effects, it has been generally assumed that the cw power emission levels were too low to result in rapid device degradation due

* This research was partially supported by NASA, Langley Research Center, Hampton, Virginia, under Contract NAS1-11421 and the Office of Naval Research, Arlington, Virginia, under Contract N00014-73-C-0335.

to facet damage ("catastrophic degradation")[7] well known to occur in pulsed diode operation when the power density at the emitting facet reaches values in the MW/cm^2 range. For pulsed laser operation, it is experimentally found[8] that the power level at which facet damage occurs decreases approximately with pulse length t as $t^{-1/2}$. Hence, it is reasonable to suppose that cw laser operation will induce damage at substantially lower levels than usually observed in pulsed diode operation.

The object of the present work was to explore the power density limits in practical cw laser diodes. We show by comparing cw and pulsed diode operation that facet failure (uncoated facets) occurs at significantly reduced levels in cw operation. In addition, long term cw operation tests were conducted with similar lasers at initial power emission levels which did not exceed ⅓ the maximum value from a given diode. Many thousands of hours of relatively stable operation are obtained under conditions of constant current operation.

Experimental Results

The diodes studied, designed for emission in the 8100–8300 Å spectral range, were $Al_xGa_{1-x}As-Al_yGa_{1-y}As$ double heterojunction structures, with typically $x \approx 0.1$ in the recombination region and $y \approx 0.3$ in the adjoining p- and n-type regions, as discussed previously.[5] The material was grown on (100) GaAs substrates by liquid-phase epitaxy using the thin-melt technique[9] in a horizontal growth apparatus. The width of the n-type recombination region was 0.2–0.3 μm. Germanium was used to dope the p-type regions, while Sn or Te were used for the n-type regions. The recombination region was not deliberately doped, and has an electron concentration of 5×10^{16} cm^{-3}. To minimize the thermal and electrical resistance of the diodes, the distance between the recombination region and the surface of the diode was typically 2–3 μm and included a highly doped p-type GaAs layer to improve the ohmic contact.

The combination of a very narrow recombination region and appropriate Al concentrations in the two surrounding layers of the diodes completely confines the carriers to the recombination region, but only partially confines the radiation there.[10] This is reflected in the far-field radiation pattern, which shows a full beam width value (at half-power) of 33 to 38° in the direction perpendicular to the junction plane, indicative of an "effective aperture" greater than the recombination region width.

The active diode area was defined by a SiO_2-isolation stripe-contact process[11] with stripe widths of 13, 50, and 100 μm. The diodes

were mounted p-side down on copper heat sinks using indium as the solder.

Fig. 1 shows the current dependence of the light emitted by typical wide stripe diodes used in this study. For the 50 μm stripe width, the indicated threshold current is 0.76 A and the cw power output peaks at 90 mW (one side) at a current of 1.3 A. The saturation in the power output is due to the increase with current of the junction temperature and the consequent increase in the threshold current density. The position of the maximum in the power output depends on the thermal and electrical resistance of the diode, and on the threshold current density dependence on temperature.[12] Comparative studies were conducted with the 100-μm-wide stripe diodes in which the power output saturates at ~ 100 mW (~ 200 mW total emission) at a current of 1.6 A as shown in Fig. 1(b).

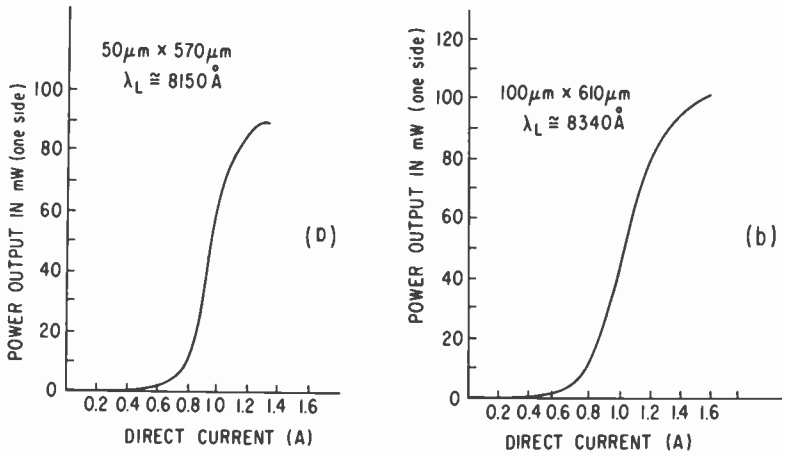


Fig. 1—Power emission from one laser facet as a function of direct current for two lasers selected from different wafers operating cw at room temperature: (a) diode with 50-μm-wide stripe contact; (b) diode with 100-μm-wide stripe contact.

The facet damage effects were studied in the 13- and 50-μm-wide stripe diodes where cw-operation damage could be induced within relatively short periods of time (on the order 1 hr) when the diodes were operated at their peak optical output. The diode facets were uncoated and the damage experiments were conducted in a normal laboratory ambient with the heat sink temperature kept constant at ~ 20°C, with provisions made to ensure that the diode facets were dry.

First, we consider facet damage seen in a 50-μm-wide stripe laser initially emitting 90 mW from one facet. When the emitted power had decreased to ¼ of its initial value, the damaged region on the facet shown in Fig. 2 was observed. This region is 25 μm wide and occurs in the central region under the stripe contact where the current and, hence, the power emission is at a maximum. (Two other samples

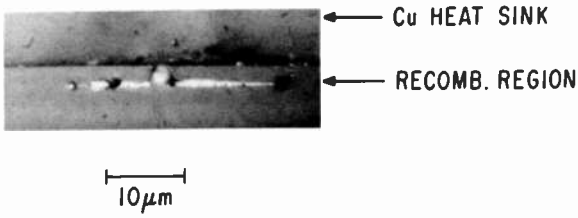


Fig. 2—Optical micrograph of the facet damage following operation at the 90 mW emission level from a 50-μm-wide stripe diode. The extent of the damaged region is ~25 μm in the central region of the stripe width.

tested failed at power levels within 10 mW of the above value.) Facet damage could not be induced under similar test conditions in less efficient diodes operating at the same current of 1.3 A (for example, diodes emitting only 70 mW instead of 90 mW), suggesting a correlation between facet damage and the power emission level. Other observations were consistent with this hypothesis. No facet damage was induced in the 100-μm-wide stripe diodes where a power level as high as 110 mW was emitted from a region about twice as wide.

In the 13-μm-wide diodes, facet damage occurred at power levels between 23 and 35 mW (in four samples tested). As shown in Fig. 3, the damaged region extends over a width of 5–6 μm in the plane of the junction and is thus similar in this respect to the damage in the wider diodes, which extends over only part of the stripe. Consistent with the damage observed, the central portion of the stripe contact has the highest optical power density as shown by examination of the near-field emission of Fig. 4.* Note in the scanning-electron-micrograph of Fig. 3(b), that the mechanical damage spreads substantially beyond the recombination region in a direction *normal* to the junction plane.

In order to obtain a further basis of comparison, 13-μm-wide stripe diodes similar to those studied above were operated with pulse lengths of ~100 nsec† (1 kHz) and 400 nsec (250 Hz) under conditions

* We are indebted to H. S. Sommers, Jr. for these data.

† Gaussian pulse shape, 100 nsec width at half-intensity point.

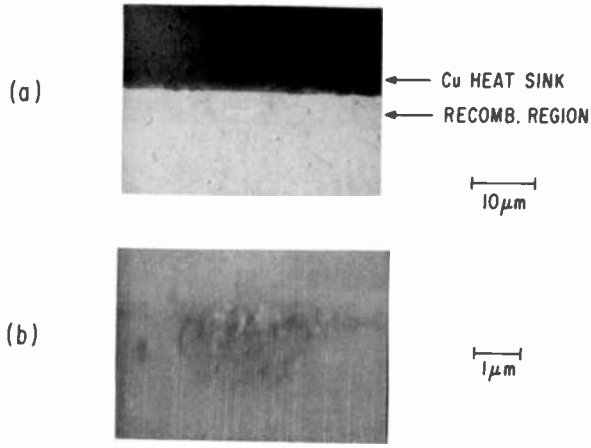


Fig. 3—Facet damage in a 13-μm-wide stripe diode: (a) optical micrograph showing damage extending over a distance of 5–6 μm in the stripe center; (b) scanning-electron-microscope micrograph showing that the damaged region extends substantially in the direction perpendicular to the junction plane.

to induce catastrophic damage. We found that of three diodes tested, the damage limits were 220, 230, and 260 mW with 100 nsec pulse length, while with the 400 nsec pulse, damages occurred at an average value of 90 mW. These damage power level values are therefore significantly greater than those found under cw operation.

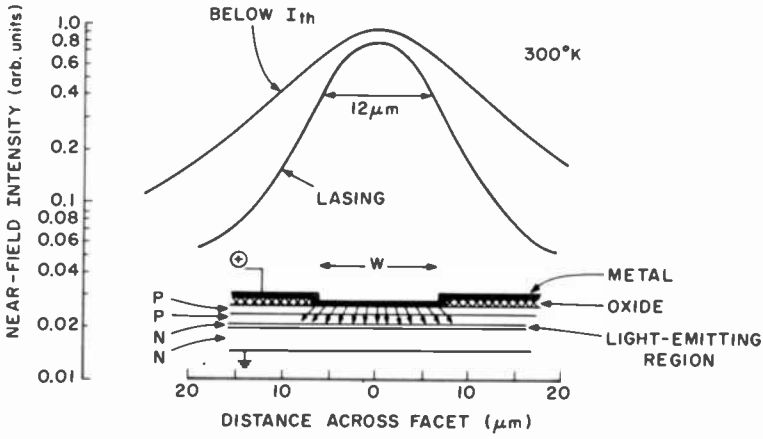


Fig. 4—Near-field intensity distribution of a 13-μm-wide stripe laser diode below and just above lasing threshold. Note the peak intensity at the center of the stripe region. (The two curves are shifted for convenience.)

The damage results described above are phenomena that occur in a relatively short time period. It is obviously of practical interest to determine the power level that can be sustained for long term operation, although a definitive study of such effects must involve a large number of samples and testing over periods of years. Preliminary tests were conducted with the diodes operating in a dry air atmosphere to minimize facet erosion accelerated by moisture.[5] Both 13-

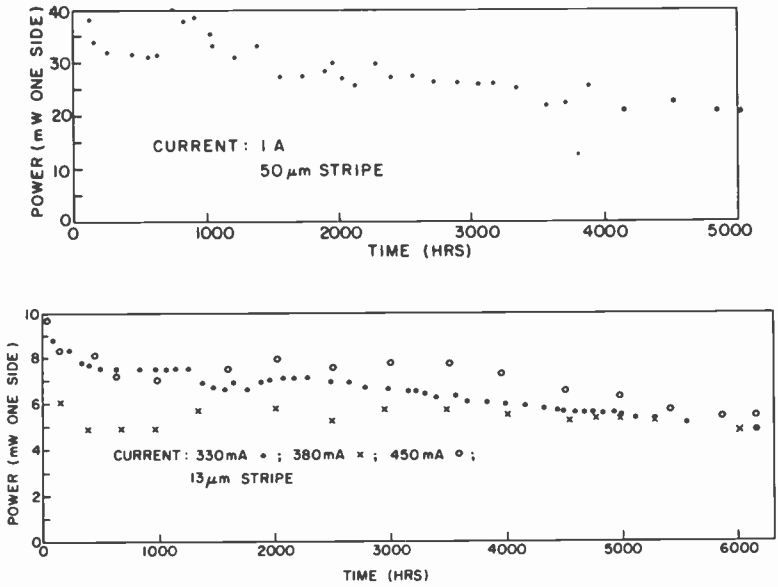


Fig. 5—Power emitted from one facet of cw (AlGa)As lasers operating at room temperature at constant current: (a) diode with 50-μm-wide stripe operating at a drive current of 1 A; (b) diodes with 13-μm-wide stripe operating at 330 mA, 380 mA, and 450 mA. The fluctuations in the power output are partially due to slight changes in the ambient temperature, which reversibly changes the threshold current density. (Uncoated laser facets.)

μm-wide stripe lasers and 50-μm-wide lasers were subjected to extended constant-current life tests at initial power levels not in excess of one-third the initial maximum attainable from a given device. The output from one facet as a function of operating time at a fixed current is shown in Fig. 5. The 50-μm-wide stripe diode is emitting between 22 and 33 mW from one facet for more than 5000 hrs with a drive current of 1 A. The 13-μm-wide lasers are emitting between 5 and 7 mW for more than 6000 hrs. (The fluctuations in the power

output seen in Fig. 5 are due to slight changes in the ambient temperature, which affect the threshold current and are reversible.) The power reduction seen in some of the diodes can be partly recovered by an increase in the diode current.

Discussion

It is evident that the damage level will scale with the extent of the emitting region, but the relationship between the observed power level for facet damage and the actual optical flux density is difficult to establish with high accuracy in stripe-contact diodes because of the graded lateral and perpendicular radiation distribution. For the 13-μm stripe diodes, the near-field pattern shows that the half-power point is reached at the edge of the stripe (Fig. 4), with the radiation being most intense in the central region where the damage is observed. However, stimulated emission occurs over a significant distance beyond the stripe width because of the lack of effective lateral current confinement. Considering the initial near-field distribution and power range at which damage was observed (23–35 mW), the power level in the 6 μm damaged region is estimated to be between 2 and 3 mW/μm of emitting region.

For the 50 μm-wide stripes, we noted that the damage occurred in the central 25 μm of the device. We also noted that the power reduction was typically about 75% when damage of the type shown in Fig. 2 was observed. Taking the value of 90 mW at the damage point, the power concentrated in that 25 μm region can be assumed, as a first approximation, to represent 75% of 90 mW, or 68 mW, and hence to correspond to 2.7 mW/μm of emitting region, a value within the above estimate for the 13-μm stripe diodes.

With regard to the power density, the best estimate of the effective perpendicular width of the emitting regions based on the far-field beam width of these devices (33–38°), is about 1 μm (but not less than 0.7 μm). Hence, the optical power density is $2 - 4.2 \times 10^5$ W/cm^2. These values can be compared to the much higher power level for facet damage estimated[13] for 100-nsec pulse-length measurement of $4 - 8 \times 10^6$ W/cm^2. This large difference between the cw and 100-nsec-pulse-operated power densities for failure is also consistent with our measurement results for pulsed operation described above.

We wish to emphasize that in all cases of facet damage, the effect can occur at lower than expected power levels in the presence of dust particles on the laser facet, gross metallurgical flaws and, as pointed out earlier, when operation occurs in a relatively moist ambient.[5] Facet protection is, therefore, very desirable. Furthermore, the use of

antireflective films (such as Al_2O_3 or SiO) on the facet are known to improve the catastrophic damage limit by as much as a factor of 3 in the case of pulsed diode operation.[14] A similar improvement is expected for cw laser operation, and preliminary results do show a significant improvement in our cw devices.* The above results should not, therefore, be taken as indicative of the maximum power that can be reliably achieved from stripe-contact lasers having the indicated geometry.

Fig. 6 shows preliminary life data at constant current for a group of four 13-μm-wide stripe lasers with facet coatings. (The heat sink temperature was maintained at 22°C ± 1°). Little or no evidence of degradation is seen in the time indicated. It is possible that the slow reduction in the power emission seen in the diodes with uncoated facets operating at constant current is due to facet damage, possibly starting at tiny flaws on the emitting facets.

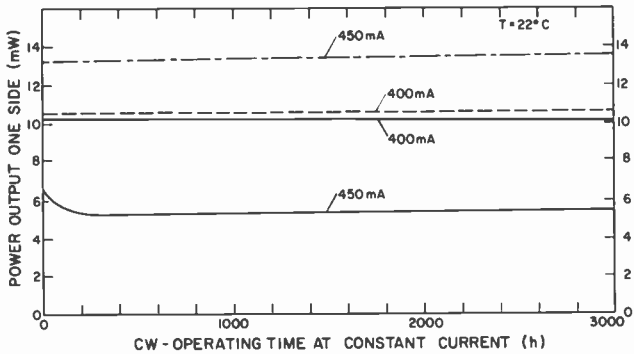


Fig. 6—Power emitted from one facet of four lasers with dielectric facet coating operating cw with a heat sink temperature of 22 ± 1°C. The fixed current of operation is shown for each diode. The stripe width was 13 μm.

It is encouraging to note that useful power output over many thousands of hours is possible with the relatively simple construction used for the present diodes. Long term operating-life data (in excess of 1000 hrs) were previously reported for diodes using other stripe contact formation techniques, including selective diffusion,[3] etched mesas,[6] and proton bombardment.[2] The reported data generally indicate that substantial current increases were needed to maintain rea-

* The dielectric coating experiments were performed in collaboration with M. Ettenberg and H. F. Lockwood.

sonably constant power output, an indication of degradation occurring in the course of operation.

Conclusions

A study has been made of operating conditions leading to facet damage (catastrophic degradation) of (AlGa)As laser diodes operating cw at room temperature (emission wavelength of 8100–8300 Å). It has been shown that facet damage occurs when the power level in the highest intensity region of the uncoated emitting facet reaches an estimated value of 2–3 mW/μm. Based on a reasonable approximation of the extent of emitting region in the direction perpendicular to the junction plane, the estimated optical power density is 2–4.2 $\times$ 10^5 W/cm^2. This value is several times lower than observed under short-pulse operation (100 nsec) for the same type of device.

It has also been shown that similar diodes have operated for periods of time in excess of 5000 hrs at initial power values that did not exceed ⅓ the damage limit. The maximum reduction in the output of diodes with uncoated facets at constant current was below 50%. (An increase in the current can recover part of the power reduction.) The cause for the observed reduction in some of the diodes tested is as yet undetermined, but could be due to incipient facet damage originating at tiny facet flaws. Preliminary results using dielectric facet coating indicate that the damage limit is substantially increased (as has previously been observed in diodes designed for pulsed operation). Facet protection is believed to be important, for very long term stable operation, as a means of eliminating facet damage.

Acknowledgments

We are indebted to H. F. Lockwood and M. Ettenberg for collaboration on the dielectric film work; to D. P. Marinelli, V. M. Cannuli, D. B. Gilbert, H. Kowger, and M. Harvey for technical assistance; and E. R. Levin for the scanning electron microscopy.

References:

[1] A comprehensive discussion of the literature until 1973 was presented by H. Kressel and H. F. Lockwood, "A Review of Gradual Degradation Phenomena in Electroluminescent Diodes," *J. de Physique, C3*, Suppl. Vol. 35, p. 223, 1974.

[2] R. L. Hartman and R. W/ Dixon, "Reliability of DH Lasers at Elevated Temperatures," *Appl. Phys. Lett.* **26**, p. 239, 1975, and references therein.

[3] H. Yonezu, K. Kobayashi, K. Minemura, and I. Sakuma, "GaAs-$Al_xGa_{1-x}As$ Double Heterostructure Laser for Optical Fiber Communication System," *International Electron Devices Meeting, IEEE, Tech. Digest*, p. 324, 1973.

[4] M. Ettenberg, H. Kressel, and H. F. Lockwood, "Degradation of $Al_xGa_{1-x}As$ Heterojunction Electroluminescent Devices," *Appl. Phys. Lett*, **25**, p. 82, 1974.

[5] I. Ladany and H. Kressel, "The Influence of Device Fabrication Parameters on the Gradual Degradation of (AlGa)As CW Laser Diodes," *Appl. Phys. Lett.*, **25**, p. 708, 1974.

[6] O. Nakada, N. Chinone, S. Nakamura, H. Nakashima, and R. Itoh, "Continuous Operation Over 2500 Hours of Double Heterostructure Laser Diodes with Output Powers More than 80 mW," *Japan, J. Appl. Phys.*, **13**, p. 1485, 1974. (By private communication, the authors have indicated 40mW/face emission powers.)

[7] H. Kressel and H. P. Mierop, "Catastrophic Degradation in GaAs Injection Lasers," *J. Appl. Phys.*, **38**, p. 5419, 1967.

[8] Extensive data are given by P. G. Eliseev, "Degradation of Injection Lasers," in *Semiconductor Light Emitters and Detectors*, A. Frova, ed., North-Holland Publishing Co., Amsterdam, 1973, p. 338. [*J. Luminescence*, Vol. 7, 1973].

[9] H. F. Lockwood and M. Ettenberg, "Thin Solution Multiple Layer Epitaxy," *J. Crystal Growth*, **15**, p. 81, 1972.

[10] H. Kressel, J. K. Butler, F. Z. Hawryio, H. F. Lockwood, and M. Ettenberg, "Mode Guiding in Symmetrical (AlGa)As-GaAs Heterojunction Lasers with Very Narrow Active Regions," *RCA Review*, **32**, p. 393, 1971.

[11] J. C. Dyment, "Hermite-Gaussian Mode Pattern in GaAs Junction Lasers," *Appl. Phys. Lett.*, **10**, p. 84, 1967.

[12] R. W. Keyes, "Thermal Problems of the CW Injection Laser," *IBM J. Research and Develop.*, **15**, p. 401, 1971.

[13] B. W. Hakki and R. Nash, "Catastrophic Failure in GaAs Lasers," *J. Appl. Phys.*, **45**, p. 3907, 1974.

[14] M. Ettenberg, H. S. Sommers, Jr., H. Kressel, and H. F. Lockwood, "Control of Facet Damage in GaAs Laser Diodes," *Appl. Phys. Lett.*, **18**, p. 571, 1971.

Mobility of Current Carriers in Silicon-on Sapphire (SOS) Films*

S. T. Hsu and J. H. Scott, Jr.

RCA Laboratories, Princeton, N. J. 08540

Abstract—A new method has been developed to measure the mobility of current carriers in thin semiconductor films. A deep-depletion MOSFET is used for the mobility measurements. It is shown that by measuring the gate capacitance and the transconductance in the linear region as a function of gate bias voltage, the mobility of the majority current carriers can be obtained as a function of distance from the surface of the film. The possible sources of errors and methods of overcoming these errors are also discussed. Experimental data on mobilities were obtained from n- and p-type SOS films with film thickness varying from 0.4 to 2 μm. It was found that the mobility of current carriers near the surface of thicker films is limited by ionized impurity scattering. The mobility of current carriers in thinner films and in thicker films decreases exponentially with distance from the Si-SiO_2 interface. However, in thicker films, there is a region near the interface where the mobility is independent of depth. The mobility, presumably, is limited by dislocation scattering. The method described also measures the mobility of current carriers in accumulated surface layers. When the accumulated charge density is larger than 10^{12} cm^{-2}, the mobility of accumulated surface charge carriers is approximately equal to that in an inverted surface with the same current-carrier density.

1. Introduction

Silicon-on-sapphire (SOS) films have been used extensively in the recent years for the fabrication of high-speed metal–oxide–semiconduc-

* The research reported in this paper was sponsored by the Wright Patterson Air Force Base, Dayton, Ohio under contract F3615-73-C-1117.

tor (MOS) integrated circuits. Several authors[1–4] have proposed methods of measuring the electrical properties of thin semiconductor films. Those methods, either destructive or nondestructive, measure the average current carrier mobility. In addition, those methods require the preparation of special samples. In this paper we propose a new method to measure the carrier mobility of thin semiconductor films as a function of distance from the surface. The measurement is nondestructive and is performed on a deep-depletion MOS field-effect transistor (MOSFET). This measurement, therefore, may be taken from one of the deep-depletion MOS transistors (MOST's) in the integrated circuits. If no deep-depletion MOST is available in the circuit, a sample may be prepared by exactly the same process as that used for the MOS integrated circuit.

2. Theory

Consider an ideal deep-depletion, n-type, thin-film MOSFET as shown in Fig. 1 (the calculation is also directly applicable to p-type

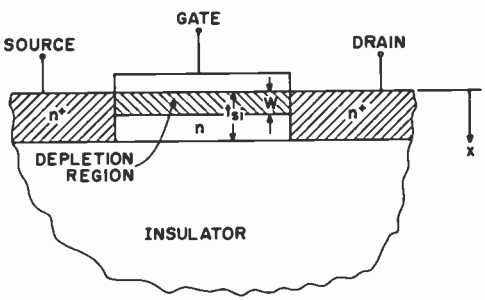


Fig. 1—Ideal deep-depletion thin-film MOSFET.

films). The gate of the FET is biased such that the film is partially depleted. The thickness of the semiconductor film is t_{si}, the width of the depletion region is W, and the length and the width of the gate are L and Z, respectively. The conductance between the source and drain is:

$$G = \frac{Z}{L} \int_{W}^{t_{si}} Q(x)\mu(x)dx \qquad [1]$$

where x is the distance between the plane under consideration and the Si-SiO_2 interface and $Q(x)$ and $\mu(x)$ are the carrier density and

the carrier mobility at x, respectively. The gate voltage is given by:

$$V_G = V_{FB} - \frac{1}{C_o}\int_W^0 Q(x)dx + \frac{1}{\epsilon_s}\int_W^0\int_W^X Q(x')dx'dx \qquad [2]$$

Here, V_{FB} is the flat-band voltage, C_o is the unit area capacitance of the gate insulator, and ϵ_s is the dielectric constant of the semiconductor. A complete depletion of current carriers has been assumed in formulating Eqs. [1] and [2].

Differentiating Eqs. [1] and [2] with respect to W yields

$$\frac{\delta G}{\delta W} = -\frac{Z}{L}Q(W)\mu(W) \qquad [3]$$

and

$$\frac{\delta V_G}{\delta W} = \frac{Q(W)}{C_o} + \frac{W}{\epsilon_s}Q(W) = \frac{Q(W)}{C_g(W)} \qquad [4]$$

where

$$C_g(W) = \frac{C_o\epsilon_s}{\epsilon_s + C_oW} \qquad [5]$$

Combining Eqs. [3] and [4] we obtain the variation of source to drain conductance with respect to gate voltage.

$$\frac{dG}{dV_G} = -\frac{Z}{L}\mu(W)C_g(W). \qquad [6]$$

The mobility of current carriers at a distance W from the gate-insulator–semiconductor interface is therefore given by

$$\mu(W) = -\frac{L}{ZC_g(W)}\frac{dG}{dV_G}. \qquad [7]$$

If V_D is the dc voltage between the source and the drain of the device, the conductance, G, can be written as

$$G = \lim_{V_D \to 0}\frac{I_D}{V_D}. \qquad [8]$$

Substituting Eq. [8] into Eq. [7] yields

$$\mu(W) = -\lim_{V_D \to 0}\frac{Lg_{mo}}{ZC_g(W)V_D} \qquad [9]$$

where

$$g_{mo} = \lim_{V_D \to 0} \frac{dI_D}{dV_G}$$

is the small-signal transconductance of the deep-depletion MOST's operated in the linear region. Thus, measurement of the gate capacitance and the small-signal transconductance of a deep-depletion thin-film MOSFET as a function of dc gate bias voltage provides the carrier mobility in the thin semiconductor film at a given distance from the gate-insulator–semiconductor interface. The distance, W, was obtained from the gate capacitance versus the gate bias voltage plot as given by Eq. [5].

If we differentiate Eq. [5] with respect to W and combine the result with Eq. [4], the charge density, $Q(W)$, at W is obtained:

$$Q(W) = \frac{C_g^3}{\epsilon_s \frac{dC_g}{dV_G}} \qquad [10]$$

Eq. (10) is exactly the same as that given by Fair.[5]

The material of a thin layer of semiconductor grown on an insulator is far from perfect, and difficulties can be caused by an imperfect material. We discuss sources of errors of this mobility measurement and suggest means to overcome these difficulties. The effects of surface-state charges and the assumption of depletion approximation on the measurement of the doping concentration in MOS structures has been discussed very extensively by Nicollian et al[6] and Darwish.[7] In the measurement of carrier mobility, the same difficulties appear on the conversion of gate voltage into the width of the depletion region, W. We shall not report these discussions here.

2.1 Deep Trap States Effects

It has been shown[8–11] that there is a large density of electron trap states in SOS films. The most important trap states are located at 0.65 eV below the edge of the conduction band. These electron trap states will be occupied when the quasi-Fermi energy of electrons is larger than the trap state energy. In n-type films all deep electron trap states are occupied when the surface potential of the film is smaller than 0.6 eV. The trapped electrons have the opposite polarity of that of ionized donor states. The net charge density in the depletion region is decreased when density of trapped electrons is in-

creased. The width of the depletion region when electrons are trapped is larger than that when no electrons are trapped. However, when the surface potential approaches 0.65 eV or larger, the trap states at the surface region become unoccupied. The net charge density at that region is increased. Similarly, in p-type films at small gate biases the deep electron trap states are empty. At larger gate biases, part of the deep electron trap states are occupied, which increases the charge density in the depletion region. The width of the depletion region of both p- and n-type films increases more slowly with large gate bias voltage than it does with small gate bias voltages if the doping concentration of the film is uniformly distributed. The deep electron trap states, therefore, play an important role in determining the width of the depletion region. The time constant of these trap states was found to be of the order of a millisecond.[9,10] Consequently, the use of pulsed measurement techniques (pulse repetition rates larger than a few thousand per second) should make it possible to eliminate deep trap state effects. If the frequency of the small signal applied to the gate is smaller than a few kHz, the deep electron trap states are able to follow the signal frequency and contribute current to the source–drain terminals. The measured mobility is, therefore, larger than the actual value.

2.2 Generation-Recombination (g-r) Centers and Fast Surface State Effects

The fast surface states are effective g-r trapping centers. The effect of fast surface states on the mobility measurement are the same as that of g-r centers in the depletion region. The current generated by the g-r centers in the depletion layer and the fast surface states of the film flows through the source–drain terminals. Since the occupancy probability of these centers depends on the surface potential, the fast surface states and the g-r centers affect the small-signal transconductance of the MOSFET's. These centers also influence the capacitance measurement. If the measurement frequency is sufficiently large, the surface states and the g-r centers are unable to follow the signal frequency and their effect on the mobility measurement is not important. However, the series resistance of the films, particularly in thinner films, prevents a very high frequency measurement.

3. Experiments

The devices used in this measurement are standard deep-depletion silicon-gate SOS MOSFET's. The source and drain diffusion time

was longer than that in the standard process, however. This longer diffusion time gives a deeper diffusion, which minimizes geometrical errors. It is well known that the lateral diffusion of impurities in SOS films is much less than the vertical diffusion. Devices with nearly ideal structures, such as that shown in Fig. 1, may be obtained by use of the silicon gate process. The gate length should be long enough to minimize the effect of gate length reduction by lateral diffusion, but it should also be small enough to avoid series resistance effects on the gate capacitance measurement. The lengths of the gates of devices used in this experiment are either 0.3 mil or 0.4 mil.

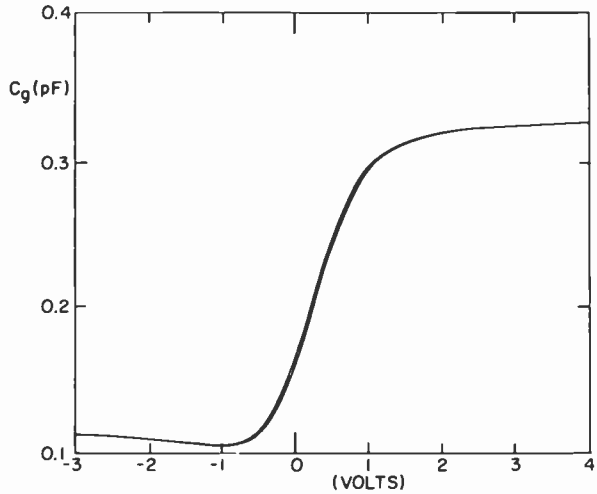


Fig. 2—Gate capacitance versus gate voltage characteristics of a 1 μm n-type deep depletion SOS MOST.

The measurement of the gate capacitance is conventional and uses a frequency of 1 MHz. At that frequency the g-r centers have negligible effect on the gate capacitance. The transconductance measurement was made by applying a dc voltage no larger than 0.1 V for the drain electrode. A small ac voltage was superimposed on the dc gate bias voltage. The ac current at the source was then plotted as a function of dc gate voltage. This plot directly yields the transconductance of the device operated in the linear region. The gate capacitance versus gate voltage (C_g versus V_g) curve of a typical n-type 1 μm film deep-depletion SOS/MOST is plotted in Fig. 2. The width and the length of the gate of this device are 7 and 0.3 mils, respectively. The carrier density obtained from this C_g versus V_g curve and Eq. [10] in-

creases toward the Si–SiO_2 interface. This is because the depletion region is not completely depleted and because of the effect of surface charges. The transconductance of the device used in Fig. 2 operated in the linear region is shown in Fig. 3 with frequency as a parameter. When the gate voltage is smaller than --0.5 V the transconductance is independent of gate bias voltage. When the gate bias voltage is larger than 1 V, the surface of the film is strongly accumulated. In the region where the gate bias voltage is larger than 0 V but smaller than 1

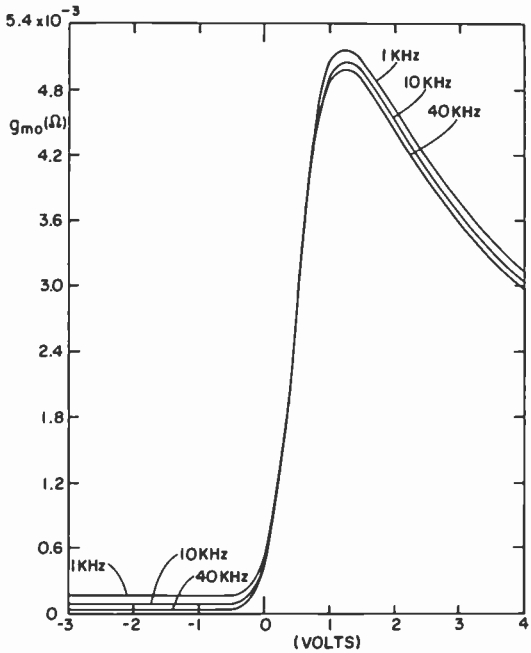


Fig. 3—Transconductance of the device used in Fig. 2 operated in the linear region as a function of gate bias voltage with frequency as a parameter.

V, the transconductance is independent of frequency if the frequency is larger than 1 kHz. Otherwise the transconductance decreases as the frequency is increased. This is due to the influence of the g-r centers and the surface states. At higher frequencies, the g-r centers and the surface states are unable to fully follow the signal frequency; therefore, the transconductance is reduced. The drain current of the device when the surface of the device is completely depleted is the maximum current that may be contributed by the g-r centers in the de-

pletion region. Fig. 3 shows that the transconductance is independent of gate bias voltages below −0.5 V. This indicates that the surface inversion layer has negligible effect on the transconductance. At $V_G < -1$ V the surface of the device is strongly inverted.

Although the transconductance contributed by g-r centers in the depletion region is small, as seen from Fig. 3, it is important to determine the frequency response of the g-r centers. To do so, we plotted the transconductance of the device as a function of frequency when biased at $V_G = -3$ V in Fig. 4. The transconductance is independent of frequency when the signal frequency is smaller than 1 kHz and larger than 30 kHz. Between these two frequencies, the transconductance decreases with the increase of frequency. The lower and the upper cutoff frequencies are equal to 2 and 20 kHz, respectively. To minimize the g-r center effect, the measurement frequency must be larger than 20 kHz.

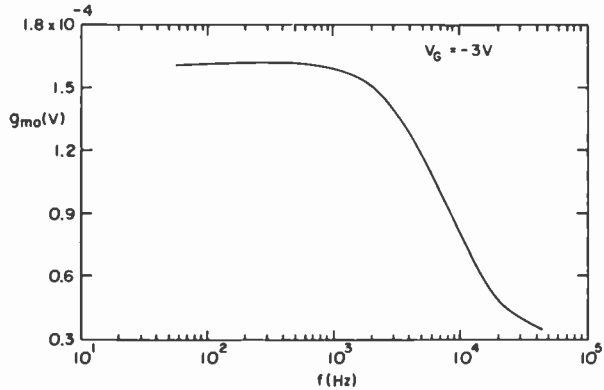


Fig. 4—Frequency response of the transconductance of the MOST used in Fig. 2 when $V_G = 3$ V.

The width of the depletion region was calculated from Eq. [5] and the capacitance data in Fig. 2. The results are plotted in Fig. 5. The maximum width of the depletion region was found to be slightly larger than 0.4 μm. The surface of the silicon film is inverted at gate voltages smaller than −1 V. The inversion layer is unable to absorb all the signal applied to the gate. Part of the electric field lines due to small signal applied to the gate terminate at g-r centers in the depletion region or terminate at the edge of the undepleted neutral region. Consequently, there is a small transconductance even when the surface is strongly inverted. When the gate bias voltage is larger than 1

V, the width of the depletion region is smaller than 0.02 μm. The surface potential is equal to a few millivolts, which can only slightly deplete the surface of the silicon. The depletion approximation is, therefore, invalid.

The mobility obtained from the above measurements is plotted in Fig. 6, curve (c). The doping concentration of the film when the surface of the film is approximately intrinsic is equal to 2×10^{16} cm^{-3}. Near the Si–SiO_2 interface the electron mobility is approximately equal to that in a bulk material with the same doping concentration. The mobility decreases exponentially with W.

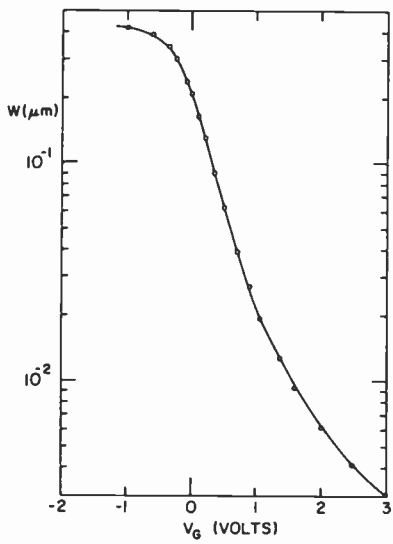


Fig. 5—The width of the depletion region as a function of gate bias voltage of the device used in Fig. 2.

Eqs. [6] and [9] show that as long as the drain current can be modified by a small signal voltage applied to the gate, the mobility of current carriers can be measured. Therefore, the mobility of current carriers in an accumulated layer can be obtained without altering the measurement procedures. The electron mobility in the accumulated layer of the FET used in previous figures is plotted in Fig. 7 as a function of charge density in the accumulated layer. This electron mobility decreases as the number of accumulated carriers is increased. It is seen that when the carrier density in the accumulated layer is larger than 10^{12} cm^{-2} the mobility of electrons in the accumulated layer is

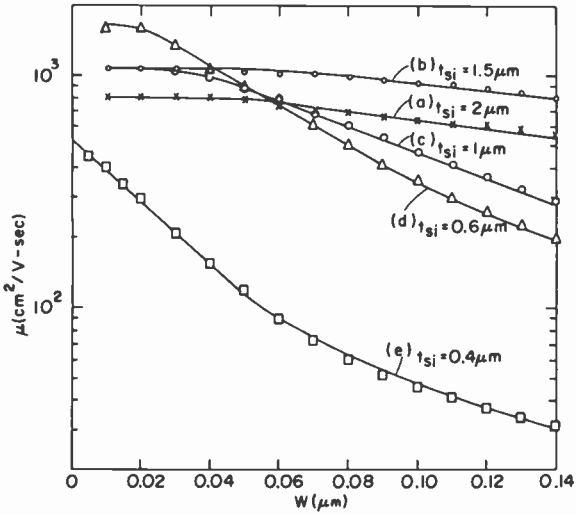


Fig. 6—Electron mobility in n-type silicon films.

approximately equal to that in a surface inversion layer of bulk silicon with the same surface carrier density.[13]

The transconductance of a deep depletion n-type SOS MOST made on a 2-μm film is shown in Fig. 8. The capacitance curve is similar to that of Fig. 2. In contrast to the 1-μm-film devices, the transconductance increases with signal frequency when the surface of the silicon is inverted. The frequency dependence of the transconductance when the gate bias voltage is equal to −3 V, i.e., the surface is

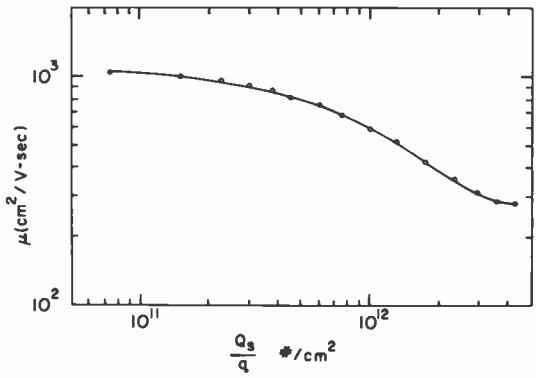


Fig. 7—Mobility of electrons in the accumulated surface of the device used in Fig. 2 as a function of the accumulated charge density.

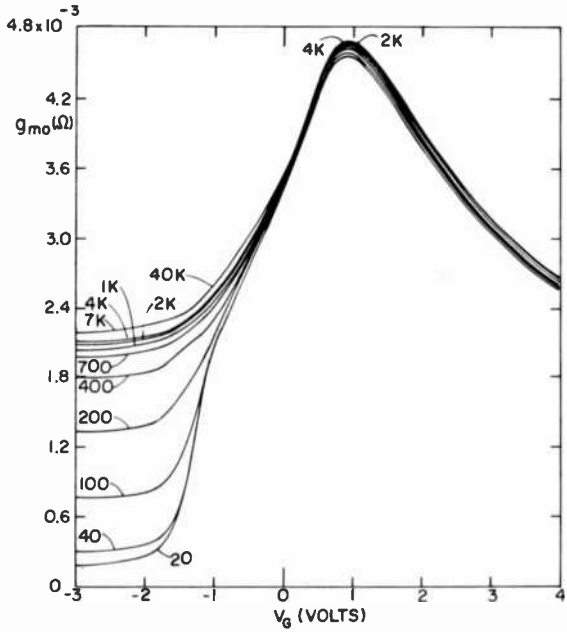


Fig. 8—Transconductance as a function of gate bias voltage of a SOS MOST made on a 2-μm n-type film.

strongly inverted, is shown in Fig. 9. The transconductance increases with frequency in range from 30 to 700 Hz. Outside of this range the transconductance is independent of frequency. This type of characteristic is expected only when the conductivity of the film at the neu-

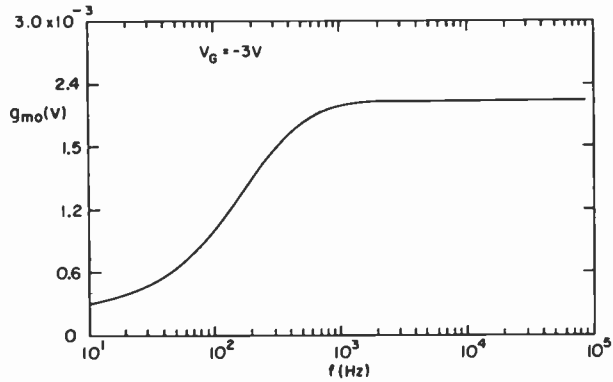


Fig. 9—Frequency dependence of the transconductance of the device used in Fig. 8 when the surface is strongly inverted.

tral region is large. The modulation of the conductivity of the neutral region is larger than the small-signal conductivity due to the g-r centers in the depletion region and the surface states.

The electron mobility of the device used in Fig. 8 is shown by curve (a) of Fig. 6. The electron mobility is approximately 800 cm^2/V-sec near the $Si–SiO_2$ interface. When W is larger than 0.04 μm, the electron mobility also decreases exponentially with distance. The spatial dependence of electron mobility in this device is very small, however. From the C_g versus V_G curve the doping concentration of this film when the surface is nearly intrinsic is found to be 10^{17} cm^{-3}.

Curves (b), (d), and (e) of Fig. 6 are the measured mobility of n-type deep depletion MOST's made on 1.5-μm, 0.6-μm and 0.4-μm SOS films, respectively. The doping concentration of these films when the surface of the silicon films are nearly intrinsic are 2×10^{16}, 2×10^{15}, and 4×10^{15} cm^{-3}, respectively.

Experimental data showed that the transconductance and the depletion width characteristics of SOS deep depletion MOST's made on p-type films have the same shapes as those of n-type devices. We shall not discuss them in detail. The hole mobility of p-type SOS films with film thickness equal to 0.6, 1.5, and 2 μm are plotted in Fig. 10. The doping density when the surface is nearly intrinsic are equal to 3×10^{15}, 7×10^{16}, and 10^{16} cm^{-3} for 2-μm, 1.5-μm, and 0.6-μm films, respectively. The general shape of these three curves is the same as that of electron mobility. Near the $Si–SiO_2$ interface, where the hole mobility of 2-μm and 1.5-μm films is independent of distance, the hole mobility is equal to that of a bulk silicon wafer with comparable doping concentrations.

4. Conclusions and Discussions

It is clearly seen from Figs. 6 and 10 that the mobility of current carriers in a thinner SOS film decreases exponentially with depth from the $Si–SiO_2$ interface. Near the silicon–sapphire interface the mobility of current carriers is very low. Ipri[9] viewed this low-mobility region as a nonconductive layer. The mobility of current carriers in thicker films, however, is independent of depth near the $Si–SiO_2$ interface. As the distance from the interface is increased, both electron and hole mobility decreases exponentially with depth. The experimentally observed electron and hole mobilities in SOS films within 0.2 μm of the $Si–SiO_2$ interface can be expressed by the following equation:

$$\frac{1}{\mu} = \frac{1}{\mu_I} + \frac{1}{\mu_D}, \qquad [11]$$

where $\mu_D = \mu_{DO}\, e^{-\alpha W}$ and α has a dimension of cm^{-1}. The value of α decreases as the thickness of the silicon film is increased. The limiting mobility, μ_I, was found to be equal to the ionized impurity limited mobility, as is the case in a bulk silicon wafer. It is quite possible that the component μ_D is caused by dislocations. An edge dislocation in a Si crystal creates a dangling bond per unit cell. Each dangling bond is a deep electron trap state with energy located near the middle of the energy gap.[12] The trapped electron will be surrounded by positive charges. This type of dislocation in n-type materials is a very effective electron scattering center. For p-type materials, holes are scattered by the dislocation strain field. Therefore, dislocations in a very thin silicon film grown on a sapphire substrate can be the dominant scattering source that limits the mobility of both electrons and holes.

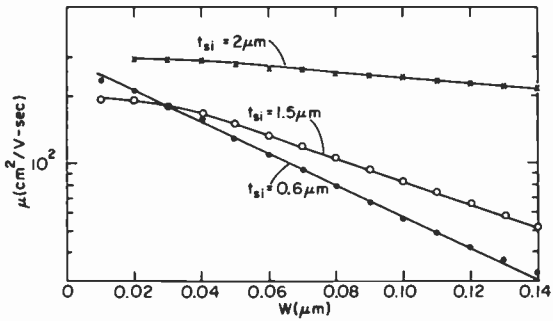


Fig. 10—Hole mobility in p-type silicon films.

Accordingly, the density of this type of dislocation must increase toward the silicon–sapphire interface. This agrees with what was generally observed. As the thickness of the film is increased, the dislocation density decreases and the dislocation scattering process is less important. Therefore, near the surface of a thicker film, the mobility of electrons and holes is limited by the ionized impurity scattering process.

At this time we have no satisfactory explanation of why the distance between the silicon–sapphire interface and the point where the mobility changes from ionized impurity-limited mobility to dislocation-scattering-limited mobility increases as the thickness of the film is increased.

It was also shown that the new measurement technique revealed the mobility of current carriers in an accumulated surface. The experimental results showed that when the accumulated charge density is larger than 10^{12} cm^{-2}, the mobility of electrons or holes in the accumulation layer is approximately equal to that in an inversion layer with the same charge density.

Acknowledgments

The authors wish to express their appreciation to J. M. Cartwright for his assistance in device fabrications and electrical measurements. They are also indebted to W. C. Schneider for supplying some of the devices used in this investigation.

References:

[1] D. J. Dumin and P. H. Robinson, "Electrically and Optically Active Defects in Silicon On Sapphire Films," *J Crystal Growth*, ,"**3**, p. 214 (1968).

[2] E. C. Ross and G. Warfield, "Effects of Oxidation on Electrical Characteristics of Silicon on Sapphire Films," *J. Appl. Phys.*, **40**, p. 2339 (1969).

[3] J. Mercier, "Properties of Low Temperature Heteroepitaxial Thin Silicon Films Grown in Helium Atmosphere," *J. Electrochem. Soc.*, **118**, p. 962 (1971).

[4] A. C. Ipri, "Variations in Electrical Properties of Silicon Films on Sapphire Using the MOS Hall Technique," APL 20 #1 pp 1 (1972).

[5] R. B. Fair, "A Self-Consistent Method for Estimating Non-Step Junction Doping Profiles from Capacitance-Voltage Measurements," *J. Electrochem. Soc.*, **118**, p. 971 (1971).

[6] E. H. Nicollian, M. H. Hanes and J. R. Brews, "Using the MIS Capacitor for Doping Profile Measurements with Minimal Interface State Error," *IEEE Trans. Electron Devices*, **ED-20**, p. 38 (1973).

[7]. Y. Darwish, "Measurement of the Mobility and Concentration of Carriers in Diffused Zones in Si with a Gate Controlled Structure," *IEEE Trans. Electron Devices*, **ED-21**, p. 397 (1974).

[8] S. T. Hsu, "Low Frequency Excess Noise in SOS MOST's," Noise Symposium, University of Florida, Gainsville, Florida, Dec. 1972.

[9] S. T. Hsu, "Trapping Noise in SOS MOS Field Effect Transistors Operated in the Linear Region," (to be published).

[10] S. T. Hsu, J. H. Scott, Jr., and W. E. Ham, "Anomalous Characteristics of SOS MOST's," International Electron Devices Meeting, Wash. D.C., Dec. 3–5, 1973.

[11] S. T. Hsu, "Characteristics of SOS MOST's," Government Microelectronics Application Conference, University of Colorado, Meeting, Boulder, Colo. June 1974.

[12] See Frank J. Blatt, *Physics of Electronics Conduction in Solids*, Chapter 6, McGraw-Hill Book Company, New York, N. Y., (1968).

[13] O. Leistiko, Jr., A. S. Grove and C. T. Sah, "Electron and Hole Mobility in Inversion Layers on Thermally Oxidized Silicon Surfaces," *IEEE Trans. Electron Devices*, **ED-12**, p. 248 (1965).

Three-Filter Colorimetry of Color-Television Picture Tubes*

G. M. Ehemann, Jr.

RCA Picture Tube Division, Lancaster, Pennsylvania

Abstract—This paper presents the derivation and description of a technique for measuring color-picture-tube chromaticity and luminous efficiency. The derivation is based on Grassmann's equation and the linear model of a three-filter colorimeter. Three colorimeter channel readings are taken of each of the three primary fields, which are sequentially excited in a 4 × 4-inch raster. The colorimeter filters are simple red, green, and blue band-pass filters. Complete colorimetric information is obtained from the nine channel readings of a tube under test after a known standard tube is used to calibrate the colorimeter. The standard is calibrated conventionally using a spectroradiometer and a foot-candle meter. A computer program is used to produce the following information:

(1) Chromaticities x, y of each primary field of the tube being tested.
(2) Efficiencies (fL/μA) of each primary field.
(3) Required currents in the primary fields of the tube under test, which are needed to produce any specified color mixture. The system efficiency in fL/μA is also printed out for each specified color mixture.

1. Introduction

The colorimetric description of a radiation field consists of the quantities X, Y, Z. These quantities are determined by mathematical procedures that sum the individual spectral contributions after applying standard[1] weighting functions $\bar{x}$, $\bar{y}$, $\bar{z}$ to the spectral energy distribution (SED) of the radiation to be described. In a physical sense, there

* This paper was presented at the Spring 1974 meeting of the Electrochemical Society.

are two fundamental ways of doing colorimetry, both invoking the same mathematical formulas. Fundamental procedure A consists of dispersing the spectrum of the radiation with a diffraction grating or prism and numerically applying the weighting functions to it. Fundamental procedure B consists of passing the undispersed radiation through filters whose spectral transmissions are carefully matched to the mathematical weighting functions. A group of three or four filters can thus be carefully made to match the three fundamental functions $\bar{x}$, $\bar{y}$, $\bar{z}$, thereby separating the radiation into a parallel arrangement of its colorimetric components X, Y, Z.

All colorimetry, including the three-filter technique presented here, must ultimately be based on one or the other of the fundamental procedures. Our three-filter technique is calibrated from procedure A, and thereafter eliminates the need to disperse, record, and operate mathematically upon spectra. The three filters are off-the-shelf items and are merely required to pass light in the broadly defined red, green, and blue regions. Their cost of a few dollars is to be compared with a few hundred dollars for the precise $\bar{x}$, $\bar{y}$, $\bar{z}$ matching set required in procedure B. The only price one pays for the simple three-filter method described in this paper is the restriction that tubes tested with these filters must have as basic phosphors the same materials as in the standard tube that is used to calibrate the filters. Direct application requires neither spectral dispersion nor use of a filter having a spectral transmission curve carefully matched to a mathematical formula.[7]

Relative simplicity is not the only virtue of the method presented here. It is better suited for color-picture-tube colorimetry than other known techniques. The latter include some commercial devices using procedure B. These incorporate (at additional cost) electronic logic behind the four filters to compute chromaticities x and y, where

$$x = \frac{X}{X + Y + Z}, \qquad y = \frac{Y}{X + Y + Z}$$

The method described here also computes chromaticities after the three primary fields of the color picture tube are sequentially viewed. The chromaticity calculations of the new method are part of an open-loop process designed to give system colorimetry as well as colorimetry of the primary fields. The term "closed-loop" is used to describe colorimeters capable of determining x, y of single fields only, i.e., lacking the capacity to store these values for three sequential fields and to subsequently operate upon the resulting number array. The aspect of system colorimetry is neglected in commercial units which,

with closed-loop electronic logic behind the $\bar{x}$, $\bar{y}$, $\bar{z}$ filters, yield less information than desired by the system engineer. For this reason the present method was developed. It yields X, Y, Z values for each primary field, and groups these nine quantities into a system matrix that completely characterizes a sample tube. The system matrix generates the following data: (I) Chromaticities x, y of each primary field of a tube; (II) luminous efficiency of each field; (III) individual field currents in μA required to produce any specified color mixture; and (IV) system efficiency in fL/μA. McCoy[2] has already used three-filter colorimetry to obtain (I) and (II). His work could easily have produced (III) and (IV) as well by using the technique described in this paper. R. B. Platt[3] and J. W. Mirsch[4] have also produced (I) from a three-filter colorimeter. Platt also obtained (II) using a foot-candle meter (fundamental procedure B) and assumed bogie values of (I) to calculate (III) and (IV). The fact that the reported technique for finding (III) and (IV) is new can be explained by past lack of emphasis on system colorimetry as opposed to individual or specific field colorimetry,[7] or to the algebraic processes involved, which require implementation on a computer.

The new method has been tested in the RCA Spectroradiometer Laboratory at Lancaster where it now is standard procedure. It is reasonable to expect that before long hundreds of color picture tubes will be measured each day by the method presented here.

2. Fundamental Procedure for Expressing Color of Radiation

The fundamental procedure is to multiply the spectral energy distribution(SED) of the radiation by standard tristimulus functions $\bar{x}$, $\bar{y}$, $\bar{z}$[1] and integrate over the region of definition, namely between 4000 and 7000 Å. This procedure results in the three quantities X, Y, and Z.

$$X = \int_{4000}^{7000} \bar{x} \cdot \mathrm{SED} \cdot d\lambda$$

$$Y = \int \bar{y} \cdot \mathrm{SED} \cdot d\lambda \qquad [1]$$

$$Z = \int \bar{z} \cdot \mathrm{SED} \cdot d\lambda$$

Figs. 1, 2, and 3 show the respective tristimulus functions $\bar{x}$, $\bar{y}$, $\bar{z}$ and the closest corresponding color phosphor SED. The arrangement of the six curves suggests that, for the red phosphor, X is the dominant colorimetric quantity, for the green Y is dominant, and for the blue Z is dominant.

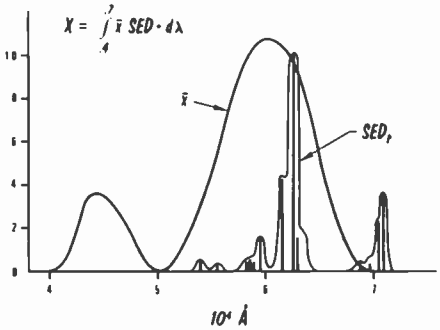


Fig. 1—Fundamental procedure for *X*.

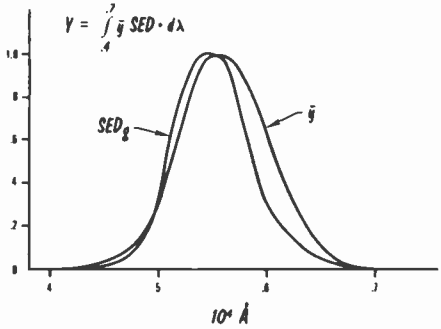


Fig. 2—Fundamental procedure for *Y* (luminance).

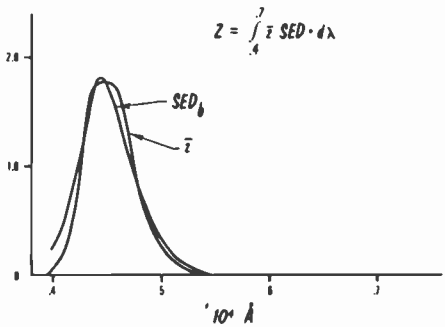


Fig. 3—Fundamental procedure for *Z*.

The quantity Y is proportional to the luminance of the radiation, as $\bar{y}$ is the photopic curve of the human eye.

It is convenient to represent the color or radiation by the three quantities x, y, Y where

$$x = \frac{X}{X + Y + Z}$$
$$y = \frac{Y}{X + Y + Z} \qquad [2]$$

The two-dimensional representation of a color by the quantities x, y, independent of luminance Y, is often used on a planar diagram commonly called the CIE[5] diagram. The quantities x, y are called chromaticities. Chromaticity and luminance, therefore, completely specify a color, as do the fundamental quantities X, Y, Z. Note that

$$\frac{X}{Y} = \frac{x}{y}, \qquad \frac{Z}{Y} = \frac{1 - x - y}{y}. \qquad [3]$$

Fig. 4 shows the $x - y$ representation of three phosphor primaries as points on the CIE chromaticity diagram.

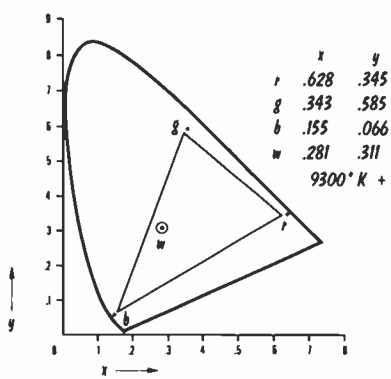


Fig. 4—1931 CIE diagram—two-dimensional representation.

Luminance Y is often expressed in units of foot-lamberts, or lumens emitted per square foot of screen surface. This quantity is measured by a luminous intensity (foot-candle) meter placed a distance d_0 from the screen, as shown in Fig. 5.

3. The Sequential Field Test

The new method is called the sequential field test (SFT), a term referring to the individual fields in a color picture tube made from three phosphor materials emitting in the red, green, and blue spectral regions. A three-filter colorimeter separates the light emitted by each field into three parts by means of color-separation filters. Filters transmitting predominantly red, green, and blue light are the choices for the colorimeter filters. An example of a valid choice of filters would be Kodak #26 (red), #61 (green), and #47 (blue). The colorimeter mechanical structure is basically a black rectangular box, open

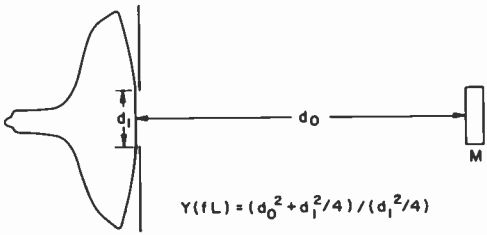


Fig. 5—Measurement of luminance Y from ft-candle meter M placed at distance d_0 from screen. Distance d_1 represents diameter of circular baffle placed in front of TV faceplate.

at one end where contact is made with the glass face of the picture tube. The opposite end supports the three color-separation filters and, behind them, three photocells. Digital panel meters, one for each photocell, provide the display, which characterizes the light distribution passed through the filters. The numbers so displayed are commonly called channel readings (see Figs. 6 and 7).

The three-filter colorimeter is calibrated by a standard tube whose individual r, g, b fields, as well as a known mixture of the three fields resulting in a white, have been measured by a fundamental procedure yielding in each case the chomaticities x, y and, in the case of the white mixture, the luminance Y_w as well. The data consist of nine numbers shown in Table 1.

The white field (w) is obtained by mixing the individual fields in proportion to the respective electron-beam currents, often referred to as "standard currents," and the white field is usually the "standard white" closely approximating a 9300°K black-body radiator. The SED of a typical white field is shown in Fig. 8. Also see Table 2,

Table 1—Chromaticities of Standard Fields

Field	x	y	Y
r	x_r	y_r	
g	x_g	y_g	
b	x_b	y_b	
w	x_w	y_w	Y_w

where three additional numbers I_r, I_g, I_b have been introduced, making a total of twelve so far.

For computational purposes the twelve quantities listed in Tables 1 and 2 will be combined and grouped as follows:

Group A_{33} (3 × 3 matrix, composed of the six individual field chromaticities of the standard tube)

$$A_{33} = \begin{vmatrix} \dfrac{x_r}{y_r} & \dfrac{x_g}{y_g} & \dfrac{x_b}{y_b} \\ 1 & 1 & 1 \\ \dfrac{1 - x_r - y_r}{y_r} & \dfrac{1 - x_g - y_g}{y_g} & \dfrac{1 - x_b - y_b}{y_b} \end{vmatrix} \qquad [4]$$

Group B_{31} (3 × 1 matrix, composed of the standard white color mixture x_w, y_w, Y_w)

$$B_{31} = \begin{vmatrix} X_w \\ Y_w \\ Z_w \end{vmatrix} = Y_w \cdot \begin{vmatrix} \dfrac{x_w}{y_w} \\ 1 \\ \dfrac{1 - x_w - y_w}{y_w} \end{vmatrix} \qquad [5]$$

Group C_{31} (3 × 1 matrix, composed of the three full raster currents which, as shown in Table 2, produced the standard white mixture B_{31})

Table 2—Field Mixture Currents for White

Full raster currents, μA					
Red	Green	Blue	x	y	Y
I_r	I_g	I_b	x_w	y_w	Y_w

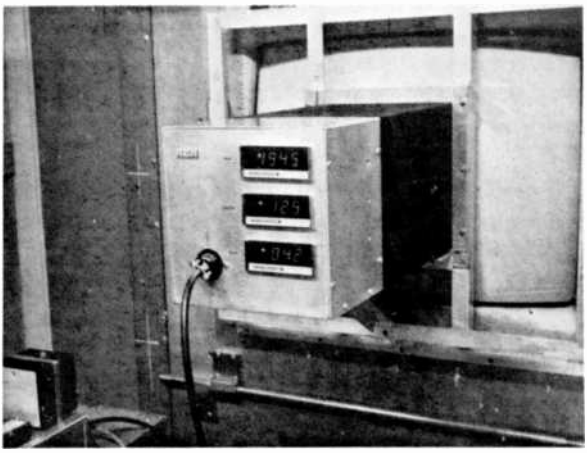


Fig. 6—Three-filter colorimeter displaying channel readings of red 4 × 4-inch raster.

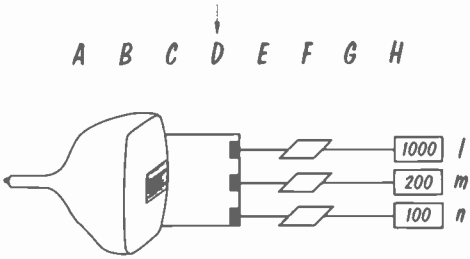


Fig. 7—Schematic of tube-colorimeter system showing (a) tube under test, (b) color separation filters and selenium photocells, (c) operational amplifiers, (d) digital panel meters displaying channel readings of red 4 × 4-inch raster.

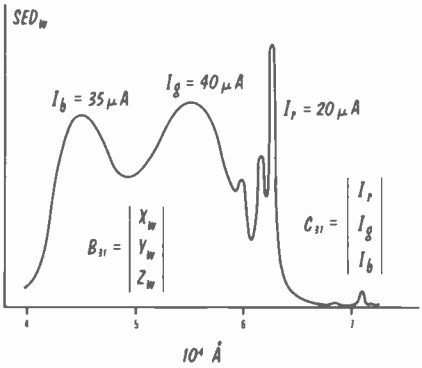


Fig. 8—Spectral energy distribution described by matrix B_{31}, and produced by C_{31}.

$$C_{31} = \begin{vmatrix} I_r \\ I_g \\ I_b \end{vmatrix} \qquad [6]$$

Group diag C_{33}^{-1} (3 × 3 diagonal matrix, whose diagonal elements are the reciprocals of the elements of C_{31})

$$\text{diag } C_{33}^{-1} = \begin{vmatrix} \frac{1}{I_r} & 0 & 0 \\ 0 & \frac{1}{I_g} & 0 \\ 0 & 0 & \frac{1}{I_b} \end{vmatrix} \qquad [7]$$

Recalling Eq [3] and the remarks in Section 2 concerning Figs. 1, 2, and 3, it becomes apparent that matrix A_{33} is well conditioned in the following sense: For each column and each row except the middle row, the largest element is on the diagonal. This desirable feature minimizes computational errors in solving simultaneous equations represented by matrix products involving A_{33}, as found in Eq. [8] below.

From matrices A, B, C several important field luminance matrices can be computed, which we list here:

$$E_{31} \equiv \begin{vmatrix} Y_r \\ Y_g \\ Y_b \end{vmatrix} = A_{33}^{-1} \cdot B_{31} \text{ (3 × 1 matrix of individual field luminances } Y_r, Y_g, Y_b \text{ (fL) which produce the standard white mixture } B_{31}\text{).} \qquad [8]$$

$$F_{31} \equiv \begin{vmatrix} F_r \\ F_g \\ F_b \end{vmatrix} \equiv \begin{vmatrix} Y_r/I_r \\ Y_g/I_g \\ Y_b/I_b \end{vmatrix} = \text{diag } C_{33}^{-1} \cdot E_{31} \text{ (3 × 1 matrix of individual field efficiencies, fL/}\mu\text{A, of the standard fields)} \qquad [9]$$

$$\text{diag } F_{33}^{-1} = \begin{vmatrix} I_r/Y_r & 0 & 0 \\ 0 & I_g/Y_g & 0 \\ 0 & 0 & I_b/Y_b \end{vmatrix} \text{ (diagonal 3 × 3 matrix derived from } F_{31}\text{)} \qquad [10]$$

The following equation uses expanded notation for matrices A_{33}, B_{31}, and E_{31} and shows how the three are related by Eq. [8]:

$$Y_r \begin{vmatrix} x_r/y_r \\ 1 \\ \dfrac{1 - x_r - y_r}{y_r} \end{vmatrix} + Y_g \begin{vmatrix} x_g/y_g \\ 1 \\ \dfrac{1 - x_g - y_g}{y_g} \end{vmatrix}$$

$$+ Y_b \begin{vmatrix} x_b/y_b \\ 1 \\ \dfrac{1 - x_b - y_b}{y_b} \end{vmatrix} = \begin{vmatrix} X_w \\ Y_w \\ Z_w \end{vmatrix} = Y_w \begin{vmatrix} x_w/y_w \\ 1 \\ \dfrac{1 - x_w - y_w}{y_w} \end{vmatrix} \qquad [11]$$

The above is equivalent to Eq. [8] except matrix A_{33} is inverted and carried to the right hand side of Eq. [8]. Eq. [11] is the direct form of Grassmann's color-mixing law,[6] which is really three separate equations. The first equates X_w of the color mixture to the sum $X_r + X_g + X_b$ of the components, the second equates Y_w to the sum of $Y_r + Y_g + Y_b$ while the third equates Z_w to $Z_r + Z_g + Z_b$, using Eq. [3].

The second luminance matrix, F_{31} is obtained by premultiplying E_{31} by the diagonal matrix diag C_{33}^{-1} according to Eq. [9]. This operation gives unambiguously the desired form of "fL divided by μA" for each standard field. Premultiplying by a diagonal matrix separately scales each row of the second matrix, hence each element of E_{31} in fL is separately scaled into the efficiency, fL/μA, of the respective field by dividing by its full raster current.

The third luminance matrix, diag F_{33}^{-1} is formed from F_{31} in the same way diag C_{33}^{-1} was obtained from C_{31}. Generating a diagonal matrix from a column matrix is merely an exercise in notation, but it results in forms both relevant and useful to the algebraic model of three-filter colorimetry. Whereas diag C_{33}^{-1} was used as a row multiplier to obtain F_{31}, diag F_{33}^{-1} will be used as a column multiplier in Eq. [15] to produce colorimeter channel readings per fL/μA.

The next step is to use the standard tube, from which matrices A, B, C and the field luminance matrices E, F, diag F were derived, to calibrate the three-filter colorimeter. Calibration is a systematic process wherein the individual fields of the standard tube are excited sequentially, and their r, g, b channel readings recorded. The red field, reduced to a 4 × 4-inch raster, is viewed first, followed by green and blue under controlled conditions of current and anode potential. The latter is usually fixed at 25 kV, and the test current adjusted to pro-

vide readily measurable channel readings. This degree of freedom in the selection of the test current in the 4 × 4-inch raster does not affect subsequent calculations which are based on channel readings per test current.

In Table 3 the r, g, b channel readings of the standard red field are represented by l_r, m_r, n_r etc., and the test currents are represented by

Table 3—Colorimeter Calibration Data from Standard Tube

Test Current in 4 × 4 inch raster	Channel Readings r	g	b
TC_r in red field	l_r	m_r	n_r
TC_g in green field	l_g	m_g	n_g
TC_b in blue field	l_b	m_b	n_b

TC_r, TC_g, TC_b. For computing purposes, the calibration data is subsequently grouped in the matrix form,

$$D_{33} = \begin{vmatrix} l_r & l_g & l_b \\ m_r & m_g & m_b \\ n_r & n_g & n_b \end{vmatrix} \cdot \begin{vmatrix} \frac{1}{TC_r} & 0 & 0 \\ 0 & \frac{1}{TC_g} & 0 \\ 0 & 0 & \frac{1}{TC_b} \end{vmatrix} \qquad [12]$$

Matrix D_{33} is a 3 × 3 matrix of channel readings produced by sequentially exciting each of the fields of the standard tube by one μA in the 4 × 4-inch raster. Postmultiplying by the diagonal matrix in Eq. [12] causes each column of channel readings l, m, n to be divided by the respective test current TC.

The calibration of the three-filter colorimeter is now complete. The power of the SFT lies in the use of matrices A, B, C, D in place of the fundamental mathematical formulas, to operate upon data for the tube under test and derive complete colorimetric information. Tube under test data must consist of the twelve numbers indicated in Table 4. These twelve numbers are grouped into matrix D_{33}' according to Eq. [13] in the same way D_{33} was obtained from the standard tube data appearing in Table 3. Symbols with primes such as D_{33}' will carry the same quantities and definitions of the original symbol to tubes under test. Symbols without primes refer to the standard tube.

$$D_{33}' = \begin{vmatrix} l_r' & l_g' & l_b' \\ m_r' & m_g' & m_b' \\ n_r' & n_g' & n_b' \end{vmatrix} \cdot \begin{vmatrix} \frac{1}{TC_r}' & 0 & 0 \\ 0 & \frac{1}{TC_g}' & 0 \\ 0 & 0 & \frac{1}{TC_b}' \end{vmatrix} \qquad [13]$$

Complete characterization of a tube under test consists of X, Y, Z values of each of its three fields. These nine quantities can be grouped into a single system matrix, which we will denote by H_{33}'. A

Table 4—Colorimeter Test Data from Test Tube

Test Current for 4 × 4 inch raster	Channel Readings r	g	b
TC_r' in red field	l_r'	l_g'	l_b'
TC_g' in green field	m_r'	m_g'	m_b'
TC_b' in blue field	n_r'	n_g'	n_b'

transformation ultimately derived from matrices A, B, C, D is used to obtain H_{33}' from D_{33}'. This transformation is directly stated in terms of A, D, and diag F_{33}^{-1}.

$$H_{33}' = A_{33}' \cdot F_{33}' \qquad [14]$$

where

$$F_{33}' = (D_{33} \cdot \text{diag } F_{33}^{-1})^{-1} \cdot D_{33}' \qquad [15]$$

Matrix F_{33}' is the color equivalent luminance matrix that describes each individual field of the tube being tested as an equivalent mixture of the three fields of the standard tube. The equivalent amount of each standard field "present" in the test field is expressed in units of luminance per μA, i.e., the same units as diag F_{33}. For example, suppose the red field of the tube being tested produced in its 4 × 4 inch raster the channel readings l_r', m_r', n_r'. Its color equivalent luminance mixture is symbolized by F_{rr}', F_{gr}', F_{br}', and its fundamental colorimetric components by X_r', Y_r', Z_r'. If, however, the red field of the tube under test should be identical to the standard red, its X, Y, Z would equal X_r, Y_r, Z_r of the standard red; and according to Grassmann's law of equivalence, its color equivalent luminance mixture would be represented by F_r, 0, 0, since the standard red field is so de-

Table 5—SFT Description of Red Test Field

Test Current	TC_r'
Channel Readings	$l_r', m_r', n_r' = TC_r' \times$ 1st column of D_{33}'
Equivalent Luminance Mixture	$F_{rr}', F_{gr}', F_{br}'$ = 1st column in F_{33}'
Fundamental Color Components	X_r', Y_r', Z_r' = 1st column in H'

Table 6—SFT Description of Green and Blue Test Fields

Green	Blue
TC_g'	TC_b'
$l_g', m_g', n_g' = TC_g' \times$ 2nd column in D	$l_b', m_b', n_b' = TC_b' \times$ 3rd column in D_{33}'
$F_{rg}', F_{gg}', F_{bg}'$ = 2nd column in F_{33}'	$F_{rb}', F_{gb}', F_{bb}'$ = 3rd column in F_{33}'
X_g', Y_g', Z_g' = 2nd column in H_{33}'	X_b', Y_b', Z_b' = 3rd column in H_{33}'

scribed according to Eqs. [9] and [10]. The colorimeter channel readings would also equal those produced by the standard red field. Tables 5, 6, 7, 8 list the general description of standard and test fields according to the SFT. The first two lines in each of these tables con-

Table 7—SFT Description of the Standard Red Field

Test Current	TC_r
Channel Readings	$l_r, m_r, n_r = TC_r \times$ 1st column in D_{33}
Equivalent Luminance Mixture	$F_r, 0, 0$ = 1st column in diag F_{33}
Fundamental Color Components	$X_r, Y_r, Z_r = Y_r \times$ 1st column in A_{33}

Table 8—SFT Description of the Standard Green and Blue Fields

Green Field TC_g	Blue Field TC_b
$l_g, m_g, n_g = TC_g \times$ 2nd column in D	$l_b, m_b, n_b = TC_b \times$ 3rd column in D_{33}
$0, F_g, 0$ = 2nd column in diag F_{33}	$0, 0, F_b$ = 3rd column in diag F_{33}
$X_g, Y_g, Z_g = Y_g \times$ 2nd column in A_{33}	$X_b, Y_b, Z_b = Y_b \times$ 3rd column in A_{33}

tain symbols for recorded tube colorimeter data, whereas the last two lines contain symbols of calculated quantities stored in matrices F_{33}' and H_{33}', respectively.

In agreement with previous work of McCoy,[2] the following equations can be written relating the quantities appearing in Table 5 to the corresponding quantities in Tables 7 and 8. Note the results: a description of the red test field as a color equivalent luminance mixture of the three standard fields, and finally the fundamental description of the red test field in terms of X, Y, Z. First, consider the color equivalent luminance mixture equation. This is really a set of three equations stating the color equivalence condition in terms of channel readings. It is based on the hypothesis that equivalent color fields produce identical channel readings.

$$\frac{F_{rr}'}{F_r}\frac{1}{TC_r}\begin{vmatrix} l_r \\ m_r \\ n_r \end{vmatrix} + \frac{F_{gr}'}{F_g}\frac{1}{TC_g}\begin{vmatrix} l_g \\ m_g \\ n_g \end{vmatrix}$$

$$+ \frac{F_{br}'}{F_b}\frac{1}{TC_b}\begin{vmatrix} l_b \\ m_b \\ n_b \end{vmatrix} = \frac{1'}{TC_r}\begin{vmatrix} l_r' \\ m_r' \\ n_r' \end{vmatrix} \qquad [16]$$

The unknowns F_{rr}', F_{gr}', F_{br}' are found by solving the simultaneous Eqs. [16]. Observe in Eq. [16] the red (first) column of D_{33} appears explicitly as

$$\frac{1}{TC_r}\begin{vmatrix} l_r \\ m_r \\ n_r \end{vmatrix}$$

and is scaled by $1/F_r$, the green column of D_{33} is scaled by $1/F_g$ and the blue column by $1/F_b$.

This indicates post multiplication of D_{33} by diag F_{33}^{-1}. In terms of D, diag F_{33}^{-1}, and D' we symbolize Eq. [16] by

$$D_{33} \cdot \text{diag } F_{33}^{-1}\begin{vmatrix} F_{rr}' \\ F_{gr}' \\ F_{br}' \end{vmatrix} = D_{3r}' \qquad [16a]$$

and its solution for the equivalent luminance mixture by

$$F_{3r}' = \begin{vmatrix} F_{rr}' \\ F_{gr}' \\ F_{br}' \end{vmatrix} = \left| D_{33} \cdot \text{diag } F_{33}^{-1} \right|^{-1} \cdot D_{3r}' \qquad [16b]$$

where D_{3r}', the first column in D_{33}, also appears as the right hand side of Eq [16]. Note the similarity of Eq. [16b] to Eq. [15]. The latter represents nine equations, three for each test field, whereas Eq. [16b] represents the three equations for the red test field only.

Having found F_{3r}', the equivalent luminance mixture of a single field, the final step in McCoy's[2] procedure was to compute the fundamental colorimetric components of that field. Our techniques are identical, and consist of writing

$$F_{rr}' \begin{vmatrix} x_r/y_r \\ 1 \\ \dfrac{1 - x_r - y_r}{y_r} \end{vmatrix} + F_{gr}' \begin{vmatrix} x_g/y_g \\ 1 \\ \dfrac{1 - x_g - y_g}{y_g} \end{vmatrix} + F_{br}' \begin{vmatrix} x_b/y_b \\ 1 \\ \dfrac{1 - x_b - y_b}{y_b} \end{vmatrix}$$

$$= \begin{vmatrix} X_r' \\ Y_r' \\ Z_r' \end{vmatrix} = H_{3r}' = Y_r' \begin{vmatrix} x_r'/y_r' \\ 1 \\ \dfrac{1 - x_r' - y_r'}{y_r'} \end{vmatrix} \qquad [17]$$

which is another set of three simultaneous equations, again Grassmann's equations, stating the conservation law with respect to X, Y, Z. The second equation in [17] states that the luminance of the mixture color (Y_r' in this case) is the sum of the luminances of all components, namely $F_{rr}' + F_{gr}' + F_{br}'$. The first and third equations in [17] state the conservation of X and Z respectively.

Note that matrix A appears in Eqs. [17], which can be symbolized as follows:

$$H_{3r}' = \begin{vmatrix} X_r' \\ Y_r' \\ Z_r' \end{vmatrix} = A_{33} \cdot F_{3r}' \qquad [17a]$$

Note the similarity between Eq. [17a] and Eq. [14]. The latter represents nine equations, giving X, Y, Z for each test field, whereas Eq. [17a] yields X_r', Y_r', Z_r' of the red test field only.

The derivation of Eqs. [16b] and [17a] is based on the linear algebraic model of the three-filter colorimeter and upon Grassmann's equations. It is hypothesized that a test field is colorimetrically equivalent to the sum of individual standard fields which when added give channel readings identical to those of the test field. The concept of color equivalent luminance mixture of the standard fields that describe a test field is based on this hypothesis. McCoy has specified the condition under which the equivalence concept is valid: the test fields must be made from the same set of color phosphors as were used in the standard fields. This restriction does not mean that, given a set of color phosphors, one color field may not be contaminated by phosphor particles of the different colors. If the red field of a tube is heavy in blue phosphor contamination, its filter readings by the SFT will accurately detect this condition and express it in quantitative chromaticity terms.

The transformation from channel readings to X, Y, Z via equivalent luminance mixtures of standard fields is described, for a single test field, by Eqs. [16b] and [17a]. This has been the stopping point of all previous work in three-filter colorimetry. The SFT goes further, not in concept but in practice, in emphasizing test field sequencing for color picture tubes. The complete characterization of a tube consists of X, Y, Z of each of the three test fields r′, g′, b′. This would indicate the recording of channel readings for the green and blue test fields, and repeating the operations of Eqs. [16b] and [17a] to get X_g', Y_g', Z_g' and X_b', Y_b', Z_b' in the same way as already indicated for the red. The SFT does in fact accomplish this. The result is matrix H_{33}', a 3×3 matrix containing X, Y, Z of each field of a tube under test. This matrix, from which the system performance is found for any desired system output B_{31}', has dimensional units of fL/μA. The SFT provides through matrix H_{33}' the system efficiency in fL/μA for any desired color mixture B_{31}' of its three fields.

$$\text{System fL}/\mu\text{A} = \frac{Y_w'}{I_w'} \qquad [18]$$

where Y_w' is the luminance of the desired color mixture, and $I_w' = I_r' + I_g' + I_b'$ is the total full-raster screen current required to produce the color mixture B_{31}'.

$$B_{31} = Y_w' \begin{vmatrix} x_w'/y_w' \\ 1 \\ \dfrac{1 - x_w' - y_w'}{y_w'} \end{vmatrix} \qquad [19]$$

The individual field currents (μA) in the tube under test are determined by

$$C_{31}' = \begin{vmatrix} I_r' \\ I_g' \\ I_b' \end{vmatrix} = H_{33}^{-1'} \cdot B_{31}' \qquad [20]$$

The current sum I_w' is inserted into Eq. [18] to yield the system efficiency in fL/μA.

The first column of H_{33}' is given by Eq. [17a] and consists of X, Y, Z of the red field of the test tube. The second column of H_{33}', equal to X_g', Y_g', Z_g' is obtained from the product of matrices A_{33} and F_{3g}', where the latter is the equivalent luminance mixture of the standard fields with respect to the green test field. Similarly for the blue test field, the equivalent luminance matrix F_{3b}' is also multiplied by A_{33} to yield X_b', Y_b', Z_b', the third column of H_{33}'. We can therefore symbolize H_{33}' by

$$H_{33}' = A_{33} \cdot F_{33}' \qquad [14]$$

where F_{33}' is a 3×3 matrix composed of the three column vectors F_{3r}, F_{3g}', F_{3b}'. Each of the latter is a transformation of one of the column vectors D_{3r}', D_{3g}', D_{3b}' which compose the 3×3 matrix D_{33}'. Eq. [16b] shows the transformation between F_{3r}' and D_{3r}', the red test field equivalent luminance mixture vs its channel readings per unit test current. The entire transformation connecting all three F_{3f}', D_{3f}', where $f' = r', g', b'$, pairs can be expressed separately in Eq. [16b] for each pair, or can likewise be stated in SFT symbolism as follows:

$$F_{33}' = (D_{33} \cdot \text{diag } F_{33}^{-1})^{-1} \cdot D_{33}'$$

which is Eq. [15].

As Y_r', the luminance per μA of the red test field, is the second element in H_{3r}' (Eq. [17]), so the middle row of H_{33}' contains the fL/μA Y_r', Y_g', Y_b' of the three test fields. Recalling that F_{31} of the standard tube was composed of fL/μA of the standard fields, one can write the direct analogy for the tube being tested.

$$F_{31}' \equiv \begin{vmatrix} Y_r' \\ Y_g' \\ Y_b' \end{vmatrix} = \text{middle row of } H_{33}' \qquad [21]$$

Fig. 9 shows the expanded notation for the matrix multiplication resulting in $H_{33}' = A_{33} \cdot F_{33}'$ and in particular the computation of the middle row of H_{33}'. Horizontal lines are drawn through each row of A_{33} and a vertical line through each column of F_{33}'. Nine intersection

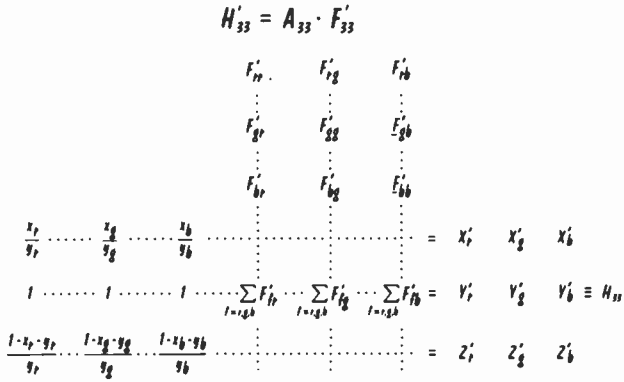


Fig. 9—Expanded notation for the matrix operation $H_{33}' = A_{33} \cdot F_{33}'$.

points result in the lower right box, each corresponding to an element of H_{33}'. The value of each element is the product of the row of A and column of F_{33}' whose respective lines form the intersection. A similar procedure is understood to apply for all matrix products. The mnemonic method in Fig. 9 enables one to quickly visualize the effects of pre- and post multiplication by a diagonal matrix.

Note that each term in the middle row of H_{33}' is merely the sum of terms in the corresponding column of F_{33}'. The luminance Y_f' in each test field f' is the sum of the standard field equivalent luminances, F_{3f}', which match the test field in accordance with Grassmann's law.

Test field efficiencies in lumens per watt are given by

$$G_{31}' = \alpha F_{31}' \qquad [22]$$

where α equals the total area of the phosphor screen in square feet di-

vided by the product of screen potential in megavolts times the combined transmission of the shadow-mask, matrix, and glass faceplate.

A computer program associated with the SFT was used for the calculations. The SFT requires matrix multiplication and matrix inversion at several stages, and could hardly be implemented without the computer. Time-sharing computers and especially programmable calculators, on the other hand, have made the SFT very practical.

4. Advantages of the Sequential Field Test Using a Three-Filter Colorimeter

The first advantage of the SFT is the open-ended system colorimetry it generates. The required input I_r', I_g', I_b' to the system is obtained by specifying the chromaticity and luminance of any desired color mixture and operating upon these numerical quantities by matrix $H_{33}^{-1\prime}$. Comparing desired luminance to total required input results in the system efficiency in fL/μA.

Other advantages of the SFT relate to the ease of its execution from the viewpoint of the test set operator. It is very systematic, requiring only one beam current adjustment at any given stage, while allowing the freedom of varying the test current to obtain readable colorimeter channel deflections. The exact dimensions of the 4 × 4 inch raster are not critical, and a single standard tube can be used to calibrate the colorimeter which then can measure tubes of any size or shape.

Acknowledgment:

The able assistance of W. J. Zettlemoyer in establishing the new technique is gratefully acknowledged.

References:

[1] A. C. Hardy, Handbook of Colorimetry, MIT Press (1936).

[2] D. S. McCoy, Private Communication

[3] R. B. Platt, Private Communication

[4] J. W. Mirsch, Private Communication

[5] Commission Internationale de l'Eclairage [Consult Ref. (1)].

[6] D. B. Judd and G. Wyszecki, *Color in Business, Science, and Industry*, p. 49, John Wiley and Sons, Inc., N.Y., (1973).

[7] R. Q. Vogel of Sylvania has done extensive work in picture tube colorimetry, and in particular has developed techniques using instrumentation directly employing the fundamental procedures (disper-

sion, $\bar{x}$, $\bar{y}$, $\bar{z}$ spectral transmission filters). Refs. (8) and (9) are two of his publications. On page 70 of Ref. (8) is his statement of the system colorimetry approach: "If the colors of each of the primaries are measured, as well as any arbitrary combination of the three primaries, the currents needed for . . . any color in the tube gamut can be readily calculated."

[8] R. Q. Vogel, "Luminance Measurement Techniques for Color Television Picture Tubes," *IEEE Trans. on Broadcast and TV Receivers,* **13,** p. 66, Nov. 1967.

[9] J. P. O'Donnell and R. Q. Vogel, "Instrumentation for Determining Proper White Balance when Measuring Luminance of Color Television Picture Tubes," *IEEE Trans. on Broadcast and TV Receivers,* **15,** p. 171, July 1969.

A Study of Second-Harmonic-Extraction Trapatt Amplifiers for X-Band Operation*

W. R. Curtice, J. J. Risko, and P. T. Ho

RCA Laboratories, Princeton, New Jersey

Abstract—An experimental study of X-band (8.2 to 12.4 GHz) pulsed amplifiers constructed using p-type second-harmonic-extraction silicon Trapatt devices is presented. Low-dielectric-constant substrates are used to fabricate simple microstrip circuits for use with packaged Trapatt devices. Amplifier gain, efficiency, bandwidth, average rf power output, and noise properties are evaluated and a theoretical analysis is presented for the device and circuit interaction.

1. Introduction

Pulsed Trapatt operation in X-band (8.2 to 12.4 GHz) has been reported by many researchers, but little information is available on the operating characteristics of X-band Trapatt amplifiers. The purpose of this paper is to describe the experimental characteristics and analysis of one particular type of X-band Trapatt amplifier—the second-harmonic-extraction amplifier. It will be shown that the devices and circuits are easier to construct than are fundamental-frequency Trapatt amplifiers, and that simple microstrip rf circuits can be used instead of the conventional coaxial circuit.

Excellent dc-to-rf conversion efficiency has been obtained for fundamental Trapatt oscillators operating at low duty factor between 7

* The work discussed in this paper was sponsored in part by the Air Force Systems Command, United States Air Force, Wright-Patterson Air Force Base, Ohio, under Contract No. F33615-73-C-1076.

and 9 GHz.[1-3] For example, Fong, Ying, and Lee[3] obtained 29.5% efficiency at 8.4 GHz using Trapatt diodes having a novel doping profile. This is about twice that that can be achieved with presently available second-harmonic devices.

The performance achieved by others for X-band reflection-type Trapatt amplifiers should be noted. Gibbons and Grace[4] obtained 7 dB gain with a saturated power output of 19 W at 29% efficiency at 8.6 GHz using a bias pulse width of 0.2 μs and a duty factor less than 1%. Cox, Rucher, and Gsteiger[5] have constructed fundamental-frequency Trapatt amplifiers that operate between 8.0 and 8.5 GHz with 7.7 W to 33.9 W peak rf power output and 12 to 16.1% efficiency at low duty factor. Risko et al.[6] have reported both oscillator and amplifier results for X-band harmonic-extraction Trapatts. The amplifier efficiency was low and the best oscillator efficiency was 17.5% at 8.8 GHz.

The purpose of the work reported here was to develop second-harmonic-extraction Trapatt amplifiers with at least 10 W output power in X-band and to study their characteristics. The amplifier performance reported earlier in this study[7] showed that output power between 10 and 20 W with efficiency slightly larger than 10% and a 3-dB bandwidth of about 200 MHz could be achieved with abrupt-junction, p-type diodes in microstrip circuits. Similar results have now been achieved with graded-junction, p-type devices. The experimental characteristics of these amplifiers is described and a simple method of amplifier analysis is presented.

2. Device Design and Fabrication

The Trapatt amplifiers described here operate with the input and output frequency being twice the fundamental Trapatt frequency. Therefore, the device design (i.e., the impurity doping, depletion layer width, area, etc.) is significantly different from that of a fundamental Trapatt device for the same operating frequency.

Clorfeine et al.[8] estimate that the optimum depletion layer width at breakdown, W_{BK}, for an abrupt-junction Trapatt device should be approximately, in microns,

$$W_{BK} = \frac{7}{F}, \qquad [1]$$

where F is the fundamental Trapatt frequency in GHz. Kawamoto et al.[9] showed that a graded-junction device will operate best with depletion width about 50% larger. Therefore, using 9 GHz as the design

frequency, the fundamental Trapatt would have $W_{BK} \approx 0.8$ μm if abrupt junction and $W_{BK} \approx 1.2$ μm if graded. A second-harmonic device would have $W_{BK} \approx 1.6$ μm if abrupt and $W_{BK} \approx 2.4$ μm if graded, since the fundamental Trapatt frequency for this case is 4.5 GHz. The extremely narrow depletion-layer width required for the 9-GHz fundamental devices makes it necessary that the width and doping of the epitaxial layer be known to a higher degree of accuracy in order to fabricate the proper depletion width. As will be demonstrated by the experimental results, the depletion widths of the second-harmonic-extraction devices are not very critical.

The impurity doping in the epi-layer will determine the punch-through voltage, the breakdown voltage, and affect the rf performance. Punch-through is said to occur when sufficient reverse voltage bias is applied such that an epi-layer just becomes fully depleted. Impatt's generally are designed for punch-through at breakdown, whereas Trapatts operate best with punch-through occurring at much lower voltage than the breakdown voltage. Such devices are said to have a high punch-through factor. Evans[10] has discussed the influence of punch-through factor upon dc-to-rf conversion efficiency, operating bias current, and overvoltage (i.e., the amount by which the device voltage exceeds the breakdown value during the Trapatt cycle). Evans shows that a high punch-through factor is required in order to keep the required value of overvoltage low enough to be realizable in experimental circuits.

The punch-through voltage, V_{PT}, may be estimated for an abrupt-junction device made from a uniformly doped epi-layer. Integration of Poisson's equation for this case yields

$$V_{PT} = \frac{e}{2\epsilon} N_B W_{PT}^2, \qquad [2]$$

where N_B = impurity doping of the epi-layer,
e = charge on an electron,
ϵ = dielectric constant of silicon,
W_{PT} = depletion-layer width at punch-through

Since the value of W_{PT} will not be greatly different from that of W_{BK}, V_{PT} is mainly determined by N_B. For the case of a graded-junction device, V_{PT} will be less than predicted by this equation because N_B is effectively less.

Clorfeine[11] has discussed the choice of J_o, the bias current density, and the choice of $N_B W_{BK}$ products for abrupt-junction devices. For high average power, a good choice is

$$N_B W_{BK} \approx 5 \times 10^{15}, \qquad [3]$$

and

$$J_o \approx J_c, \qquad [4]$$

where N_B = impurity doping per cm^3,
W_{BK} = depletion width at breakdown, in μm,
$J_c = eN_B v_s$ = critical current,
v_s = saturated drift velocity of majority carrier.

Since the depletion-layer width of a fundamental X-band Trapatt is one-half that of the second-harmonic device, the fundamental X-band Trapatt should have about twice the doping value and twice the current density of a second-harmonic device.

Again, the graded-junction device should be considered to be of similar design but with a smaller effective value of N_B. Then Eq. [3] suggests a larger depletion-layer width and a lower operating J_o for a graded-junction device. These features of the graded-junction device result in an increased rf impedance, which eases the circuit-design requirements and promotes larger peak power output.

Once J_o is chosen, the operating bias power density can be calculated. The rf peak power output is then simply the estimated efficiency times the device area times the bias power density. However, the choice of device area does affect the efficiency, as Fong et al.[3] have shown. The pulse heating of the device is also influenced by the choice of area (see Ref. [9] and the amplifier bandwidth may also be affected. Furthermore, Trapatt reflection amplifiers differ fundamentally from Trapatt oscillators with respect to their starting mechanisms. Thus, the choice of J_o/J_c and N_B cannot be made wholly on the basis of these guidelines for oscillators. Later sections of this paper show that J_o strongly influences the "time delay," which is the time between the device voltage rising above threshold and the formation of an avalanche shock front (and a voltage maximum). The time delay of an amplifier is also influenced by the rf input signal, the impurity doping, the avalanche region width, and other factors. One contribution of this study is a clearer understanding of the factors affecting operation of Trapatt amplifiers, based upon an analysis of experimental amplifiers.

Table 1 presents the starting-wafer characteristics and final (diffused) wafer characteristics for the four p-type second-harmonic wafers. The characteristics of the starting wafers were obtained from C-V data on test diodes made with a shallow phosphorus diffusion. The characteristics of the final wafers were obtained from finished Tra-

Table 1—Second-Harmonic Wafer Characteristics

	Starting Wafer		Diffused Wafer						
	Epi-Thickness (μm)	Doping (atoms/cm^3)	Contact Pot. ϕ (V)	Initial Slope of C-V	Junction Type	V_{PT} (V)	V_{BK} (V)	W_{BK} (μm)	Junction Depth (μm)
RK-1	4.84	4.8×10^{15}	0.75	0.48	Abrupt	14.	55.	2.13	1.5
JR-27	5.18	6.4×10^{15}	0.50	0.38	Graded	15.	58.	1.45	1.6
JR-28	5.31	8.8×10^{15}	0.50	0.37	Graded	14.	52.	1.85	1.4
JR-37	5.97	3.0×10^{15}	0.58	0.38	Graded	8.	65.	1.79	2.0

patt diodes. These wafer characteristics are also tabulated in Table 1. Fig. 1 shows the C-V data for the Trapatt diodes. Notice that wafer RK-1 has an abrupt junction, but the other three wafers have graded junctions. All wafers have a depletion width at breakdown close to 2.0 μm and are heavily punched-through at breakdown. An estimate of the junction depth for each wafer is also given in Table 1.

The devices are fabricated in a conventional manner. Initially, the wafer is mechanically reduced in thickness. The wafer is then metallized across the surface with successive evaporations of Ti, Pd, and Au. A gold heat sink is plated on the metallization. The silicon is then

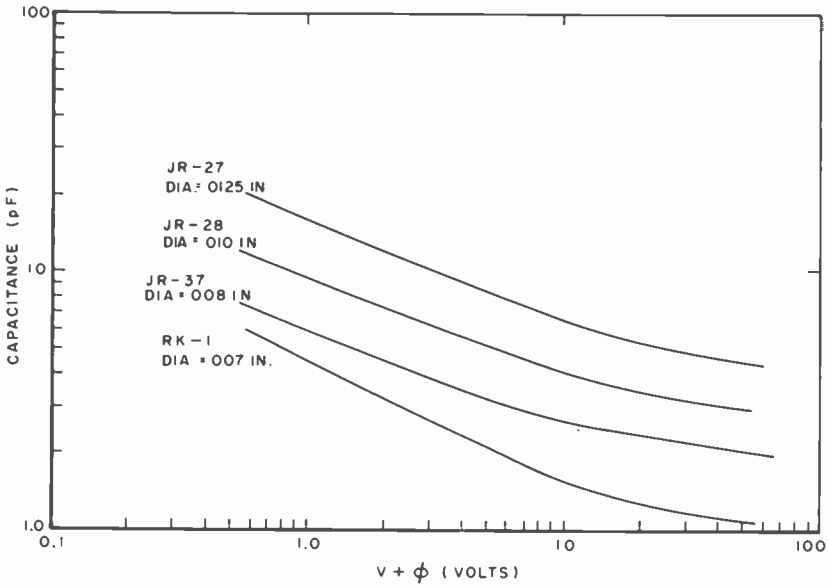


Fig. 1—Capacitance-voltage data for four p-type Trapatt wafers.

reduced to final device thickness by chemical etching. Ti-Pd-Au is next evaporated on the substrate side. Photoresist is then used to define the contact geometry. The exposed gold on the substrate side is etched away and the photoresist is then removed. Photoresist is next applied to define the contact areas to be built up by plating. The build-up of the contact areas forms a mushroom-like heat cap that helps dissipate transients during pulsed operation. Kawamoto et al.[9] show that with silicon device thicknesses of about 0.5 mil, the heat capacitor allows about 54% longer bias pulse for the same bias power and pulse temperature rise.

The photoresist around the heat caps is removed, and the heat caps are then used as a mask against the etchant used to define the indi-

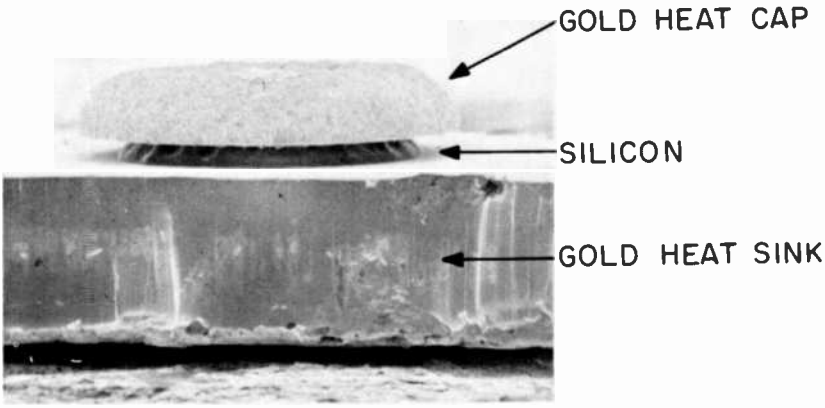


Fig. 2—10-mil-diameter mesa Trapatt with heat capacitor, as viewed with scanning electron microscope.

vidual diodes. The individual diodes are then separated by cutting through the gold heatsink with a sharp blade mounted on a micromanipulator. Fig. 2 shows a completed device with heat capacitor. The diodes are mounted in threaded pill packages using gold–tin preforms and four lead wires are bonded to it. The standard package used (designated No. 921) has a package capacitance of 0.37 pF and the lead inductance is approximately 0.3 nH.

3. RF Circuit and Test Equipment

The rf microstrip circuit design used for most tests is shown in Fig. 3. The circuit boards are 0.040 inch thick Duroid material (relative di-

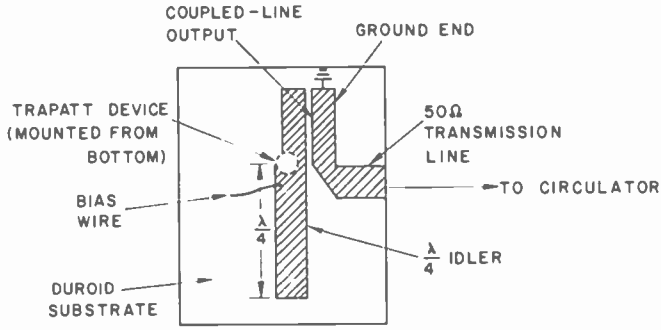


Fig. 3—Top view of microstrip circuit board.

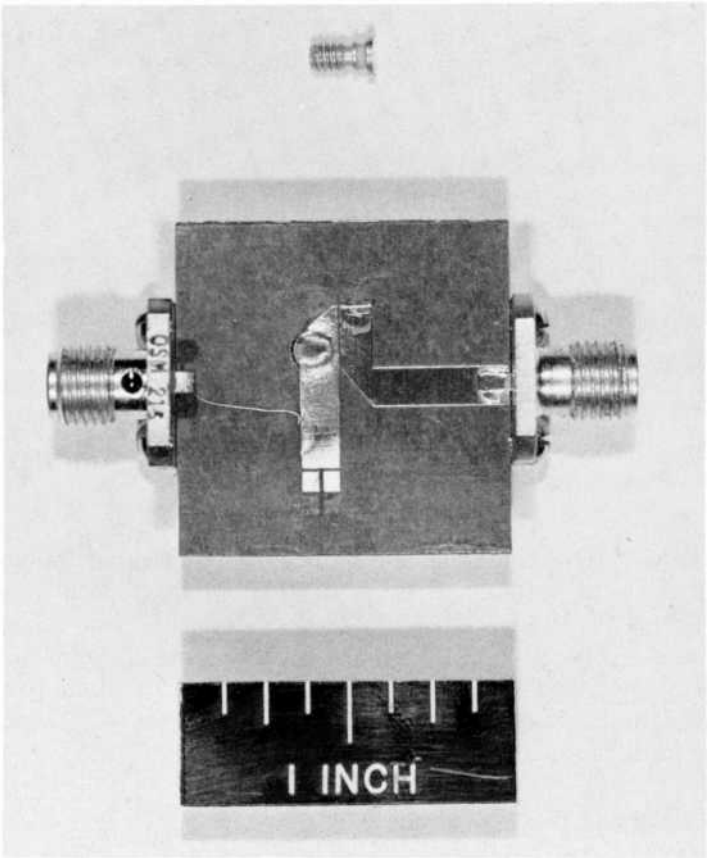


Fig. 4—Trapatt amplifier and device package.

electric constant ≈ 2.2) and have copper jackets. The rf output circuit consists of a coupled-line output transformer and is similar in design to that described by Rosen et al.[12] for S-band Trapatt operation. The even- and odd-mode impedances are typically about 70 ohms and 30 ohms, respectively, and the transformer length is approximately λ/4 at the output frequency. The circuit requires tuning by means of copper and alumina chips secured to the top of the circuit board for optimum rf performance. Each circuit board is soldered to a block of "OFHC" copper that serves as a heat sink for the Trapatt device. Fig. 4 shows a finished amplifier.

The Trapatt device in the rf circuit sees an open-circuited (resonant) microstrip line on one side and the coupled-line transformer on the other side. The impedance presented to the device can be calcu-

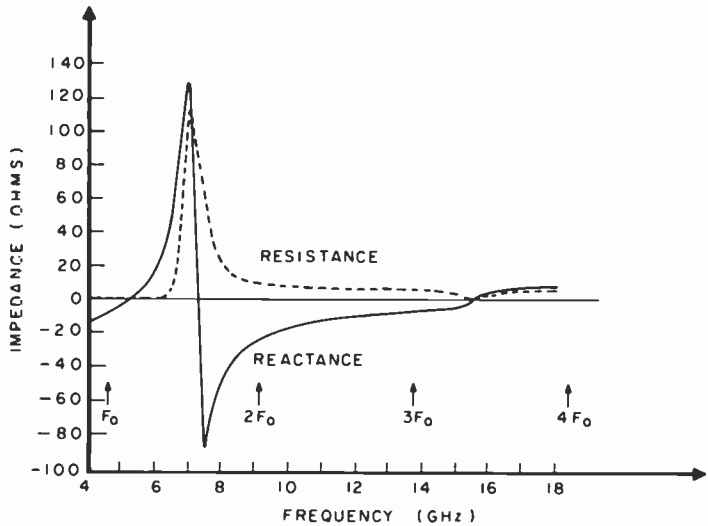


Fig. 5—Theoretical circuit impedance for coupled-line Trapatt circuit for even-mode impedance of 70.5 ohms, odd-mode impedance of 24.9 ohms, coupled line length (rf air line) of 0.5 cm, open stub line length (rf air line) of 1.45 cm, and output load of 50 ohms.

lated theoretically for specified values of resonant line length, transformer line length, even- and odd-mode impedances of the coupled-line section, and load impedance on the output line. A digital computer program was developed for this purpose. Fig. 5 shows an example of such a computation for parameters of one experimental circuit. The figure presents the calculated impedance as a function of rf fre-

quency at the device position. For present values of package parameters, Trapatt operation would be expected, with F_o being the fundamental Trapatt frequency. Calculations are shown up to $4F_o$. Fig. 6 shows the actual impedance of the circuit measured using an automatic network analyzer system. Notice that the general behavior of resistance and reactance is similar. The experimental circuit that was considered here did operate as a Trapatt amplifier at 9.2 GHz with device JR-28-14 producing about 11 W power output with 4.1 dB gain and 5.5% efficiency.

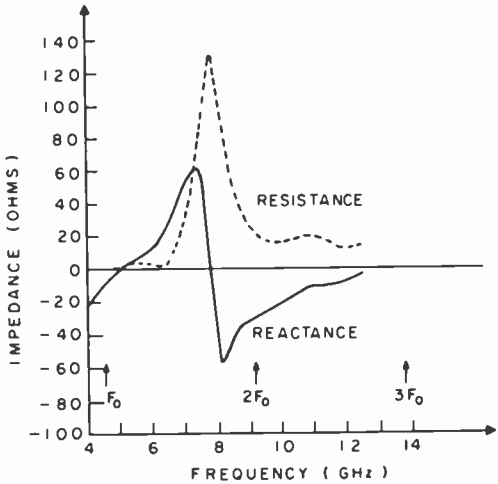


Fig. 6—Measured circuit impedance presented to the Trapatt diode.

The impedance characteristic of this circuit should be viewed in the following manner. The impedance at F_o results principally from being below the $\lambda/4$ resonant frequency of the open stub, which then causes capacitive loading to the diode. The impedance at $2F_o$ consists of the 50-ohm load transformed through the coupled-line transformer and a capacitive reactance term due to being below the $\lambda/4$ frequency of the transformer. Both the experimental and the theoretical impedance measurements are consistent with this interpretation. The circuit analysis just described has provided a means for evaluating the effect of circuit design changes without constructing the circuits and has been a valuable design tool.

Most rf tests were performed using the Trapatt device package screwed into a 0.125-inch stud which then was inserted into a hole in

the copper block and secured when the top of the device touched the microstrip line. This technique enabled any number of devices to be tested in the same circuit and, in addition, made possible the measurement of circuit impedance as seen by the device. For this test a special coaxial probe was inserted into the hole, and impedance data were measured using the automatic network analyzer.

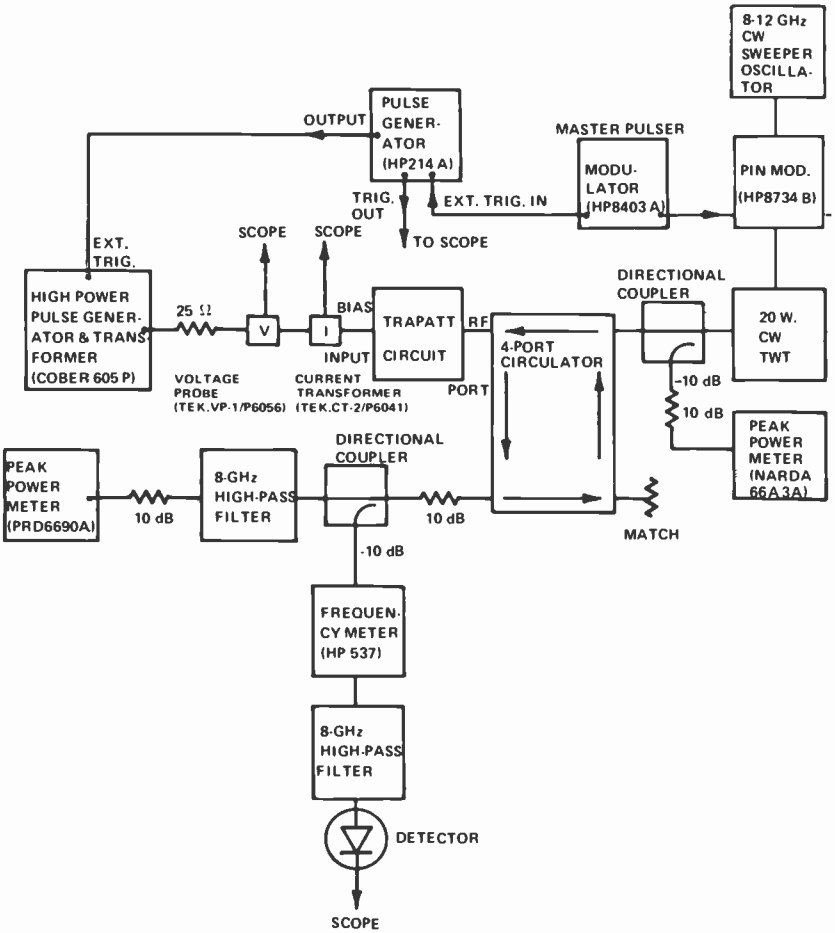


Fig. 7—Equipment layout for pulsed amplifier tests.

Fig. 7 shows the equipment layout for the amplification tests. The input and output power meters were calibrated to provide power data referred to the Trapatt rf port. Thus, circulator losses are not included in the data presented. The one-way circulator loss is typically

0.4 dB. Note that a dc blocking capacitor is not required at the rf port of the amplifier because of the use of a coupled bar rather than a coaxial structure. The 8-GHz high-pass filter is required because under some tuning conditions significant power output results at the fundamental Trapatt frequency (i.e., 4 to 5 GHz). However, no significant output power could be detected for harmonics above 10 GHz (i.e., $3F_o$, $4F_o$, etc.).

4. Amplifier Performance

The amplifiers described generally behaved in one of two different ways. They either operated as over-saturated amplifiers where an increase in rf input power caused a decrease in rf output power, or as under-saturated amplifiers, where an increase in rf input power caused an increase in rf output power. In both cases the input signal could be varied at least 3 dB without any problems in the rf output (i.e., distortion, pulse break-up, etc.). Fig. 8 shows a typical example

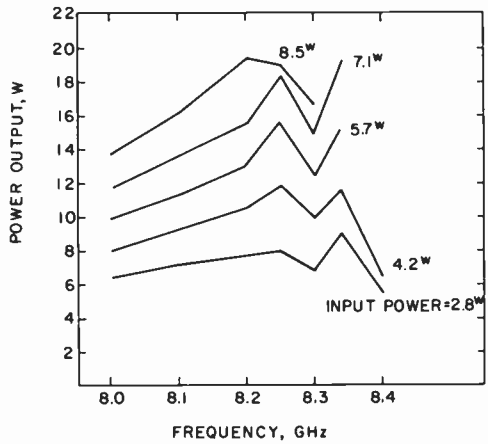


Fig. 8—Power output as a function of frequency for an amplifier at several values of rf input power (current = 2.5 A, voltage ≈ 50 V, mesa diameter = 0.007 inch, pulse width = 0.5 μs, pulse repetition rate = 1 kHz).

of the rf behavior for an under-saturated amplifier for narrow pulse width (0.5 μs). Data was not obtained below 8.0 GHz, but it appears that the bandwidth of operation is not changed much as the rf input power is increased. A noisy output signal results if the rf input power is outside the range of values shown in the figure. The bandwidth is generally determined by the end of the pulse breakup at the high-fre-

quency side and just by the low rf power output (3 dB down) on the low-frequency side.

Fig. 9 shows rf power output and bandwidth data for an over-saturated amplifier tested under similar conditions but with different circuit tuning. Notice here that the rf output power decreases with increase of rf input power and that the amplification bandwidth increases strongly with increase of rf input power. Here operation below 0.22 W of rf input power results in a very noisy rf output pulse. Thus, the under-saturation condition cannot be achieved. This type of behavior is very much like injection-locked oscillator behavior. That is, the amplifier is probably tuned for a load condition that is nearly proper for oscillation without an input signal.

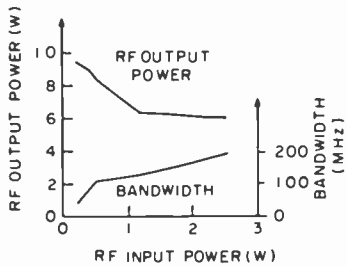


Fig. 9—RF power output and bandwidth vs. rf input power for Trapatt amplifier operated with fixed tuning and pulse width = 0.5 μs, repetition rate = 1 kHz, I_o = 2.5 A, V_o = 37 to 42 V, V_{BK} = 53.5 V.

In general, both types of behavior were observed, and either may be used for power amplification. Similar values of efficiency and bandwidth were observed for both types of saturation conditions.

Table 2 presents a summary of amplifier measurements for 0.5-μs pulse width and with a low duty factor. Test data for all four operating wafers are shown. The efficiency shown is calculated as the rf output power minus the rf input power divided by the bias power. The bandwidth is the 3-dB value unless otherwise specified. A variety of device diameters were tested, and many circuit improvements were made in the circuit during this study. The best performance with 10 to 12 W rf output power is achieved with device diameters from 7 to 9 mils. Greater than 10% efficiency can be achieved with about 200 MHz bandwidth and greater than 6 dB gain. No improvements in rf performance is seen for the smaller (1290) device package compared with that of the standard (921) package. Calculation shows that the bias current densities are generally lower for the larger-diameter de-

vices, which is favorable for achieving operation with much longer bias pulse length.

The circuit was easily tunable. Several RK-1 devices about 0.007 inch in diameter were shown to be tunable for an amplification bandwidth of 100 to 200 MHz anywhere between 8.0 GHz and 10.0 GHz. The operating bias and rf input power were held approximately constant for these tests. Similar tunability was demonstrated on devices of larger diameter, although more care (and time) was required. The same result was found for devices made from wafers JR-27, JR-28 and JR-37. In general, no optimum frequency was found for devices

Table 2—RF Performance of Second-Harmonic-Extraction Trapatt Amplifiers for Pulse Width = 0.5 μs and Repetition Rate = 1 kHz

Device	Mesa Dia. (mils)	V_{BK} (V)	V_o (V)	I_o (A)	Output P_{RF} (W)	Input P_{RF} (W)	Gain (dB)	Frequency (GHz)	Efficiency (%)	Bandwidth (MHz)
RK-1F-1	7.1	54.	42.5	2.6	15.0	4.5	5.2	8.43	9.5	350
RK-1F-6	6.8	54.	45.	2.5	16.0	4.0	6.0	9.43	10.7	180
RK-1F-30*	7.8	51.	42.	3.55	13.3	1.1	10.7	9.85	8.2	170
RK-1F-36	8.6	53.5	46.	1.7	12.0	2.1	7.6	8.10	12.6	50
RK-1F-36	8.6	53.5	46.	1.9	11.0	1.8	7.9	8.44	10.5	260
RK-1F-38	9.5	53.5	41.	2.35	10.4	3.2	5.1	8.10	7.5	160†
JR-27-1	7.0	60.	50.	2.5	14.0	4.5	4.9	8.72	7.6	30
JR-27-6	11.5	60.	52.	3.5	21.0	7.6	4.4	8.16	7.4	60
JR-28-7	8.7	62.	55.	2.4	14.6	1.7‡	9.3	8.10	8.7	150
JR-37-34	6.1	61.	49.	1.9	12.1	2.05	7.7	8.96	11.3	160†

* Smaller device package used (No. 1290).
† 1-dB bandwidth.
‡ Bad amplitude distortion.

from any given wafer. Any given device could be made to operate with about the same efficiency and bandwidth over a frequency range of 1.0 to 1.5 GHz.

Devices made from wafers JR-27 and JR-28 did not exhibit as good rf performance as the RK-1 devices. JR-27 amplifiers exhibited quite narrow bandwidths and poor efficiency except at very large current density. JR-28 amplifiers usually showed amplitude modulation distortion in their rf output signals. This effect seems to be related to the large doping depletion-width products of these devices. The product does exceed Clorfeine's[11] maximum design limits. Device JR-28-14 mentioned earlier shows a small but noticeable amount of this distortion. Operation at low current density did eliminate this effect.

Good rf performance was achieved with devices made from wafer JR-37. The combination of the graded junction and the lower impurity doping (as compared with JR-27 and JR-28) definitely improved the rf performance. Table 2 shows that efficiency above 10% was achieved with good bandwidth and gain. Fig. 10 shows data for rf output as a function of frequency for several values of bias current. Good efficiency is achieved even for the low bias condition.

Several RK-1 devices were tested as fundamental Trapatt oscillators. It was observed that the maximum conversion efficiency as a fundamental oscillator is about twice the efficiency as a second-harmonic-extraction oscillator. For example, device RK-1F-10 produced 25 W at 4.3 GHz with 19% efficiency but a maximum of 10 W at 9.55 GHz with 8.9% efficiency. It is not clear whether this efficiency reduction is a fundamental property, as Snapp[13] suggests, or due to circuit difficulties at the higher frequency.

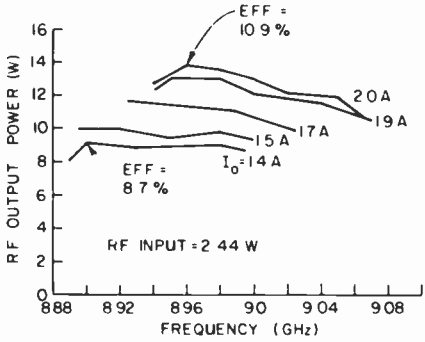
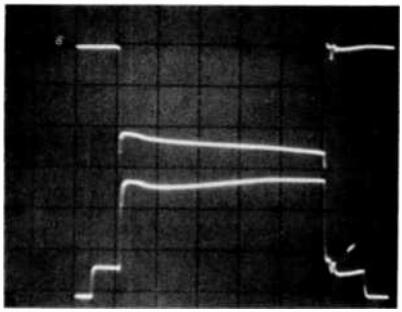


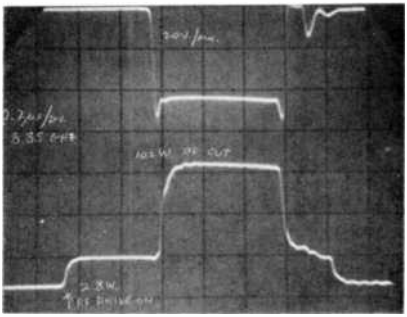
Fig. 10—RF power output vs. frequency for amplifier JR-37-32, fixed tuned during tests and for bias pulse = 0.5 μs, pulse repetition rate = 1 kHz, device diameter = 0.0074 inch, voltage = 48 to 56 V, V_{BK} = 61.2 V.

Many amplifiers were operated with bias pulse length longer than 0.5 μs. In general, the device efficiency and amplification bandwidth were markedly reduced with bias pulse lengths above 1 μs. The best efficiency for long-pulse operation was achieved with an RK-1 amplifier. This device did not have a heat capacitor but could be operated with a long bias pulse, because Trapatt mode could be obtained with good efficiency with very low bias power density (about 240 kW/cm^2). Over 10 W was obtained at 8.1% efficiency for 5.0 μs pulse width with 5.6 dB gain and 40 MHz bandwidth with amplifier RK-1F-36. Fig. 11 shows an oscilloscope picture of the bias voltage and the detected rf

output for this amplifier. The shoulders on the detected rf presentation are due to reflection of the rf input signal. Notice that the bias voltage rises during the 5.0-μs pulse although the current is maintained nearly constant. This is evidence of the influence of pulsed heating upon the operating point of the Trapatt device. The second oscilloscope picture in Fig. 11 shows a shorter pulse length and shows that there is no observable lead edge jitter for the same tuning condition.



(a) BIAS VOLTAGE (UPPER WITH 20V/div) AND DETECTED RF OUTPUT (LOWER) vs. TIME (1μs/div)



(b) BIAS VOLTAGE (UPPER) AND DETECTED RF OUTPUT (LOWER) vs TIME (0.2μs/div)

Fig. 11—Bias voltage and rf output pulse vs. time for amplifier RK-1F-36 operating at 8.85 GHz, I_o = 1.9 A, rf output = 10.2 W, and rf input = 2.8 W.

A number of amplifiers were tested at high duty factors with a water-cooled block and the data is tabulated in Table 3. For devices with thermal resistance less than 20°C/W it was found possible to obtain 0.6 to 0.94 W of average rf output power. The bandwidth of amplification in present circuits does not appear to reduce greatly with

increased duty factor produced by high repetition rates. However, the device efficiency is reduced by increased duty factor.

A microwave noise-measuring instrument* was used for measurement of noise modulating sidebands. Using calibrating AM sidebands produced by a p-i-n modulator, it was found that the lowest AM sidebands observable with pulsed rf of 0.7 μs duration and 50 kHz repetition rate were about −60 dB per kHz bandwidth. This is equivalent to −90 dB per Hz bandwidth. For FM measurements, a transmission-type cavity wavemeter preceded the noise-measuring instrument and functioned as a discriminator (i.e., to convert FM to AM) by adjusting its resonant frequency to be just slightly different from the signal frequency. Unfortunately, this type of system is not very sensitive to FM. The minimum observable peak frequency deviation was about 3 kHz per kHz bandwidth for modulation rates of 2 kHz to 40 kHz using an rf pulse of 0.7 μs duration and 50 kHz repetition rate.

Table 3—High-Duty-Factor Tests for Pulse Width ≈ 0.5 μs and Heatsink Temperature ≈ 15°C

Device	Thermal Resistance (°C/W)	I_o (A)	V_o (V)	RF Input (W)	RF Output (W)	Frequency (GHz)	Duty Factor (%)	Average RF Output (mW)
RK-1F-6	26.0	2.5	51.	3.0	13.6	8.91	3.1	420.
JR-37-MI-10	15.2	3.2	46.	3.2	12.8*	9.22	5.0	640.
JR-37-MI-11	17.2	3.5	57.	2.4	15.0	9.14	6.25	938.

* Bandwidth = 190 MHz.

A Trapatt amplifier using a diode from wafer JR-37 was operated at 11.6 W peak rf power output, 5.6 dB gain, at 8.9 GHz with a bias pulse length of 0.7 μs and a pulse repetition rate of 50 kHz. The AM and FM sidebands from 2 kHz to 40 kHz were below the measurement sensitivity of the equipment. When the amplifier was operated in a detuned condition or with improper bias pulse, the resulting AM and FM modulation was measurable.

5. Trapatt Circuit Analysis

A technique for Trapatt amplifier analysis will now be discussed. The results are in reasonable agreement with experimental data and it is expected that the analysis will give insight to achieving broader amplification bandwidth by circuit changes.

* Spectra-Electronics, Inc., Model SE200A.

Fig. 12 shows the model assumed for the Trapatt amplifier containing a packaged diode. The Trapatt diode is modeled, as suggested by Evans,[14] as a pulsed current generator in shunt with C_o, its depletion-layer capacitance, and both are in series with a resistor R_o that represents diode and package losses. The current generator is taken to be of the form suggested by Carroll and Crede,[15]

$$I_p(t) = A\left(\frac{1 + \cos\omega t}{2}\right)^N, \qquad [5]$$

where $\omega = 2\pi F_o$ = fundamental (angular) Trapatt frequency,
t = time,
N = factor determining sharpness of the current pulses, and
A = amplitude factor that is adjusted so that the average value of $I_p(t)$ is I_o, the bias current.

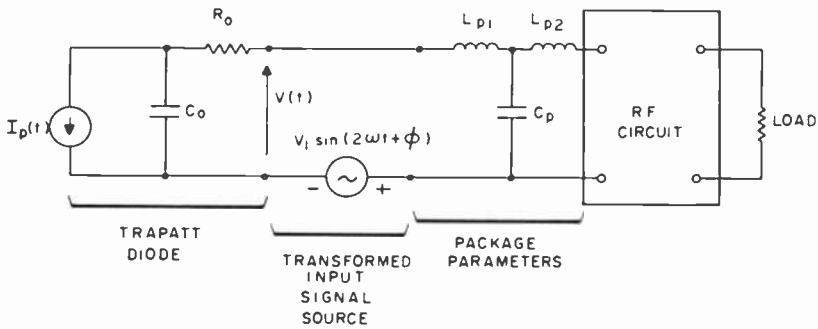


Fig. 12—Model assumed for the second-harmonic-extraction Trapatt amplifier.

The second-harmonic signal source of amplitude V_1 and phase φ is the equivalent input signal source transformed through the rf circuit. The input signal power is calculated as the available (or incident) input power and the output power is calculated as the power reflected to the generator. This is consistent with the experimental procedure, since a circulator is always used between the signal source and the Trapatt amplifier.

The Trapatt voltage waveform, $V(t)$, can be found for any $I_p(t)$ if the rf circuit impedance characteristic has been measured and V_1 and φ are specified. Let $Z_c(\omega)$ represent the circuit impedance as seen by the packaged device and let $Z_c{}^T(\omega)$ represent $Z_c(\omega)$ transformed to the Trapatt current source (i.e., transformed through L_{p1}, L_{p2}, C_p, R_o, and C_o). $Z_c(\omega)$ can be measured and $Z_c{}^T(\omega)$ can be calculated once

packaging and device parameters are determined. The pulsed current (Eq. [5]) then can be represented by a Fourier series as

$$I_p(t) = \sum_{n=0,\pm 1,\ldots}^{n=\pm\infty} C_n e^{jn\omega t}. \qquad [6]$$

The diode voltage waveform is then

$$V(t) = V_o - \sum_n C_n Z_c^T(n\omega) e^{jn\omega t} + V_1' \sin(2\omega t + \varphi'), \qquad [7]$$

where $Z_c^T(n\omega)$ is the value of $Z_c^T(\omega)$ at (angular) frequency $n\omega$, V_o is the value of bias voltage, and V_1' and φ' are signal source parameters transformed to the Trapatt terminals. For the calculations presented, a value of $n = 4$ was used.

This analysis was applied to the case of an oscillator by setting $V_1 = 0$. Using measured values of C_o, R_o, I_o, V_o, and packaging parameters, acceptable voltage waveforms were found only for $N = 6$. In addition, good Trapatt waveform showing overvoltage and voltage drop at proper times was found only for an output frequency of 8.01 GHz ± 50 MHz, with rf output power of 10 W. The experimental oscillator operated with 8.5 W output at 7.98 GHz, which is in reasonably good agreement.

This technique was next applied to an amplifier operating around 8.1 GHz. This frequency was picked for the reason that the circuit impedance could be accurately measured through $3F_o$ for this case, and the impedance at $4F_o$ can be measured, but with less accuracy. Device RK-1F-38 in a type 921 package, was used with the following operating conditions:

I_o	= 2.35A
V_o	= 41 V
V_{BK}	= 53 V
Pulse Width	= 0.5 μs
Pulse Repetition Rate	= 1 kHz

The value of R_o was determined from Q measurements with the diode biased below breakdown in the manner described by De Loach[16] and using the automatic network analyzer. Packaging parameters were also determined, and measurements were made of the experimental circuit impedance after tuning for amplification around 8.1 GHz.

Fig. 13 shows the diode voltage waveform $V(t)$ calculated from the impedance data ($n = 4$) and for $N = 8$ (in Eq. [5]), $2F_o = 8.1$ GHz, and rf input power = 1.02 W. A 2 π phase change occurs in one funda-

mental Trapatt period, which is $1/F_o$. Notice that $V(t)$ with the input signal present is a good Trapatt waveform showing overvoltage and voltage drop at proper times. The pulsed current waveform is also shown in the figure for reference. The rf power output calculated for this case is 12 W.

No other values of N produced an acceptable waveform for this value of input power. The waveform for $V_1 = 0$ is also shown for comparison. It is not acceptable because $V(t)$ exceeds V_{BK} near $\omega t = \pi/2$,

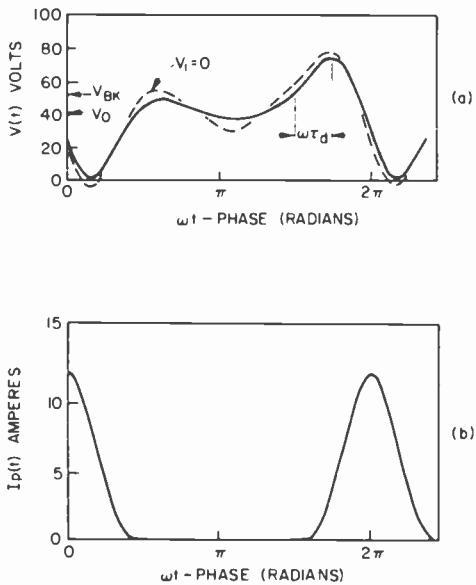


Fig. 13—Calculated Trapatt amplifier waveforms for (a) diode voltage and (b) pulsed current source for amplifier RK-1F-38 with I_o = 2.35 A, V_o = 41 V, F_o = 4.05 GHz, C_o = 2.25 pF, R_o = 0.16 ohm, $L_{p1} = L_{p2}$ = 0.16 nH, C_p = 0.35 pF, 921 package, V_1 = 11.3 V, and φ = 2.42 radians.

which would cause premature avalanche.[8] In addition, $V(t)$ goes below zero, which would interfere with the plasma-extraction process.

Fig. 14 shows the computer-generated results and also the experimental data. The predicted and measured rf power output values are in good agreement. Furthermore, an increase in rf input power causes a reduction in rf output in the computer results as well as in the experiment. Thus, the operating characteristics of the amplifier are reasonably duplicated by the model and therefore the model should be helpful to evaluate changes in circuit design.

Another important operating parameter that has been identified is τ_d, the time delay. The time delay is the time it takes the avalanche

shock front to form, and is the time between $V(t) = V_{BK}$ and $V(t) = V_{MAX}$, the maximum voltage value; τ_d is illustrated in Fig. 13. The time delay is of fundamental importance in achieving Trapatt operation, and an estimate of τ_d for a given diode is useful for evaluating computer-generated waveforms for that diode.

Evans[10] gives an estimate of τ_d as

$$\tau_d = \frac{\epsilon}{\lambda J_{OB}} \ln\left[1 + \frac{\lambda J_{OB}}{\epsilon}\left(t_d + \frac{\overline{W}}{2v_s} \ln\frac{J_{OB}}{J_s}\right)\right]. \qquad [8]$$

where ϵ = dielectric constant of silicon,
λ = coefficient in the assumed relationship for ionization rate of $A_0 e^{\lambda E}$,
E = electric field,
J_{OB} = current density at $V(t) = V_{BK}$,
$\overline{W}$ = avalanche region width,
v_s = saturation velocity,
and J_s = reverse saturation current density.

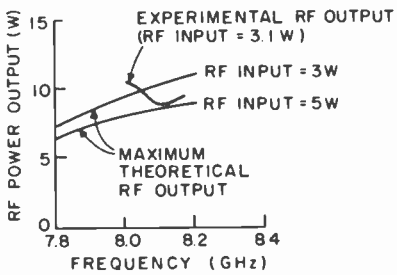


Fig. 14—Experimental and theoretical rf power output as a function of frequency for amplifier RK-1F-38.

Values of λ, v_s, $\overline{W}$, and J_s were estimated for the Trapatt device tested from the known values of breakdown voltage, impurity doping, and operating temperature. The electric field values in the depletion layer at breakdown were estimated to be from 2 to 4 × 10^5 V/cm. The parameter λ was evaluated by matching the assumed form of ionization rate to experimental rates for this range of electric field values, giving $\lambda = 1.67 \times 10^{-5}$ cm/V. The avalanche region width was estimated to be about 0.7 μm, and J_s was taken to be 1 A/cm^2. Using these values, Eq. [8] can be solved by trial and error to give a value of τ_d = 20 ps for J_{OB} = 5 kA/cm^2, and a value of 13 ps for J_{OB} = 10 kA/

cm^2. τ_d does not vary strongly with change in the other parameters, including temperature. Thus, as J_{OB} is increased, τ_d is shortened, and it is easier to produce Trapatt operation. J_{OB} is principally due to the displacement current at breakdown. For the amplifier of Fig. 13, J_{OB} is about 3.7 kA/cm^2, which is much less than J_c (i.e., $eN_B v_s$ = 7.7 kA/cm^2). The average bias current density is 5.1 kA/cm^2 and is also less than J_c. Fig. 13 gives a value of τ_d = 29 ps, which is in reasonable agreement with Eq. [8], which predicts 25 ps for this value of J_{OB}.

Inspection of the calculated $V(t)$ waveforms shows that the circuit does not permit as much overvoltage or as large a value of τ_d on the high-frequency side of the amplification band as compared with that on the low-frequency side. No valid Trapatt waveforms were found above 8.3 GHz.

5. Conclusions

It has been shown possible to construct reflection-type X-band pulsed amplifiers with approximately 10 W peak power output using second-harmonic p-type Trapatt devices with simple microstrip circuits. The maximum amplifier power conversion efficiency obtained with an abrupt-junction device was 12.6%, and 11.5% was obtained with a graded-junction device. The best efficiency for graded junctions was produced by devices from the wafer of lowest impurity doping (JR-37). A further reduction in doping may be beneficial. It is also possible that the substrate out-diffusion present in all wafers did reduce the achievable efficiency. Grace et al.[1] have found this to be an important factor for n-type Trapatts, and the out-diffusion in p-type material is considerably larger. Less out-diffusion could be achieved by starting with p-type epitaxial layers thinner than used here and using a shallower diffusion or ion implantation to form the junction. Devices of n-type should also have less out-diffusion.

The Trapatt amplifier analysis presented shows that the impedance characteristic of the circuit determines the voltage waveform and amplifier bandwidth. Since reasonable agreement is obtained between the model and experimental results, it is believed that this model will enable circuit improvements to be made in a systematic manner.

Present p-type Trapatt amplifiers are shown capable of about 200 MHz bandwidth, 6 to 8 dB gain, 10 to 20 W output power with 10 to 11% efficiency from 8 to 10 GHz for low duty factor and 0.5 μs pulse width. A reduction in efficiency and bandwidth occurs for increased pulse width. Likewise, efficiency was found to be reduced at large

duty factor but bandwidth was not affected very much. An average rf power output of 0.94 W was achieved with 15 W peak power at 9.14 GHz.

Acknowledgments

The authors are indebted to J. B. Klatskin for assistance in fabrication of the Trapatt devices and to E. C. McDermott for construction of the microstrip circuits.

References:

[1] M. I. Grace, H. Kroger, and J. Telio, "Improved Performance of X-band TRAPATT's," *Proc. IEEE*, **60**, No. 11, p. 1443, Nov. 1972.

[2] R. S. Ying and T. T. Fong, "C-Band Complementary TRAPATT Diodes," *Proc. IEEE*, **62**, No. 2, p. 1974, Feb. 1974.

[3] T. T. Fong, R. S. Ying, and D. E. Lee, "Ion-Implanted X-Band IMPATT/TRAPATT Back to Back Diodes," *Proc. IEEE*, **61**, No. 5, p. 1044, July 1973.

[4] G. Gibbons and M. I. Grace, "High-Efficiency Avalanche-Diode Oscillators and Amplifiers in X-Band," *Proc. IEEE*, **58**, No. 3, p. 512, March 1970.

[5] N. W. Cox, C. T. Rucker and K. E. Gsteiger, "X-Band Trapatt Amplifier," 1974 IEEE S-MTT International Microwave Symp., Atlanta, Ga., *IEEE Catalog No. 74CH0838-3MTT*, p. 367, June 1974.

[6] J. J. Risko, J. Thomas, H. J. Prager and K. K. N. Chang, "I-Band (8–10 GHz) TRAPATT Diode Sources," *Elec. Lett.*, **9**, No. 24, p. 572, Nov. 20, 1973.

[7] P. T. Ho and W. R. Curtice, "A Microstrip Trapatt Amplifier for X-Band Operation," *Proc. IEEE*, **62**, No. 7, p. 1029, July 1974.

[8] A. S. Clorfeine, R. J. Ikola, and L. S. Napoli, "A Theory for the High-Efficiency Mode of Oscillation in Avalanche Diodes," *RCA Review*, **30**, No. 3, p. 397, Sept. 1969.

[9] H. Kawamoto, H. J. Prager et al., *S-Band Avalanche Diode Amplifiers*, Final Report, RCA Laboratories, Princeton, NJ, Contract No. F19628-70-C-0230, ESD-TR-74-151, Oct. 1973.

[10] W. J. Evans, "Computer Experiments on TRAPATT Diodes," *IEEE Trans. Microwave Theory and Techniques*, **MTT-18**, No. 11, p. 862, Nov. 1970.

[11] A. S. Clorfeine, "Guidelines for the Design of High-Efficiency Mode Avalanche Diode Oscillators," *IEEE Trans. Electron Devices*, **ED-18**, No. 8, p. 550, Aug. 1971.

[12] A. Rosen, J. F. Reynolds, S. G. Liu, and G. E. Therlault, "Wideband Class-C TRAPATT Amplifiers," *RCA Review*, **33**, No. 4, p. 729, Dec. 1972.

[13] C. F. Snapp, "Harmonic Synergy and the Trapped-Plasma Mode in Avalanching P^+NN^+ Semiconductor Junctions," Technical Report No. RADC-TR-71-67 for Rome Air Development Center (Griffis Air Force Base, New York), Cornell University, Ithaca, N.Y., April 1971.

[14] W. J. Evans, "Circuits for High-Efficiency Avalanche Diode Oscillators," *IEEE Trans. Microwave Theory and Techniques*, **MTT-17**, No. 12, p. 1060, Dec. 1969.

[15] J. E. Carroll and R. H. Crede, "A Computer Simulation of TRAPATT Circuits," *International J. of Electronics*, **32**, p. 273, 1972.

[16] B. C. DeLoach, "A New Microwave Measurement Technique to Characterize Diodes," *IEEE Trans. Microwave Theory and Tech.*, **MTT-12**, No. 1, p. 15, Jan. 1964.

Microwave Integrated-Circuit MIS Varactor Phase Shifter

R. L. Camisa,* B. F. Hitch,† S. Yuan,† and M. Ettenberg‡

Abstract—A digital voltage-controlled diode phase shifter that requires virtually no steady-state dc power has been developed using metal-insulator-semiconductor (MIS) varactors. Physical mechanisms and device design relationships relevant to phase-shifter performance are described. The optimal bandwidth design of a reflection-type phase shifter using an ideal two-state capacitor is reviewed. A phase-stretching technique is introduced for the design of large phase shift bits (90°, 180°) using low ratio ($C_H/C_L = 2$) capacitors. The design, fabrication, and test results of a 4-bit MIS phase shifter are presented. The power-handling capability and switching-speed characteristics of this type of phase shifter are described.

Introduction

Phase shifters using p-i-n diodes have been available for years and have reached a high state of development. Ince[1] has provided an excellent summary of the state of the art of phase shifters. The main disadvantage of p-i-n diodes is that heavy current is required with forward bias in order to achieve a short circuit. This is particularly objectionable in phased-array applications, where a computer must control many microwave modules and drivers are required to interface between the computer and phase shifters. The additional drivers add to cost and complexity of the system. One solution is the use of

* RCA Laboratories, Princeton, N.J.
† RCA Government Communications and Automated Systems Div., Camden, N.J.
‡ City University of New York, N.Y., N.Y.

reverse-bias varactor diodes in a reflection-type phase shifter;[2] although this approach reduces the current requirements, conventional varactors suffer from power-handling limitations and sensitivity to power-supply variations. Burns et al[3] discuss the use of a field-effect diode (FED) to eliminate forward conduction and to minimize phase sensitivity due to power-supply variations. They described capacitance versus voltage characteristics of the FED and have reported small-signal characteristics of phase-shift bits using this device. The present paper discusses the design and performance of a 4-bit microwave digital phase shifter using metal–insulator–semiconductor (MIS) varactors. Physical mechanisms and device design relationships are described. The optimal bandwidth design of a reflection-type phase shifter using an ideal two-state capacitor is reviewed. A phase-stretching technique is introduced for the design of large phase shift bits (90°, 180°) using low ratio (C_H/C_L = 2) capacitors. The large-signal performance of the phase shifter is described.

MIS Device Description

Fig. 1 shows the cross section of a structure consisting of metal, insulator, epitaxial semiconductor, and metal (M-i-n-n$^+$-M). When re-

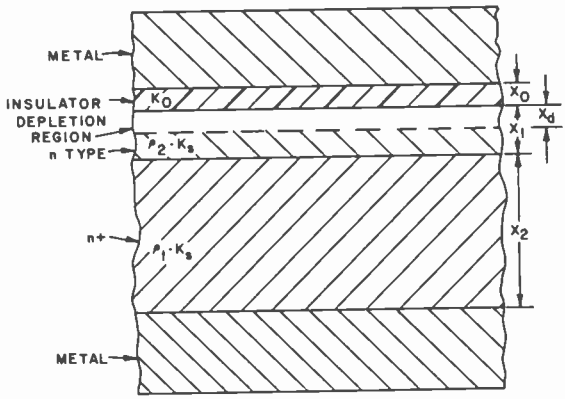


Fig. 1—Metal-insulator-semiconductor one-dimensional cross section.

verse biased, the surface of the semiconductor at the semiconductor–insulator interface is depleted of carriers, forming a series connection of fixed insulator capacitance and a variable depletion capacitance. In series with the two capacitors is a resistance that corresponds to the resistance from the edge of the depletion region to the substrate. As the reverse bias voltage is increased, the depletion layer pene-

trates deeper into the epitaxial layer until a maximum (X_{dm}) is reached. This maximum depth is limited by the formation of an inversion layer (holes → n/n+, electrons → p/p+) at the semiconductor–insulator interface. The maximum depletion depth is dependent on the resistivity of the epitaxial layer and the semiconductor used. If the inversion layer is prevented from forming, there is no limit on depletion depth.

Fig. 2 is an approximate equivalent circuit of an MIS structure operated as a surface depletion device. For this discussion, the insulator

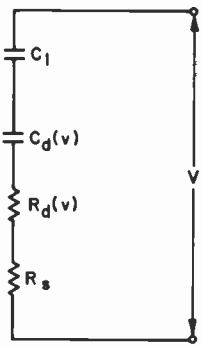


Fig. 2—MIS approximate equivalent circuit: C_1 = insulator capacitance, $C_d(v)$ = depletion capacitance, $R_d(v)$ = depletion resistance, and R_s = substrate resistance.

is assumed perfect and is characterized by a relative dielectric constant K_o. The insulator capacitance is

$$C_1 = K_0\epsilon_0\frac{A}{X_0}, \qquad [1]$$

where ϵ_o is the permittivity of free space (8.86×10^{-6} pF/μm), A is the cross sectional area (μm^2), and X_o is the insulator thickness (μm). The epitaxial semiconductor layer of thickness X_1 (Figure 1) may be characterized by the resistivity of the material ρ_1, which is a function of doping density, and the relative dielectric constant of the semiconductor (K_s). Therefore, the depletion capacitance and resistance may be described by

$$C_d(v) = \frac{K_s\epsilon_0 A}{X_d(v)}, \qquad [2]$$

$$R_d(v) = \rho_1\frac{X_1 - X_d(v)}{A}. \qquad [3]$$

The substrate resistance is dependent upon the resistivity of the substrate (ρ_2) and any geometric spreading of the fields.

$$R_s = F\rho_2 \frac{X_2}{A}, \qquad [4]$$

where F is a geometric factor and $0 < F \leq 1$.

In the surface depletion mode, an important parameter is the maximum depletion depth X_{dm}. As the semiconductor is continuously reverse biased, a voltage is reached at which further depletion of the semiconductor is not possible and an inversion layer forms at the semiconductor–insulator interface. The maximum depletion depth for an n-type semiconductor has been shown by Grove and other to be approximately

$$X_{dm} = \left(\frac{4K_s\epsilon_0\phi_F}{qN_D}\right)^{1/2}, \qquad [5]$$

$$|\phi|_F = \frac{kT}{q}\ln\frac{N_D}{n_i}, \qquad [6]$$

where k is Boltzmann's constant, T is temperature in °K, q is magnitude of electronic charge, N_D is donor density, and n_i is the intrinsic carrier concentration.

For a given semiconductor material, the relationship between maximum depletion depth and the resistivity of the epitaxial layer is given by the empirical relationship

$$X_{dm} = \alpha\rho^{\beta}, \qquad [7]$$

where α,β depend upon the material and the temperature. At room temperature, $\alpha = 0.44$ μm and $\beta = 0.465$ for silicon, and $\alpha = 1.3$ μm and $\beta = 0.477$ for GaAs.

The structure shown in Fig. 1 may be operated in a variable-capacitance mode by using a relatively low-resistivity ($\rho < 1$ ohm-cm) thin epitaxial layer and a thin oxide. The capacitance-versus-voltage characteristic is sketched in Fig. 3. As is shown by this figure, the high-frequency total capacitance versus voltage has a unique two-stage nature where, in the two separate regions, capacitance does not vary with applied voltage. The ratio of the high-state to low-state capacitance (C_H/C_L) is given by

$$\frac{C_H}{C_L} = 1 + \frac{K_0}{K_s}\frac{X_{dm}}{X_0}. \qquad [8]$$

Note that, if the epitaxial layer is not thin enough to fully deplete the

semiconductor under reverse bias, under large-signal conditions the capacitance will not follow the small-signal capacitance curve but will follow the pumped depletion-mode curve.

The nonlinear resistance exhibits behavior similar to the capacitance characteristics.

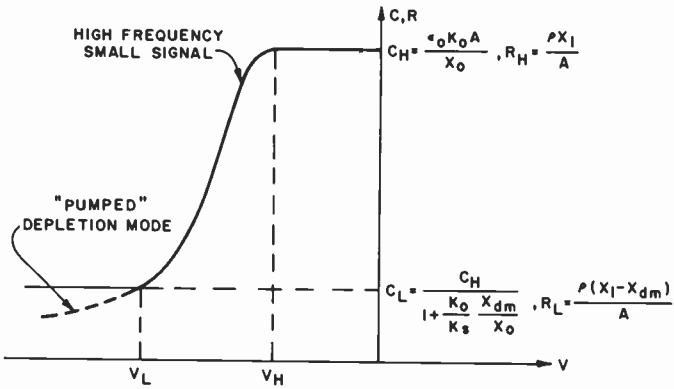


Fig. 3—Capacitance-voltage characteristics and resistance voltage characteristics of MIS devices.

MIS Varactors in Phase-Shifter Applications

To overcome the high dc current requirements of p-i-n diode phase shifters, reverse-biased junction varactors have been used. The voltage-variable capacitance becomes a voltage-variable reflection coefficient angle when terminating a transmission line. In a reflection-type phase shifter, this is translated into a voltage-variable transmission coefficient angle. This is discussed more fully later. At this point, we can see that by changing the bias, we change the phase shift and, because the device is reverse biased at all times, the steady-state current required is very small—on the order of microamperes or nanoamperes.

This approach suffers limitations, however, due to the shape of the C-V curve of a standard varactor. In Fig. 4, the varactor is biased between voltages 1 and 2, giving capacitances C_1 and C_2, which cause the phase to shift. At voltage 1, the capacitance is approximately constant with voltage. However, at voltage 2, the capacitance changes relatively rapidly with voltage. If the rf voltage swing is large, therefore, a nonlinear waveform is generated with high harmonics. Moreover, as the nonlinear portion of the curve is approached, the average

capacitance at voltage 1 changes, causing the phase to change. This phase shift is not constant with power level. Another problem is that power-supply drift or ripple will cause phase error. This is sometimes turned into an advantage, since it enables phase trimming through supply-level adjustment. However, this is not usually a desirable way to trim.

Fig. 4 shows how the MIS device overcomes these disadvantages. The *C-V* curve is approximately flat in both bias regions, which results in great reduction in the sensitivity to power-level and bias variations. That the device is forward biased at voltage 2 is not a disadvantage, since the insulating layer prevents any dc current from flowing.

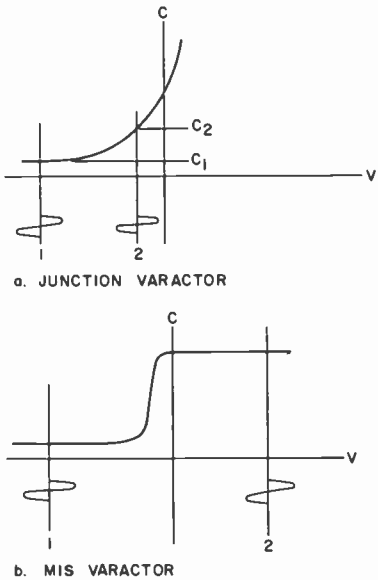


Fig. 4—C-V characteristics of junction and MIS varactors.

Design and Fabrication of MIS Phase Shifter

Loaded-Line Phase Shifter

Two types of phase shifters were considered for use with MIS devices—the loaded-line phase shifter and the reflection-type phase shifter.

The loaded-line phase shifter is attractive because of its simplicity and relatively low loss. It suffers from narrow bandwidth and from

the fact that higher phase shifts cannot be easily produced. The latter drawback is significant, because it means that two 45° bits would be needed for a 90° phase shift and four for a 180° phase shift. Such a large number of sections would negate the advantages of lower loss; hence, the loaded-line design was considered only for the 22.5° and 45° bits.

Reflection-Type Phase Shifter

The reflection-type phase shifter uses a circulator or a quadrature hybrid to convert a reflection phase shift to a transmission phase shift. The hybrid or circulator represents additional loss and complexity in comparison with the loaded-line phase-shifter. However, there is no requirement that the diode impedance be zero or infinite

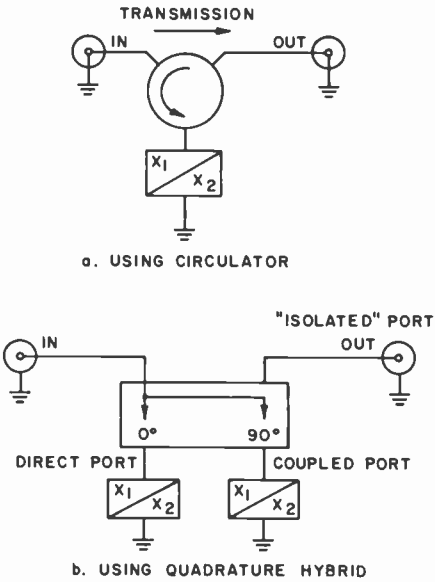


Fig. 5—Reflection type phase shifters.

in either state (except for the 180° bit), so this approach is quite suited to the finite capacitance ratio of the MIS device. The two types of reflection phase shifter are shown in Fig. 5. When the terminating device or network switches from reactance X_1 to reactance X_2, the reflection coefficient phase angle shifts from ϕ_1 to ϕ_2, for a phase shift of ϕ_2-ϕ_1. If one port of a circulator is terminated in such a device (Fig. 5a), the transmission coefficient between the remaining two ports in

the direction shown is equal to the reflection coefficient of the terminating device. Hence, a transmission phase shift of ϕ_2-ϕ_1 is produced. A 3 dB quadrature hybrid can be used, as shown in Fig. 5b, with identical terminations on the direct and coupled ports. The transmission is between the input and normally "isolated" ports. If the terminations are pure reactances (no conductive component), the only loss is that of the hybrid or circulator.

The hybrid approach was chosen for this application because of its smaller size and weight. It was previously mentioned that the 180° bit still presents a problem with the reflection-type phase shifter approach, because an infinite capacitance ratio would be required in order to switch between a short circuit and an open circuit for a 180° phase shift. One solution is the use of two cascaded 90° bits, but this causes excessive loss. An alternative is to imbed the MIS diode in an appropriate network so that the overall phase shift is 180°. A similar technique can be used for the 90° bit if insufficient capacitance ratios are available.

Optimum C_H and C_L for Lower Order Bits (22.5°, 45°)[6]

The reflection coefficient for a capacitor terminating a transmission line is

$$\Gamma = \frac{\frac{1}{j\omega c} - Z_0}{\frac{1}{j\omega c} + Z_0}. \qquad [9]$$

The phase of the reflection coefficient may be expressed as

$$\theta = 2\tan^{-1}(-\omega c Z_0). \qquad [10]$$

If the capacitor may be switched between two different values (c_1,c_2), the net differential phase shift is

$$\Delta\theta = \theta_2 - \theta_1 = 2\tan^{-1}\left[\frac{\omega c_2{}^2 Z_0 - \omega c_1 Z_0}{1 + \omega^2 c_1 c_2 Z_0{}^2}\right] \qquad [11]$$

In general it is desired to make the phase variation as small as possible across a given frequency band. This may be done by choosing parameters such that the differential phase shift slope is zero at the design center frequency. This slope will be zero at a center frequency ω_0 if

$$c_1 c_2 = \frac{1}{\omega_0^2 Z_0^2}. \qquad [12]$$

Combining Eqs. [11] and [12],

$$C_1 = \frac{1}{\omega_0 Z_0}\left[\sec\frac{\Delta\theta}{2} - \tan\frac{\Delta\theta}{2}\right] \qquad [13]$$

$$C_2 = \frac{1}{\omega_0 Z_0}\left[\frac{1}{\sec\frac{\Delta\theta}{2} - \tan\frac{\Delta\theta}{2}}\right] \qquad [14]$$

$$\frac{C_2}{C_1} = \frac{1}{\left(\sec\frac{\Delta\theta}{2} - \tan\frac{\Delta\theta}{2}\right)^2} \qquad [15]$$

Table 1 summarizes capacitance ratios and values at a design center frequency of 3.5 GHz, in a 50-ohm system.

Table 1—Summary of Capacitance Ratios and Values (f_0 = 3.5 GHz and Z_0 = 50 Ohms)

$\Delta\theta$	c_2/c_1	c_1(pF)	c_2(pF)
22.5°	1.48	0.746	1.11
45.0°	2.24	0.608	1.36
90.0°	5.85	0.376	2.20
180.0°	∞	0	∞

It should be noted that the values of capacitance obtained have a graphical significance on a reflection chart. The optimal reactance values (X_1,X_2) for a given phase shift $\Delta\theta$ fall symmetrically about the −90° phase coordinate. For two small capacitance values near the open-circuit point, the reactance curves tend to bunch together yielding a small phase shift. For two large capacitors with reactances near the short-circuit point, the differential phase shift is small, this time because the change in reactance is small. The optimum, therefore, lies halfway between the open-circuit and short-circuit point.

To determine the useful bandwidth of the phase shifter, one would plot $\Delta\theta$ as a function of frequency and determine phase flatness over the band of interest. In our case, however, a circuit analysis program was available that allows circuit simulation of the MIS varactor, including chip, parasitics, and external tuning elements.

Phase-Stretching Network and Design of 90° and 180° Bits

It has been shown that the 180° bit cannot be realized with only a nonlinear capacitor. We will show how one can increase the phase shift capability of a two-state capacitor using two lumped inductors. The price paid for the increased phase shift is phase variation with frequency. However, with a number of iterations on the circuit elements, phase variations can be minimized. It should be stressed that the approach presented here is not a direct synthesis and, therefore, optimal solutions cannot be guaranteed.

The principle of the technique used is illustrated in Fig. 6. An ideal two-state capacitor can be thought of as an ideal switch embedded in

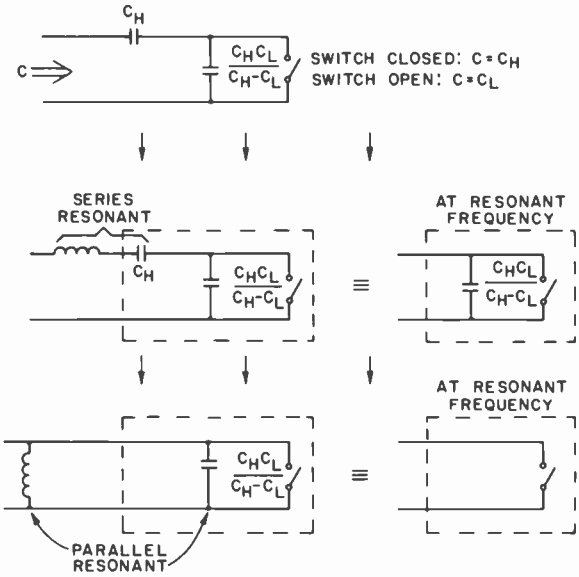


Fig. 6—Method of approximating an ideal switch with an MIS diode.

a network of two capacitors of the values shown, so that the equivalent capacitance would vary between C_H and C_L when switched. By resonating out these capacitors at a given frequency, one is left with an ideal switch at this one frequency. By proper choice of the starting capacitor values and the two inductors, one can optimize the bandwidth over which ideal switch performance is approximated. For broad bandwidth, the capacitance ratio should be as high as possible.

It is also useful to view the problem on a Smith chart as shown in Fig. 7. A diode is assumed whose reactances initially fall at points A

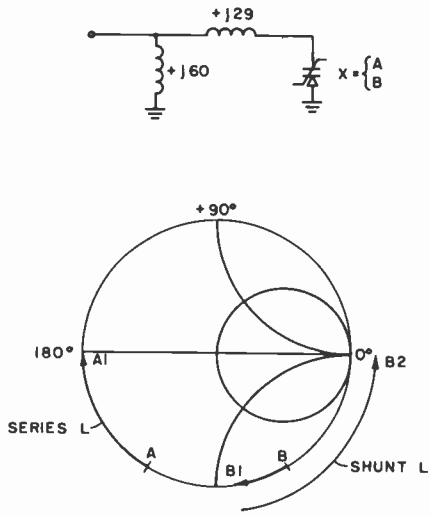


Fig. 7—Reflection coefficient plot of ideal-switch approximation using an MIS diode.

and B, which gives a phase shift of 60°. If a series inductor whose reactance is 29 ohms is added, A is shifted to A1, the short circuit point, while B moves to B1. The phase shift is now 100°. A shunt inductance sufficient to move B to B1 is now added. A1 will not move with the addition of a shunt element, since anything in shunt with a short

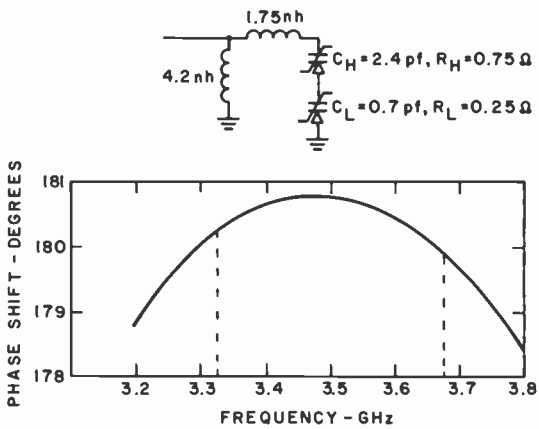


Fig. 8—Computed response of 180-degree bit assuming no package strays.

is still a short. Hence, the overall reactance is a short circuit in one state and an open circuit in the other state, and the phase difference is 180°.

Attention is next turned to achieving 180° with the devices on hand. Some typical measured C_H and C_L values for one of the MIS types fabricated are C_H = 2.4 pF and C_L = 0.7 pF. For broadest bandwidth, it was found that two of these devices should be connected in series. Such a configuration has the additional advantage of higher breakdown voltage and resulting higher power handling capability. The required inductance values and the calculated performance are shown in Fig. 8. Note that the phase error can easily be kept within ±1° over the required band for this ideal case.

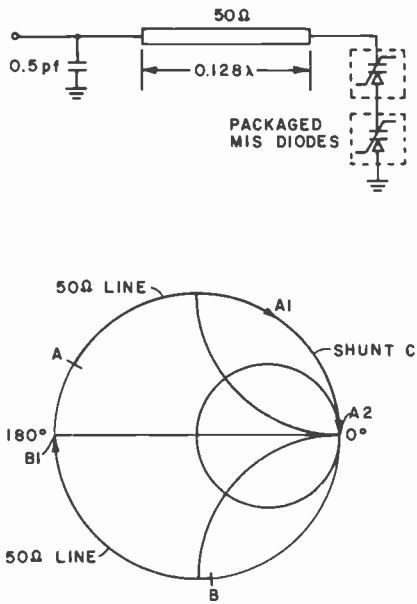


Fig. 9—Alternate method of designing 180-degree bit.

While the performance using "ideal" two-state capacitors would be very good, the problem is complicated by stray reactances when the device is mounted in a package. In fact, package strays change the diode characteristics so much that the design of Fig. 8 can no longer be used, and an alternative design must be found.

An alternative approach for the 180° bit is shown in Fig. 9. The series connection of the two packaged diodes switches between the

reactances spanned by bracket A. These two points are 123° apart. A section of 50-ohm line, in this case 0.128 λ, can be used to shift these points clockwise to bracket B while preserving the phase relationship. From here a shunt capacitance of 0.5 pF moves the rightmost point down to the 0° (or open-circuit) point. The leftmost point, at the short-circuit point on the chart, is unaffected by the shunt capacitance.

This method of designing a 180° bit has less bandwidth than the method used with ideal capacitors, but the phase error can still be kept within the design goal. In practice, it was found easiest to design by trial and error. The diodes were mounted at the end of a 50-ohm line and a capacitive stub of indium foil was varied in size and position along the line.

Since the diodes alone in Fig. 9 give a phase shift of 123°, there is more than enough phase shift for the 90° bit. The phase shift of a series pair of diodes was reduced to 90° by trimming.

Diode Loss

The transmission coefficient of a reflection phase shift bit is equal to the reflection coefficient of the devices terminating the hybrid, and reflection loss in the diode shows up as transmission loss in the bit.

An MIS diode can be seen as a capacitance C in series with a small resistance R_S, both of which change when the device switches state. The reflection loss in either state is calculated as follows:

$$Z = R_s - j\frac{1}{\omega c}$$

$$\rho = \frac{Z_0 - Z_0}{Z + Z_0} = \frac{R_s - j\frac{1}{\omega c} - Z_0}{R_s - j\frac{1}{\omega c} + Z_0}$$

$$= \frac{R_s^2 - Z_0^2 + \frac{1}{\omega^2 c^2} - j\frac{2Z_0}{\omega c}}{(R_s + Z_0)^2 + \frac{1}{\omega^2 c^2}} \qquad [16]$$

$$|\rho| = \left[\frac{\left[R_s^2 - Z_0^2 + \frac{1}{\omega^2 c^2}\right]^2 + \frac{4Z_0^2}{\omega^2 c^2}}{\left[(R_s + Z_0)^2 + \frac{1}{\omega^2 c^2}\right]^2}\right]^{1/2}$$

$$\text{Loss} = 20 \ \text{Log}_{10}|\rho|$$

Fig. 10 shows the effect of R_s on diode loss for a 45° bit using the

ideal capacitance values, C_H = 1.36 pF and C_L = 0.61 pF. For simplicity, R_s is assumed to be the same in both diode states. The loss, however, is greater in the high-capacitance state, so this is the loss shown on the plot. Since the loss calculated is diode loss only, the loss of the hybrid and other external circuitry must be added. If 0.4 dB is assumed for all non-diode losses, then the loss of 45° bit using ideal capacitance diodes of 1 ohm series resistance is 0.4+0.24 = 0.64 dB.

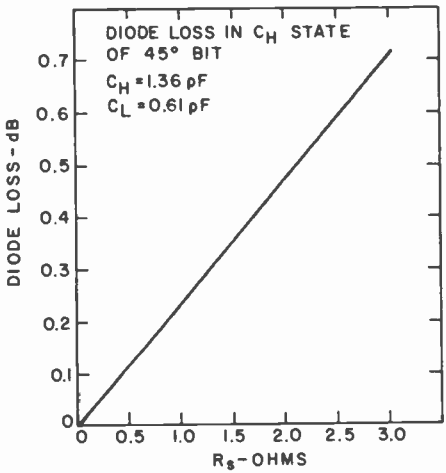


Fig. 10—Effect of R_s on diode loss.

Power-Handling Capability

In calculating the theoretical power-handling capability of MIS diodes, the power into the hybrid is assumed to split exactly in half and travel to the diodes. Each port (direct and coupled) of the hybrid is then thought of as a generator of impedance R_o and voltage $[2P_{IN}R_o]^{1/2}$, where R_o is the characteristic impedance of the system and P_{IN} is the power into the hybrid. The voltage across the capacitor is found by simple network analysis;

$$V_c = V\left[\frac{\dfrac{1}{\omega^4 c^4} + \dfrac{R_0^2}{\omega^2 c^2}}{\left(R_0^2 + \dfrac{1}{\omega^2 c^2}\right)^2}\right]^{1/2} \qquad [17]$$

The maximum rf voltage swing allowed across the diode is determined by examining the nonlinear portion of the C-V curve. The av-

erage capacitance will change in such a direction as to decrease the phase shift if this region is entered. Furthermore, the nonlinearity will generate harmonics. Thus the knee of the curve represents one limit on voltage swing; the other limit is the breakdown voltage in either state. If the bias point is halfway between these limits, then the allowed voltage swing is maximized. The theoretical power-handling limit of a 45° bit using MIS diodes has been calculated to be 2.9 watts.

Diode Design and Fabrication

As mentioned previously, there is a trade-off between diode series resistance and capacitance ratio in the selection of epitaxial resistivity. For this application, a resistivity of 0.36 ohm-cm (on a 0.005 ohm-cm n^+ substrate) was chosen as a compromise between loss and large phase-shift and because of its ready availability. The maximum depletion depth for this material is 0.27 μm. For minimum loss, the epitaxial layer should ideally be no thicker than this. However, such small thicknesses are not easy to achieve and 0.6 μm was used because of availability.

Silicon dioxide was thermally grown in dry oxygen at 1000°C to thicknesses of 600 Å, 1000 Å, and 1200 Å in order to provide a range of breakdown voltages and capacitance ratios. In later devices, annealing the oxides at 900°C in nitrogen provided some improvement in breakdown voltage. On top of the oxide, 11,000 Å of hyperpure aluminum was vacuum deposited and then photoetched to provide arrays of dots. Two dot sizes were used—1 mil diameter and 2 mil diameter. The n^+ side was etched away to 3–4 mils to reduce loss caused by series resistance. Then the wafers were scribed and diced into individual diodes.

The chips were mounted in the packages by heating to 450°C in a nitrogen atmosphere. Connection from the diode dot to the package was accomplished by ultrasonic wire bonding.

Experimental Results

A network was used to measure loss, phase shift, and VSWR. The data are all taken at 100 mW power level. Bias levels of +15 V and −20 V were used. The measured data for each of the four bits is plotted in Figs. 11 through 14. Note that the phase error for each bit is 4° or less except for the 180° bit where it is 9°. The maximum cumulative phase error for four cascaded bits would be about 10° peak at

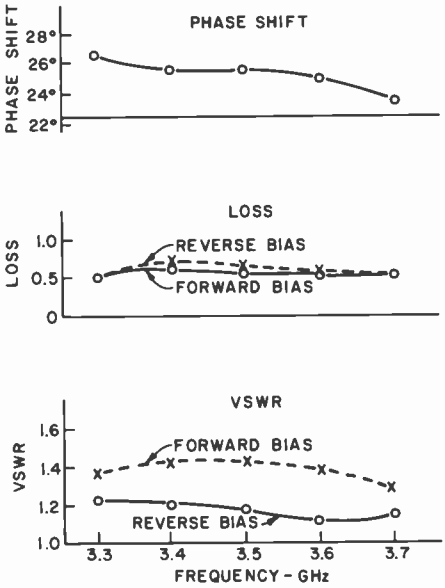


Fig. 11—22.5-degree-bit performance.

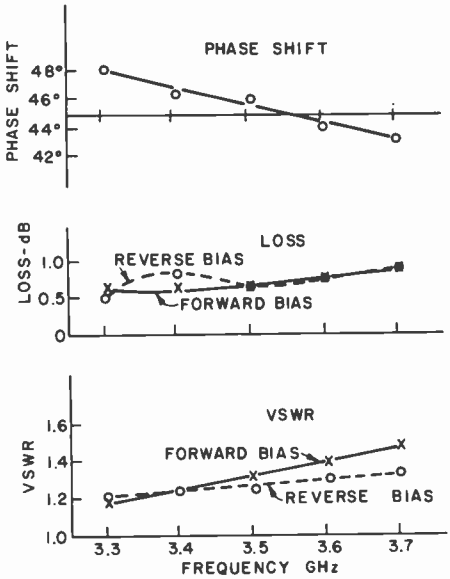


Fig. 12—45-degree-bit performance.

3.55 GHz if the effects of VSWR were ignored. The maximum loss can be estimated as about 4 dB at 3.7 GHz, again ignoring the effects of mismatch between bits.

Since the computer calculations showed that higher loss occurs in the forward-bias state, the higher loss at reverse bias in most bits was surprising. Since it was usually accompanied by higher reverse-bias

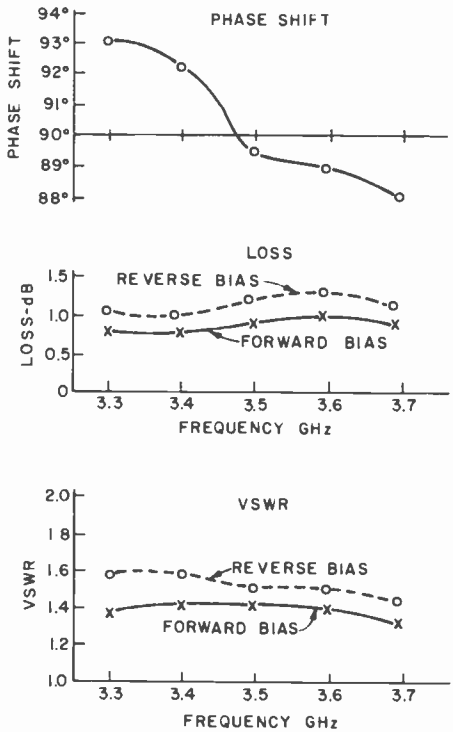


Fig. 13—90-degree-bit performance.

VSWR, the cause is most likely mismatch. The diodes in each pair may have tracked less well at reverse bias, or the hybrids may have had poorer performance as the impedances terminating them moved toward the open circuit point, as they would in reverse bias. Measurements made on the interdigitated hybrids tend to confirm this.

One of the most important parameters is the phase-shifter switching speed, i.e., the time required for the phase to be varied from one state to another. Typical bias-waveform and phase-shift versus time curves are shown in Fig. 15. The rise time (T_r) is of the order of nanoseconds. This settling time depends upon the microwave power level. There is such a thing as a settling time in an MIS varactor, because of

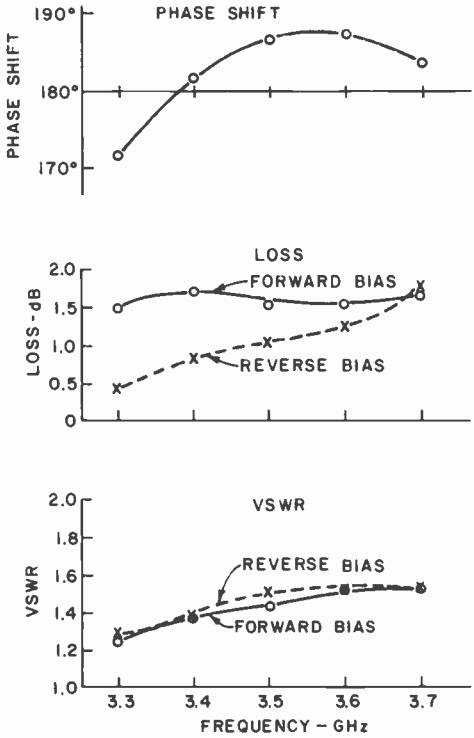


Fig. 14—180-degree-bit performance.

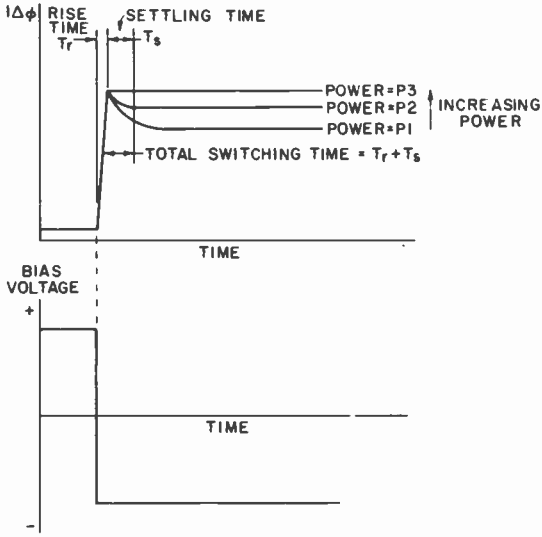


Fig. 15—Instantaneous phase shift and bias waveforms.

the formation of an inversion layer. When the bias voltage is switched suddenly, the surface depletes to a depth X_d in time T_r. But then the surface starts to invert; when this occurs the depletion depth is limited to a maximum value $X_{dm} < X_d$. The time that it takes to go from X_d to X_{dm} is the settling time T_S. Fig. 16 shows switching speed ver-

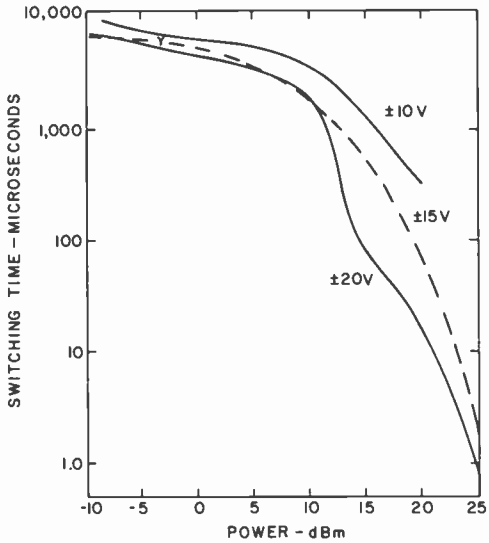


Fig. 16—Switching time versus power into phase shifter.

sus applied power level for various bias voltages. For small signals, the switching speed is of the order of milliseconds. For +25 dBm the speed is of the order of 0.6 μsec. In general, the larger the bias voltage the faster the switching speed.

The individual bits were cascaded to produce a four-bit 360° phase shifter with the following overall performance:

Frequency Range	3.325 to 3.675 GHz
Maximum Insertion Loss	5.2 dB
Maximum Phase Error	14.9° peak, 9.61° rms
Maximum Phase Deviation for ±10% Bias Voltage Change	+1.8° to −2.2°
Bias Voltage	+12.5 to −17.5 V
Maximum Bias Current	1.7 μA
Maximum Bias Power	33 μW
Power Shifter	100 mW
Maximum Power Capability	500 mW

Conclusions

A comparison of design goals and measured results is as follows:

	Design	Measured
Frequency	3.5 GHz	3.5 GHz
Bandwidth	10%	10%
Phase Shift	0°–360°	0°–360°
Insertion Loss	3 dB Max.	5.2 dB
Phase Error	10° rms Max.	9.6° rms
Diode Forward Bias	0 mA	0 mA
Phase sensitivity for ±10% Bias Voltage Variation	±5° Max.	+1.8°, −2.2°
Power	20 W cw	100 mW cw
Switching Time	1 μsec	50 nsec

All the design goals were met except two—insertion loss and power-handling capability. Such results are quite encouraging, since this is a first attempt coming at a very early stage in the development of MIS diodes. Further developments efforts will certainly result in considerable improvement.

References:

[1] W. J. Ince, "Recent Advances in Diode and Ferrite Phase Technology for Phased-Array Radars, Part 1," *Microwave J.*, **15**, No. 9, p. 36, Sept. 1972; Part II: **15**, No. 10, p. 31, Oct. 1972.

[2] J. F. White, "Review of Semiconductor Microwave Phase Shifters," *Proc. IEEE*, **56**, p. 1924, Nov. 1968.

[3] Richard W. Burns, Russel L. Holden, and Raymond Lang, "Low Cost Design Techniques for Semiconductor Phase Shifters," *IEEE Trans. Microwave Theory and Tech.*, **MTT-22**, No. 6, p. 675, June 1974.

[4] J. F. White, "High Power, PIN Diode Controlled, Microwave Transmission Phase-Shifters," *IEEE Trans. Microwave Theory and Tech.*, **MTT-13**, p. 233, March 1965.

[5] F. L. Opp and W. F. Hoffman, "Design of Digital Loaded-Line Phase-Shift Networks for Microwave Thin-Film Applications," *IEEE Trans. Microwave Theory and Tech.*, **MTT-16**, No. 7, p. 462, July 1968.

The Conversion Efficiency of Ideal Shockley P-N Junction Photovoltaic Converters in Concentrated Sunlight

F. Sterzer

RCA Laboratories, Princeton, New Jersey

Abstract—The conversion efficiency of ideal Shockley p-n junction photovoltaic converters is calculated as a function of sun concentration ratio. It is shown that the conversion efficiency is significantly larger in highly concentrated sunlight than in unconcentrated sunlight.

Introduction

Solar energy arrives at the earth at the enormous rate of about 1.5×10^{18} kW-hours per year, but the power densities associated with this energy are relatively low.[1] Outside the earth atmosphere the solar irradiance averages only 1353 watts/m^2, and on a sunny day the earth surface receives an irradiance of only about 1 kW/m^2. By comparison, the power densities in high-power microwave transistors often reach 10^5 kW/m^2, i.e., a factor of 100,000 larger than the maximum terrestrial solar irradiance. Since modern semiconductors can handle orders of magnitude greater power densities than the solar irradiance, it is possible to design p-n junction photovoltaic energy converters that can operate with highly concentrated sunlight.[2,3] The use of highly concentrated sunlight sharply reduces the area of the p-n junction converter required for a given amount of electrical output, and therefore represents one of the most promising approaches to low-cost generation of electricity from sunlight.[4]

The fundamental limits of the efficiency with which unconcentrated sunlight can be converted into electrical energy in p-n junction photovoltaic converters have been explored by several authors,[5–7] but little has been published so far on what these limits might be for photovoltaic converters operating with concentrated sunlight. In the present paper upper limits of efficiency are calculated as a function of sun-concentration ratio for p-n junction photovoltaic converters that obey the Shockley diode equation. It is shown that for ideal Shockley p-n junction converters, the conversion efficiency is significantly larger in highly concentrated sunlight than in unconcentrated sunlight.

Voltage Factor, Curve Factor and Characteristic Factor

Fig. 1 shows the I-V characteristics of a p-n junction photovoltaic diode that obeys Shockley's equation. When the diode is in the dark,

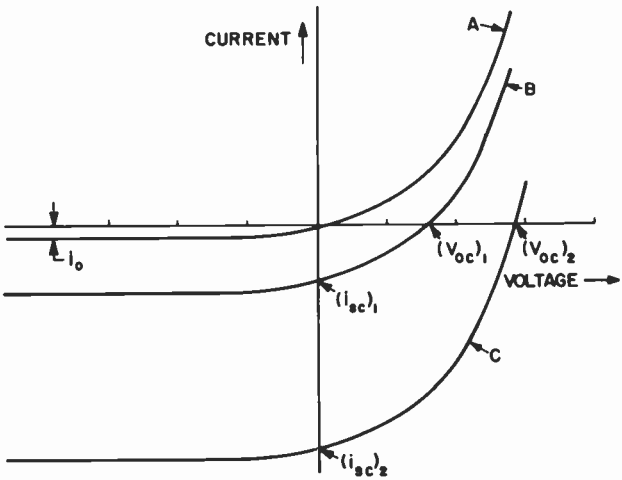


Fig. 1—Current-voltage characteristics of a p-n junction photovoltaic diode in the dark (curve A), and under two different levels of illumination (curves B and C).

the I-V characteristics is as shown in curve A of Fig. 1 and can be written

$$\begin{aligned} i_j &= i_o(e^{\Lambda V} - 1) \\ &= q(N_c N_v \exp\{-E_g/kT\})^2 \left[\frac{1}{N_A}\left(\frac{D_n}{\tau_n}\right)^{1/2} + \frac{1}{N_D}\left(\frac{D_p}{\tau_p}\right)^{1/2}\right] \end{aligned} \quad [1]$$

where i_j is the junction current, i_o is the reverse saturation current, Λ

$= q/kT$, (q = electronic charge, k = Boltzmann's constant, T = absolute temperature), V is the voltage across the junction, N_c *and* N_v are respectively the effective densities of states in the conduction and valence bands, E_g is the bandgap energy, N_A and N_D are, respectively, the acceptor and donor densities, D_n and D_p are respectively the diffusion constants of minority electrons and holes, and τ_n the τ_p are respectively the minority carrier lifetimes of electrons and holes. When the diode is illuminated, a photocurrent i_{sc} is generated in the junction, and the I-V characteristic is shifted as shown in curves B and C of Fig. 1. The current flowing through a load resistor placed across an illuminated diode is given by

$$i_L = i_o(e^{\Lambda V} - 1) - i_{sc} \qquad [2]$$

If the load across the diode is a short circuit, a current i_{sc} will flow through the load. If the load is an open circuit, a voltage

$$V_{oc} = \frac{\ln\left(\frac{i_{sc}}{i_o} + 1\right)}{\Lambda} \qquad [3]$$

will be developed across the diode.

The photocurrent generated in an ideal Shockley photodiode can be written as

$$i_{sc} = N_g q \qquad [4]$$

where N_g is the incident flux of all photons with energies equal to or greater than the bandgap energy E_g of the semiconductor from which the diode is fabricated. Eq. [4] assumes that every incident photon with energy equal to or greater than E_g produces one electron–hole pair in the diode, and that all electron–hole pairs that are generated are collected in the junction of the diode. The minimum incident power required to produce a short circuit current i_{sc} is given by

$$(P_{in})_{min} = N_g E_g = N_g q V_g = i_{sc} V_g, \qquad [5]$$

where V_g is the bandgap voltage.

The power delivered by an illuminated Shockley photodiode to a resistive load placed across it is from Eq. [1].

$$P_L = V i_L = V[(i_o e^{\Lambda V} - 1) - i_{sc}] \qquad [6]$$

Using elementary differential calculus, one can show that P_L will be a

maximum if the load resistance assumes the following value

$$R_{Lmp} = \frac{V_{mp}}{i_{Lmp}} = \frac{e^{-\Lambda V_{mp}}}{i_o \Lambda}, \qquad [7]$$

where the subscript mp denotes the condition for maximum power transfer. Using this notation, the maximum power that the diode can deliver to the load can be written

$$P_{Lmp} = i_{Lmp} V_{mp}. \qquad [8]$$

The conversion efficiency of a photovoltaic diode for the case where the incident radiation is at frequency $\nu = E_g/h$ (h = Planck's constant) is called the characteristic factor of the diode. Using Eqs. [5] and [8], we can write the following expression for the characteristic factor

$$\frac{P_{Lmp}}{(P_{in})_{min}} = \text{CH.F.} = \frac{i_{Lmp} V_{mp}}{i_{sc} V_g} = \left(\frac{V_{oc}}{V_g}\right)\left(\frac{i_{Lmp} V_{mp}}{i_{sc} V_{oc}}\right) \qquad [9]$$

The term (V_{oc}/V_g) in Eq. [9] is a measure of how closely the open circuit voltage of the diode approaches the bandgap voltage and is referred to as the voltage factor. The term $(i_{Lmp}\ V_{mp})/(i_{sc}\ V_{oc})$ in Eq. [9] is a measure of how closely the third quadrant I-V characteristics of the diode approaches the ideal rectangle defined by i_{sc} and V_{oc}. This term is referred to as the curve factor of the diode.

In an ideal Shockley photodiode, all three factors increase with increasing levels of illumination. Fig. 2 shows the normalized I-V characteristics of a Shockley photodiode for ratios of short circuit to reverse saturation current of 10^{11} and 10^{14}. The ratio of 10^{11} is typical for a silicon diode under air mass zero illumination, while the ratio of 10^{14} corresponds to 1000 times concentration of air mass zero illumination. Note that the 1000 times concentration leads to a significant increase in open circuit voltage, and also to a more nearly rectangular I-V characteristics. The quantitative dependence of the voltage, curve, and characteristic factors on the ratio i_{sc}/i_o are plotted in Fig. 3 for ideal Shockley diodes.

The dependence of the I-V characteristics of actual photovoltaic diodes on the level of illumination deviates, of course, from the ideal behavior assumed in Figs. 2 and 3. Still, there is good agreement between the voltage factor of Fig. 3 and the experimentally measured voltage factor of n- and p-type silicon photodiodes for sun concentra-

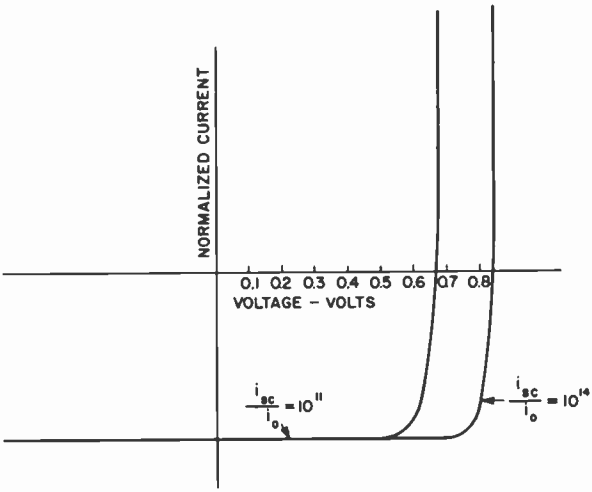


Fig. 2—Normalized current-voltage characteristics of an illuminated ideal Shockley photovoltaic diode for two values of the ratio of short-circuit current to reverse saturation current.

tion ratios of up to[8] 1000, the actual measured open-circuit voltages of p-type diodes under 1000-sun illumination being significantly larger than the maximum saturated photovoltage for p-type silicon diodes that had been predicted earlier.[9] The measured curve factors are significantly smaller than the values shown in Fig. 3, a large part of the discrepancy being due to the parasitic series resistance of present diodes.

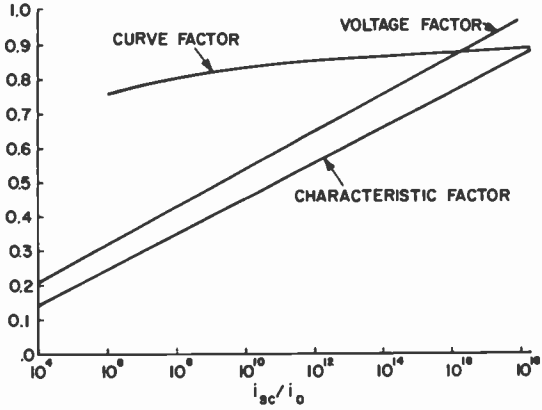


Fig. 3—Curve factor, voltage factor, and characteristics factor as a function of the ratio of short-circuit current to reverse saturation current for an ideal Shockley photovoltaic diode.

Efficiency Limitations Caused by the Spectral Distribution of the Solar Irradiance

The efficiency of photovoltaic converters is significantly degraded because sunlight is not monochromatic, but rather contains photons with a wide range of energies.[6] Photons with energies smaller than

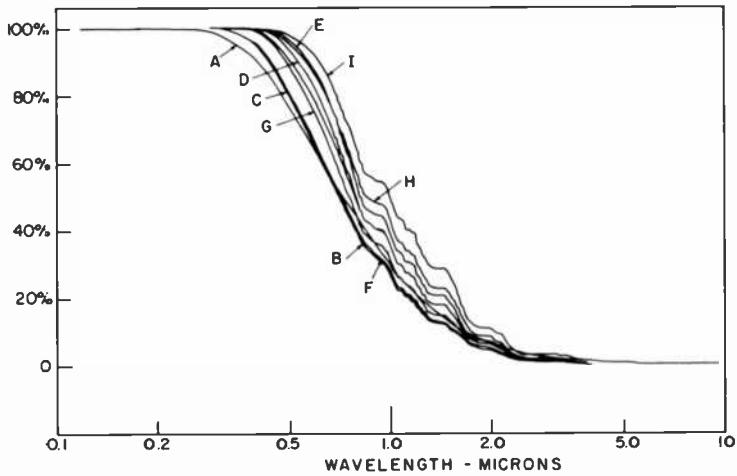


Fig. 4—Percent of solar irradiance with wavelength longer than λ for several values of air mass and atmospheric pollution. The wavelength exponent is 0.66; H_2O = 20 mm, O_3 = 3.4 mm. Values of air mass and β (the Angstrom β coefficient) for the various curves, which were calculated from data given in Ref. [1], are as follows:

Curve	Air Mass	β	Curve	Air Mass	β
A	0		F	7	
B	1	0.085	G	7	0 170
C	1	0 170	H	10	0 085
D	4	0.085	I	10	0.170
E	4	0 170			

the bandgap energy of the photodiode do not contribute to the conversion process, because their energies are insufficient to generate electron–hole pairs in the diode. Photons with energies greater than the bandgap energy of the diode can generate electron–hole pairs, but the fraction of their energies greater than the bandgap energy is not used in the conversion process.

Figs. 4 through 6 are plots that allow one to compute the percent of unusable solar irradiance for a diode with a given bandgap energy. Fig. 4 shows the percent of solar irradiance with wavelength shorter than λ as a function of λ. Fig. 5 shows the percent of excess photon energy as a function of $\lambda_g = hc/E_g$ (c = velocity of light). The sum of Figs. 4 and 5 is the percentage of solar irradiance not available for

conversion to electricity in the photovoltaic converter. This sum is plotted in Fig. 6 as a function of λ_g.

Efficiency of Ideal Shockley Photovoltaic Diodes

The efficiency of an ideal Shockley photodiode made from a given semiconductor material can be calculated for any value of i_{sc} from

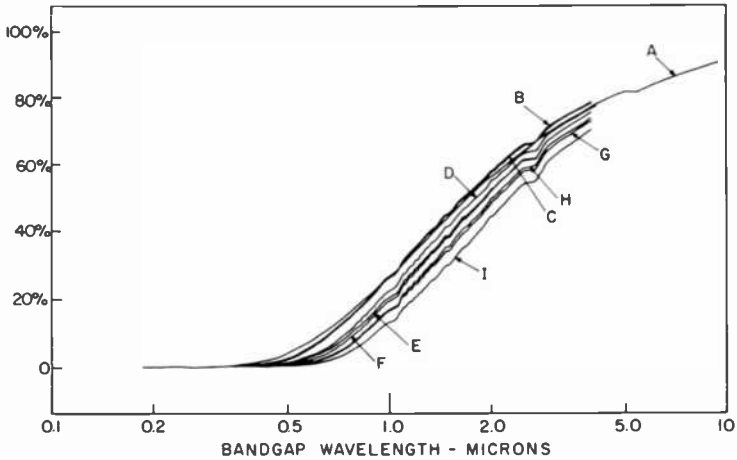


Fig. 5—Percent excess photon energy as a function of bandgap wavelength *(hc/E$_g$)* for several values of air mass and atmospheric pollution (α = 0.66). Air mass and β values for the curves are as given for Fig. 4 (all curves were calculated from data given in Ref. [1]).

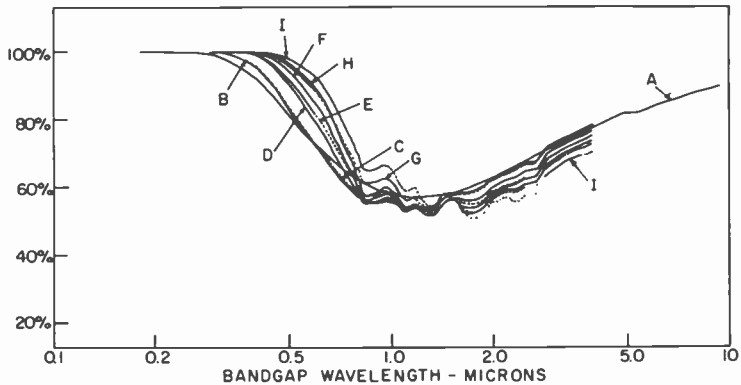


Fig. 6—Percent of solar irradiance not available for conversion to electricity in photovoltaic converters as a function of bandgap wavelength *(hc/E$_g$)* for several values of air mass and atmospheric pollution (α = 0.66). Air mass and β values for the curves are as given for Fig. 4 (all curves were calculated from data given in Ref. [1]).

Figs. 3 and 6 and Eq. [1]. Fig. 7 shows, for example, the calculated efficiency as a function of i_{sc}/i_o of a silicon Shockley diode under air mass zero illumination. Note that for $i_{sc}/i_o = 10^{11}$, which is a typical value for no concentration, the efficiency is about 22%. If the sunlight is concentrated by a factor of 1000 ($i_{sc}/i_o = 10^{14}$), the efficiency rises

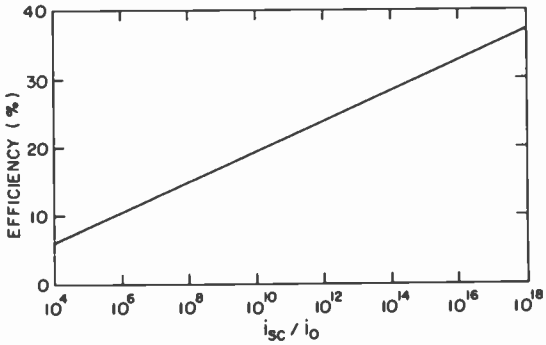


Fig. 7—Efficiency as a function of the ratio of short-circuit current to reverse saturation current for an ideal silicon Shockley photovoltaic diode under air-mass-zero illumination.

to about 28.5%. Similar improvements in efficiency with increased concentration ratio occur for all values of air mass and also for diodes fabricated from semiconductor materials other than silicon.

Acknowledgment

The author wishes to thank S. G. Liu and L. S. Napoli for helpful discussions and B. S. Perlman for help in computing Figs. 4, 5, and 6.

References:

[1] Matthew P. Thekaekara, "Data on Incident Solar Energy," Symp. on Solar Energy Utilization and Panel Discussion on Solar Energy Programs and Progress, Wash. D.C., April 30, 1974.

[2] L. W. James and R. L. Moon, "GaAs Concentrator Solar Cell," Appl. Phys. Lett., **26**, No. 8, April 15, 1975.

[3] R. H. Dean et al. "Silicon Solar Cells for Highly Concentrated Sunlight," *RCA Review*, **36**, p. 324, June 1975.

[4] E. L. Ralph, "A Plan to Utilize Solar Energy as an Electric Power Source," *Conf. Record 8th IEEE Photovoltaic Specialists Conf.*, Wash., D.C., Aug. 1970, IEEE Cat. No. 70C, 32ED.

[5] P. Rappaport, "The Photovoltaic Effect and its Utilization," *RCA Review*, **20**, p. 373, Sept. 1959.

[6] M. Wolf, "Limitations and Possibilities for Improvement of Photovoltaic Energy Converters, Part I: Considerations for Earth's Surface Operation," *Proc. IRE*, **48**, p. 1246, July 1960.

[7] *Solar Cells—Outlook for Improved Efficiency*, National Academy of Sciences, Wash., D.C., 1972.

[8] S. G. Liu, Private Communications.

[9] John F. Parrot, "The Saturated Photovoltage of a P-N Junction," *IEEE Trans. Electron Devices*, **ED-21**, p. 39, Jan. 1974.

Silicon Solar Cells for Highly Concentrated Sunlight

R. H. Dean, L. S. Napoli, and S. G. Liu

RCA Laboratories, Princeton, N. J. 08540

Abstract—Silicon solar cells yield 8% overall conversion efficiency at 800–1600 × peak concentration. Electrical sheet resistance and thermal and high-injection-level effects are not serious problems with proper cell design.

Introduction

Ralph[1] has suggested that the effective cost of silicon solar cells might ultimately be reduced to about $150/kilowatt by using high energy concentration ratios. Beckman, et al[2,3] constructed a parabolic mirror concentration system that produced a peak intensity concentration ratio of about 313 ×, and an average intensity concentration ratio of about 275 ×. Their 36-cm^2 cell area generated 40 watts for a maximum array efficiency of 4.4%. Since most of the reflected light missed the cells, the maximum system efficiency, defined as the maximum net output power divided by the total beam radiation (0.090 W/cm^2) incident on their 1.8-meter-diameter concentrator, was 1.53%. They projected[3] that improvements in the concentrator and cells should lead to an overall system efficiency of about 4%.

It has been standard practice[2–5] to apply metal grids to the top surface of the cells to reduce the adverse effects of sheet resistance in the surface layer of the cell. In trying to determine the proper grid size to use with a specific concentrated system, one is led to a simple scaling law that shows how the grid size should be changed as the concentration is increased.

In addition to the electrical problem associated with sheet resistance, there are problems associated with heat transfer and local

heating of the silicon chip due to thermal spreading of the concentrated energy. We have treated these problems separately and show that with proper design, solar-cell temperature can be controlled without resorting to forced cooling. With these design rules as guides, we have carried out experiments with concentration ratios of 1200 to 1600 × on solar cells especially designed for high concentration ratios. The results are considerably more encouraging than Beckman's.[3] The efficiency was 8% and the degradation due to thermal effects or high current density appeared to be minimal.

Theory

A. Electrical Considerations

The sheet resistance of the top layer of a solar cell is ρ_{se}. Assume that the grid surrounds a nearly circular free area. If current is generated uniformly over this area, a simple integration shows that the average resistance through which this current must flow is

$$\langle R \rangle = \frac{\rho_{se}}{4\pi}. \qquad [1]$$

This electrical resistance is independent of radius and concentration ratio. The voltage drop in the surface layer is

$$V_s = J_e \langle R \rangle l_{uc}^2 \qquad [2]$$

This voltage, which is proportional to cell area l_{uc}^2, is directly subtracted from the intrinsic cell voltage and degrades the efficiency. As the cell area increases and the concentration factor M increases, the efficiency degrades. The unit cell, therefore, must decrease in size as the concentration factor increases. The overall solar-cell area, however, can be made much larger provided a metal-grid ohmic contact is placed at unit cell intervals. The metal grid must not cover a significant percentage of the area, so as not to shadow much of the cell, and it cannot be excessively thick. As the solar cell increases in size, while grid spacing remains constant, the resistance of the metal grid itself tends to limit the efficiency. Thus the solar cell must be of limited size or be terminated in a super grid (i.e., a superimposed coarser grid structure) at finite intervals. Fig. 1 summarizes the effects of silicon sheet resistance and metal-grid resistance on solar-cell efficiency. In the example, the semiconductor sheet resistance is 100 ohms/square, and the metal grid is gold, one μm thick, covering 5% of the area of

the solar cell. The ultimate efficiency without degradation is postulated to be 14% and the degraded efficiency is

$$\eta = 14\left(1 - M l_{uc}^2\left[1 + 6 \times 10^{-3}\left\{\frac{l_{sc}}{l_{uc}}\right\}^2\right]0.446\right) \qquad [3]$$

where l_{uc} = the linear dimension of the uncovered silicon unit cell, or the distance between grid lines,

l_{sc} = the linear dimension of the gridded solar cell, or the distance between super grid lines.

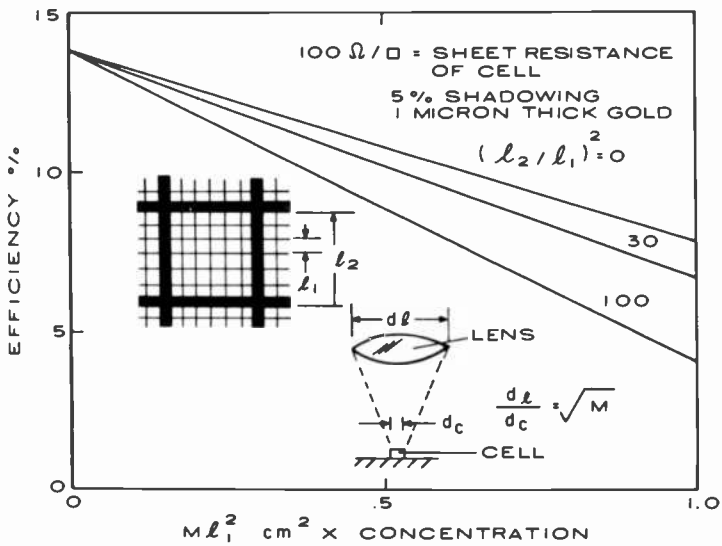


Fig. 1—Degradation of solar-cell efficiency due to semiconductor sheet resistance and metal-grid resistance as a function of concentration ratio, grid spacing, and super grid spacing.

Ohmic contact resistance does not come into play unless the product of the ohmic contact current density and the contact resistance is a significant portion of 0.5 volt. The degradation in efficiency resulting from this is shown in Fig. 2. Note that at 100 × and 5% contact area there is no significant degradation with a contact resistance as high as 10^{-4} ohm/cm^2.

B. Thermal Considerations

A practical system for photovoltaic energy conversion using concen-

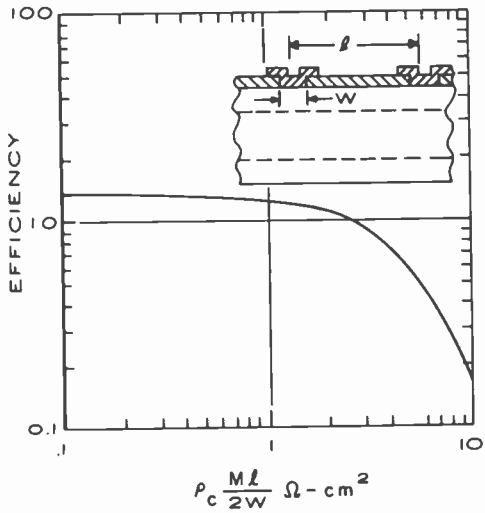


Fig. 2—Degradation of solar-cell efficiency due to ohmic contact resistance.

tration must consist of an array of power units each of which is comprised of a lens, a solar cell, and enclosure. The enclosure serves four functions: (1) protection, (2) a heat sink, (3) a thermal radiator, and (4) a means to convect heat. The thermal problem associated with such a system is briefly outlined in Fig. 3. The power incident from

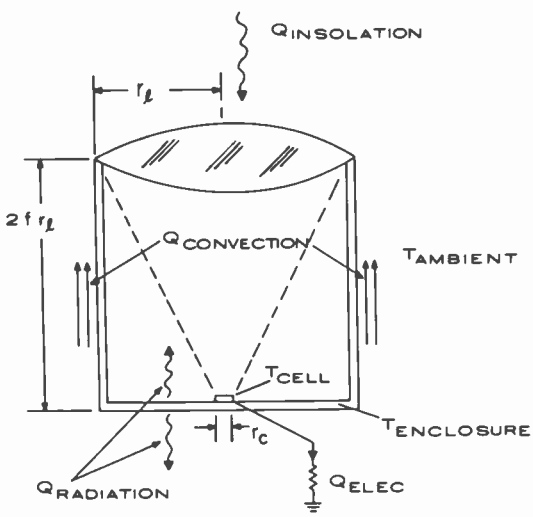


Fig. 3—Thermal schematic of solar-cell enclosure.

the sun (insolation) is converted into electricity and heat. The heat leaves by means of radiation and convection. Convection takes place along the sides and upward, and occurs naturally at atmospheric pressure subjected to the gravitational forces at sea level. Radiation occurs upward and downward but not sideways since adjacent units inhibit this direction of radiation. The temperature of the enclosure rises to an equilibrium point where the radiated power plus the convected power equals the incident power less that which is converted into electricity. Since all of the power incident on the lens is focused on the solar cell, the temperature of the solar cell rises above that of the enclosure according to the following relation:

$$T_{cell} - T_{Encl} = (Q_{inc} - Q_{elec})\left(\frac{t_{si}M}{\sigma_{si}} + \frac{r_l^2}{4\sigma_{HS}t_{HS}} + \frac{r_l^2}{\sigma_{HS}t_{HS}}\ln M\right), \quad [4]$$

where Q_{inc} and Q_{elec} are insolation power and electric power, t_{si} and t_{HS} are the silicon and heat sink thickness, σ_{si} and σ_{HS} are the thermal conductivities, r_l is the lens radius, and M is the concentration ratio.

The power leaves the unit according to the following relation:

$$Q_{inc} - Q_{elec} = 2\epsilon(T_{Encl}^4 - T_{Amb}^4) + 4(f)1.4\frac{(T_{Encl} - T_{Amb})^{1.25}}{(2r_l f)^{1/4}}. \quad [5]$$

In Eq. [4], the first term in the brackets represents the temperature drop in the silicon, and the second and third terms are due to the thermal spreading resistance in the heat sink. In Eq. [5], the first part is the Stephan–Boltzmann equation for radiation, and the second part is derived from a handbook relation for natural convection,[6] where f is the ratio of the length to the diameter.

Clearly the problem has been simplified by assuming a constant enclosure temperature for convection and radiation, and an independent rise in heat-sink temperature due to thermal spreading in the vicinity of the solar cell. A more rigorous solution would be unwieldy and would add little to the results.

This simplified solution does allow one to understand the problem in separable parts. For example, Eq. [5], the radiation and convection equation, holds whether or not concentration is used, and f, the aspect ratio, must be of the order of, or greater than, unity if convection is to be employed in the removal of heat. The only effect of concentration is in the elevation in temperature in the vicinity of the solar cell. Eq. [5] also shows that, for terrestrial application, convection partly makes up for the reduced radiation due to a finite temperature ambient.

The enclosure temperature resulting from equilibrium radiation and convection has been plotted as a function of aspect ratio in Fig. 4 for conditions of 43°C ambient (110°F). For an aspect ratio f of zero, no natural convection takes place, all the heat must leave by radiation, and the temperature is 95°C (i.e., the temperature rise is 52°C). This would be the case if the enclosure were compressed so that the

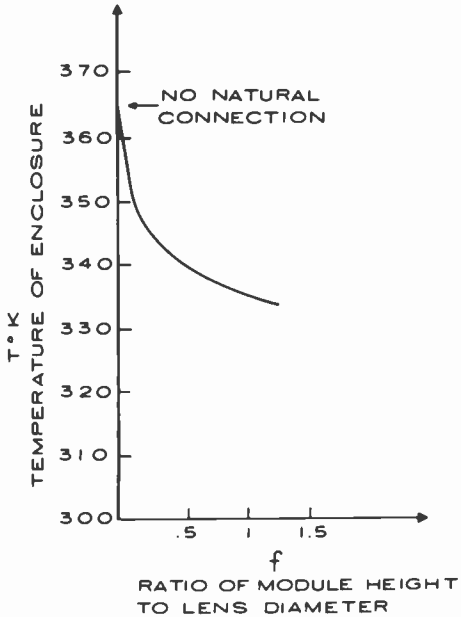


Fig. 4—Solar-cell enclosure temperature showing the effect of side-wall convection.

lens (or plastic or glass shield) was in contact with the heat sink and solar cell and was parallel to the ground. This condition inhibits natural convection. As the enclosure expands, thereby allowing a significant portion of the heat to be carried away by convection, the temperature reduces to 63°C. It must be emphasized that these calculations are based on the total absence of wind and are worst-case solutions only.

The temperature rise versus concentration ratio due to thermal spreading resistance is shown in Fig. 5, with lens radius as a parameter. As can be seen from the figure, the temperature rise is not a strong function of the concentration ratio, but is a strong function of

lens radius. The example used to plot Fig. 5 employed an aluminum heat sink 1.5 mm thick supporting a silicon solar cell 0.1 mm thick. The diameter of the heat sink is equal to that of the lens. Even at high concentration ratios of 1000, only 1°C temperature rise occurs in the solar cell. This figure is not terribly surprising, because even at 1000 suns, the power density incident upon the solar cell is three orders of magnitude lower than the power density of typical microwave power transistors. Also, at 1000 suns, no significant temperature rise occurs in the heat sink until the lens radius reaches 5 cm (a diameter of approximately 4 inches).

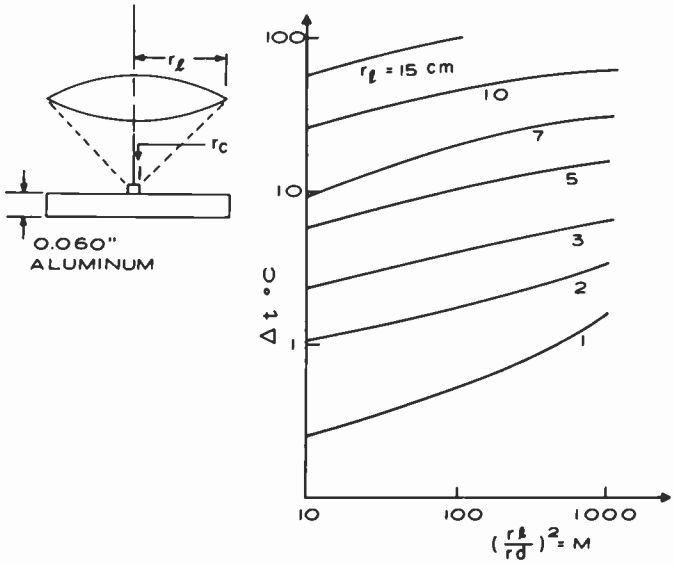


Fig. 5—Temperature drop due to local thermal spreading in the vicinity of the solar-cell chip as a function of concentration ratio and lens radius.

To sum up the theoretical calculations, we will consider the example of an array of modules consisting of a lens 10 cm in diameter on top of a cylindrical enclosure 20 cm high (f = 2). The enclosure in this example is aluminum 1.5 mm thick with a solar cell 3.3 mm in diameter mounted on the aluminum heat sink at the bottom of the enclosure. The cell has a grid spacing of 0.16 mm and a grid width of 4 μm, a super grid spacing of 1.6 mm and a super grid width of 16 μm. Efficiency and power are summarized in Table 1.

Table I—Efficiency and Power of Solar-Cell Array.

	Efficiency	Power Watts
Fundamental Solar Cell at 20°C	14%	1.4
At 48°C Ambient	12.5%	1.25
At 68°C Enclosure Temperature* (Radiation and Convection Equilibrium)	11.4%	1.14
At 84°C Silicon Temperature (Thermal Spreading)	9.8%	0.98
Including I^2R loss in silicon	8.8%	0.88
Including I^2R loss in metal grid	8	0.8
Including Grid shadowing	7.7%	0.765

* This would be the ultimate terrestrial efficiency under these ambient conditions with no wind and *without* energy concentration.

Experiment

Our experiments were conducted with p^+-n-n^+ silicon cells. Two configurations were tried. These two configurations have essentially the same base regions—50-μm-thick epitaxial layers with doping density of 10^{14} cm^{-3}. The surface layers and grid contact geometries are different, however.

The surface layer on NSO2 was formed by diffusing p^+ with a final surface concentration of 5×10^{19} cm^{-3}. An SiO_2 antireflection coating was grown and holes were opened in the oxide for a grid pattern. Extra doping was added to the exposed silicon by ion implantation for an expected surface concentration of 1×10^{20} cm^{-3}. After annealing, an ohmic-contact grid was formed on the top surface by evaporating chromium+gold into the oxide openings. The wafer was thinned, and the other chromium+gold ohmic contact was deposited on the back.

The surface layer on NSO4 was formed by ion implantation only with a surface doping density of 5×10^{20} cm^{-3}. An SiO_2 antireflection coating was grown, and annealing of the ion-implanted layer occurred during the oxidation process. Holes were opened in the oxide, and an ohmic contact grid was formed on the top surface by evaporating chromium+gold into the openings in the oxide. After thinning, the other chromium+gold ohmic contact was deposited on the back.

The different cell configurations were compared by observing the *I–V* characteristics as the optical intensity was increased from below

normal sunlight intensity to almost the maximum intensity produced by our concentration system. At normal sunlight intensity, both NSO2 and NSO4 cells exhibited a sharp knee in the fourth quadrant. As the sunlight intensity increased, short-circuit current and open-circuit voltage increased monotonically for both kinds of cells. The knee of NSO4 remained in the fourth quadrant, as desired, but the knee of NSO2 shifted over into the third quadrant (see Fig. 6). This behavior indicates that concentration results in significant series resistance problems in NSO2, but not in NSO4. Wafer NSO4 was somewhat easier to fabricate than wafer NSO2, even though it required a relatively heavy ion-implanted dosage. Thus, the process employed to fabricate NSO4 is inherently superior.

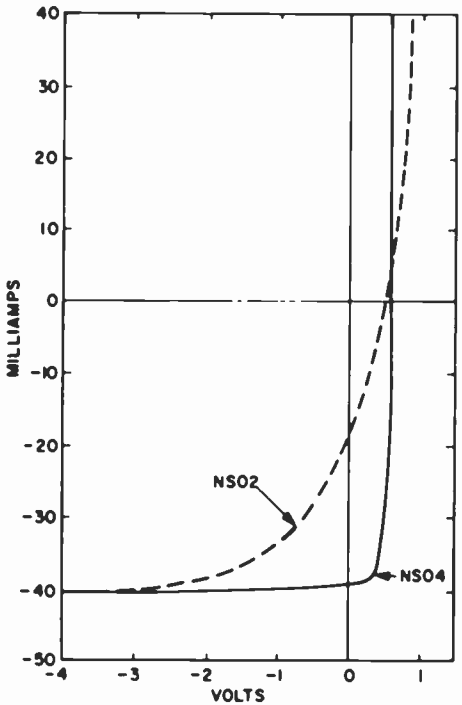


Fig. 6—Current–voltage characteristics of solar cells from two wafers at high concentration of sunlight intensity.

The measured surface resistance of the final surface layer of NSO4 is 13.6 ohms/square, a factor of seven less than that used to calculate the curves of Fig. 1. The grid spacing was 90 μm so that Ml^2 at 1000

× is 0.08; thus only a negligible efficiency reduction is due to the surface-resistance losses.

For the concentration experiments, we used the four-cell test fixture shown in Fig. 7. The base of the test fixture was made of 0.16-cm printed-circuit board with 0.01-cm copper on both sides. The copper on one side was patterned so that four cells could be connected in series. Each cell was mounted on an approximately 1.9-cm-square pad of copper designed to allow heat to fan out over the surface. Four injection-molded acrylic lenses were supported 5.0 cm above the surface of the circuit board, with the centers of the lenses directly above the centers of the silicon chips. The lenses were plano-convex spherical lenses, with a 5.0-cm focal length and 2.0-cm diame-

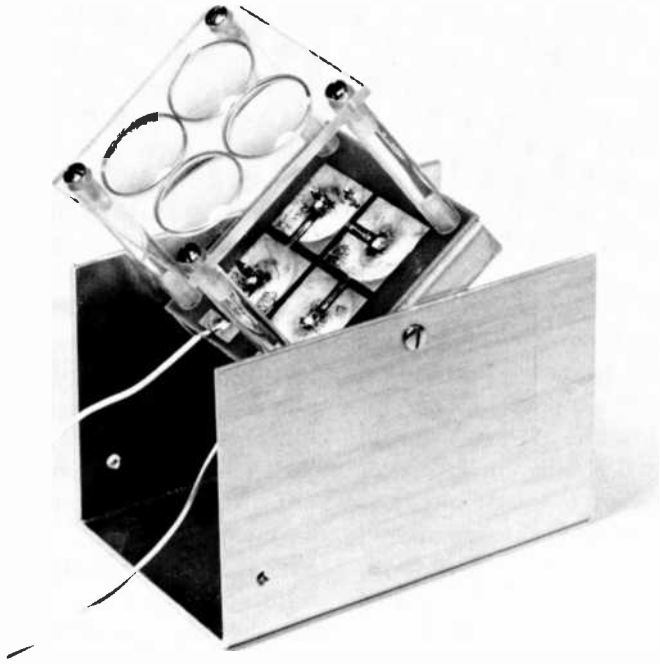


Fig. 7—Experimental four-cell test fixture.

ter. They were mounted with the curved side away from the lenses for most of the experiments, because this provided the least amount of aberration. Since the sun subtends an angle of 0.01 radian, the theoretical image diameter is 0.05 cm, corresponding to a lens area to sun-

spot area concentration of 1600 ×. The visually observed image diameter is approximately 0.07 cm, corresponding to an area concentration of 800 ×. The silicon chips are 0.16 cm square, corresponding to a lens area to cell area concentration of about 120 ×. The oversizing made the chips easy to register during the mounting process.

The copper cladding to the circuit board is 0.075 mm thick. Copper has a conductivity twice that of aluminum; thus, its equivalent thermal sheet resistance is ten times that of the aluminum used in the calculated curves of Fig. 4. For the 1-cm-radius lens at 1000 × concentration, Fig. 4 predicts a temperature rise of 1.6°C, 1°C of which is the thermal drop in the silicon; thus 0.6° is due to the 1.5-mm-thick aluminum. Therefore, the temperature drop in the copper-cladded circuit board used in the experiment was 6°C. Most of the temperature rise is associated with the convective and radiative transfer from the two sides of the circuit board to the air and the surroundings.

The array of four series cells of NSO4 was tested outdoors on a clear day in New Jersey, at 1:00 p.m. in mid-April. The solar intensity was approximately 0.090 W/cm^2. The array of four cells developed 2.0 volts across a 43-ohm load resistor, for a current of 0.047 A. Thus, the actual output power was 0.093 W. The total solar power incident on the clear area of lenses was approximately 1.150 W, giving an overall efficiency of approximately 8%. Antireflection coatings on the lenses, a better antireflection coating on the cell, and a smaller cell area would all contribute to making the overall efficiency higher.

The peak concentration in this experiment was probably over 1000 ×. A step recovery measurement indicated an average carrier lifetime of 3 μsec. Assuming that the concentrated photon flux produces electrons at the rate of 10^{20} cm^{-2} sec^{-1}, in a depth of 0.5×10^{-2} cm, the 3-μsec lifetime indicates that the average electron density in the base is about 6×10^{16} cm^{-3}. This is more than two orders of magnitude higher than the base doping density. Spectral response measurements indicated that, compared to a conventional solar cell, excessive losses were occurring at both short and long wavelengths. This suggests that the efficiency might be improved by reducing the thickness of the surface layer and increasing the thickness of the base.

Conclusions

We have presented simple rules for designing solar cells for highly concentrated sunlight. Electrical sheet resistance and thermal problems can be made almost arbitrarily small with a sufficiently small chip size. High doping density at the surface is required to minimize

contact resistance. Using four 0.16-cm-square silicon cells positioned at the focal points of four 2.0-cm-diameter plano-convex acrylic lenses, we obtained 93 mW of dc power under a clear sky in New Jersey, for an approximate over-all efficiency of 8%. The experiment demonstrates that peak solar concentrations of order 1000 × can be used with silicon solar cells without series degradation from electrical sheet resistance, thermal, or high-injection-level effects, provided the cells are properly designed.

Acknowledgments

The authors are indebted to F. Sterzer for suggesting this work and to H. Huang for many stimulating discussions. J. F. Kaminski provided processing assistance. The ion implantation was made by C. P. Wu. The lifetime and spectral measurements were made by V. Dalal. The plastic lenses were provided by D. F. Hinchman of U.S. Precision Lens, Inc., Cincinnati, Ohio.

References:

[1] E. L. Ralph, "A Plan to Utilize Solar Energy as an Electric-Power Source," *Conf. Record 8th IEEE Photovoltaic Specialists Conf.*, Wash, D.C., Aug. 1970, IEEE, Cat. No. 70C, 32 ED.

[2] W. A. Beckman, P. Schoffer, W. R. Hartman, Jr., and G. O. G. Lof, "Design Considerations for a 50-Watt Photovoltaic Power System Using Concentrated Solar Energy," *Solar Energy*, **10**, p. 132 (1966).

[3] W. A. Beckman and P. Schoffer, "An Experimental High Flux Solar Power System," *Proc. 20th Annual Power Sources Conf.*, May, 1966, Univ. of Wis. Engineering Experimental Station Reprint No. 916.

[4] M. Wolf, "Limitation and Possibilities for Improvements of Photovoltaic Solar Energy Converters," *Proc. IRE*, **48**, p. 1246 (1960).

[5] P. A. Berman, "Design of Solar Cells for Terrestrial Use," *Solar Energy*, **11**, p. 180 (1967).

[6] Baumeister and Marks, *Standard Handbook for Mechanical Engineers*, 7th Ed., McGraw Hill, New York.

[7] R. K. Yasai and L. W. Schmidt, "*Conf. Record 8th IEEE Photovoltaic Specialists Conf.*, p. 110, Aug. 1970.

Review on Industrial Applications of High-Power Laser Beams III*

I. P. Shkarofsky

RCA Ltd., Ste. Anne de Bellevue, Canada

Abstract—High-power laser interactions with matter as related to industrial applications are reviewed. Special emphasis is placed on results with CO_2 lasers, since they provide the high powers necessary for these applications. Theory and experimental results are presented on welding, drilling, microprocessing, cutting, and other physical and chemical changes induced by laser heat. Tabulated and plotted results are given that illustrate current devices in operation.

1. Introduction

We review here the use of lasers, mainly the CO_2 laser, for material processing, such as welding, drilling, trimming, scribing, cutting, etc. All these applications utilize the transfer of heat to the material by means of laser energy. Focusing spots down to 50 μm diameter are obtainable. Most surfaces, especially if polished, initially reflect a large portion (95–100%) of the CO_2 laser beam (see Table 1 taken from Banas[1] and Harvey[2]). However, as the surface temperature increases, the reflectivity decreases and the absorptivity increases appreciably. Although the reflectance of most metals increases with wavelength and hence should be worse at 10.6 μm with respect to the visible, the

* This paper is the third in a series; Part I, "Review of Gas-Breakdown Phenomena Induced by High Power Lasers," by I. P. Shkarofsky appeared in the March 1974 issue of *RCA Review*, Part II, "Review on High-Power-Laser Damage to Materials," by A. K. Ghosh," appeared in the June 1974 issue of *RCA Review*. The original manuscript for this paper was received in January 1974.

initial high power available from CO_2 lasers compensates for this disadvantage.

The material properties of importance are surface reflectivity, volume absorption, volume thermal diffusivity, melting point of the material and its oxides, ignition temperature and latent heat, resistance to thermal shock, surface irregularity, volume homogeneity, and material thickness. The laser parameters of concern are wavelength, power, focal spot diameter, pulse duration, repetition frequency, and whether or not gas assistance is present.

Table 1—Reflectance of Materials at 10.6 μm.[1,2]

Material	As Received	Polished
316 stainless steel	61	91
416 stainless steel	48	90
Al	88	98
Cu	88	98
4130 steel	15	—
Mo	88	97
Ta	79	96
Ag		99
Rhodium		96
Au		99

2. Theory

The quantity of heat q necessary to impart to a surface, including possible phase transitions, is

$$q = \rho V(c_p \Delta T + L_f + L_v), \qquad [1]$$

where c_p is the specific heat of the substance and ΔT is the rise in temperature from the initial to the final value. [If ρc_p differs for the solid or liquid state, one should write the summation over the parameters instead.] Also ρ is the density and V is the volume. In passing through the melting stage, we add the contribution L_f, the latent heat of fusion, and in passing through the vaporization stage, we add L_v, the latent heat of vaporization.

The diffusion of heat within the substance is described by the heat conduction equation, namely

$$\nabla^2 T - \frac{1}{\kappa}\frac{\partial T}{\partial t} = -\frac{A}{K} \qquad [2]$$

where T is the temperature, t is the time, ∇^2 is the spatial Laplacian, K is the thermal conductivity, A is the source of heat per unit volume per unit time present *within* the material and κ is the thermal diffusivity given by

$$\kappa = \frac{K}{\rho c_p}. \qquad [3]$$

Materials with low κ values (such as stainless steel and Rene 41 alloy)

Table 2—Heating of Metals and Appropriate Constants.[3]

Metal	κ (10^{-4}m²/s)	K (cal/s,cm,°C)	c_p (cal/g,°C)	ρc_p (cal/cm², °C)	T_0 (°K)	ΔH (J/μg × 10^{-3})
Al	0.91	0.53	0.215	0.58	1900	2.11
Be	0.42	0.35	0.45	0.84	2700	8.95
Cr	0.20	0.16	0.11	0.79	2600	2.04
Cu	1.14	0.94	0.092	0.82	2370	1.08
Au	1.18	0.71	0.031	0.60	2400	0.37
Fe	0.21	0.18	0.11	0.87	2800	2.11
Mo	0.51	0.34	0.066	0.68	4300	1.79
Ni	0.24	0.22	0.105	0.94	2700	1.74
Pd	0.24	0.17	0.058	0.70	2300	0.76
Pt	0.24	0.16	0.031	0.67	3700	0.67
Si	0.53	0.20	0.162	0.38	2300	3.64
Ag	1.71	1.0	0.056	0.59	1900	0.54
Ta	0.23	0.13	0.034	0.56	3270	0.48
Sn	0.38	0.15	0.054	0.40	1900	0.47
Ti	0.082	—	—	—	3000	2.25
W	0.62	0.40	0.033	0.64	3700(T_m)	0.50
Zn	0.41	0.27	0.092	0.66	1000	0.42

are difficult to weld (which requires good heat diffusion and fusion penetration) but are easy to drill, since they retain their heat leading to vaporization. Table 2 (taken from Charschan,[2] pps. 156, 157, 193) lists values of κ, K, c_p and ρc_p for various metals. Note that ρc_p does not vary much for various metals, so that a high c_p metal such as Al doesn't conduct heat much differently from others. Nickel has the largest ρc_p value although its K value is small. Consequently, it absorbs energy for a longer time and can be welded. Also given in the

table is the temperature T_0 equal to $(T_m + 2T_v)/3$ where T_m is the melting temperature and T_v the vaporization temperature. Then the enthalpy

$$\Delta h = \int_{298}^{T_0} c_p dT$$

is given, which is the amount of energy required to be absorbed by a unit mass of material at constant pressure to have its temperature raised from room temperature (298°K) to T_0. The laser has to be capable of applying energies per mass of the order of magnitudes indicated in the table.

Various solutions of Eq. [2] use the following assumptions: (1) the laser duration is longer than the electron collision time (10^{-12} to 10^{-14} sec) within the material so that the laser energy is instantaneously converted to heat; (2) K, ρ and c_p are independent of T; (3) heat reradiation from the material is negligible. Several solutions for particular cases are now given.

Case (1): Material surface in xy plane maintained at a uniform temperature, valid for a wide laser beam. Also assume that

$$A = \alpha I_o \exp(-bz),$$

where b is the inverse of the penetration depth or half the skin depth, α is the absorption coefficient, and I is the beam intensity absorbed (not incident) by the material with $I_0 = I(0,t)$. Then the solution of

$$\frac{\partial^2 T}{\partial z^2} - \frac{1}{\kappa}\frac{\partial T}{\partial t} = -\frac{\alpha I_o}{K}\exp(-bz),$$

with $T(z,0) = 0$, is[4,5]

$$\begin{aligned} T(z,t) = {} & \frac{2I_0\alpha}{Kb}(\kappa t)^{1/2}\ \mathrm{ierfc}\left(\frac{z}{2(\kappa t)^{1/2}}\right) - \frac{I_0\alpha}{Kb^2}\exp(-bz) \\ & + \frac{I_o\alpha}{2Kb^2}\exp(b^2\kappa t + bz)\mathrm{erfc}\left[b(\kappa t)^{1/2} + \frac{z}{2(\kappa t)^{1/2}}\right] \\ & + \frac{I_o\alpha}{2Kb^2}\exp(b^2\kappa t - bz)\mathrm{erfc}\left[b(\kappa t)^{1/2} - \frac{z}{2(\kappa t)^{1/2}}\right] \end{aligned} \qquad [4]$$

where

$$\mathrm{ierfc}(y) \equiv \frac{2}{\sqrt{\pi}}\int_y^\infty (p - y)e^{-p^2}dp$$

and

$$\operatorname{erfc}(y) = \frac{2}{\sqrt{\pi}} \int_y^\infty e^{-p^2} dp \xrightarrow[y \to \infty]{} \frac{e^{-y^2}}{y\sqrt{\pi}}$$

If $\alpha = b$ and they are sufficiently large, one obtains from Eq. [4]

$$T(z,t) = \frac{2I_o(\kappa t)^{1/2}}{K} \operatorname{ierfc}\left(\frac{z}{2(\kappa t)^{1/2}}\right)$$

$$= \frac{2I_o}{K}\left[\left(\frac{\kappa t}{\pi}\right)^{1/2} \exp\left(-\frac{z^2}{4\kappa t}\right) - \frac{z}{2}\operatorname{erfc}\left(\frac{z}{2(\kappa t)^{1/2}}\right)\right] \qquad [5]$$

In particular at $z = 0$

$$T(0,t) = \frac{2I_o}{K}\left(\frac{\kappa t}{\pi}\right)^{1/2} \qquad [6]$$

Case (2): The general solution in Eq. [4] can be extended by the convolution theorem to include the case where the rate of heat production $A(z,t)$ is also a function of time. Let T' be the solution in Eq. [4]. Then

$$T(z,t) = \int_z^\infty \int_0^t \frac{I(\tau)}{I_o}\left[\frac{\partial^2}{\partial t \partial z'} T'(z',t - \tau)\right] dz' d\tau$$

$$= \int_z^\infty \int_0^t \frac{I(t - \tau)}{I_o}\left[\frac{\partial^2}{\partial \tau \partial z'} T'(z',\tau)\right] dz' d\tau. \qquad [7]$$

For example, applying this to Eq. [5] yields

$$T(z,t) = \frac{1}{K}\left(\frac{\kappa}{\pi}\right)^{1/2} \int_0^t \frac{I(t - \tau)}{\tau^{1/2}} \exp\left(-\frac{z^2}{4\kappa\tau}\right) d\tau. \qquad [8]$$

Fig. 1 (taken from Ready[5] or Charschan,[3] p. 163) shows numerical values of the temperature rise versus time and depth for an assumed laser pulse shape. Very large gradients (up to 10^6 deg/cm) and rates of temperature increase (10^{10} deg/sec) are evident. Ready[5] also finds that a shorter pulse containing higher power density leads to a higher temperature which, for some applications, provides an advantage of Q-switched pulses over longer pulses.

Case (3): Absorption and laser heating assumed to occur entirely in the xy surface plane (finite absorption length omitted). A semi-infinite body is considered with $T = 0$ as $z \to \infty$ and $\partial T/\partial z = 0$ at $z = 0$. Radial heat conduction is included as well as conduction into depth of material. The solution of $\nabla^2 T = \kappa^{-1}\partial T/\partial t$ is then (see Charschan,[3] p. 253 and ref. [6])

$$T(r,z,t) = \int_0^a \int_0^t \frac{q(r',\tau)r'}{4\pi^{1/2}\kappa^{3/2}(t-\tau)^{3/2}}$$
$$\times \exp\left[-\frac{r^2+r'^2}{4\kappa(t-\tau)}\right] I_0\left(\frac{rr'}{2\kappa(t-\tau)}\right)\left\{\exp\left[-\frac{(z-f(\tau))^2}{4\kappa(t-\tau)}\right]\right.$$
$$\left. + \exp\left[-\frac{(z+f(\tau))^2}{4\kappa(\tau-\tau)}\right]\right\} dr'd\tau. \qquad [9]$$

Here I_0 is the modified Bessel function, $q(r')\rho c_p$ is the instantaneous amount of heat liberated within the laser spot at $t = 0$ over a disc of radius a, and $q(r',\tau)$ extends this to a continuous source integrated from $t = 0$. As the laser penetrates into the material, we assume that the position of this disc is known and given by $z' = f(\tau)$.

Consider the following example of Paek and Gagliano.[6] They obtain $z' = 0.53\ \tau^{0.83}$ as the fit to their data (see Charschan,[3] p. 253) in drilling a 3.2 mm thick alumina with a ruby laser 4 J, 1.5 ms, 1.67×10^8 W/cm^2 and $2a = 0.5$ mm. Substituting this $f(\tau)$ into Eq. [9] and numerically integrating the equation, yields curves of $T(z,t)$ versus r. From these curves one can also estimate the tangential plane stress at a distance r, namely,

$$\sigma_\theta' = \frac{\alpha' E'}{r^2}\left[\int_a^r Trdr - Tr^2\right], \qquad [10]$$

where α' is the coefficient of linear thermal expansion and E' is Young's modulus. They find that shorter pulse lengths give lower stresses, a condition desirable in hole drilling. Of course, if the object is to produce radial damage, higher stresses can be generated by longer pulses, since this allows time for further heat diffusion.

We now give a particular solution of Eq. [9]. Let $f(t) = 0$ (no drilling) and $\rho c_p q(r) = I \exp(-r^2/a^2)$ (viz. a Gaussian beam) and extend the upper limit a to ∞. We then obtain

$$T(r,z,t) = \int_0^t \frac{I(t-\tau)\kappa^{1/2}a^2}{K(\pi\tau)^{1/2}(4\kappa\tau + a^2)}$$
$$\times \exp\left[-\frac{z^2}{4\kappa\tau} - \frac{r^2}{4\kappa\tau + a^2}\right] d\tau. \qquad [11]$$

Here I is the power per unit area at the center of the Gaussian spot. Eq. [11] agrees with Ready[4] (p. 76), who also supplies computer curves for various $I(\tau)$ and plots T versus z/a, $4\kappa t/a^2$ for various r/a. Note that Eq. [11] reduces to Eq. [8] as $a \to \infty$. *For* $r = z = 0$ and constant I_0, Eq. [11] yields

$$T(0,0,t) = \frac{I_o a}{K\sqrt{\pi}} \tan^{-1}\left(\frac{4\kappa t}{a^2}\right)^{1/2}.$$

For a cw laser with $t \to \infty$, this reduces to

$$T = I_o a\sqrt{\pi}/2K.$$

Equating $T = T_m$, the melting temperature, and t to the pulse length, provides for a given $a^2/4\kappa$, a representative value of I_0, that induces melting in various materials. Similarly, letting $T = T_v$ gives the maximum allowed I_0 that prevents vaporization in welding.

Case (4): Consider the heating of a thin sheet. Suppose that throughout the thickness D of the medium, the temperature input is constant in the z-direction and extends over a finite radial region corresponding to complete laser penetration. Then if H is the surface conductance of the material at the laser boundary, and if T_0 is the surrounding material temperature at $t = 0$, we have to solve

$$\frac{1}{r}\frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial r^2} - \frac{1}{\kappa}\frac{\partial T}{\partial t} - k^2(T - T_o) = 0,$$

where $k^2 = 2H/KD$ so that $\kappa k^2 = 2H/c_p\rho D$. The solution for a Gaussian beam is (see Ready,[4] p. 85)

$$T - T_o = \int_o^t \frac{a^2\kappa I(t - t')}{KD(a^2 + 4\kappa t')} \exp\left(-\kappa k^2 t' - \frac{r^2}{4\kappa t' + a^2}\right)dt'. \qquad [12]$$

This can be compared with Eq. [11]. Differentiating Eq. [12], we can derive the integrated heat flux out of the surface at radius r:

$$-K\int dS\mathbf{n}\cdot\nabla T = -K\int 2\pi r\frac{\partial T}{\partial r}dz = 4\pi\kappa(ra)^2$$
$$\times \int_o^t \frac{I(t - t')}{(a^2 + 4\kappa t')^2} \exp\left(-\kappa k^2 t' - \frac{r^2}{4\kappa t' + a^2}\right)dt'$$
$$\longrightarrow 4\pi\kappa r^2 D(T - T_o)/(4\kappa t + a^2) \qquad [13]$$

if I is instantaneous in time. Figure 2 (taken from Ready,[4] p. 86) provides computer solutions of Eq. [12] as functions of r/a and $k^2a^2/4$ for a source I constant in time. As an example, Ready considers a CO_2 laser having 1 W/cm^2 and $a = 1$ cm heating a plastic sheet with $H = 5$

$\times 10^{-4}$ W/cm^2-deg, $\kappa = 0.005$ cm^2/sec, $K = 0.0025$ W/cm-deg and $D = 0.01$ cm. After the beam is on for 1 sec, we obtain $4KDT/(I_0a^2) = 0.018$ or $T = 180°$C at $r = 0$ using the parameters $4\kappa t/a^2 = 0.02$ and $k^2a^2/4 = 10$. The temperature is still increasing linearly with increasing time and the plastic will eventually be melted even with this low-power CO_2 laser.

Case (5): Consider again a thin sheet with a constant cylindrical heat flux Q source over radius a (rather than the Gaussian beam considered previously), dissipating heat into the surrounding medium. Here Q (heat/time-area) is given by $\int Q dt = -K\partial T/\partial r$. We now solve the equation $\partial^2T/\partial r^2 + r^{-1}\partial T/\partial r - \kappa^{-1}\,\partial T/\partial t = 0$ for $r > a$, subject to $T = 0$ as $r \to \infty$. The solution is given by Carslaw and Jaeger[7] (p. 338), namely,

$$T = -\frac{2Q}{\pi K}\int_0^\infty (1 - e^{-\kappa u^2 t})\frac{[J_0(ur)Y_1(ua) - Y_0(ur)J_1(ua)]}{u^2[J_1^2(ua) + Y_1^2(ua)]}\,du, \qquad [14]$$

where J and Y are Bessel functions. For small $\kappa t/a^2$, we obtain

$$T = \frac{2Q}{K}\left(\frac{\kappa a t}{r}\right)^{1/2}\left\{\text{ierfc}\left[\frac{(r-a)}{2(\kappa t)^{1/2}}\right] - \frac{(3r+a)(\kappa t)^{1/2}}{4ar}\,i^2\text{erfc}\left[\frac{(r-a)}{2(\kappa t)^{1/2}}\right]\right\} \qquad [15]$$

and for large $\kappa t/a^2$, we obtain

$$T = \frac{Qa}{2K}\left[\ln\frac{4\kappa t}{Cr^2} + \frac{a^2}{2\kappa t}\ln\frac{4\kappa t}{Cr^2} + \frac{1}{4\kappa t}\left(a^2 + r^2 - 2a^2\ln\frac{a}{r}\right)\right]$$

$$\xrightarrow[r=a]{} \frac{Qa}{2K}\left[\ln\frac{4\kappa t}{Ca^2} + \frac{a^2}{2\kappa t}\ln\frac{4\kappa t}{Ca^2} + \frac{a^2}{2\kappa t}\right] \qquad [16]$$

where $C = 1.781$.

Gonsalves and Duley[8] apply Eqs. [14] to [16] to a cw CO_2 laser, drilling thin 302 stainless steel sheets. They write $Q = P\epsilon/2\pi aD$ where D is the sheet thickness, P is the incident power and ϵ is the emissivity at 10.6 μ into the surrounding medium. Due to the formation of an oxide layer, they find that the value ϵ(air) = 3.2 ϵ(vacuum) = 0.16 gives the best fit for steel. Letting $T = T_m$, the melting temperature, in Eq. [16] yields the required power P_m. Experimental plots are given by Gonsalves and Duley[8] and reproduced in Ref. [9]

(Figure 3 showing power versus time and sheet thickness). Fig. 4 in Ref. [9] gives normalized theoretical curves of melting time versus κ/a^2 and $12.57KT_mD/\epsilon P_m$.

Case (6): Consider a thick semi-infinite medium with a constant circular heat flux Q source (heat/time-area) over a radius a on a xy-surface, dissipating heat radially and into the medium. The solution of $\nabla^2 T - \kappa^{-1}\,\partial T/\partial t = 0$ for $r > 0$, $z > 0$ is (see Carslaw and Jaeger,[7] p. 264)

$$T = \frac{aQ}{2K}\int_0^\infty J_0(\lambda r)J_1(\lambda a)\left\{\frac{e^{-\lambda z}}{\lambda}\operatorname{erfc}\left[\frac{z}{2(\kappa t)^{1/2}} - \lambda(\kappa t)^{1/2}\right]\right.$$
$$\left. - \frac{e^{\lambda z}}{\lambda}\operatorname{erfc}\left[\frac{z}{2(\kappa t)^{1/2}} + \lambda(\kappa t)^{1/2}\right]\right\}d\lambda. \qquad [17]$$

In particular at $r = 0$ but $z \neq 0$

$$T(z,\ r=0) = \frac{2Q}{K}(\kappa t)^{1/2}\left\{\operatorname{ierfc}\frac{z}{2(\kappa t)^{1/2}}\right.$$
$$\left. - \operatorname{ierfc}\left[\frac{1}{2}\left(\frac{z^2 + a^2}{\kappa t}\right)^{1/2}\right]\right\}. \qquad [18]$$

Here one writes $Q = P\epsilon/\pi a^2$, and the above relations are also considered by Duley and Gonsalves[8] for calculating the maximum depth of welding in a semi-infinite sheet. As $t \to \infty$, Eq . [18] yields

$$T(r = 0,\ z,\ t \to \infty) = \frac{P\epsilon}{\pi a^2 K}[(z^2 + a^2)^{1/2} - z]. \qquad [19]$$

Equating $T = T_m$ (melting temperature required for welding) and $z = z_m$ gives

$$z_m = \frac{a}{2}\,\frac{T_v^2 - T_m^2}{T_m T_v},$$

where the maximum surface temperature near $z = 0$ is limited by the vaporization value $T_v = P_v\epsilon/\pi aK$. This gives the maximum P that can be used. The minimum P is obtained from $T_m = P_m\epsilon/\pi aK$. It is worth nothing as well that the cw CO_2 laser gives larger z_m or deeper welds than those found by Cohen[10] for pulsed ruby or Nd systems, since for short pulses, T_v is reached too rapidly precluding deep welds.

Case (7): In welding or drilling where melting or vaporization occurs, one should allow for the fact that the interface surface is moving. Cohen[10] investigates a one-dimensional solution with a wide laser beam causing the solid boundary to move inward upon melting. The boundary conditions are $T_1 = T_2 = T_m$ at $z = Z$, $t > 0$, where 1 refers to the solid and 2 to the liquid material, located between $z = 0$ and $z = Z$, and Z is the position of the interface. Also conservation of energy requires that

$$K_2 \frac{\partial T_2}{\partial z} - K_1 \frac{\partial T_1}{\partial z} = \rho L_f \frac{\partial Z}{\partial t} \text{ at } z = Z, \qquad [20]$$

where L_f is the latent heat of fusion. He then assumes a spatially and temporally uniform heat flux input so that at $z = 0$, $t > 0$, we have $I = -K_1 \partial T_1/\partial z$. Furthermore, in a semi-infinite medium, we require that $T_2 \rightarrow T_\infty$ and $\partial T_2/\partial z \rightarrow 0$ as $z \rightarrow \infty$. With these conditions, the heat equation: $(\partial T_{1,2}/\partial t) - \kappa_{1,2} \partial^2 T_{1,2}/\partial z^2 = 0$ is solved in both media. The computer solutions are given in Charschan[3] (p. 166) and Ready[4] (p. 93). Some results have already been shown in Figs. 1 and 2 of [9] and illustrate the previous observation that for short high intensity pulses, T_v is reached too rapidly. One can apply the same theory given above to vaporization, with T_v and L_v replacing T_m and L_f and with no liquid phase, so that Eq. [20] is replaced by

$$K_2 \frac{\partial T_2}{\partial z} - I = \rho L_v \frac{\partial Z}{\partial t} \text{ at } z = Z. \qquad [21]$$

Case (8): We now consider vaporization (no liquid phase), where the actual laser interaction surface is at $z = Z$. Here the more elaborate approach of Dabby and Paek[11] for a one dimensional, or plane beam, is considered. They substitute in Eq. [2], the expression

$$A = bI \exp[-b(z - Z)] \text{ for } z > Z,$$

where Z is the position of the propagating surface and $Z = 0$ at $t = 0$. Because of energy conservation, we have

$$\rho L_v \frac{\partial Z}{\partial t} = K\left(\frac{\partial T}{\partial z}\right)_{z=Z}, \qquad [22]$$

where L_v is the latent heat of vaporization. Other boundary condi-

tions are $T = T_v$ at $z = Z$ and $T = 0$ at $z = \infty$ (a semi-infinite medium). They also approximate the initial temperature profile by

$$T = T_v(1 + qz)e^{-qz} \text{ at } t = 0,$$

where q is an adjustable constant. The heat conduction equation for a moving surface, namely

$$\frac{\partial^2 T}{\partial z^2} + \frac{\dot{Z}}{\kappa}\frac{\partial T}{\partial z} - \frac{1}{\kappa}\frac{\partial T}{\partial t} = -\frac{I}{K} b \exp[-b(z - Z)] \qquad [23]$$

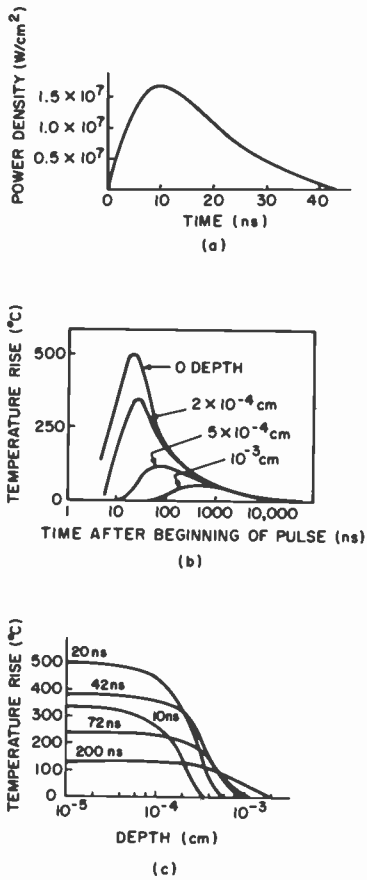


Fig. 1—(a) Laser pulse shape used to obtain results[3,4] shown in b and c. (b) Calculated temperature rise as a function of time, with depth as a parameter (tungsten sample initially at 0°C when a laser pulse of the shape and intensity shown above is absorbed at the surface); (c) calculated temperature as a function of depth, with time as a parameter (copper sample, same conditions as for b).

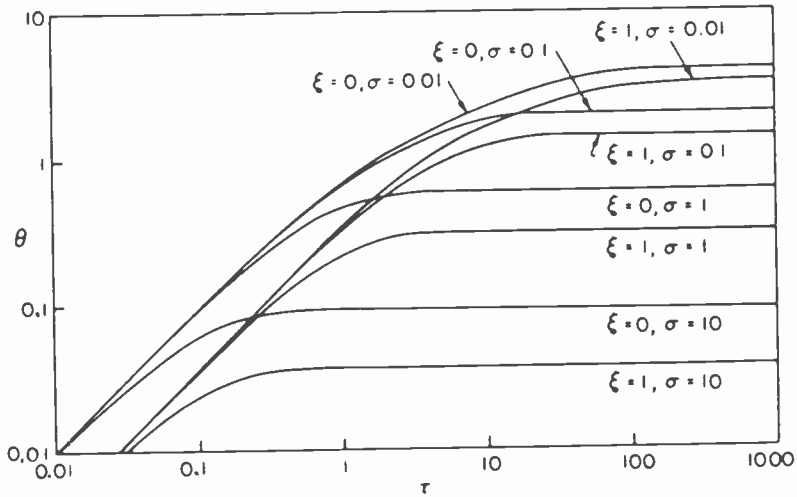


Fig. 2—Generalized time–temperature curves,[4] θ versus τ, for heating thin sheets: $\theta = 4KDT/I_0a^2$, $\tau = 4\kappa t/a^2$, $\xi = r/a$, and $\sigma = k^2a^2/4$.

is then solved subject to the previous boundary conditions. The analytic solutions are given in Dabby and Paek.[11] A sample calculation, taken from Charschan[3] (p. 170) is shown here in Fig. 3. An important observation is that the temperature inside the material can be greater than the propagating surface temperature T_v. This is due to the fact that surface cooling by vaporization causes the maximum tempera-

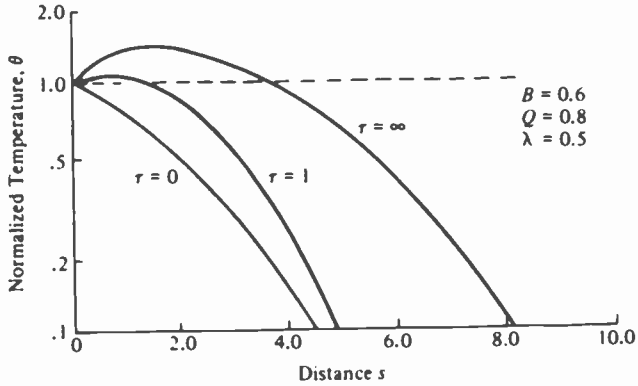


Fig. 3—Normalized temperature versus distance[3,11] with normalized time τ as a parameter. Here $\theta = T/T_v$, $s = Ic_p(z-Z)/KL_v$, $\tau = c_pI^2t/\rho KL_v^2$, $B = KL_vb/Ic_p$, $Q = KL_vq/Ic_p$, and $\lambda = c_pT_v/L_v$.

ture to lie below the surface. If this occurs, then the material may vaporize or even explode below the surface, leading to additional removal of material. Experimentally, this seems to have been observed in Cd, Bi, and Pb for which the parameter

$$B \equiv KL_v b/(Ic_p) \qquad [24]$$

is less than or of order of one; but not in Ag, Cu, and Fe for which $B \gg 1$. If for the latter three, B is decreased by increasing I, the position of the temperature peak shifts towards the surface and the effect becomes unimportant.

For large t, or in a steady state, the analytic solution simplifies to

$$\frac{T}{T_v} = \exp\left[-\frac{Ic_p(z - Z)}{(L_v + c_pT_v)K}\right] - \frac{I(c_pT_v + L_v)}{T_v[Kb(c_pT_v + L_v) - Ic_p]}\left\{\exp[-b(z - Z)] - \exp\left[-\frac{Ic_p(z - Z)}{K(L_v + c_pT_v)}\right]\right\}. \qquad [25]$$

The surface moves at a constant velocity given by

$$\dot{Z} = \frac{I}{\rho(c_pT_v + L_v)}. \qquad [26]$$

Suppose that the pulse time τ_p is longer than both the vaporization time and the time required to reach a steady state. Also if no subsurface explosions occur, the depth drilled is given by

$$d = \frac{1}{\rho(L_v + c_pT_v)}\int_0^{\tau_p} I(t)dt = \frac{\epsilon_T}{\rho(L_v + c_pT_v)}, \qquad [27]$$

where ϵ_T is the pulse energy density. This upper limit is equivalent to having all the energy go into raising the temperature to the boiling point and then supplying L_v. Sample calculations (see Charschan,[3] p. 174) were done for a Nd laser, $I_0 = 7.4 \times 10^8$ W/cm^2 peak, and constant for a time interval of 110 ns and with linear rise and fall times of 40 ns for the pulse. The times for vaporization and to achieve a steady state are between 1–10 ns for various metals and thus a steady state solution is valid. Calculated and observed values of d agree very well and are between 16–40 μm.

3. Experimental Results on Welding

The advantages in the use of lasers are as follows:

[1] Adjacent heat-sensitive elements remain unaffected, which is important in micro-welding.
[2] Material properties remain unaffected except for the heat-treated zone.
[3] No physical contact means that inaccessible regions may be reached.
[4] Welding (e.g., of broken wires) can be performed through enclosures and in many environments.
[5] Materials difficult to weld otherwise can be joined by lasers.
[6] Negligible distortion and shrinkage occurs upon welding.
[7] Wires can be welded without removal of insulation.
[8] Closed circuit TV inspection is possible.
[9] Beam can be transmitted an appreciable distance in air to the working piece without attenuation.
[10] The beam optics can direct the beam to the desired spot.
[11] No x-rays are generated.
[12] The process lends itself to automation.
[13] Speeds are comparable to those of standard techniques.

In welding, the intensity and duration of the laser beam have to be adjusted to induce melting and avoid vaporization. There is a critical maximum possible flux density. The intensity required to overcome the initial surface reflectivity can, however, be greater than that necessary to sustain and move the melting surface. Proper laser control can be a problem for very highly reflecting materials such as gold. Coating the surface with an absorption layer (e.g., SiO_2) or working with a certain degree of surface roughness eases the problem. Also metals with high vapor pressures have a tendency to vaporize and are difficult to weld. Besides metals, thermoplastics and quartz, for example, can be joined with lasers.

The laser pulse length should be greater than the thermal time constant $D^2/4\kappa$ where D is the sheet thickness. The welding process is better as the pulse repetition rate and pulse duration increase, pointing to the advantage of a cw laser, such as the high power CO_2 laser. Nonetheless, repetitively pulsed lasers are more economical and the initial pulse can serve to break down the surface reflectivity. Studies have shown[3] that a time exponentially decaying pulse (rather than square) shape is favored for welding, initially to overcome the high reflectivity and then to continue to induce melting but not vaporization. Q-switched lasers are not suitable for melting materials, since the surface temperature is raised too high. The previous theory (see

Table 3(a)—Experimental Results on Welding with CW CO_2 Lasers

Material	Weld Thickness (mm)	Rate (cm/sec)	Depth/Width	Laser Power (kW)	Focal Spot Diam. Approx. (mm)	Type of Weld	Ref.
304 Stainless Steel	17	1.06	6.5	16	—	—	*Laser Focus* p.32, May 1973
304 Stainless Steel	6.35	8.47	6.5	16	—	—	
302 Stainless Steel	20.3	2.1	5	20	1	butt	12
302 Stainless Steel	12.7	4.2	5	20	1	butt	12
302 Stainless Steel	8.9	1.27	3	8	1	butt	12
302 Stainless Steel	6.35	2.11	7	3.5	0.7	butt	14
302 Stainless Steel	6.35	1.69	3	8	1	butt	12
302 Stainless Steel	0.254	0.42	0.25	0.25	0.2	butt	16
302 Stainless Steel	0.254	0.64	0.4	0.25	0.2	lap fillet	16
302 Stainless Steel	0.203	1.27	0.4	0.25	0.2	butt	16
302 Stainless Steel	0.127	2.12	0.25	0.25	0.2	butt	16
321 Stainless Steel	0.254	1.48	0.36	0.25	0.2	butt	16
321 Stainless Steel	0.254	0.85	—	0.25	0.2	corner	16
321 Stainless Steel	0.127	3.8	0.28	0.25	0.2	butt	16
17-7 Stainless Steel	0.127	4.65	0.28	0.25	0.2	butt	16
17-7 PH Stainless Steel	0.127	1.9	—	0.25	0.2	edge	16
Stainless Steel (Type?)	3.18	0.59	1.14	4.6	—	butt	1
Stainless Steel (Type?)	0.97	0.34	0.54	0.6	—	butt	1
Low Carbon Steel	10.16	0.42	—	4.7	0.64	—	*Lasersphere* p. 8 May 15 1973
Low Carbon Steel	1.78	14.82	—	5	—	coach	
Low Carbon Steel	1.78	10.58	—	5	—	lap	
Low Carbon Steel	1.78	8.5	—	5	—	edge,bead	
Low Carbon Steel	1.52	12.7	—	4.7	0.64	—	
Low Carbon Steel	1.19	0.34	0.63	0.65	—	butt	1
Low Carbon Steel	0.79	0.85	0.62	0.6	—	butt	1
1010 Steel	0.94	5.08	—	3.9	—	fillet	13

Table 3(a)—(continued)

Material	Weld Thickness (mm)	Rate (cm/sec)	Depth/Width	Laser Power (kW)	Focal Spot Diam. Approx. (mm)	Type of Weld	Ref.
Alloy Steel (AMS5544)	0.46	0.5	—	0.195	1.2	seam	15
Tin Plated Steel	0.305	0.85	0.48	0.25	0.2	lap fillet	16
Tin Plated Steel	0.178	1.91	0.35	0.25	—	butt	4, p. 370
Ti(6Al.4V)	0.508	1.14	—	0.25	0.2	edge	16
Ti(6Al.4V)	0.251	12.7	—	0.25	0.2	edge	16
Ti C.P.	0.251	2.12	0.45	0.25	0.2	butt	16
Ti C.P.	0.127	5.92	0.33	0.25	0.2	butt	16
TD Nickel	0.251	1.69	0.4	0.25	—	butt	4, p. 370
Nickel 200	0.127	1.48	0.28	0.25	0.2	butt	16
Nickel 200	0.127	0.76	—	0.25	0.2	edge	16
Inconel 600	0.419	1.06	—	0.25	0.2	edge	16
Inconel 600	0.254	1.7	0.56	0.25	0.2	butt	16
Inconel 600	0.102	6.35	0.4	0.25	0.2	butt	16
Monel 400	0.254	0.63	0.4	0.25	0.2	butt	16
Monel 400	0.254	1.06	—	0.25	0.2	edge	16

Fig. 1 of Ref. [9]) illustrates that welding is limited to penetration depths less than mm's for pulsed (2–10 ms) lasers and to cm's for kW cw CO_2 lasers.

Table 3a gives results on welding with cw CO_2 lasers and Table 3b with pulsed CO_2 lasers. All types of welds[1,4,12–16] have been reported with cw lasers. The results with pulsed lasers are mainly from Ulmer and Tandler.[17] Figs. 4, 5[12] and 6[13] plot experimental results for various metals. Fig. 7[18] plots results for fused quartz. Welding at a rate of 0.2–1 cm/sec with a pulsed CO_2 laser having 2–10 ms pulse duration

Table 3(b)—Welding with Pulsed CO_2 Lasers.

Material	Weld Thickness (mm)	Rate cm/sec	Av. Laser Power (watts)	Pulse Width (msec)	Time Sep. Between Pulses (msec)	Ref.
Stainless Steel	0.38	2.54	100	?	?	4, p.368
Stainless Steel	0.1	0.125	100	0.2	0.1	17
Phosphor Bronze	0.1	0.217	100	300	350	17
Cu, Ni, Zn alloy	0.1	0.125	100	65	75	17
Monel	0.1	0.125	100	70	200	17
Cu, Ni alloy	0.1	0.127	100	300	350	17

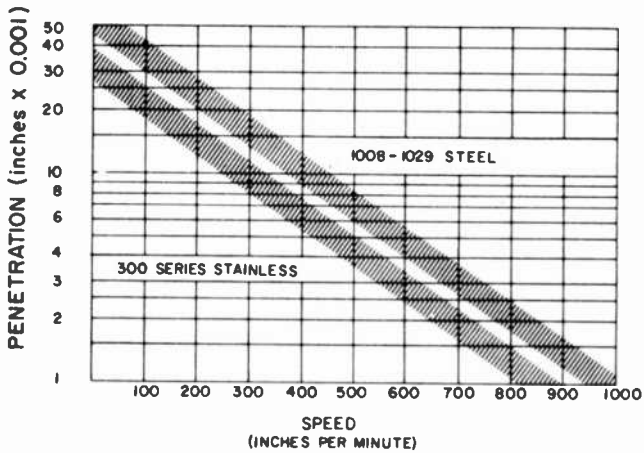


Fig. 4—Weld speed versus penetration with a 425 W cw CO_2 laser (*Photon Sources J.*, Vol. 1, May/June 1971).

requires 100–1000 pps; with sufficiently high power, one can approach the rates in Table 3a, viz., above 2 cm/sec.

At very high cw CO_2 laser powers (above 8 kW), Locke et al[12] observe plasma ejection with laser absorption in this cloud. For welding applications, this cloud can be eliminated with a gas assist[12] or by moving the work piece at a speed greater than 10 spot diameters in the time required for material expulsion.[13,14]

With a 3–8 kW cw CO_2 laser, high speed welding in thin-gauge sheet metals and deep penetration welding in metals up to 1/2 inch thick appear feasible. A high power 14 kW cw CO_2 laser is able to weld 0.76-mm-thick steel at a reasonable rate (*Laser Focus*, p. 12,

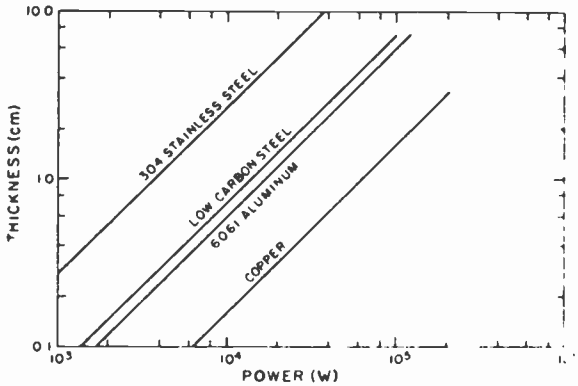


Fig. 5—Minimum power[12] for efficient welding with a CO_2 laser.

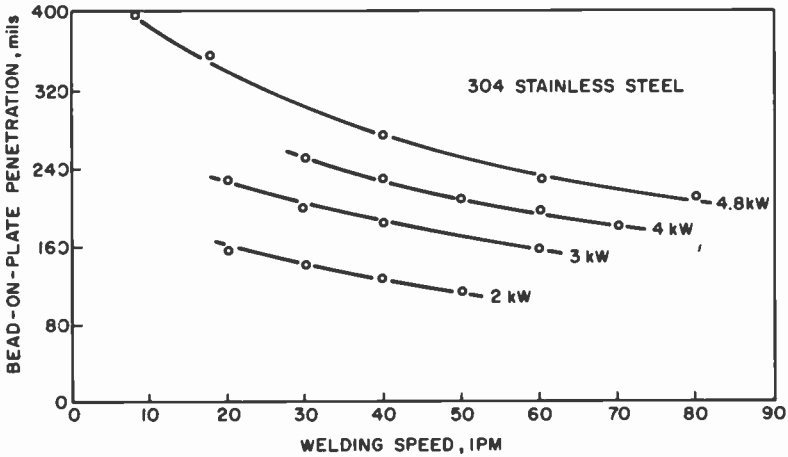


Fig. 6—Effect of welding speed[13] with a CO_2 laser on depth-of-penetration for various laser powers.

April 1972). Deep penetration welding requires 10–30 kW, and with 16 kW power, a weld as deep as 17 mm in 304 steel was obtained (*Laser Focus*, p. 32, May 1973). On the other hand, for welding fused silica and quartz, less than 100 W of CO_2 laser power is sufficient.[18,19] Microsoldering with a focused pulsed CO_2 laser is quite practical if high speed and automation are desired.[16] TEA CO_2 lasers of 100 kW, 0.4 μs pulse duration and 150 pps have been used to weld ceramics.

4. Experimental Results on Drilling and Evaporation

The advantages of lasers for drilling are:

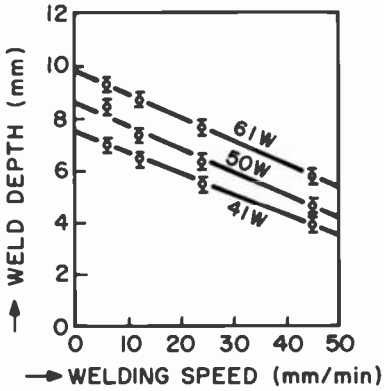


Fig. 7—Weld depth[18] as a function of welding speed and CO_2 laser-beam energy in fused silicon.

[1] No physical contact occurs and problems such as drill-bit breakage are nonexistent.
[2] Optics alignment simplifies hole location.
[3] Large depth-to-width ratio holes are achievable.
[4] Beam diameter can be varied over a wide range.
[5] Little physical force is applied to the workpiece.
[6] High speed manipulation and on-off switching of the beam are possible.

For hole drilling of non-metals, a CO_2 laser is usual. In hole drilling of metals, a pulsed laser with less than 1.5 ms pulse duration is usually employed. In drilling metals with a pulsed Nd glass laser (30 kW), it is found (see Charschan,[3] p. 231) that increasing the pulse duration from 50 to 100 μs, results in a rapid incremental increase in the ratio of molten material removed to total material removed; above 200 μs, this ratio exceeds 80%. Fig. 3 also shows that explosive removal of material occurs when the subsurface temperature exceeds that at the surface. With a 10^7 W/cm^2, 600 μs laser pulse, the depths of material removed from various metals vary between 0.5–1 mm. With a Q-switched 44 ns pulse of 10^9 W/cm^2 intensity, the depth is much shallower, between 1–4 μm. (See Ref. [4], p. 110, or Ref. [5].) For such short, high-energy pulses, shock or pressure waves induced by the laser into the material are thought[5] to be responsible for the observed depth of the drilled holes. Generally, metals with small T_v experience larger material removal at high laser flux density.

In drilling at high powers, the expelled plume of material can become large and highly excited and can appreciably absorb the laser radiation. This is not a problem for brass because of its high vapor pressure and 0.5–1 mm holes can be drilled in 70:30 brass, 2.4 mm thick, with a pulsed ruby laser having 75 J, 5 ms and 10 MW/cm^2 (Charschan,[3] p. 232). For other nonferrous materials, the depth is limited to about 0.5 mm with a single pulse, and for ferrous materials to about 1 mm depth. The laser is most competitive when one requires holes less than 0.76 mm diameter or for drilling slant holes.

Charschan[3] (p. 236) illustrates drilling Al_2O_3 with a ruby laser having $I_0 = 4 \times 10^7$ W/cm^2 and 1.48 ms pulse duration. Sixty percent of the volume of removed material is by explosive expulsion of hot solid, liquid particulate matter, and plasma after subsurface vaporization. To obtain hole diameters greater than 0.25 mm and depths greater than 1 mm, one uses repetitive laser pulses at lower power. For example, a 0.25 mm hole in 3.2 mm thick Al_2O_3 is obtained with 40 pulses of ruby energy 1.4 J, 4 MW/cm^2, 0.5 ms pulse width and 0.2 pps. Another phenomena of interest, when using short focal length lenses, is

the possibility of internal multiple reflections of the incident light within the hole, which increases the depth of drilling. Furthermore, the wall holes are straighter with no refocussing than with refocussing. To improve the hole geometry, high pressure gas can be directed onto the material, which improves the resolidified structure.

Table 4 gives results on drilling with CO_2 lasers, including a few examples on drilling metals. Asmus and Baker[21] also looked at using a cw CO_2 laser to drill 0.25 mm stainless steel plates. The cw or pulsed (including TEA) CO_2 laser is used to drill holes in nonmetals, including fused quartz and silica[8,18] (see Figs. 8 and 9), rubber,[20] alumina (Ref [3], p. 252), glass, plastics, and abrasive materials. [22,23] Glass and plastics are opaque at 10.6 μ but transparent at optical wavelengths, giving a significant advantage to the CO_2 laser. Alumina readily ab-

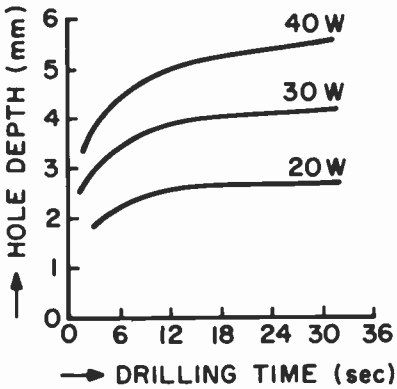


Fig. 8—Dependence[3,18] of hole depth on time drilled with a CO_2 laser for various powers in fused silica.

sorbs (up to 100%) the CO_2 laser radiation after surface breakdown. If the laser pulse is too long, the adjacent surface is also heated and there can occur an excess of resolidified material within the hole. This is eliminated by either reducing the pulse duration or by using an assisting gas jet directed into the hole. Optimum conditions[3] are given in Table 4 for drilling 0.7 mm thick Al_2O_3 with hole diameters 0.25 mm, spaced 1.64 mm apart. The laser is assisted by an air jet. A whole 31-hole Al_2O_3 substrate can be drilled in 116 sec, one hole per pulse, and the time is actually limited by the speed of the numerically controlled x-y table movement. In rubber, a matrix of 31,400 holes can be drilled[20] in only 38 minutes.

Table 4—Experimental Results on Drilling with CO_2 Lasers

Material	Thickness (mm)	Hole Diameter at two ends (mm)	Laser Characteristics	Drilling time per Hole	Holes (min^{-1})	Ref.
Stainless Steel	1.52	0.13—0.23	Not Given	0.3 sec	—	*Laser Focus*, p.26 Oct. 1968
Stainless Steel	0.65	0.15—0.25	1J, 75 W av, repet. pulsed	10 pulses	—	4, p.275
Nickel Base Superalloy	1.78	0.3	16J, 0.8ms pulse width	0.8 ms	120	1
Aluminum Oxide	15	0.075	200—400 W av, 400pps	—	—	17, p.18
Aluminum Oxide	0.7	0.25	2J per pulse, 7.5ms, 0.526pps and air assisted	7.5 ms	31	3, pp.247—252
Aluminum Oxide	0.63	0.2—0.23	75 W av, repet. pulsed	0.2 sec	—	4, p.387
Natural Gum Rubber	0.15	0.09	1 kW peak, 25pps, 60μs	—	826	20
Fused Quartz	5.8	1.6	100 W cw	3 sec	—	8
Mylar	0.254	0.3	Not Given	—	12,000	*Laser Focus*
Acrylic Material	25.4	9.5	Not Given	—	60	p.26, Oct. 1968

The use of high power cw CO_2 lasers for rock tunneling and cutting is another interesting application. The laser energy can penetrate various types of granite, basalt, limestone, dolomite, and concrete (*Optical Spectra*, p. 16, June 1971). If the molten rock forms a vitreous slag, the laser requires a gas assist. Fig. 10 gives some[1] results on boring marble rock.

5. Experimental Results on Microprocessing

5.1 Microdrilling

Interesting applications arise in microdrilling with lasers. The drilling of diamond dies for wire (0.05–1.1 mm diam.) drawing was actually

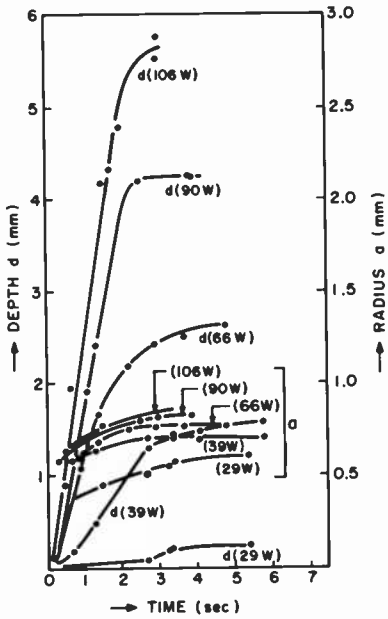


Fig. 9—Depth and radius[9] of CO_2 laser drilled holes in fused quartz as a function of time for various powers with constant focusing over an area of 1.7 mm^2.

the first successful industrial application of solid state lasers[24] (e.g., a 0.5 ms, 3 J ruby laser). The fabrication time per die was reduced from 20 hours to 1 hour and drilling time to less than 10 minutes (*Laser Focus*, p. 12, Aug. 1968). Ruby stones used in timepiece movements

are drilled with Nd/YAG lasers (0.15 J per pulse, 0.1 ms, 20 pps) (see Charschan,[3] p. 228). At BTL,[25] using a ruby laser, holes (25 μm diam.) are drilled in a silicon chip (0.25 mm thick) and 37.5–75 μm holes in 1.58 mm thick ceramic with depth-to-diameter ratios exceeding 25. These holes serve for alignment or for leads.

With the CO_2 laser, the minimum hole size is limited by the wave-

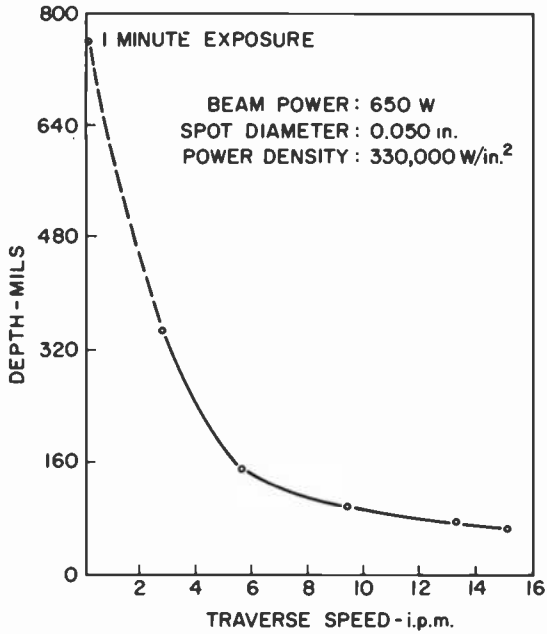


Fig. 10—Effect of CO_2 laser-beam cutting speed on penetration of Georgia marble.[1]

length to around 50 μm. Typically, it is used for drilling holes 0.13–0.26 mm, with tolerances of 13 μm. With the ruby laser, the minimum diameter depends on the material thickness. For example, if the thickness is less than 0.38 mm, the diameter can be as small as 13 μm.

Other applications for microdrilling with lasers include removing imperfections in diamonds, drilling holes for jewel watch bearings, manufacture of contact lenses, electron gun fabrication, and such diverse objects as high nickel alloys for aircraft engines, surgical needles, and rubber nipples.

5.2 Thin-Film Evaporation

In this application, a high energy laser is used to evaporate a material in vacuum, which is then deposited as a thin film on a substrate. The advantages of using lasers are as follows:

[1] The vacuum can be filled to any desired pressure with a gas transparent to the laser radiation, since the laser is outside the vacuum chamber.
[2] Contamination is not present.
[3] Selected areas of the source material can be evaporated.
[4] The distance of deposition from the evaporation region can be close.
[5] Most of the laser energy can be used efficiently for evaporation.

A 20 W cw YAG laser has been used to deposit Pt on Si at a rate of 12 Å/hr with 6-cm separation of source to substrate (Charschan,[3] p. 280). The process would be economical if the rate could be increased to 100 Å/min. Schwartz[26] has succeeded in obtaining deposition rates of 10^5–10^6 Å/sec at 5-cm separation for compounds like Sb_2S_3, ZnS, $SrTiO_3$ and $BaTiO_3$. However, this was achieved with only a single pulse of Nd/glass laser radiation having 80–150 J and 2–4 ms pulse duration, providing a thickness of 2000–3000 Å. Thicknesses for other materials (Cr, W and Ti) obtained with one 2-ms pulse were 500–1000 Å. Sublimation rather than vaporization is thought to produce the vapor for deposition. At RCA Laboratories, it was found that the pulsed ruby laser is good for III–IV compounds, such as GaP and GaAs, due to its high power density (10^8 W/cm^2). Related work has been reported by V. S. Ban of RCA Laboratories.

Groh[27] has used a 25 W CO_2 laser to evaporate films of SiO, ZnS, ZnSe, PbF_2, Na_3AlF_6, and SiO_2 on substrates of Ge, Si, or glass. Deposition rates are between 2000–4000 Å/min at a separation of 7 cm. For a SiO_2 dielectric film that absorbs well at 10.6 μ, a 60–65 W cw CO_2 laser is good. For GaP and GaAs the deposition rate is low, however, and the film is poor.

5.3 Microelectronics

An established application of lasers is for semiconductor integrated-circuit processing and microcircuit adjustments. Nd/YAG, argon ion, or CO_2 lasers are used. High repetition rates and short pulses are usually desirable. Care must be taken to assure that the substrate absorbs much less than the thin film at the particular wavelength used. Use of beam shaping optics also helps. Computer controlled laser devices for microcircuit machining are available.

5.4 Resistor Trimming

The resistance between conductor posts can be increased by reducing the cross-sectional area. This process can be accomplished with lasers by (1) vaporizing small holes (8–40 μm) inside the resistive material, or (2) altering or reshaping the cross-sectional area, (3) oxidizing the surface. This last is the most stable, since holes or cuts, unless large, tend to be subject to temperature drifts or aging. Laser trimming can be done at any stage in the circuit fabrication, even after sealing, and can adjust resistances to within 0.1–1%. Because the focal spot at 10.6 μ is larger than at 1.06 μ, CO_2 lasers are usually applied to thick- rather than thin-film resistor trimming. The initial cost and maintenance requirements are usually less for CO_2 than for Nd-YAG laser trimming systems.

5.5 Scribing

Scribing involves removal of material along a given path to facilitate subsequent fracture, e.g., separation of chips from a block. In order not to affect the adjacent devices, Q-switched or repetitively pulsed lasers are best for vaporizing without excess heating or melting. Lasers have the advantage over diamond scribers of not wearing out. Because of the higher cost of lasers, speed is essential. Both the Q-switched Nd/YAG laser with greater than 10^4 pps rates and the pulsed CO_2 lasers are used, the latter mainly for ceramics which absorb well the 10.6 μ radiation. Silicon cannot be scribed with the CO_2 laser. For scribing silicon (0.25 mm thick) with a Nd/YAG laser (10^9 W/cm^2, 0.1 μs), 70% penetration with 87% overlap of focused spots was the optimum for subsequent fracturing with the least force along the scribed line (Charschan,[3] p. 273). With 20 W, 3 ms, 333 pps of CO_2 laser power, a 0.64 mm alumina substrate can be scribed at a rate of 6 cm/sec. The holes are 0.125 mm deep, 0.1 mm wide and are spaced 0.19 mm apart. The rate can be increased to 10 cm/sec with 50 W power (Charschan,[3] p. 265). Computer-controlled systems are already in operation.

5.6 Controlled Fracturing

Controlled fracturing involves actual separation along designed paths. A laser beam is used to create thermal gradients leading to mechanical stress and fracturing. Fig. 11 gives[22] the minimum power of a CO_2 laser required to fracture alumina, 0.7 mm thick and 11.4 cm long, as a function of parting rate for various widths. Charschan[3] (p. 278) also mentions that single crystals can be cleaved along their pre-

ferred directions (e.g., 0.6-mm-thick, 8-mm-wide quartz) with 40 W of CO_2 energy.

6. Experimental Results on Cutting

Cutting is accomplished with laser beams by vaporizing or sublimating the material. A coaxial gas stream, either reactive for cutting metals, or inert for cutting ceramics, plastics, etc., can be applied to the surface as well. The advantages in laser use are (1) high cutting speeds, (2) improved product quality with smooth edges, and (3) adaptability to automation.

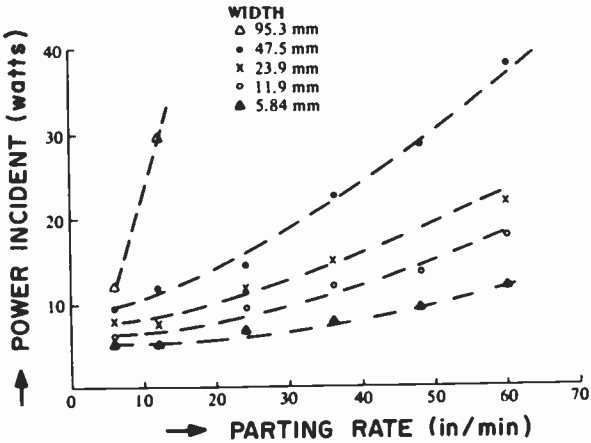


Fig. 11—CO_2 laser power versus parting rate for high alumina ceramic[3,22] (thickness 0.7 mm, length 11.4 cm, and width as indicated).

If the gas is reactive, such as oxygen, it reacts with the laser-heated material and is instrumental both in cutting and in expelling the molten material. Table 5 provides experimental results on laser cutting of metals. Most entries show O_2 gas assistance although some have CO_2 or no gas assistance. Stainless steel up to 4 mm thick has been cut[16] with a 300 W CO_2 laser, whereas titanium has been cut up to 5 cm thickness. For titanium, a laser cutter reduces the processing cost by 75% as compared to a plasma arc cutter. The effectiveness of oxygen assisted cutting, especially for Ti, was first reported[28,29] by the British Welding Research Association. Other metals that can be cut this way are low-carbon steel, iron, magnesium and zircaloy. Metals such as Cu, Ag, and Au have 99% reflectivity and a high thermal con-

Table 5—Metal Cutting with CO_2 Lasers

Material	Thickness (mm)	Cutting Rate (m/min)	Max Kerf Width (mm)	Laser Power (kW)	Gas Assistance	Ref.
321 Stainless Steel	1.52	0.38	—	0.25	O_2	4, p.378
321 Stainless Steel	1.27	0.76	0.51	0.165	O_2	17; 4, p.378; 3, p.283
304 Stainless Steel	4.75	1.27	2.0	20	none	12
302 Stainless Steel	0.76	2.4	—	0.3	O_2	16; 3, p.282
Stainless Steel (?)	44.5	0.38	—	12	none	*Optical Spectra*, p.40, May 1973
Stainless Steel (?)	12.7	0.23	—	3.8	CO_2	13
Carbon Steel C1010	3.2	0.6	1.0	0.19	O_2	3, p.283
Carbon Steel	6.35	2.29	1.0	15	none	12
Carbon Steel	3	1.0	0.5	0.3	O_2	*Product Eng* 39, p.101, 1968
Mild Steel	3.18	0.89	0.5	0.4	O_2	4, p.378
Mild Steel	1.52	1.78	0.5	0.4	O_2	4, p.378
Titanium pure	1.52	2.54	—	0.25	O_2	4, p.378
Titanium pure	0.5	15.2	0.4	0.135	O_2	3, p.283; 17
Ti(6Al.4V)	15	1.2	—	0.3	O_2	16
Ti(6Al.4V)	10	2.5	1.6	0.26	O_2	3, p.283; 17
Ti(6Al.4V)	9	1.0	2.5	0.06	O_2	4, p.378
Ti(6Al.4V)	6.4	2.8	1	0.25	O_2	3, p.283; 17
Ti(6Al.4V)	6	6	—	0.3	O_2	16
Ti(6Al.4V)	2.2	3.8	0.8	0.21	O_2	3, p.283; 17
Ti(6Al.4V)	1.3	7.6	0.8	0.21	O_2	3, p.283; 17
Zircaloy	0.5	15.2	0.5	0.23	O_2	3, p.283; 17
2219 Al Alloy	6.35	0.3	—	3.8	CO_2	13
Al	12.7	2.3	1.0	15	none	12
Inconel 718	12.7	1.27	—	16	none	*Laser Focus*, p.32, May 1973

ductivity, and the CO_2 laser, with several hundred watts, is impractical for cutting them. The results in Table 5 also indicate that in high speed cutting at very high powers (over 10 kW), a gas assist is not required.

If the gas is inert, it only serves to minimize oxidization, purge the expelled material and cool the substance, eliminating thermal shock. With the gas-assisted CO_2 laser, materials that can be cut are ceram-

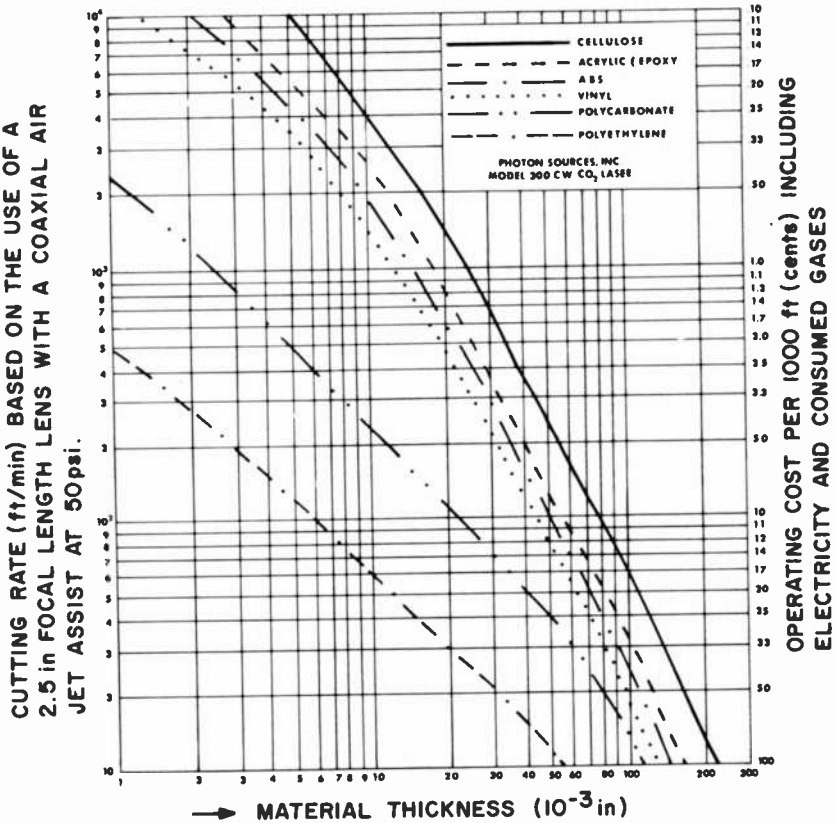


Fig. 12—Cutting rates versus material thickness with a 375 W CO_2 laser (*Photon Sources J.*, Vol. 1, July/Aug. 1971).

ics (e.g., aluminum oxide, beryllium oxide), glass, plastics, wood, cloth, nylon, rubber, wool, paper, fiberglass, leather, asbestos, etc., since these absorb well 10.6 μ radiation. Clean and stress free cuts are obtainable. Alumina cracks, due to thermal shock, unless the gas jet is present. Germanium, GaAs and most salts are however transparent at 10.6 μ. Table 6 and Fig. 12 give examples of CO_2 laser cutting of

Table 6—Cutting of Non-Metals with CO_2 Lasers

Material	Thickness (mm)	Cutting Rate (m/min)	Kerf Width (mm)	Laser Power (kW)	Gas Assistance	Ref.
Boron Fibre Composite	1.3	0.8	0.8	0.26	N_2	3, p.385; 30
Boron Epoxy Composite	8.1	1.65	1.0	15	none	12
Aluminum Oxide	1.4	0.76	0.4	0.2	N_2	17
Aluminum Oxide	0.8	1.27	—	?	N_2	17
Aluminum Oxide	0.6	1.3	0.3	0.25	N_2	3, p.285; 30
Carpet Polyester	9.5	3.0	0.5	0.2	Ar	3, p.285; 17; 30
Carpet Polyester	6.4	3.0	0.5	0.2	Ar	3, p.285; 17; 30
Plexiglass	25.4	1.52	1.52	8	none	12
Plexiglass (Clear)	9	0.25	1.0	0.2	N_2	17
Plexiglass (Dull)	9	0.65	0.8	0.2	N_2	17
Acrylic Plastic	3	2.3	0.5	0.2	N_2	17
Acrylic Plastic	2	3.8	0.6	0.2	N_2	17
Acrylic Plastic	1.59	25.6	—	?	?	4, p.387
ABS Plastic	2.5	3.8	0.8	0.24	N_2	3, p.285; 30
Fiberglass Epoxy Composite	12.7	4.57	0.64	20	none	12
Glass	9.53	1.52	1.0	20	none	12
Concrete	38	0.05	6.35	8	none	12
Limestone	38	0.13	—	3.5	none	13
Borsical Tape	0.23	30.5	—	3.9	none	13
Corrugated Board	4.6	106.7	—	3.9	none	13
Plywood	25.4	1.52	1.52	8	none	12
Plywood	20	0.25	—	0.3	yes	16
Cardboard	1.9	4.8	—	0.25	yes	16
Paper	2.2	3.8	0.6	0.2	yes	17
Paper	0.9	15	—	0.2	yes	17
Paper	0.56	60	—	0.25	yes	16
Paper	0.5	15	—	0.2	yes	17
Paper	Lightweight	300	—	?	yes	4, p.387
Sandpaper	?	305	—	3.9	none	*Laser Focus*, p.14, Apr. 1972

non-metals. We see that no gas assist is necessary for high power lasers above 3.5 kw.

Very high speed cutting of thin materials looks promising with laser powers in the range 0.5–3 kW, whereas thick materials require a 2–6 kW laser. Table 6 shows that a few hundred watts are normally required and lasers in this range are commercially available. However, for plastic and glass cutting at slower rates, 20 W lasers are sufficient and available. For paper cutting, a laser provides the advantages of no wear and no dust. It has the disadvantage that even with a few hundred watts power, the speed of cutting is limited to below that required (800–900 m/min).

Some other practical applications are computerized laser cloth cutting according to programmed patterns (which is revolutionizing this industry) and use of the laser to cut paper and cardboard patterns in the shoe industry.

7. Experimental Results on Other Physical and Chemical Changes Induced by Laser Heat

CO_2 laser light focused on a target can exceed an intensity of 10^9 W/cm^2 and provide temperatures and temperature heating rates (10^6 °C/sec) capable of inducing physical changes in configuration and chemical reactions. Very small particles can be singled out for treatment over times variable from a few nanoseconds to hours. Several applications are listed below.

7.1 Annealing (Softening) and Strengthening

Annealing or metal softening by heat treatment is done to make hardened metals reworkable. Annealing may be accomplished with lasers by applying heat to either annihilate dislocations or change the crystal structure or unstrain the lattice structure. The advantage of lasers lies in that the treated region can be localized.

Conversely, strengthening of steel and other metals (Al or Ti alloys) may be possible with Nd/glass and TEA CO_2 lasers (*Laser Weekly*, April 23, 1973). A pulsed μs laser, upon evaporating a few μm of materials, also imparts momentum and induces a shock wave into the material. This can produce defects and dislocations, thereby hardening the material. Shock hardening of weld zones is one possibility.

7.2 Quenching

At the end of a laser heating pulse, the material can cool with a gas assist at very fast rates, namely 10^3–10^5 °C/sec, unobtainable by other means. Thereby, various metals may be alloyed together to produce various crystal structures or phases and possibly even new compounds.

7.3 Fabrication of Crystals by Zone Melting

Single crystals are formed from their crystalline forms by zone melting. CO_2 lasers have distinct advantages as a source of heat in the availability of ambient atmosphere control, temperature range, and energy concentration and control. They are capable of high efficiency absorption in dielectric materials and, being in the far infrared, they do not affect pyrometer measurements. Gagliano et al[29] describe formation of a sapphire rod crystal, 1.5 mm diam., by using a CO_2 laser for zone melting. Impurities can be reduced by three zone refining passes. It has also been reported (*Optical Spectra*, p. 18, July 1972) that with a 400 W CO_2 laser, monocrystalline fibres of sapphire used to reinforce turbine engine blades to withstand high temperatures (0.05–2 mm diam.) can be drawn at rates up to 5 cm/min. Fibres of titanium carbide and yittrium trioxide have also been grown. This could not be done otherwise because of their high melting temperatures.

7.4 Photolysis, Pyrolysis and Fluorescence Spectroscopy

Photolysis involves the breakup of an excited molecule by light. UV lasers are mainly suitable for this application. Pyrolysis is the decomposition and subsequent analysis of materials. The decomposition can be done with CO_2 lasers. In fact, Pellet and Coper[30] use a low power (10^3 W/cm^2) CO_2 laser to decompose ammonium perchlorate and then, with a mass spectrometer, they measure the time sequence of products in the interaction. Pyrolysis of coal and meteorites, with their gases liberated by lasers, have also been performed. Atomic fluorescence spectroscopy can readily be performed using lasers. Because a pulsed laser has concentrated power in a narrow spectral band, it can excite virtually any element which has lines.

7.5 Isotope Separation by Laser Excitation

This application is of interest in heavy water production, in the sepa-

ration and enrichment of Ur^{235} for fission, in the separation of formaldehyde isotopes, and in the separation of oxygen and nitrogen isotopes. The Ur^{235} process is being investigated by Prof. Moore at the University of California, Berkeley. One method is to excite a single isotope and thereby discriminate between two isotopes of nearly equal weight. This could decrease the cost of Ur^{235} drastically. In the USSR, nitrogen isotopes have been separated with pulses from a TEA CO_2 laser.[31]

7.6 Space-Craft Propulsion by Lasers

Recently, the proposal has been made to use a powerful ground-based CO_2 laser to vaporize the propellant powering a spacecraft, rocket, or interceptor (*Laser Weekly*, p. 2, May 1972). NASA Lewis is exploring the feasibility of this concept.

8. Conclusions

Industrial applications of high power lasers have been reviewed, focusing our attention mainly on the CO_2 laser. First the theory on material processing is elaborated, covering up to eight different solutions of the heat conductivity equation as applied to heat diffusion in metals and the melting and vaporization interfaces. Experimental results on welding, drilling, and cutting are given and tabulated, the latter upon application of a CO_2 laser. A number of other uses are then discussed. It is remarkable to observe how quickly systems become available after the initial feasibility is proven, especially in applications for welding, drilling, cutting, microelectronics, resistor trimming, and scribing.

Acknowledgment

This work was sponsored by the Defence Research Establishment, Valcartier, Quebec, under DSS Contract Serial No. 2GR1-0325, DSS File No. 03GR.CD7020421. Part of this effort was company supported.

References:

[1] C. M. Banas, "The Role of the Laser in Material Processing," United Aircraft, Report Presented at Can. Materials and Processing Tech. Conf., Toronto, Canada, Sept. 29–Oct. 2 (1969).

[2] A. F. Harvey, *Coherent Light*, Wiley-Interscience, N.Y. (1970) p. 60.

[3] S. S. Charschan, editor, *Lasers in Industry*, Van-Nostrand Reinhold Co., N.Y. (1972).

[4] J. F. Ready, *Effects of High-Power Laser Radiation*, Academic Press, N.Y. (1971).

[5] J. F. Ready, "Effects Due to Absorption of Laser Radiation," *J. Appl. Phys.*, **36**, p. 462 (1965).

[6] U-C. Paek and F. P. Gagliano, "Thermal Analysis of Laser Drilling Processes," *Quant. Elec.*, **8**, p. 112 (1972).

[7] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*, Oxford Un. Press, N.Y. (1959).

[8] J. N. Gonsalves and W. W. Duley, "Interaction of CO_2 Laser Radiation with Solids. I. Drilling of Thin Metallic Sheets," *Can. J. Phys.*, **49**, p. 1708 (1971); "II Drilling of Fused Quartz," *Can. J. Phys.*, **50**, p. 216 (1972).

[9] A. K. Ghosh, "Review on High Power Laser Damage to Material," *RCA Review*, **35**, p. 279, June 1974.

[10] M. I. Cohen, "Melting of a Half-Space Subjected to a Constant Heat Input," *J. Franklin Inst.* **283**, p. 271 (1967).

[11] F. W. Dabby and U.-C. Paek, "High Intensity Laser-Induced Vaporization and Explosion of Solid Material," *Quant. Elec.*, **8**, p. 106 (1972).

[12] E. V. Locke, E. D. Hong, and R. A. Helia, "Deep Penetration Welding with High-Power CO_2 Lasers," *Quant. Elec.*, **8**, p. 132 (1972).

[13] C. M. Banas, A. P. Walch, and C. O. Brown, "Material Processing with CO_2 Laser," United Aircraft Rep. UAR-K125, presented at IEEE 11th Symp. on Electron, ion, and Laser Beam Tech., Colorado, May (1971).

[14] C. O. Brown and C. M. Banas, "Deep Penetration Laser Welding," United Aircraft Rep. UAR-K109, presented at Am. Welding Soc. 52nd Annual Meeting, Calif., April (1971).

[15] W. G. Alwang, L. A. Cavanaugh, and E. Sammartino, "Continuous Butt Welding Using a CO_2 Laser," *Welding J.*, **48**, p. 110S (1969).

[16] W. S. W. Tandler, "State of the Art of Laser Machining," Coherent Radiation Report (1971).

[17] W. Ulmer and W. S. Tandler, "Der CO_2-Infrarot Laser als Werkzeug," *Laser und angew. Strahlentechnik No. 1*, p. 17, March (1970).

[18] J. E. Siekman and R. E. Morijn, "The Mechanism of Welding with a Sealed-Off Continuous CO_2 Gas Laser," *Phillips Res. Rep.*, **23**, p. 367 (1968).

[19] E. K. Pfitzer and R. Turner, "Quartz Working with a CO_2 Laser," *J. Sci. Inst. (J. Phys. E.) Series 2*, **1**, p. 360 (1968).

[20] J. Longfellow, "Production of an Extensible Matrix by Laser Drilling," *Rev. Sci. Inst.*, **41**, p. 1485 (1970).

[21] J. F. Asmus and F. S. Baker, "Nonlinear Surface Phenomena Associated with Laser Beam Penetration of Metals," *10th Symp. on Electron, ion, and Laser Beam Tech.*, L. Marton, ed., San Francisco Press, p. 241 (1969).

[22] R. M. Lumley, "Controlled Separation of Brittle Materials Using a Laser," 70th Annual Meeting of Am. Ceramics Soc., Illinois, April (1968); *Ceramic Bulletin*, **48**, p. 850 (1969).

[23] R. Bakish, "Industrial Material Processing with CO_2 Lasers," *Engineering Proc.*, published by Penn. State Un., College of Eng., Pa., p. 46 (1970).

[24] J. P. Epperson, R. W. Dyer, and J. C. Grzywa, "The Laser Now a Production Tool," *Western Electric Eng.*, **10**, p. 2, April (1966).

[25] M. I. Cohen, "Laser Beams and Integrated Circuits," *Bell Lab Record*, **45**, p. 246, Sept. (1967).

[26] H. Schwarz, "Thin Films of Metals and Inorganic Compounds Vacuum Deposited by High Energy Laser," *Laser Interaction and Related Plasma Phenomena*, H. J. Schwarz and H. Hora, editors, p. 71, Plenum Press, N.Y. (1971).

[27] G. Groh, "Vacuum Deposition of Thin Films by Means of a CO_2 Laser," *J. Appl. Phys.*, **39**, p. 5804 (1968).

[28] A. B. J. Sullivan and A. T. Houldcroft, "Gas-Jet Laser Cutting," *British Welding J.*, Aug. (1967).

[29] F. P. Gagliano, R. M. Lumley, and L. S. Watkins, "Lasers in Industry," *Proc. IEEE*, **57**, p. 114 (1969).

[30] G. L. Pellett and W. R. Cofer II, "High Temperature Decomposition of Ammonium Perchlorate Using CO_2 Laser-Mass Spectrometry," paper 69-143, AIAA 7th Annual Meeting, N.Y. (1969).

[31] R. V. Ambartsumyan, V. S. Letokhov, G. N. Makarov, and A. A. Puretskii, "Separation of Nitrogen Isotopes with a Laser," *JETP Lett.*, **17**, p. 63 (1973).

Recent Papers by RCA Authors

Listing is alphabetical by name of primary author. For copies of reprints, the reader should contact the publication directly.

M. S. Abrahams and C. J. Buiocchi, "Cross-Sectional Specimens for Transmission Electron Microscopy," **J. Appl. Phys.**, Vol. 46, No. 1, p. 471, Jan. 1975.

G. A. Alphonse, "Broadband Acoustooptic Deflectors: New Results," **Appl. Optics**, Vol. 14, No. 1, p. 201, Jan. 1975.

K. Ametani, "Atomic Absorption Spectrophotometry of Titanium in the Iron Sulfides of NiAs Type," **Bull. Chem. Soc. Japan**, Vol. 48, No. 3, p. 1047, March 1975.

V. S. Ban, "Mass Spectra of Silicon Containing Industrial Gases," **Mat. Res. Bull.**, Vol. 10, p. 81, Feb. 1975.

A. E. Bell and A. D. Caplin, "Magnetic and Electrical Properties of Very Dilute Zn-3d Alloys," **J. Physics F: Metal Phys.**, Vol. 5, p. 143, Jan. 1975.

D. P. Bortfeld, "Mode-Dependent Retardation in (Ga,Al)P Waveguide Modulators," **IEEE J. Quantum Electronics**, p. 108, March 1975.

P. B. Branin and W. H. Fonger, "Wavelength Dependence of PVA-Phosphor-Dot Photohardening," **J. Electrochem. Soc.**, Vol. 122, No. 1, p. 94, Jan. 1975.

E. E. Carlson, "Anodic Proton Injection in Glasses," **J. Amer. Ceramic Soc.**, Vol. 57, No. 11, p. 461, Nov. 1974.

D. J. Channin, "Liquid-Crystal Optical Waveguides," **RCA Review**, Vol. 35, No. 4, p. 652, Dec. 1974.

T. N. Chin, R. W. Cohen, and M. D. Coutts, "Electronic Processes in Oxide Cathodes," **RCA Review**, Vol. 35, No. 4, p. 520, Dec. 1974.

A. S. Clorfeine, A. Rosen, and J. F. Reynolds, "High-Power Wide-Bandwidth TRAPATT Circuits," **IEEE J. Solid State Circuits,"** Vol. SC-10, No. 1, p. 27, Feb. 1975.

R. W. Cohen, and I. Gorog, "New Results in Image Analysis," **Soc. Photographic Sci. & Eng.**, 2nd Int'l. Conf. on Electrophotography, 1974.

L. S. Cosentino and W. C. Stewart, "A Membrane Page Composer—Further Developments," **RCA Review**, Vol. 35, No. 4, p. 539, Dec. 1974.

R. S. Crandall and B. W. Faughnan, "Measurement of the Diffusion Coefficient of Electrons in WO_3 Films," **Appl. Phys. Lett.**, Vol. 26, No. 3, p. 120, Feb. 1975.

V. L. Dalal, H. Kressel, and P. H. Robinson, "Epitaxial Silicon Solar Cell," **J. Appl. Phys.**, Vol. 46, No. 3, p. 1283, March 1975.

D. A. deWolf, "Waves in Turbulent Air: A Phenomenological Model," **Proc. IEEE**, Vol. 62, No. 11, p. 1523, Nov. 1974.

D. A. deWolf, "Propagation Regimes for Turbulent Atmospheres," **Radio Science**, Vol. 10, No. 1, p 53, Jan. 1975.

J. S. Escher and D. Redfield, "Analysis of Carrier Collection Efficiencies of Thin-Film Silicon Solar Cells," **Appl. Phys. Lett.**, Vol. 25, No. 12, p. 702, Dec. 1974.

M. Ettenberg and H. Kressel, "Heterojunction Diodes of (AlGa)As-GaAs with Improved Degradation Resistance," **Appl. Phys. Lett.**, Vol. 26, No. 8, p. 478, Apr. 1975.

M. Ettenberg, C. J. Nuese, J. R. Appert, J. J. Gannon, and R. E. Enstrom, "Metallurgical and Electroluminescence Characteristics of Vapor-Phase and Liquid-Phase Epitaxial Junction Structures of $In_xGa_{1-x}As$," **J. Electronic Mat.**, Vol. 4, No. 1, p. 37, 1975.

A. H. Firester, D. M. Hoffman, and M. E. Heller, "Application of Multilayer Dielectric Coatings to Cylindrical Substrates," **Rev. Scientific Instruments**, Vol. 46, No. 3, p. 335, March 1975.

H. Fujita and Y. Iguchi, "Reflectivity Measurements near the $L_{2,3}$ Edge of p and n-type Silicon," **Japan. J. Appl. Phys.**, Vol. 14, No. 2, p. 220, Feb. 1975.

H. Fujita and Y. Okada, "Reflectivity Measurements of Single Crystal $CdIn_2S_4$ at Room Temperature," **Japan. J. of Appl. Phys.**, Vol. 13, No. 11, p.1823, Nov. 1974.

B. Goldstein and D. Szostak, "Different Bonding States of Cs and O on Highly Photoemissive GaAs by Flash Desorption Experiments," **Appl. Phys. Lett.**, Vol. 26, No. 3, p. 111, Feb. 1975.

B. Goldstein, "Leed-Auger Characterization of GaAs During Activation to Negative Electron Affinity by the Adsorption of Cs and O," **Surface Science**, Vol. 47, No. 1, p. 143, Jan. 1975.

A. M. Goodman, "A Useful Modification of the Technique for Measuring Capacitance as a Function of Voltage," **IEEE Trans. Electron Devices**, p. 754, Dec. 1974.

A. M. Goodman, "An Investigation of the Silicon Sapphire Interface Using the MIS Capacitance Method," **IEEE Trans. Electron Devices,** Annals No. 502ED008, p. 63, Feb. 1975.

L. A. Goodman, "Liquid-Crystal Displays—Electro-Optic Effects and Addressing Techniques," **RCA Review,** Vol. 35, No. 4, p. 613, Dec. 1974.

J. M. Hammer, H. Kressel, I. Ladany, C. C. Neil, and W. Phillips, "Efficient Modulation and Coupling of CW Junction Laser Light Using ElectroOptic Waveguides," **Proc. IEEE,** Annals No. 502PR013, p. 325, Feb. 1975.

D. Hoffman, "The Structure and Properties of Thin Metal Films," **Proc. 27th Annual Frequency Cont. Symp.,** p. 85, 1974.

D. Hoffman, "Advantage of Modularity in Vacuum Systems," **Research/Development,** Vol. 25, No. 12, p. 41, Dec. 1974.

H. C. Huang, J. D. Knox, Z. Turski, J. Rosen, R. Gurney, "Ultra Wide Bandwidth, Low Loss Bulk Acoustic Delay Lines," **IEDM Tech. Digest,** 1974.

M. Inoue, "Lattice Green's Function for B-Site Lattice in Spinel," **J. Math. Phys.,** Vol. 16, No. 1, p. 111, Jan. 1975.

G. S. Kaplan and F. Sterzer, "Dual-Mode Automotive Collision Avoidance Radar," **Soc. Automotive Eng.,** No. 750087, p. 1, Feb. 1975.

H. Kiess, "Charge Carrier Transport in ZnO," **Soc. Photographic Sci. & Eng.,** 2nd Int'l. Conf. on Electrophotography, p. 124, 1974.

A. G. Kokkas, "Thermal Analysis of Multiple-Layer Strusts.," **IEEE Trans. Electron Devices,** Nov. 1974.

A. G. Kokkas, "Empirical Relationships Between Thermal Conductivity and Temperature for Silicon and Germanium," **RCA Review,** Vol. 35, No. 4, p. 579, Dec. 1974.

I. Ladany and H. Kressel, "The Influence of Device Fabrication Parameters on the Gradual Degradation of (AlGa)As CW Laser Diodes," **Appl. Phys. Lett.,** Vol. 25, No. 12, p. 708, Dec. 1974.

S. G. Liu, "Planar TRAPATT Diodes," **Proc. 1974 IEEE Electron Devices Meeting,** p. 138, 1974.

H. F. Lockwood and H. Kressel, "Multiple Layer (AlGa)As-GaAs Heterojunction Laser Diodes: Synthesis and Mode Control," **J. Crystal Growth,** Vol. 27, p. 97, 1974.

W. J. Merz, "Introductory Lecture," **Ferroelectrics,** Vol. 7, p. 31, 1974.

R. S. Merrich and K. F. Etzold, "System for Visualizing and Measuring Ultrasonic Wavefronts," **RCA Review,** Vol. 35, No. 4, p. 4, Dec. 1974.

R. S. Mezrich, K. F. Etzold and D. H. R. Vilkomerson, "Ultrasonovision," **Ultrasonics Symp. Proc.,** IEEE Cat. #74, CHO 896-ISU, 1974.

H. Nelson, "Liquid-Phase Epitaxy—It's Role in Crystal Growth Technology," **J. Crystal Growth,** Vol. 27, p. 1, 1974.

J. A. Olmstead and R. S. Ronen, "Simple, Efficient, Clean-Steam Generator for Oxidation of Semiconductor Wafers," **Rev. Sci. Instrum.,** Vol. 46, No. 1, p. 110, 1/75.

H. L. Pinch, "The Structure of Co-Sputtered Tungsten-Aluminum Oxide Granular Films," **J. Vacuum Sci. & Tech.,** Vol. 12, No. 1, p. 60, Jan/Feb. 1975.

R. J. Powell and G. W. Hughes, "Charge Injection and Trapping in Al_2O_3 Gate Insulators," **IEEE Trans. Nuclear Sci.,** p. 178, Dec. 1974.

E. B. Priestley, "Introduction to the Optical Properties of Cholesteric and Chiral Nematic Liquid Crystals," **RCA Review,** Vol. 35, No. 4, p. 584, Dec. 1974.

D. Redfield, "Disordered Semiconductors with Controllable Properties," **IEEE Trans. on Parts, Hybrids, & Pkg.,** Vol. 10, No. 4, p. 239, Dec. 1974.

P. H. Robinson and N. Goldsmith, "Silicon Epitaxial Growth Using Dichlorosilane," **J. Electronic Materials,** Vol. 4, No. 2, p. 313, 1975.

A. Rose, "The Relative Performance of Human Vision, Electronic Vision, and Photographic Film," **Soc. Photographic Sci. & Eng.,** 2nd Int'l. Conf. on Electrophotography, 1974.

J. R. Sandercock, "A Light Scattering Study of the Ferromagnet $CrBr_3$," **Solid State Commun.,** Vol. 15, p. 1715, 1974.

W. W. Siekanowicz, H. C. Huang and R. E. Enstrom, "Current-Gain Characteristics of Schottky-Barrier and p-n Junction Electron-Beam Semiconductor Diodes," **IEEE Trans. Electron Devices,** Nov. 1974.

D. L. Staebler, "Oxide Optical Memories: Photochromison and Index Change," **J. Solid State Chem.,** Vol. 12, No. 3–4, p. 177, Jan. 1975.

D. L. Staebler, "Volume Holography: Science and Applications," **Soc. Photographic Sci. & Eng.,** 2nd Int'l. Conf., on Electrophotography, 1974.

D. L. Staebler, W. J. Burke, W. Phillips and J. J. Ametani, "Multiple Storage and Erasure of Fixed Holograms in Fe-doped $LiNbO_3$," **Appl. Phys. Lett.,** Vol. 26, No. 4, p. 182, Feb. 1975.

C. W. Struck and W. H. Fonger, "Unified Model of the Temperature Quenching of Narrow-Line and Broad-Band Emissions," **J. Luminescence,** Vol. 10, p. 1, Feb. 1975.

R. A. Sunshine and R. V. D'Aiello, "Direct Observation of the Effect of Solder Voids on the Current Uniformity of Power Transistors," **IEEE Trans. Electron Devices,** p. 61, Feb. 1975.

A. Sussman, "Electrochemistry in Nematic Liquid-Crystal Solvents," **RCA Review**, Vol. 35, No. 4, p. 600, Dec. 1974.
T. Takahashi, O. Yamada and K. Ametani, "Preparation and Some Properties of Rare Earth Iron Borates, $RFe_3(BO_3)_4$," **Mat. Res. Bull.**, Vol. 10, No. 2, p. 153, Jan. 1975.
M. Toda and S. Osaka, "Temperature-Independent-Time-Delay Surface Acoustic, Wave Device Using a $LiNbO_3$-Bimetallic Plate Structure," **IEEE Trans. Ultrasonics**, Vol. SU-22, No. 1, p. 39, Jan. 1975.
L. C. Upadhyayula, S. T. Jolly, H. C. Huang and B. J. Levin, "High-Efficiency GaAs Impatt Structures," **RCA Review**, Vol. 35, No. 4, p. 567, Dec. 1974.
H. A. Weakliem, D. J. Channin and A. Bloom, "Determination of Refractive Index Changes in Photosensitive Polymer Films by an Optical Waveguide Technique," **Appl. Optics**, Vol. 14, No. 3, p. 560, March 1975.
P. K. Weimer, "Image Sensors for Solid State Cameras," **Advances in Electronics and Electron Physics**, Vol. 37, p. 181, 1975.
C. P. Wen, Y. S. Chiang and E. J. Denlinger, "Multi-Layer Epitaxially Grown Silicon IMPATT Diodes at Millimeter-Wave Frequencies," **J. Electronic Materials**, Vol. 4, No. 1, p. 119, 1975.
W. Wettling, M. G. Cottam and J. R. Sandercock, "The Relation Between One-Magnon Light Scattering and the Complex Magneto-Optic Effects in YIG," **J. Phys. C: Solid State Phys.**, Vol. 8, p. 211, 1975.
R. Williams, "Behavior of Ions in SiO_2," **J. Vac. Sci. and Tech.**, Vol. 11, No. 6, p. 1025, Nov/Dec. 1974.
R. Williams, "Mobile Fluoride Ions in SiO_2," **J. Appl. Phys.**, Vol. 46, No. 2, p. 695, Feb. 1975.
P. J. Wojtowicz, "Lyotropic Liquid Crystals and Biological Membranes: The Crucial Role of Water," **RCA Review**, Vol. 35, No. 4, p. 667, Dec. 1974.

Patents Issued to RCA Inventors First Quarter, 1975

January

T. C. Abrahamsen and J. D. Zaccal Combined Check-Out Counter and Canopy Therefor (D234,037)
B. H. Buckley Rotator System Including a Remote Drive Motor and a Local Indicator-Control Motor (3,860,859)
F. Caprari Optical System (3,860,335)
R. H. Dawson High Frequency Voltage-Variable Capacitor (3,860,945)
F. C. Easter, R. H. Alres, and E. A. Goldberg Regulated Power Supply Including Forward Feed (3,863,140)
M. Ettenberg and S. L. Gilbert Method of Making a Semiconductor Device (3,862,859)
R. A. Gange Permanent Holographic Recording Medium (3,861,914)
J. B. George Video Amplifier Circuit for Use with Synchronous Detectors (3,862,361)
P. E. Haferl Vertical Deflection Circuit (3,863,106)
M. B. Knight, Composite Transistor Circuit (3,863,169)
W. F. Kosonocky and B. F. Williams Dynamic Control of Blooming in Charge Coupled Image-Sensing Arrays (3,863,065)
T. R. Krebs Stable Bonded Barrier Layer-Telluride Thermoelectric Device (3,859,143)
E. P. McGrogan, Jr. Detector for Repetitive Digital Codes (3,863,215)
L. S. Napoli and W. F. Reichert Semiconductor Devices and Methods of Making The Same (3,861,024)
F. Okamoto Etching Solution for Silver (3,860,423)
L. Pessel and A. N. Gardiner Method of Making a Joined Metal Structure (3,862,488)
W. L. Seibert and P. C. Wilmarth Tube Filament Power Supply (3,863,164)
A. H. Sommer Method of Making a III–V Compound Electron-Emissive Cathode (3,858,955)
H. Staras and J. Shefer Harmonic Radar Detection and Ranging System for Automotive Vehicles (RE28,302)
S. A. Steckler Current Output Frequency and Phase Comparator (3,863,080)
H. A. Stern Method of Forming Vitreous Enclosures for Liquid Crystal Cells (3,862,830)
A. J. Stoeckert and J. M. Hunt Semiconductor Mounting Devices Made by Soldering Flat Surfaces to Each Other (3,860,949)
L. N. Thibodeau and L. E. Smith Television Scanning Linearity Device (3,863,184)
J. L. Vossen, Jr. RF Sputtering Apparatus and Method (3,860,507)

February

A. A. A. Ahmed Current Amplifier (3,868,581)
A. A. A. Ahmed Fractional Current Supply (3,867,685)
A. A. A. Ahmed High-Input-Impedance Amplifier (3,864,641)
G. A. Alphonse and G. E. Bodeep Acousto-Optic Devices and Process for Making Same (3,867,108)
W. H. Barkow Inline Electron Gun Having Magnetically Permeable Plates for Enhancing Convergence of Electron Beams (3,866,080)
C. D. Boltz, Jr. Control Signal Apparatus for Video Disc Player (3,864,733)
J. D. Carnes Radiation Sensing Arrays (3,864,722)
J. D. Cavanagh Spacecraft Attitude Control System (3,866,025)
E. H. Del Rio Fingerprint Display System Utilizing a Stored Fingerprint (3,865,488)
J. G. Endriz Grating Tuned Photoemitter (3,867,662)
K. H. Fischbeck Replaceable Fluid Cartridge Including Magnetically Operable Fluid Jet Devices (3,864,685)

K. H. Flechbeck Printing Apparatus (3,864,696)
A. K. Ghosh and H. J. Moody Microwave Spectrometer (3,866,118)
L. Goldman and T. E. Dunn Time Measuring System (3,868,687)
P. E. Haferl Amplifier Which Consumes a Substantially Constant Current (3,868,537)
P. E. Haferl AF Amplifier Having Constant Current Consumption (3,868,582)
J. B. Halter Triangular Piezoelectric Transducer for Recording Video Information (3,865,997)
G. C. Jung Four-Port Junction Circulator Having Larger Diameter Conductive Post Contacting Gyromagnetic Post (3,866,149)
W. M. Lee Spectrally Sensitized Electrophotographic Materials (3,867,144)
R. E. Marks Battery Connecting Assembly (3,864,172)
P. Nyul High Density Light Emitting Diode Array (3,867,666)
J. T. O'Neil, Jr., A. Pellos, A. H. Simon and F. G. Nickl Machine for Reading Article Carrying Coded Indicia (3,864,548)
J. I. Pankova Electroluminescent Semiconductor Display (3,864,592)
J. I. Pankove Method for Diffusing Impurities into Nitride Semiconductor Crystals (3,865,655)
B. E. Patrusky and A. Mack Trunk Formatter (3,868,481)
W. Phillips Novel Lithium Niobate Single Crystal Film Structure (3,867,012)
A. E. Ruvin Doppler Tracker Receiver (3,866,227)
O. H. Schade, Jr. Matching of Semiconductor Device Characteristics (3,863,331)
A. Schwartzmann, V. A. Mikenas, J. J. Thomas and K. K. N. Chang Microwave Oxcillator or Amplifier Using Parametric Enhanced TRAPATT Circuits (3,868,588)
P. R. Scott, Jr. Motion Damping Apparatus (3,868,557)
T. M. Shrader Cathode-Ray tube Having an Internal-External Magnetic Shield and Degaussing Combination (3,867,668)
P. K. Weimer Charge-Transfer Display System (3,866,209)
P. K. Weimer Circuit for Amplifying Charge (3,867,645)
R. Williams and A. H. Willis Method for Recording Fingerprints (3,867,165)
J. E. Wojslawowicz Monostable Switching Circuit (3,867,651)
R. W. Young Electron Discharge Device Having a Hollow Conductor Integral with the Envelope Thereof (3,864,590)

March

M. E. Adams and R. Feryszka RF Switching Circuit (3,872,325)
A. A. A. Ahmed Comparator-Keyed Oscillator (3,873,853)
J. G. Amery and R. W. Jorgenson Signal Translating Apparatus for Composite Signal Subject to Jitter (3,872,497)
E. F. Belohoubek and A. Presser Microwave Transistor Carrier for Common Base Class A Operation (3,869,677)
J. P. Bingham Line Sequential Color Television Recording System (3,871,019)
C. D. Boltz, Jr. Speed Correction System for a Video Disc Playback System (3,873,764)
W. J. Burke Lithium Niobate Hologram Readout Using Continuous Incoherent Noise Erasing Light (3,873,179)
D. J. Carlson and S. E. Hilliker Pickup Apparatus for Capacitive Video Disc Players with Transmission Line Tuned Circuit (3,872,240)
K. K. N. Chang Enhanced Efficiency Diode Circuit (3,875,535)
A. G. F. Dingwall Complementary Field Effect Transistor Differential Amplifier (3,870,966)
G. B. Dodson III Radiation Sensing Circuit (3,872,329)
R. S. Eggelbrecht Method of Electrostatic Recording on Electrically Insulating Films by Non-Wetting Electrically Conductive Liquids (3,872,480)
R. D. Faulkner Electron Discharge Device Including an Electron Emissive Electrode Having an Undulating Cross-Sectional Contour (3,875,441)
J. J. Gibson and H. G. Schwarz Apparatus and Method for Measuring the Signal to Noise Ratio for a Periodic Signal (3,875,328)
H. H. Girvin Support and Focus Structure for Photomultiplier (3,873,867)
G. D. Hanchett Liquid Crystal Protection Circuit (3,873,888)
R. E. Hanson and H. E. Fineman Starter System Fault Detector (3,870,954)
S. E. Hilliker Video Disc Transmission Line and Stylus RF Return Systems (3,872,265)
R. H. Hughes In-Line Electron Gun (3,873,879)

P. B. Korda Recovery of Normally Illegible, Magnetically Recorded Information (3,869,721)
M. A. Leedom Rotary Stylus Reshaper (3,873,098)
M. A. Leedom Universal Stylus and Pivot Coupling (3,873,783)
D. E. Mahoney Multiple Transistor Microwave Amplifier (3,869,678)
F. J. Marlowe Digital Storage Tube Target Structure (3,873,873)
R. C. Palmer Capacitive Protection Coupling Element for Stylus Electrode Discharge (3,873,782)
S. S. Perlman and J. H. McCusker Signal Conversion Circuits (3,873,906)
D. H. Pritchard Color Information Translating Systems (3,872,498)
C. A. Rypinski and J. E. Pera Circuit for Applying Data Signals Across a Microphone Input Circuit (3,870,958)
J. C. Schopp and F. R. Stave Disc Playback System with Speed Control of a Belt Drive (3,873,765)
F. R. Stave Stylus Control Apparatus for a Video Disc Record Player (3,870,835)
F. R. Stave Record Contact Spindle (3,871,663)
B. K. Taylor Spring Loaded Stylus Arm Pivot (3,873,762)
R. D. Thompson Cathode-Ray Tube with Radiation-Emitting Index Strip-Like Areas (3,875,328)
L. A. Torrington Pickup Arm Control for Video Disc Player (3,870,320)
P. K. Weimer Charge Transfer Decoders (3,873,851)
C. F. Wheatley, Jr. Circuit with Adjustable Gain Current Mirror Amplifier (3,873,933)
R. W. Widmer Method for Producing Aluminum Holographic Masters (3,875,026)
J. A. Wilber Chrominance Signal Correction (3,871,020)
J. A. Wilber Chrominance Signal Correction (3,873,990)
D. J. Woywood and S. L. Corsover Electronic Technique for Making Multichannel, Spatial-Carrier-Encoded Recordings (3,869,705)

AUTHORS

R. Casanova Alig received the B.A. degree from Wabash College in 1963, and the M.S. and Ph.D. degrees in physics from Purdue University in 1965 and 1967. He did theoretical studies of acoustic transport phenomena in metals for his dissertation. In 1967 Dr. Alig joined RCA Laboratories as a member of the Physics and Chemistry of Solids Group in the Physical Electronics Laboratory. He has recently done theoretical studies of the optical spectra of impurities and color centers in crystals. During the academic year 1970–71 he was a visiting professor at the Instituto de Fisica e Quimica, Sao Carlos, S.P., Brasil; in addition, since 1967 he has given courses at Mercer County Community College, LaSalle College, and Princeton University. He is a member of Phi Beta Kappa, Sigma Xi, and the American Physical Society.

Gerard A. Alphonse attended New York University where he obtained a B.S.E.E. in 1958 and a M.S.E.E. in 1959, after doing research in the field of superconductivity. He joined the RCA Laboratories in June 1959 and worked toward the development of superconductive memories. In 1963 he attended the Polytechnic Institute of Brooklyn under an award from RCA Laboratories, obtaining his Ph.D. in Electrophysics in 1967. He has done research on photochromic memories, acousto-optics, and more recently on holographic storage in photorefractive media. Dr. Alphonse received an Outstanding Research Award from RCA in 1960 for his work in superconductivity. Besides his work at RCA, he is a physics instructor in the Evening Division of La Salle College in Philadelphia.

He is a senior member of the IEEE, a member of Eta Kappa Nu, Tau Betta Pi, the American Institute of Physics, and the Optical Society of America.

Raymond L. Camisa received his B.E.E., M.E.E., and Ph.D. degrees from the City College of the City University of New York in 1965, 1969, and 1975, respectively. From 1965 to 1967 he was a member of the technical staff at the RCA Advanced Communications Laboratory, New York. While there he worked on microwave filters, low-noise parametric amplifiers, and microwave integrated circuit techniques. From 1967 to 1970 he was with Wheeler Laboratories, Inc., Great Neck, L.I., where he was part of a group developing a microwave integrated circuit receiver for IFF applications. Specifically, he developed various MIC components including low-noise transistor amplifiers, frequency multipliers and filters. From 1970 to 1974 he was a part time lecturer at the City University of New York teaching courses in electromagnetic theory, electronics, and microwave measurements. At the University he also worked as a graduate research assistant investigating the use of MIS varactors in microwave networks. During the same period, he was a consultant to Wheeler laboratories and the RCA Advanced Communications Laboratory. In 1974 Dr. Camisa joined the Microwave Technology Center at RCA Laboratories, Princeton, N.J. His responsibilities include research on GaAs field-effect-transistor device technology and linear amplifier development.

Walter R. Curtice received the B.E.E., M.S., and Ph.D. degrees from Cornell University, Ithaca, New York, in 1958, 1960, and 1962, respectively. He joined the Microwave and Power Tube Division of Raytheon Co. in 1962 as a Senior Research and Development Engineer. In August of 1967, he became Visiting Assistant Professor of Electrical Engineering at the University of Michigan and was appointed Associate Professor in August 1969. He became a Member of the Technical Staff of the RCA Laboratories' Microwave Technology Center, Princeton, N.J., in January of 1973, where he is participating in programs on transferred electron devices and pulsed Trapatt devices.

Dr. Curtice is a Senior Member of the IEEE and a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.

Raymond H. Dean received his B.S. in E.E. with highest distinction from the University of Kansas in 1958. From 1958 to 1960 he monitored the design and development of electro-mechanical fusing and firing components as a technical liaison officer at the Defense Atomic Support Agency in Albuquerque, New Mexico. In 1960 he entered MIT on a N.S.F. fellowship, and one year later he received his M.S. in E.E. For the next two years he was engaged as a professional engineer in the design of high-velocity air conditioning systems for commercial applications. In 1963 he joined the staff at Midwest Research Institute and participated in research on MnBi magnetic thin films. The following year he entered Princeton University on an N.S.F. fellowship, majoring in solid-state device physics and minoring in plasma physics. His PhD dissertation, completed in June, 1968, treated the transient double-injection of an electron–hole plasma in germanium. At RCA, he participated in research on GaAs bipolar transistors and in the development of a solid-state traveling-wave amplifier fabricated on a thin epitaxial layer of n-type GaAs. Dr. Dean left RCA in May 1974. He is a member of the IEEE and Sigma Xi.

George M. Ehemann, Jr., received the BE in Engineering Physics from Cornell University in 1964. That year he joined RCA's Chemical and Physical Laboratory at Lancaster, Pa., where he has been working on color measurement techniques and optical property studies of color television screens. In 1965, while on leave of absence he earned the ME degree in Engineering Physics at Cornell University.

Morris Ettenberg received a B.A. and M.S. in Education from City College of New York in 1935 and 1936, respectively, and a Ph.D. degree in Physics from New York University in 1950. After working as a radar engineer at the U.S. Naval Shipyard in Brooklyn during World War II, he joined the Electronic Tube Division of the Sperry Gyroscope Company doing research and development on Klystrons and traveling wave tubes. In 1958 he went to the Polytechnic Institute of Brooklyn and during the academic year 1961–62 he held a Fulbright Lectureship as visiting professor at the Technion–Israel institute of Technology in Jaifa, Israel. In 1963 he joined the faculty of the Electrical Engineering Department as Professor at the City College of New York. His current research interests are in microwave circuits and solid state devices.

Benjamin F. Hitch received his B.S.E.E. degree from the University of Tennessee in 1967, and joined RCA's Communication Systems Division (CSD) the same year. From 1967 to 1970, Mr. Hitch worked primarily on HF and power amplifiers and related equipment. Since 1970 his work has been in the microwave integrated circuits area. In 1972 he was co-recipient of a CSD technical excellence award for his work on the SHF Small Sattelite Terminal Program. For the last two years he has worked on the applications of MIS devices, especially to phase shifters.

Pang-Ting Ho received his B.S. degree in Electrical Engineering from National Taiwan University in 1967 and his M.S. degree in Electrical Engineering from Princeton University in 1969. He is working toward his Ph.D. degree at Rutgers University. In 1969, Mr. Ho joined the Microelectronic Technology group of RCA in Somerville, N.J. He was primarily concerned with the characterization of microwave power transistors and the design of MIC power amplifiers for the phased-array radars. He also engaged in the design of microstrip directional couplers, filters, and broadband impedance transformers. In 1972 he transferred to the Microwave Technology Center, RCA Laboratories, Princeton, N.J., where he has been working on S-band, X-band Trapatt Oscillators, and amplifiers.

Mr. Ho is a member of IEEE and Eta Kappa Nu.

Sheng T. Hsu received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1958; the M.S.E.E. degree from National Chiao-Tung University, Hsienchu, Taiwan, in 1960; and the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, in 1966. From 1966 to 1970 he was with Fairchild Semiconductor Research and Development Laboratory, Palo Alto, California, as a Member of the Technical Staff. From 1970 to 1972 he was an Assistant Professor of the Department of Electrical Engineering, University of Manitoba, Winnipeg, Man., Canada. Since 1972 he has been a Member of the Technical Staff at RCA Laboratories, Princeton, N.J., where he is doing research on semiconductor devices.

Henry Kressel received the B.A. degree in 1955 from Yeshiva University; the M.S. in 1956 and the M.B.A. in 1959 from Harvard University; and the Ph.D. in 1965 from the University of Pennsylvania. He has been with RCA since 1959.

As a member of the RCA Solid State Division he contributed to the development of high frequency planar silicon transistors and led a group responsible for the development of high power varactor diodes subsequently used for the Lunar Excursion Module Communication System. He joined the RCA Laboratories, Princeton, N.J., in 1966 and became Head of the Semiconductor Optical Devices Research Group in 1969. He pioneered in the field of (AlGa)As-GaAs heterojunction devices, in particular laser diodes, and has been actively engaged in the study of devices and luminescent processes in various III-V compound materials. He is the recipient of three RCA Research Achievement Awards as well as a 1974 David Sarnoff Medal for Outstanding Technical Achievement, RCA's highest award. He is presently Head, Semiconductor Device Research, with responsibility in the area of silicon power and III-V compound devices. Dr. Kressel is a Fellow of the IEEE, and a Fellow of the American Physical Society./W

Ivan Ladany obtained his B.S. and M.S. degrees from Northwestern University. In 1953, he joined the Naval Research Laboratory in Washington, D.C., where he worked briefly in the field of underwater sound, spending most of his time in studies of p-n junctions and junction devices. Since coming to RCA Laboratories in 1966, he has worked in GaP, GaAlP and GaAs luminescent diode research. In 1969 he was awarded an RCA Laboratories Achievement Award for his contributions to GaP electroluminescence. More recently he has worked on improved GaS infrared diodes, green emitting GaP LED's and III-V compound growth on insulating substrates using liquid phase epitaxy. At present, he is devoting most of his time to injection laser development.

Mr. Ladany is a member of the American Physical Society, the IEEE and Sigma Xi.

Shing-Gong Liu received his B.S. degree in electrical engineering from Taiwan University, Taipei, Taiwan, in 1954, the M.S. degree in electrical engineering from North Carolina State College, Raleigh, North Carolina, in 1958, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, Calif., in 1963. From 1958 to 1959 he worked with the IBM Laboratories, Poughkeepsie, New York. From 1959 to 1963, he was a research assistant at the Hansen Microwave Laboratories, Stanford University, where he worked in the field of microwave ferrites. He joined RCA Laboratories, Princeton, N.J., in 1963, and has since worked principally in the areas of semiconductor microwave devices.

Dr. Liu is a member of Phi Kappa Phi, Sigma Xi, and the American Physical Society.

Louis S. Napoli received his B.S., in 1959 and the M.S. in 1961 in Electrical Engineering, both from Rutgers University. He has pursued further studies in plasma physics at Princeton University and in electro-physics at the Polytechnic Institute of Brooklyn. Since joining the technical staff of RCA Laboratories in 1959, he has specialized in research relating to microwave phenomena in gaseous plasmas, solid-state microwave devices, and microwave integrated circuits. He is presently Head of the Microwave Components Technology Group in the Microwave Technology Center at RCA Laboratories. His work in collaboration with Dr. George Swartz on amplification at 24 GHz by the interaction of an electron beam with a cesium plasma was cited by Industrial Research Magazine as one of the 100 most important achievements in 1963. He is the recipient of RCA Laboratories Achievement Awards in 1963, 1965, and 1968.

Mr. Napoli is a member of Sigmi Xi, Tau Beta Pi, and Eta Kappa Nu, and IEEE.

William Phillips received an A.B. degree in physics from Columbia University in 1958. He did his graduate work at the Carnegie-Mellon University, receiving an M.S. degree in physics in 1961 and a Ph.D. in electrical engineering in 1964. His thesis research was concerned with the crystal growth and optical properties of potential laser materials. Dr. Phillips joined the RCA Laboratories in 1964, and from 1964 to 1966 was engaged in research on the synthesis and properties of crystalline materials for laser applications. In 1966 he became involved in work on the synthesis and evaluation of improved photochromic materials, particularly rare-earth-doped calcium fluoride, and sodalite, and in 1970 he became involved in research on improved materials for the storage of volume phase holograms, notably iron-doped lithium niobate. He is currently working on diffused lithium niobate-tantalate films for integrated optical waveguide applications.

Dr. Phillips has received four RCA Laboratories Achievement Awards and shared in an IR-100 Award for development of the cathodochromic CRT. He is a member of the American Physical Society, the IEEE, and Sigma Xi.

John J. Risko received his B.S.E.E. from Newark College of Engineering in 1972 and is currently continuing his studies there toward a masters degree in electrical engineering. He joined the Microwave Research Laboratory at RCA Laboratories Princeton, N.J., in 1962, where he was assigned to the microwave semiconductor device group working on GaAs and GaSb tunnel diodes and varactor Diodes. From 1965 to date he has worked on silicon devices, including Impatts, Trapatts, and more recently Baritt structures. He became a member of the technical staff at RCA Laboratories in 1972 upon completion of his work toward the B.S.E.E. degree.

Mr. Risko received a David Sarnoff Research Laboratories Achievement Award for "team effor leading to the development of S-Band Trapatt Amplifiers" in 1972.

Issie P. Shkarofsky graduated in 1952 from McGill University, Montreal with a B.Sc. degree and first class honours in physics and mathematics. In the following year, 1953, he obtained his M.Sc. degree, conducting his research at the Eaton Electronics Research Laboratory, McGill University, in the fields of microwave optics and antennas. He then joined the microwave tube and noise group at the Eaton Electronics Research Laboratory, McGill University, and received his Ph.D. degree in 1957 with a thesis on modulated electron beams in space-charge-wave tubes and klystrons. After grduation, he joined the Research Labratory of RCA Limited, Montreal, where he has participated in research on microwave diffraction, millimeter waves, obstacle gain, laser applications, electromagnetic wave interaction with plasmas, plasma kinetics, and plasmas in space. In 1973 he became a R&D fellow at the RCA Research and Development Laboratories, presently in the Physical Electronics Laboratory. In the field of plasma and space studies his particular interest has been in the following topics: plasma transport coefficients, collisional effects in plasmas (slightly, partially and strongly ionized), Boltzmann and Fokker-Planck theory and appropriate expansions, bremsstrahlung, magnetohydrodynamics, re-entry plasma physics, generalized Appleton-Hartree equation for the ionosphere, laboratory simulation of geophysical phenomena such as the interaction between the solar wind and the magnetosphere and the sheaths around satellites, cyclotron harmonic resonances and related dispersion effects, diagnostics of plasmas by laser scattering, nonlinear mixing of plasma modes, the accuracy of Langmuir probes on satellites, VLF sheath admittance of antennas in the ionisphere, high power laser interaction with matter, turbulence and fusion.

Dr. Shkarofsky is a member of the Canadian Association of Physicists, of the American Physical Society, the American Geophysical Union, and of the National Research Council Advisory Committee on Physics. He is a professeur invité du Centre de Recherche INRS-Energie, Un. du Québec, and he is also the co-leader of the Task Force on Toroidal Machines for Project Fusion Canada.

David L. Staebler received his B.S.E.E., with distinction, and M.S.E.E. degree from the Pennsylvania State University in 1962 and 1963, respectively, and his Ph.D. in Electrical Engineering from Princeton University in 1970. His doctoral research involved optical studies of photochromic CaF_2, and resulted in the identification of an important class of photochromic color centers in CaF_2. He joined RCA Laboratories in 1963, and worked with narrow-band phosphors, laminated ferrite memories, and injection diode lasers. In 1964 he investigated oxidation-reduction processes in gamma-irradiated rare-earth-doped fluorides, and in 1965 he began studies

of photochromic behavior in CaF_2. Since receiving his degree in 1970, Dr. Staebler has worked on electric field coloration of transition-element doped titanates, and on the study of electro-optic materials for holographic storage, particularly iron-doped lithium niobate. He is presently with the Instituto de Fisica e Quimica de Sao Carlos, Brasil, on a one year leave of absence from RCA Laboratories. He has received RCA achievement awards for team research in photochromic materials (1967) and in hologram storage media (1972).

Dr. Staebler is a member of Eta Kappa Nu, Sigma Tau, Tau Beta Pi, and the IEEE.

Joseph H. Scott, Jr., received his A.B. degree in chemistry from Lincoln University in 1957. The following year he attended the Graduate School of Chemistry at Howard University, and he has since done graduate work in Electrical Engineering at Howard. Mr. Scott joined the RCA Solid State Division in Somerville, N.J. in 1959, doing research and development work on semiconductor devices. In 1967, he transferred to RCA Laboratories, Princeton, N.J., where in 1970, he was named Head of the Integrated Circuit Technology and Application Research Group. In 1974, he was appointed Director, Integrated Circuit Technology. Mr. Scott received an RCA Laboratories Outstanding Achievement Award in 1967, and in 1969 he received two Outstanding Achievement Awards in the same year for work on two separate research teams. In 1973, he was given the David Sarnoff Award for "outstanding team research leading to a new class of integrated semiconductor arrays."

Mr. Scott is a member of Sigma Xi and the Electrochemical Society and a senior member of the IEEE.

Fred Sterzer received his B.S. degree in Physics from the College of the City of New York in 1951, and his M.S. and Ph.D. degrees in Physics from New York University in 1952 and 1955, respectively. From 1952 to 1953, he was employed by the Allied Control Corporation in New York. During 1953 and 1954, he was an instructor in Physics at the Newark College of Engineering, Newark, N.J., and a research assistant at New York University. Dr. Sterzer joined RCA in 1954 and is now Director of the Microwave Technology Center of RCA Laboratories, Princeton, N.J. His work has been in the field of microwave spectroscopy, microwave tubes, light modulators and demodulators, and microwave solid state devices.

Dr. Sterzer is a member of Phi Beta Kappa, Sigma Xi, the American Physical Society, and is a Fellow of the IEEE.

S. Yuan received the BSEE from the University of California at Berkeley in 1952 and the MSEE from Columbia University in 1956. From 1953 to 1955, he worked for RCA on the development of broadcast equipment, and from 1957 to 1963 he was a senior engineer with the electronics division of Curtis Wright Corporation and a member of the electronic research laboratory staff at Columbia University. From 1963, when he rejoined RCA, to 1966, his work was in the area of solid-state rf techniques. He became a Leader, Advanced Communications Laboratory, Government Communications and Automated Systems Division, Somerville, N.J., in 1966, with responsibility for the development of circuits such as uhf amplifiers, uhf/vhf voltage-controlled oscillators, phase comparators, and crosspoints using beam-lead diodes for switching matrices. He is the co-inventor of a balanced mixer circuit that can alleviate high-order intermodulation and crossmodulation distortion. His work on frequency multipliers has resulted in a theoretical analysis for predicting the efficiency of an XB multiplier as a function of the nonlinear coefficient and in an experimental model of the multiplier which yields a 20% efficiency "octupier" (500 MHz to 4 GHz).

Mr Yuan is a member of the Chinese institute of Engineers, Tau Beta Pi, Eta Kappa Nu, and Phi Tau Phi.

STATEMENT OF OWNERSHIP

Statement of Ownership, Management and Circulation (Act of October 23, 1962. Section 4369, Title 39, United States Code.)

1. Date of filing: September 25, 1974. 2. Title of publication: RCA REVIEW. Frequency of issue: Quarterly. 4. Location of known office of publication: RCA Laboratories, Princeton, New Jersey, 08540. 5. Location of headquarters or general business offices of the publishers: RCA Laboratories, Princeton, New Jersey, 08540. 6. Names and addresses of publisher and editor: Publisher, RCA Laboratories, Princeton, New Jersey, 08540. Editor, R. F. Ciafone, RCA Laboratories, Princeton, New Jersey, 08540. 7. Owner: RCA Corporation, 30 Rockefeller Plaza, New York, New York 10020; Cede & Company, Box 20, Bowling Green Stn., New York, New York 10004; Merrill, Lynch, Pierce, Fenner & Smith, Inc., 70 Pine Street, New York, New York 10005; Cudd & Company, c/o Chase Manhattan Bank, Box 1508, C.S.S., New York, New York 10008; Martin B. Seretean, c/o Coronet Industries, Box 1248, Dalton, Georgia 30720; Burl Jackson Bandy, c/o Coronet Industries, Box 1248, Dalton, Georgia 30720; TNOM & Company, Box 577, Central Station, St. Louis, Missouri 63188.

I certify that the statements made by me above are correct and complete.

Ralph F. Ciafone, Editor