RCA Review

September 1969 Volume 30 No. 3

PRINTED IN U.S.A.

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

RCA REVIEW

a technical journal

Published quarterly by
RCA RESEARCH AND ENGINEERING
in cooperation with all subsidiaries and divisions of the
RCA CORPORATION

VOLUME 30 SEPTEMBER 1969 NUMBER 3

CONTENTS

PAGE

A Theory for the High-Efficiency Mode of Oscillation in Avalanche Diodes 397
A. S. CLORFEINE, R. J. IKOLA, AND L. S. NAPOLI

Characteristics of a Sealed-Off He^3-Cd^{114} Laser 422
J. R. FENDLEY, JR., I. GOROG, K. G. HERNQVIST, AND C. SUN

Low-Radiation-Noise He-Ne Laser 429
K. G. HERNQVIST

The Acoustoelectric Effects and the Energy Losses by Hot Electrons—Part IV Field and Temperature Dependence of Electronic Transport 435
A. ROSE

Ionospheric Phase Distortion and Faraday Rotation of Radio Waves .. 475
T. MURAKAMI AND G. S. WICKIZER

Spectral Analysis of Turbulent Wakes 504
D. A. DEWOLF

Bistatic Clutter in a Moving Receiver System 518
E. G. MCCALL

Miniature Microstrip Circulators Using High-Dielectric-Constant Substrates 541
B. HERSHENOV AND R. L. ERNST

RCA Technical Papers 544

Patents 546

Authors 549

RCA REVIEW is regularly abstracted and indexed by *Abstracts of Photographic Science and Engineering Literature, Applied Science and Technology Index, Bulletin Signalétique des Télecommunications, Chemical Abstracts, Electronic and Radio Engineer, Mathematical Reviews, and Science Abstracts* (I.E.E.-Brit.).

RCA REVIEW

BOARD OF EDITORS

A THEORY FOR THE HIGH-EFFICIENCY MODE OF OSCILLATION IN AVALANCHE DIODES*

By

A. S. Clorfeine, R. J. Ikola, and L. S. Napoli

RCA Laboratories
Princeton, N. J.

Summary—High-efficiency oscillations, previously observed in avalanche diodes and simulated with a computer, are explained by means of an analytic theory in which all the relevant physical processes are clearly displayed and tied together. It is shown how efficiencies of 50-60% can be achieved.

On the basis of the theory, design formulas are derived. Relations involving the choice of width and doping density of the active layer, diode area, bias current density, and fundamental circuit impedance are established. Predictions of the theory are seen to be in good agreement with reported experimental results.

I. Introduction

THIS PAPER presents a theory for the high-efficiency mode of oscillation in avalanche diodes. Since the theory is analytic, all the relevant physical processes are clearly displayed and tied together. It is demonstrated how efficiencies of 50-60% can be achieved and how oscillators can be designed for optimum performance.

The high-efficiency mode was discovered in silicon by Prager, Chang, and Weisbrod.[1] The mode was considered "anomalous" in that, first, the oscillation efficiencies were markedly higher than had been previously observed with avalanche diodes and, second, the oscillation frequencies were considerably lower than could be expected from

* This work was supported by the Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Ohio, under AF Contract No. F33615-68-C-1688, Project No. 4460.

[1] H. J. Prager, K. K. N. Chang, and S. Weisbrod, "High-Power, High-Efficiency Silicon Avalanche Diodes at Ultra-High Frequencies," *Proc. IEEE* (Letters), Vol. 55, p. 586, April 1967.

transit-time theory.[2,3] Subsequently, Johnston, Scharfetter, and Bartelink observed similar oscillations with germanium diodes[4] and were able to simulate these oscillations with efficiencies of up to 26%[5] on a computer. The computer simulations provided insight in two respects. First, they demonstrated that the high-efficiency mode could be achieved without, for example, such effects as filament formation or some fortuitous doping irregularity; also, the simulations revealed the existence of a high-density, slow-moving electron–hole plasma during part of the oscillation cycle.

What was still required, however, was a theory—one that could overcome the limitations of a simulation approach. Such limitations include (a) the need to first establish oscillations in the laboratory so that parameters may be supplied to the computer, (b) the substantial computer time required, (c) the obscuring of the relevant physics by waveform properties that are germane only to the particular nonoptimum diode-circuit combination being simulated, and, finally, (d) the difficulty in applying the simulation procedure to the task of designing and optimizing oscillators.

The approach taken in the present paper is, first, to develop a detailed analytic theory of the oscillation physics, and then to consider the question of what diode and circuit parameters are necessary to obtain these oscillations. The answers are seen to follow directly from the analytic work. More specifically, Section II gives a detailed description of the formation of a plasma in an avalanche diode (the plasma is an important ingredient in the theory, in contrast to Hoefflinger's postulated operation[6]). The analytic results obtained for the plasma formation are then used to calculate the response of the diode to certain applied currents (Section III). It is shown that much of the important physics for high-efficiency oscillations follows from a consideration of the diode response to a step current. Next, the criteria for a successful model of the high-efficiency mode are discussed and waveforms consistent with these criteria are introduced and analyzed. The circuit impedances necessary to obtain these waveforms must, of course, be realizable; these can be calculated. Finally, in Sections IV

---

[2] W. T. Read, Jr., "A Proposed High-Frequency Negative-Resistance Diode," *Bell System Tech. Jour.*, Vol. 37, p. 401, March 1958.

[3] H. K. Gummel and D. L. Scharfetter, "Avalanche Region of Impatt Diodes," *Bell System Tech. Jour.*, Vol. 45, p. 1797, Dec. 1966.

[4] R. L. Johnston, D. L. Scharfetter, and D. J. Bartelink, "High-Efficiency Oscillations in Germanium Avalanche Diodes Below the Transit-Time Frequency," *Proc. IEEE* (Letters), Vol. 56, p. 1611, Sept. 1968.

[5] D. L. Scharfetter, "High Efficiency Operation of Impatt Diodes," International Electron Devices Meeting, Washington, D. C., Oct. 1968.

[6] B. Hoefflinger, "Recent Developments on Avalanche Diode Oscillators," *Microwave Jour.* Vol. 12, p. 101, March 1969.

and V, design formulas are presented and predictions of the theory are summarized. These predictions are seen to be in good agreement with reported experimental results.

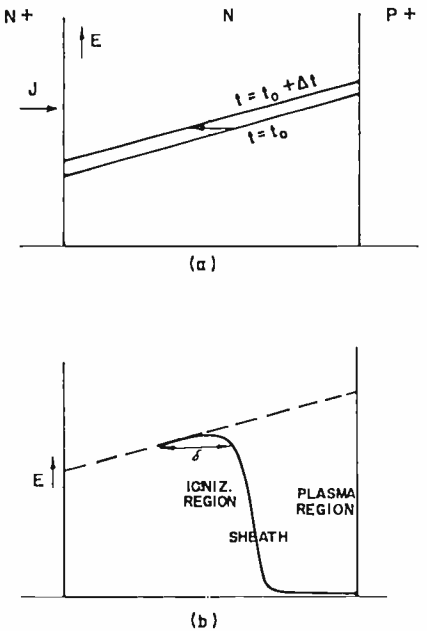


Fig. 1—Plasma formation in avalanche diodes: (a) electric field prior to ionization and (b) electric field subsequent to ionization.

II. Plasma Formation in Avalanche Diodes

A. Wave Propagation

Figure 1(a), in order to be specific, illustrates an abrupt junction p+-n-n+ silicon diode. The results, however, to a large extent, should be applicable to a graded junction as well. In addition, the analysis is appropriate, for example to a germanium p+-p-n+ structure, provided the words hole and electron are everywhere interchanged.

Consider the behavior of the electric field when the diode is under the influence of an external current density, J. At a time, $t = t_o$, the electric field, E, is assumed to have "punched through" the n-n+ boundary; however, it is not yet large enough to produce significant ionization. In the absence of such ionization, there is negligible con-

duction current, so that the displacement current everywhere is J. At a fixed point in the diode, the electric field will increase in a time interval, Δt, by

$$\Delta E = \frac{J}{\epsilon}\Delta t, \tag{1}$$

where ϵ is the dielectric constant. The slope of the electric field is given by Poisson's equation,

$$\frac{\Delta E}{\Delta z} = \frac{qn_d}{\epsilon}, \tag{2}$$

where $q = 1.6 \times 10^{-19}$ coulomb and n_d is the ionized doping density. The distance that a point of constant electric field moves in Δt is found from Equations (1) and (2) to be $\Delta z = (J/qn_d)\Delta t$. Thus, such a point moves with a wave velocity,

$$v_w = \frac{\Delta z}{\Delta t} = \frac{J}{qn_d}. \tag{3}$$

For sufficiently large J, the wave velocity can exceed the maximum velocity of particles that result from subsequent ionization. It is, in fact, such a condition that leads to the most interesting results and to which attention is now exclusively turned.

B. The Effect of Ionization

The dashed line in Figure 1(b) represents the electric field that would exist at a later time if there were no ionization. When the field exceeds a critical value, E_c, however, electron–hole pairs are created by impact ionization (the concept of a "critical field" is an approximate one, the consequences of which will be examined subsequently). The wave front (represented by the point for which $E = E_c$) moves toward the n-n+ boundary at a wave velocity that exceeds the particle velocities; therefore, the created carriers lag behind the advancing field front. Since the generated electrons drift in the same direction as the wave, while holes drift in the opposite direction, the density of electrons in the immediate wake of the wave front is higher than the density of holes. This creates an electron-rich region that serves to reduce the field. If a sufficient number of electrons are generated to overcome the

positive doping density, n_d, the slope of the field profile reverses and the field reaches a very low value. At low fields, the electrons and holes are substantially slowed, leading to the formation of a dense plasma.

For purposes of analysis, three distinct regions are considered: (1) the ionization region, in which electron–hole pairs are created and particle velocities are saturated. (2) the sheath, in which the field is too low for ionization but sufficiently large for velocities to reach saturation, and (3) the plasma, in which the field is sufficiently low that velocities are determined by low-field mobilities. There is a transition region between the sheath and plasma, in which the velocity is not simply expressed; however, this region is so short that ignoring it introduces negligible error. Finally, it is assumed throughout this paper that the defect density is sufficiently low for recombination to be neglected, i.e., relevant times are short compared to recombination times. For carefully made diodes, this approximation is quite good, even in the high-density slow-moving plasma.

C. *Ionization Region*

The continuity equations for electrons and holes are

$$\frac{\partial n}{\partial t} = v_{ns}\frac{\partial n}{\partial z} + \alpha v_{ns} n + \beta v_{ps} p, \tag{4}$$

$$\frac{\partial p}{\partial t} = -v_{ps}\frac{\partial p}{\partial z} + \alpha v_{ns} n + \beta v_{ps} p, \tag{5}$$

where n and p are the densities, v_{ns} and v_{ps} are the saturation velocities, and α and β are the ionization coefficients of electrons and holes, respectively. We look for a solution for which the ionization region (and the sheath) is moving with the wave velocity, v_w; therefore, we define the moving coordinate $x = z + v_w t$ (the positive sign is appropriate, since the wave is moving in the negative z direction). The particle densities are solely functions of x, and so we may reduce the partial differential equations to ordinary differential equations with the transformations

$$\frac{\partial}{\partial t} = v_w\frac{d}{dx}, \quad \frac{\partial}{\partial z} = \frac{d}{dx}. \tag{6}$$

It is also convenient to define a critical current density $J_c = q v_{ns} n_d$, a

normalized current density $j = J/J_c = v_w/v_{ns}$, and a saturation velocity ratio $r = v_{ps}/v_{ns}$. J_c is the current density above which the wave travels faster than the generated electrons. When Equations (6) are substituted in Equations (4) and (5) we obtain the reduced continuity relations

$$\frac{dn}{dx} = \frac{\alpha n + r\beta p}{j - 1}, \tag{7}$$

$$\frac{dp}{dx} = \frac{\alpha n + r\beta p}{j + r}. \tag{8}$$

Since α and β are functions of electric field, Equations (7) and (8) together with Poisson's equation form a coupled set of three equations. To facilitate their solution, we shall assume (1) that $r\beta p \ll \alpha n$, i.e., ionization by holes can be neglected (for germanium, ionization by electrons is neglected) and (2) that $\alpha = 0$ for $E < E_c$ and α is constant for $E \geq E_c$.* While the first assumption is not necessary for an analytic solution of Equations (7) and (8), its use is convenient; and, in any case, it is well justified for silicon and germanium, since not only is β significantly smaller than α but $p < n$ everywhere.** Assumption (2) might appear to be rather severe; however, as we shall see, a specific procedure in which the constants α and E_c are chosen leads to results that are in excellent agreement with more rigorous computer computations.

With assumptions (1) and (2), the solutions to Equations (7) and (8) are

$$n = n_o \exp\left\{\frac{\alpha x}{j - 1}\right\} \tag{9}$$

and

$$p = \left(\frac{j - 1}{j + r}\right) n_o \left[\exp\left\{\frac{\alpha x}{j - 1}\right\} - 1\right] + p_o, \tag{10}$$

* A more accurate representation of α is one in which $\alpha = 0$ for $E < E_c$ and α is proportional to $E - E_c$ for $E > E_c$. This form, as pointed out to us by L. S. Nergaard of RCA Laboratories, also results in an equation which is integrable.

** It is interesting to note that for a dc analysis of avalanche breakdown, assumption (1) is not justified, since there is a significant region for which $p \gg n$.

where the left end of the ionization region has been chosen as the origin for x; n_o and p_o are the (extremely small) densities at the origin. The electric field in the ionization region is found by substituting Equations (9) and (10) in Poisson's equation and integrating. The result, after dropping negligibly small terms, is

$$E = E_c + \frac{q}{\epsilon}\left\{ n_d x - \frac{(j-1)(1+r)}{\alpha(j+r)} n_o \left[\exp\left\{ \frac{\alpha x}{j-1} \right\} - 1 \right] \right\}. \quad (11)$$

The location of the peak field is found by demanding that the derivative of Equation (11) vanish. Inserting the value of x so determined in Equation (11) shows the peak field to be

$$E_{pk} = E_c + \frac{q n_d}{\epsilon \alpha} (j-1)(\ln\gamma - 1) \quad (12)$$

where

$$\gamma = \left(\frac{j+r}{1+r} \right) \frac{n_d}{n_o}.$$

D. *Number of Ionizations Required to Turn the Field Around*

Referring again to Figure 1(b), δ is defined as the distance for which $E > E_c$. Each electron created at the wavefront produces αl ionizations, where l is the distance traversed by the electron while exposed to an ionizing field. Considering the motion of the electron relative to the ionization wave, it is easily shown that $l = \delta/(j-1)$. Thus, in order to "turn the field around" (i.e., reduce E to a value below E_c), one electron must encounter

$$m = \frac{\alpha\delta}{j-1} \quad (13)$$

ionizing collisions. To find m in terms of more basic parameters, we set $E = E_c$ in Equation (11) at $x = \delta$. The result is the transcendental equation

$$\gamma m = e^m, \quad (14)$$

which, by virtue of the fact that $m \gg 1$, has the approximate solution

$$m \cong \ln [\gamma \ln \gamma]. \tag{15}$$

For purposes of illustration, we will evaluate m for the sample parameters $j = 2.2$, $n_d = 10^{15}$ cm^{-3}, and $n_o = 10^{10}$ cm^{-3}, leaving a more detailed discussion of these parameters until later. For this case, $m = 14.5$. Hence, the number of ionizations necessary to turn the field around greatly exceeds that required for dc breakdown (for the latter case it is known[7] that if $\beta = \alpha$, $m = 1$; for actual values of β, m is somewhat higher). We may expect, therefore, that peak fields will significantly exceed those required for dc breakdown.

E. The Sheath

The reduced continuity Equations (7) and (8) are also applicable in the sheath, except that, by definition of the sheath, both ionization coefficients are taken as zero. Thus, both particle densities in the sheath are position independent and are those that prevail at $x = \delta$. This substitution in Equations (9) and (10) shows that the sheath particle densities are

$$n_s = n_o e^m \tag{16}$$

and

$$p_s \cong \left(\frac{j-1}{j+r} \right) n_o e^m. \tag{17}$$

Poisson's equation indicates that the electric field decreases linearly in the sheath as

$$E = E_c - \frac{q}{\epsilon} (n_s - p_s - n_d)(x - \delta). \tag{18}$$

The sheath thickness, δ_s, is the distance over which the field decreases from E_c to a very small value.

$$\delta_s \cong \frac{E_c}{|dE/dx|} = \frac{\epsilon E_c}{q(n_s - p_s - n_d)}. \tag{19}$$

[7] K. G. McKay, "Avalanche Breakdown in Silicon," *Phys. Rev.*, Vol. 94, p. 877, May 1954.

F. Choice of α and E_c

The analytic solution for the ionization region was given in terms of two as yet unspecified constants, E_c and α. Here we wish to outline a procedure for choosing these constants. Only by making an effective choice can we hope that our results will be quantitatively accurate.

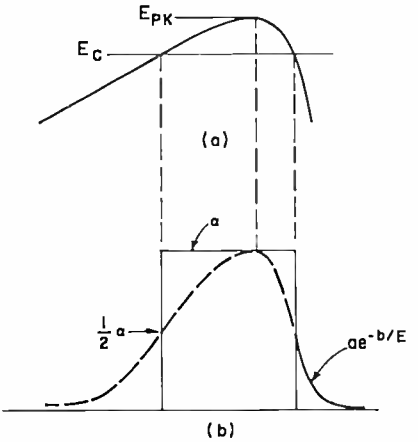


Fig. 2—Determination of ionization coefficient and critical field: (a) electric field versus position and (b) ionization coefficient versus position.

In Figure 2(a), the electric field profile is sketched. The dashed curve of Figure 2(b) represents the actual ionization coefficient, given by the relation $\alpha = ae^{-b/E}$ (a and b are constants of the material) consistent with the field profile of Figure 2(a). The solid curve of Figure 2(b) represents the approximation used in our analysis. Specifically, the criteria for choosing α and E_c are (1) that α be the actual ionization coefficient at the peak field, E_{pk}, and (2) that E_c be the field for which the actual ionization coefficient is half that at the peak. Mathematically, these conditions are written

$$\alpha = \exp\left\{\frac{-b}{E_{pk}}\right\} \tag{20}$$

and

$$a \exp\left\{\frac{-b}{E_c}\right\} = \frac{1}{2}\alpha \tag{21}$$

Along with Equation (12), these restrictions permit a unique determination of α and E_c. In Figure 3, the analytic results so obtained are plotted and compared with computer results for the sample set of parameters indicated. The abscissa, x', is the distance from the peak field. For the computer results, Equations (7), (8), and Poisson's equation were solved with accurate representations for the ionization coefficients of both holes and electrons. The most important results of

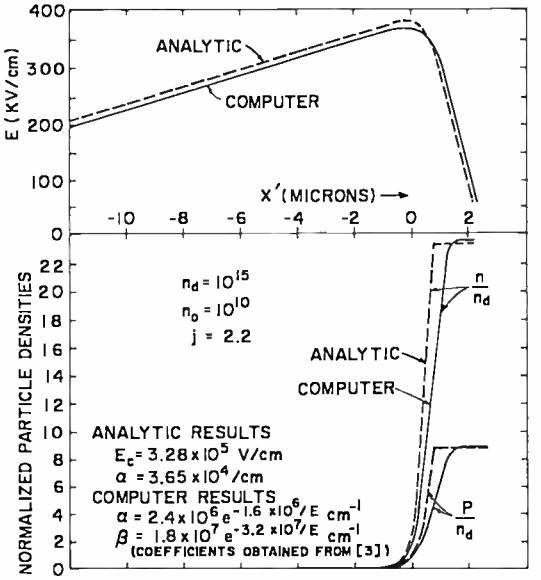


Fig. 3—Comparison of analytic and computer results.

the analytic treatment—those for the peak field and sheath particle densities—are seen to be in particularly good agreement with the computer results. Comparisons for other sample parameters also show excellent agreement.

G. The Plasma

In this section we calculate the electric field and particle densities and velocities in the plasma. The profiles of Figure 4 are stationary relative to the moving coordinate system (although, of course, the carriers are not). Because of this, and since ionization and recombination in the sheath and plasma are negligible, one may equate the electron currents in the two regions, relative to the moving coordinate

system. A similar procedure may be carried out for the hole currents. Thus,

$$(v_{iv} - v_{ns})n_s = (v_{iv} - \mu_n E_p)n_p, \tag{22}$$

and

$$(v_{iv} + v_{ps})p_s = (v_{iv} + \mu_p E_p)p_p, \tag{23}$$

where n_p and p_p are the plasma densities and μ_n and μ_p, are the low-field mobilities of electrons and holes, respectively. E_p denotes the electric

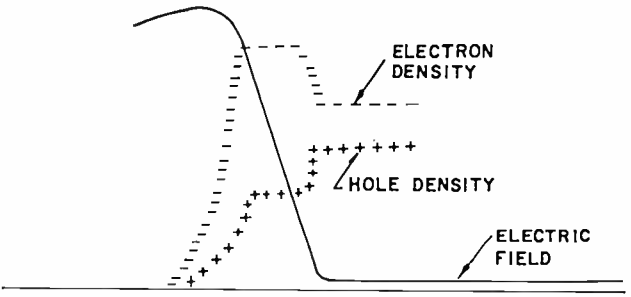


Fig. 4—Particle densities and electric field in sheath and plasma.

field in the plasma. In addition, we invoke plasma neutrality, i.e.,

$$n_p = p_p + n_d. \tag{24}$$

Equations (22) through (24) constitute a set that can be solved for E_p. Dropping negligibly small terms, one obtains

$$E_p = \frac{j^2 v_{ns}}{(j + r)(\mu_p + \mu_n)(p_s/n_d) + j(\mu_n - \mu_p)}. \tag{25}$$

The plasma particle densities may then be found from Equations (22) and (23) in terms of the now known plasma field.

$$n_p = \left(\frac{v_{ic} - v_{ns}}{v_{iv} - \mu_n E_p}\right) n_s, \tag{26}$$

and

$$p_p = \left(\frac{v_w + v_{ps}}{v_w + \mu_p E_p} \right) p_s. \tag{27}$$

Since the slow-moving plasma electrons are actually moving faster, relative to the high-field wave, than are the saturated-velocity sheath electrons, the electron density is lower in the plasma than in the sheath. The same is not true for holes. For the previously used sample set of parameters (and taking $v_{ns} = 1.0 \times 10^7$ cm/sec, $r = 1.0$, $\mu_n = 1220$ cm^2/volt-sec and $\mu_p = 460$ cm^2/volt-sec) we obtain $E_p = 1.04$ kV/cm, $n_p/n_d = 11.7$, and $p_p/n_d = 10.7$. Thus, the plasma field is about 400 times lower than the peak field (note that this low value is consistent with the use of low-field mobilities in the plasma analysis). The plasma electron velocity is $v_{np} = \mu_n E_p = 1.27 \times 10^6$ cm/sec, and the hole velocity, $v_{pp} = \mu_p E_p = 0.48 \times 10^6$ cm/sec. These greatly reduced velocities result in the relatively low frequencies that characterize the high-efficiency mode.

III. Diode Response to Applied Currents

A. Step Response

On the basis of the physics discussed in Section II and additional considerations to be developed here, we shall be able to describe and calculate the diode's response, first, to a step in current, and later, to a more interesting ac current input. Prior to the application of the step current, the diode is assumed to be biased at a voltage V_A that is slightly less than the breakdown value. It is assumed that punch-through has occured, but significant ionization has not. A double integration of Poisson's equation, taking the total charge density as n_d, relates the voltage to the field E_m at the n-p+ boundary,

$$V_A = (E_m - \frac{qn_d W}{2\epsilon}) W. \tag{28}$$

As discussed in the preceding section, the effect of the current-density step is to raise the field and launch an ionization wave that propagates through the diode leaving a plasma in its wake. This sequence of events is depicted in Figure 5(a). The peak voltage is approximately $V_B = [E_{pk} - qn_d W/(2\epsilon)]W$ and occurs after a charging time $t_{AB} = \epsilon (E_{pk} - E_m)/J$. (There are corrections to these terms that can be calculated from the results of the preceding section; however, these

corrections are rather small). E_{pk} is given by Equation (12). During the very rapid traversal through the diode by the wave, there is a slight movement of the low-velocity particles in the plasma. Also, as the wave exits the diode, there is a small perturbation in the "feeding" of the plasma at the left end of the diode. Neglecting these effects (which, one can show, can be quite small), a minimum voltage, $V_C = E_p W$, is

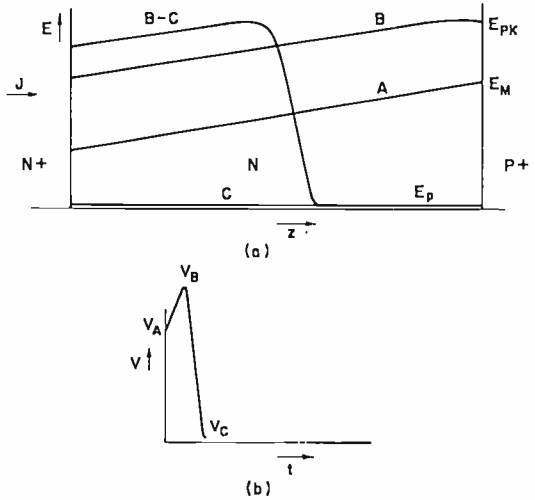


Fig. 5—Diode response to a step current: (a) electric field versus position and (b) voltage versus time.

reached, when the plasma fills the entire diode. The voltage fall time is that required for the peak field and sheath to exit the n region. Since the wave is moving at a velocity, v_w, the fall time is $t_{BC} = (W + \delta_s)/v_w$. The rise and fall of the voltage are shown in Figure 5(b).

We now turn our attention to the extraction of the plasma, with the current still maintained at J. The particle densities and electric field existing when the plasma fills the diode are sketched in Figure 6(a). The situation that prevails at some time during the ensuing extraction period is shown in Figure 6(b). Under the influence of the low plasma field, electrons drift to the left, leaving an electron-deficient region near the n-p+ boundary. Because of the electron deficiency in this region, the field is larger. As a result, the hole velocity in this region is larger, and, therefore, the "residual" hole density, p_r, is lower than in the plasma. The transition between the plasma and the rising field region is illustrated as being abrupt. Actually, particle diffusion

and the continuous variation in hole mobility with field results in a rounding effect. Approximate calculations indicate that this rounding effect is slight. In the following treatment, we neglect diffusion and assume particle velocity in the rising-field region to be saturated.

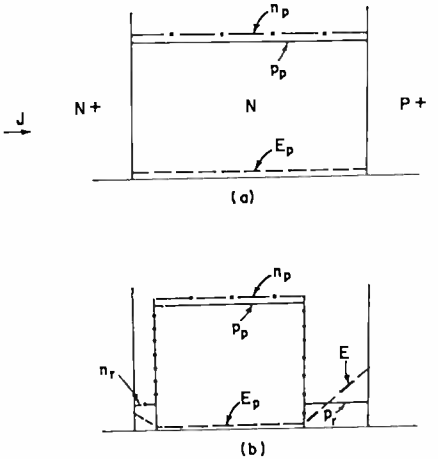


Fig. 6—Particle densities and electric field (a) at the start of the extraction period and (b) during the extraction period.

Since there is no hole accumulation at the moving boundary between the plasma and the rising-field region, we may equate plasma and residual hole currents relative to the moving boundary:

$$p_p(v_{pp} + v_{np}) = p_r(v_{ps} + v_{np}). \tag{29}$$

We can then solve for p_r in terms of quantities already calculated:

$$p_r = \left(\frac{v_{pp} + v_{np}}{v_{ps} + v_{np}}\right) p_p. \tag{30}$$

Similar considerations at the other moving boundary show that the residual electron density is

$$n_r = \left(\frac{v_{np} + v_{pp}}{v_{ns} + v_{pp}}\right) n_p. \tag{31}$$

Having solved for the residual densities and utilizing the fact that the velocities of the right and left boundaries are those of plasma electrons and holes, respectively, we can integrate Poisson's equation twice to obtain the time variation of the voltage during the extraction period. The result, which is included in the voltage-response sketch (Figure 8(a)) is

$$V_{CD}(t) = \frac{q}{2\epsilon}\left[(n_r - n_a) v_{pp}^2 + (p_r + n_a) v_{np}^2 \right] (t - t_c)^2 + V_C. \quad (32)$$

The plasma width diminishes to zero when the moving boundaries meet. Since the total width swept out by these boundaries is W,

$$v_{np} t_{CD} + v_{pp} t_{CD} = W. \quad (33)$$

Hence, the extraction period is

$$t_{CD} = \frac{W}{v_{np} + v_{pp}}. \quad (34)$$

For a plasma field of 1 kV/cm, the reduced particle velocities result in an extraction time that is six times greater than the transit time based on a saturated velocity.

The situation at the end of the extraction period, when the moving boundaries have coalesced, is illustrated in Figure 7(a). In the ensuing residual extraction period, the residual carriers move apart with saturated velocities (Figure 7(b)) and out of the n region (Figure 7(c)) while the field rises. It is easily shown that residual extraction lasts for a period $t_{DE} = W/(1 + r')v_{ps}$ (where $r' = \mu_p/\mu_n$) after which the voltage is

$$V_E = V_C + \frac{qW}{\epsilon}\left[(n_a + p_r) v_{np} (t_{CD} + t_{DE}) - n_a \frac{W}{2} \right]. \quad (35)$$

We have assumed during this discussion that fields in the extraction and residual extraction periods are not large enough to produce additional ionization. The avoidance of "premature avalanche" must be considered in the design of oscillators and is discussed later.

With the passage of an additional time, $t_{EA} = \epsilon(V_A - V_E)/(WJ)$, the voltage is again at V_A and the cycle repeats itself, as shown in Figure 8(a). Thus, we conclude that the response of the diode to a (sufficiently

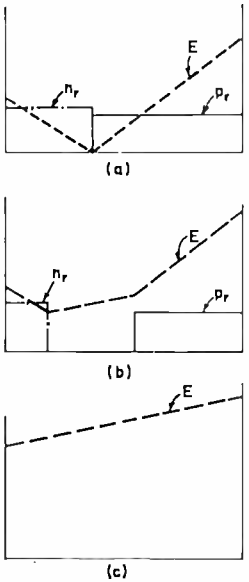


Fig. 7—Particle densities and electric field (a) at the start of residual extraction period, (b) during residual extraction period, and (c) at the end of the residual extraction period.

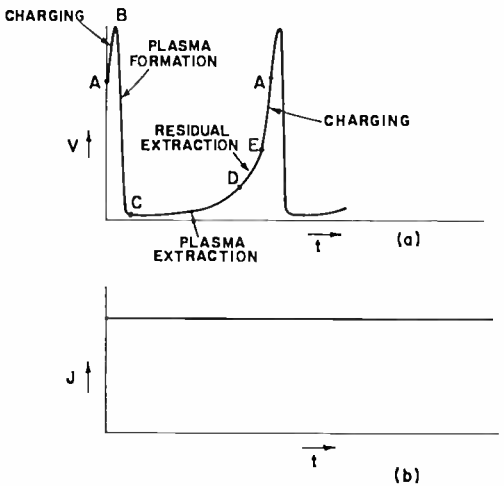


Fig. 8—(a) Oscillatory voltage response to (b) a step-current excitation.

large) step in current density is oscillatory with a large voltage modulation. It is evident, however, that no ac power can be extracted from the device if the circuit current is a constant. We now turn our attention to the consideration of ac waveforms that characterize high-efficiency oscillations.

B. Current Waveform for High-Efficiency Oscillations

A successful model must accomodate a time-varying periodic current that meets the following criteria: (1) the end of the cycle produces a diode field and densities that are the same as those at the beginning

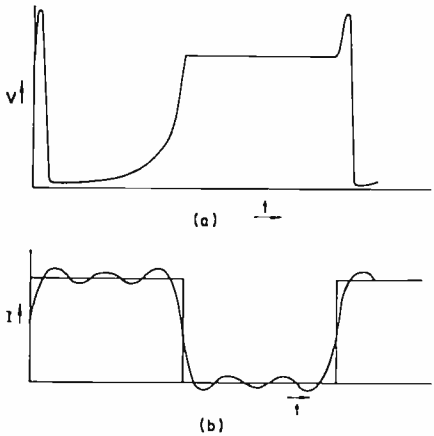


Fig. 9—(a) Voltage response to (b) a square wave of current.

of the cycle (otherwise the steady state is not being described), (2) the current and resultant diode voltage are consistent with passive external circuitry as well as with the diode physics, and (3) predictions of the model are in accord with experiment. In considering the step response of the diode, it was noted that an extended period of low voltage (the plasma-extraction period) accompanied the large current. If, now, a period of high voltage can be coupled with a low current, ac power can be efficiently extracted from the diode.

A simple current waveform, which we now consider, is the square wave of Figure 9(b) (ignore, for the moment, the rippled square wave also shown). The first part of the solution is as described earlier for the step current. Now, however, the current decreases to zero when the voltage returns to V_A. At this point, there is negligible conduction current, and, since the net current is zero, the displacement current is

zero. Hence, the voltage remains fixed at V_A, i.e., the voltage is maintained at a high value while the current is zero. The cycle repeats itself when the current again turns on. Thus, it is clear that for a square-wave of external current, the diode will efficiently convert dc power to rf power. The square wave of current, with off and on times equal, can be reasonably well approximated by a fundamental, a third, and a fifth harmonic, as shown in Figure 9(b). A consideration of the finite rise time and the ripples indicates that these should produce no first-order changes in the results.

In order for this current and the resultant diode voltage to be realizable with an external circuit that (except for the bias supply) is both passive and linear, the diode large-signal impedance (i.e., complex ratio of diode voltage to current) at each of the three frequencies must

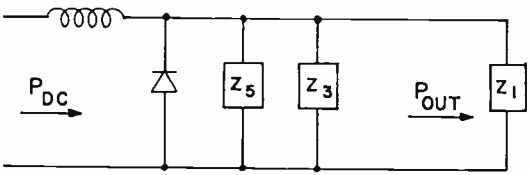


Fig. 10—Equivalent circuit for a high-efficiency oscillator.

not have a positive real part. This being the case, the required waveforms can be supported by the circuit of Figure 10. Z_1, Z_3, and Z_5 are branches with impedances equal to the negative of the diode impedance at the fundamental, third, and fifth harmonics, respectively. Each branch is characterized by a high Q and, therefore, is assumed to be essentially an open circuit at harmonics other than the one to which it is tuned.

The results of a sample calculation for a 0.025-inch-diameter silicon diode the n region of which is 8 microns long and doped with 10^{15} donors per cm^3 are now given. A square wave for which the current density is 3200 A/cm^2 ($j = 2.0$) for one-half cycle and zero for the other, was used to calculate the voltage. This corresponds to a bias current of 5 amperes. A Fourier analysis of the waveforms was carried out and the impedances and efficiency calculated. The above criterion for circuit realizability is found to be satisfied. The dc voltage is noted to have dropped to 63% of the breakdown value. At a fundamental frequency of 840 MHz, the diode impedance is $(-17 - j1)$ ohms, the efficiency is 54%, and the output power is 330 watts. Similar calculations for a wide variety of parameters consistent with the design rules established in the next section show efficiencies from 50 to 60%.

C. Other Waveforms

The gross features characterizing the waveforms of Figure 9 are (1) the formation of a low-voltage plasma and its extraction over an extended period of high current and (2) a subsequent extended period of low current in which stored energy maintains the voltage at a high level.

It is evident that these features can prevail, to a large extent, with waveforms differing in detail from those of Figure 9. To the extent that the above features are realized, however, one may expect efficiencies to approach those predicted here. Consider, for example, the waveforms of Figure 11. The minimum current is non-zero, and, conse-

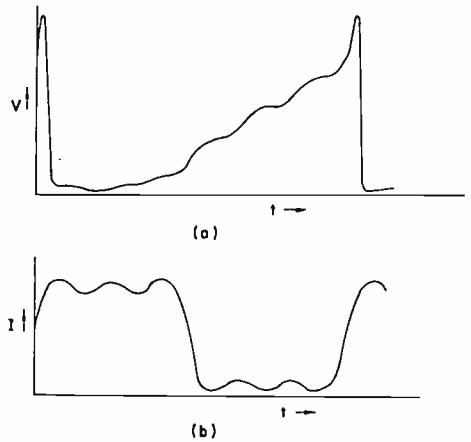


Fig. 11—Another possible set of high-efficiency waveforms: (a) voltage and (b) current.

quently, the corresponding voltage rises in a saw-tooth-like manner. Calculated efficiencies for this situation are slightly lower than those for the previous case considered. The waveforms reported in computer simulations[5] are characterized by the above features to a lesser extent; thus, efficiencies are lower still. Finally, it should be noted that high and low current states need not have equal duration, though, for extensive departures from equality, efficiencies may be expected to suffer.

D. The Role of Transit-Time Oscillations

It was reported[6] that conventional transit-time oscillations appear to trigger the high-efficiency oscillatory state. Such would be expected

[5] A. S. Clorfeine and R. D. Hughes, "Characterization of Anomalous Avalanche Diodes," IEEE Informal Conference on Active Microwave Effects in Bulk Semiconductors, New York, Jan. 1968.

on the basis of the theory presented here. If a current generator with sufficiently small rise time were available, then no additional starting mechanism would be required. In the absence of such generators, however, it is the transit-time oscillations that provide the rapid current increase necessary to trigger the high-efficiency mode.

As regards the role of transit-time oscillations in the steady state of the high-efficiency mode, we note that at least one important harmonic may lie within what is considered to be the conventional transit-time range for the diode. Thus, variations at a transit-time frequency in the current (and also voltage) waveform, as shown by the rippled curve of Figure 9(b), may be expected. However, it does not seem useful to interpret these variations as a transit-time oscillation that, each cycle, triggers the plasma. Rather, these higher frequencies are simply viewed as very essential harmonics that serve the function of wave shaping.

IV. Design Considerations

Realization of near-optimum efficiencies requires proper design of the diode and the circuit and proper bias excitation. In this section, we utilize our theoretical results to develop a number of design formulas. These formulas should be considered first-order, since approximations have been made to keep them simple and, therefore, easily applicable.

A choice of the bias current is discussed first, and then attention is turned to the design of the diode. The oscillation frequency for a given diode is tunable over an appreciable range, the center of which is determined primarily by the width, W, of the high-resistivity region. A properly designed diode must "punch through" prior to breakdown. A more stringent condition regarding the degree of punch-through is developed, however, and results in an optimum n_dW product and, thus, a choice of doping density for a given depletion width. The diode area is determined by the output power required. A general result relating the two is obtained. Finally, a formula for the required fundamental circuit impedance is derived.

A. Bias Current

We have emphasized that the type of operation described here requires that the current density be greater than a critical value $qv_{ns}n_d$ or, equivalently, that the normalized current density j be greater than unity. The detailed calculations indicate that for values of j only slightly greater than unity, there is a marked departure from the wave-

form of Figure 9(a), and performance may be expected to deteriorate. This effect is not a problem if we choose j in the range 1.5–2.5. As the current is increased further, a point will be reached where premature avalanche occurs (this is discussed in more detail later). By properly choosing other diode parameters, a larger value of j can be accommodated, but there seems to be no particular advantage in this procedure. We shall choose as our nominal design density, $j = 2$ or, equivalently, $J = 2qv_{ns}n_d$. Since the dc current density is half that of the peak value, the desired bias current is

$$I_{DC} \approx qv_{ns}n_dA. \tag{36}$$

B. Tunability

By summing the times required for each portion of the cycle in Figure 9, we obtain a relation for the frequency of oscillation;

$$f \cong \left\{ \frac{(j+r)(j-1)m}{(1+r)j^2} + \frac{4c}{j} \right\}^{-1} \frac{v_{ns}}{2W}, \tag{37}$$

where c is a function of m and j which, to first order, may be taken as unity. The quantity m has been interpreted as the number of ionizations required to turn the field around, and is given by Equation (15) and, indirectly, by the density ratio n_d/n_o. The quantity n_o is the density of electrons prior to the rapid rise in current. Its minimum value, which is determined by carrier lifetime, can be exceedingly low; however, n_o can be varied over orders of magnitude, by virtue of the multiplication associated with the pre-breakdown voltage, V_A. Thus, a slight variation in V_A can significantly change the frequency of oscillation, since multiplication is a very sensitive function of voltage near the breakdown value. Of course, V_A is only controlled indirectly, namely, by means of the circuitry. Consequently, as the circuit is tuned, i.e., as the proper impedances for the fundamental and relevant harmonics are established at a slightly different set of frequencies, the corresponding changes within the diode are primarily those of the voltage and the electron density in the low-current portion of the cycle. These, in turn, determine the plasma field and extraction time. An inspection of Equation (37) indicates that for $j = 2$, a variation of n_d/n_o from 10^2 to 10^{10} results in f changing from 0.04 v_{ns}/W to 0.12 v_{ns}/W (actually, the range that can be covered while maintaining high efficiency depends somewhat on the diode parameters). Note that a large tuning

range has been demonstrated experimentally.[9] Since the conventional transit-time range may extend well below the frequency considered optimum by Read,[2] 0.5 v_{ns}/W, the frequency for high-efficiency operation may be an appreciable fraction of that at which the same diode can oscillate in the transit-time mode.

C. *Width of the High-Resistivity Region*

We take, as the design frequency (or, in the case of tunable operation, the center of the high-efficiency range over which the diode is to operate), the geometric mean of the above two frequencies, 0.07 v_{ns}/W. Thus, our first design rule pertains to the choice of W,

$$W = 0.07 \frac{v_{ns}}{f}. \tag{38}$$

A 700-MHz diode, for example, should be fabricated with an n region ~10 microns wide.

D. *Doping Density*

A diode that is tailored so that avalanche occurs at nearly the same voltage at which the electric field punches through to the low-resistivity n+ region gives excellent results in the transit-time mode.[10] From Poisson's equation, it is seen that the above condition occurs when

$$n_d W = \frac{\epsilon E_B}{q}, \tag{39}$$

where E_B is the field at the n-p+ boundary at breakdown. We have assumed in this paper that punch-through occurs, so that Equation (39) represents an upper limit for the $n_d W$ product of the diodes under consideration. We shall see, however, that an even smaller upper limit must be imposed if premature avalanche is to be avoided; thus, a diode tailored to operate in the transit-time mode as described above

[9] P. A. Levine and S. G. Liu, "Tunable L-Band High-Power Avalanche-Diode Oscillator," IEEE International Solid-State Circuits Conf., Phila., Pa., Feb. 1969.

[10] T. Misawa, "Microwave Si Avalanche Diode with Nearly-Abrupt-Type Junction," *IEEE Trans. on Electron Devices*, Vol. ED-14, p. 580, Sept. 1967.

cannot be expected to operate optimally in the high-efficiency mode. Such was, indeed, found experimentally by others.[11]

From the results of the preceding section, one can show that the field at the n-p+ boundary at the end of the residual extraction period is

$$E_W \cong \frac{q}{\epsilon}\left[\frac{1+\frac{j}{r}}{1+r'}\right] n_d W. \tag{40}$$

To avoid premature avalanche, E_w must be less than E_B, or, from Equation (40),

$$n_d W < \left[\frac{1+r'}{1+\frac{j}{r}}\right]\frac{\epsilon}{q} E_B. \tag{41}$$

Even for $j = 1$, the coefficient of $\epsilon E_B/q$ (for silicon) is 0.69, or considerably less than the value of unity that characterizes the transit-time diode discussed above. If we permit premature avalanche to commence for $j = 2.5$, it will be safely avoided for our design choice of $j = 2$. Substituting $j = 2.5$ in Equation (41) results in

$$n_d W = 0.4\,\frac{\epsilon}{q} E_B. \tag{42}$$

Smaller values of $n_d W$ may also be used; however, a diode with a lightly doped n region might be more difficult to fabricate (e.g., it may be less amenable to epitaxial technology). Also, more detailed considerations suggest that for values of $n_d W$ that are too low, a full plasma may not form.

The breakdown field, E_B, which appears in Equation (42), is a weak function of n_d[12]; for silicon, it increases from 3×10^5 V/cm to 4×10^5 V/cm as n_d is increased from 10^{15} cm^{-3} to 10^{16} cm^{-3}. Inserting $E_B = 3.5 \times 10^5$ V/cm in Equation (42) gives $n_d W \cong 9 \times 10^{11}$ cm^{-2}.

[11] R. S. Ying, R. G. Mankarious, and D. L. English, "High-Efficiency Anomalous Mode Oscillations from Silicon Impatt Diodes at 6 GHz," IEEE International Solid-State Circuits Conf., Phila., Pa., Feb. 1969.

[12] S. L. Miller, "Ionization Rates for Holes and Electrons in Silicon," *Phys. Rev.*, Vol. 105, p. 1246 (1957).

Thus, we have a design rule that permits a choice for the doping density. By way of example, the 700-MHz 10-micron diode previously mentioned should be doped to a level of slightly less than 10^{15} cm^{-3}.

E. Diode Area

The diode area is determined by the power output required, within, of course, the limits imposed by impedance, thermal, and technological considerations. It is clear from the voltage waveform of Figure 9(a), that the average voltage, V_{DC}, is less than the breakdown value, V_b; calculations indicate that over a wide range of near optimum oscillations, $V_{DC} \cong 0.64\, V_b$. Also, it follows from Equation (42) that the average field $< E >$ at breakdown is 80% of the field at the n-p+ boundary, so that the breakdown voltage is

$$V_b = < E > W = 0.8\, E_B W. \tag{43}$$

Equations (36) and (43) show that the required dc power is

$$P_{dc} = V_{dc} I_{dc} = 0.64\, V_b I_{dc} \cong 0.5\, q v_{ns} E_B (n_d W) A. \tag{44}$$

For a 50% conversion efficiency, the power at the fundamental is $P_1 = 0.5\, P_{dc}$. Using the result obtained from Equation (42) for the $n_d W$ product in Equation (44), we obtain

$$P_1 \cong 0.1\, \epsilon v_{ns} E_B^2 A \tag{45}$$

from which the required area can be obtained.

For $E_B = 3.5 \times 10^5$ V/cm, the output available from a properly designed silicon diode is, according to Equation (45), approximately 0.8 watt per square mil.

F. Fundamental Impedance

For the approximate square wave of Figure 9(b), the fundamental current amplitude is $I_1 = (4/\pi) I_{dc}$. The impedance at the fundamental is almost entirely real, and so it follows that for a 50% efficiency oscillation (i.e., $P_{dc} = 2P_1 = V_1 I_1$), $V_1 = (\pi/4)\, V_{dc}$. Since $V_{dc} \cong 0.64\, V_b$, we find that the required circuit impedance at the fundamental is

$$Z_1 = \frac{V_1}{I_1} = \frac{(\pi/4)\,(0.64\, V_b)}{(4/\pi) I_{dc}} \cong 0.4 \frac{V_b}{I_{dc}}. \tag{46}$$

For a 0.025-inch-diameter silicon diode with an n-region 8 microns wide and a doping density of 10^{15} cm^{-3}, biased according to Equation (36) and operated optimally, we find that $Z_1 = 15$ ohms.

V. Concluding Remarks

In this paper, we have presented a large-signal analytic theory for high-efficiency oscillations in avalanche diodes. The important periods of the oscillation have been identified, described physically and mathematically, and pieced together. Also, we have seen how the theory leads to a set of design formulas.

Finally, the theory presented here is able to predict features of the high-efficiency mode that can be compared with reported experimental data. Specifically, the theory predicts (a) the magnitude of bias-voltage drop that accompanies mode onset, (b) a bias-current range for optimum performance, (c) the proper doping density for a given width, (d) the width for a given frequency range, and (e) maximum efficiency. Each of these predictions shows good agreement with experimental results.

CHARACTERISTICS OF A SEALED-OFF He^3-Cd^{114} LASER

By

J. R. Fendley, Jr.,* I. Gorog,† K. G. Hernqvist,† and C. Sun‡

Summary—The construction and performance of a sealed-off He-Cd laser are described. Cataphoretic techniques are used to confine the metal vapor to a region of the discharge tube separate from the electrode and window regions. The laser, which has an active region 41 cm long and 1.6 mm in diameter, yields 12 mW at 4416 Å (TEM_{00}-mode) and 6 mW at 3250 Å (multimode). Life test, noise, cadmium transport characteristics, and effects of a transverse magnetic field are described.

The construction of this tube is only slightly more complex than that of a He-Ne laser.

Introduction

When metal-vapor lasers were discovered it was first feared that the technological problems of such devices would be immense. Hot, corrosive metal vapors seemed incompatible with sensitive optical surfaces such as windows and mirrors. It is shown in this paper, however, that by the proper use of cataphoretic techniques, metal vapors can be confined to a restricted region of the laser tube separate from the optical components. The application of these techniques to a He-Cd laser is described and the characteristics of such a sealed-off laser are presented. Finally, the effects of a transverse magnetic field on power output and output spectra are discussed.

Cataphoretic Transport and Confinement

Consider a gaseous discharge column such as shown schematically in Figure 1. Assume that the discharge takes place in a high-ionization-

* RCA Electronic Components, Lancaster, Pennsylvania.
† RCA Laboratories, Princeton, New Jersey.
‡ RCA Electronic Components, Princeton, New Jersey.

potential gas such as helium. Assume further that this carrier gas is seeded with a low-ionization-potential vapor, e.g., cadmium, of such a low density that the electron temperature is only slightly reduced from that of the helium discharge alone. At moderate current densities the degree of ionization of the helium may be very low. The cadmium vapor will have a higher degree of ionization, and, consequently, each cadmium atom spends a substantial part of the time in an ionized

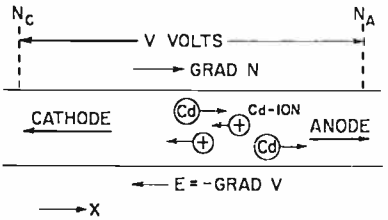


Fig. 1—Diagram illustrating the cataphoretic effect.

state. A net transport of cadmium towards the cathode will result—the cataphoresis effect. The particle transport, Γ_e, due to the electric field in the discharge is described by the mobility relation

$$\Gamma_e = N^+_{Cd}\mu_i E, \tag{1}$$

where N^+_{Cd} = cadmium ion density,

μ_i = cadmium ion mobility in helium,

E = electric field.

The diffusion of cadmium due to the density gradient is

$$\Gamma_D = -D \text{ grad } N_{Cd} \tag{2}$$

where D = diffusion coefficient of cadmium in helium,

N_{Cd} = density of cadmium atoms.

Two cases may be distinguished—cathaphoretic transport and cathaphoretic containment. The first results when the back diffusion is reduced by freezing out the cadmium at the cathode end of the discharge column. The net rate of cadmium transport is then given in Equation (1).

In the second, the density gradient is allowed to build up until

equilibrium is attained ($\Gamma_D = -\Gamma_c$). Under the assumption that the degree of ionization of 1%, a density ratio N_C/N_A of the order of reduced (using the Einstein relation and taking for the ion diffusion constant $D_+ \cong D_{\text{neutral}}$) to

$$\frac{N_A}{N_C} = \exp\left(-\frac{eV}{kT_+}\frac{N^+_{Cd}}{N_{Cd}}\right) \tag{3}$$

where N_A andN_C are the cadmium densities at two points along the discharge column separated by a distance V/E. For example, for a potential difference of 50 volts, an ion temperature of 600°K and a degree of ionization of 1%, a density ratio N_C/N_A of the order of 15,000 may be maintained.

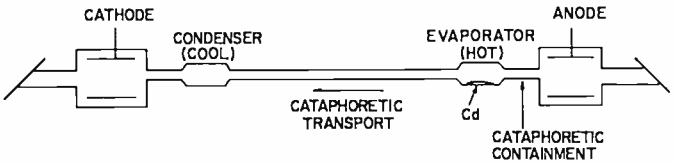


Fig. 2—Cataphoretic effects in sealed-off He-Cd laser.

Application of Cataphoresis to the He-Cd Laser

Figure 2 shows schematically how cataphoretic containment and transport may be applied to a He-Cd laser. Assume that cadmium is introduced at the point marked evaporator in Figure 2. Cataphoresis prevents the cadmium vapor from drifting toward the anode; as shown by Equation (3), the cadmium vapor density is substantially lower at the anode than at the vapor source. Cadmium is, however, transported towards the cathode, at a rate given by Equation (1). If the condenser region shown in Figure 2 is kept cool, cadmium will condense there and be prevented from reaching the cathode region. The cadmium vapor density in the main discharge column is determined by (but not identical to) the cadmium vapor density at the evaporator. In this way the cadmium vapor is confined to the main discharge region without reaching the cathode or the anode regions or the Brewster windows at the ends of the laser tube. In order to minimize the amount of cadmium needed (a matter of importance if the expensive enriched isotope of cadmium is used), the tube may be made symmetrical with

the role of cathode and anode and evaporator and condenser being reversed after a certain length of time, which is determined by the initial supply of cadmium and the transport rate.

Characteristics of a Sealed-Off He^3-Cd^{114} Laser Tube

Figure 3 is a photograph of an He^3-Cd^{114} laser tube whose design is based on the principles discussed in the previous section. It uses

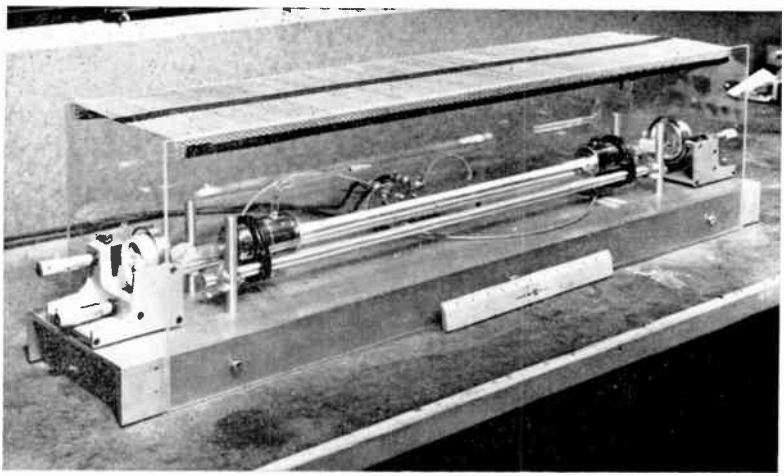


Fig. 3—Sealed-off He-Cd laser.

cold molybdenum cathodes. The main bore is a ceramic liner inside Pyrex* tubing. The evaporator is heated by the discharge, and its high temperature is maintained by introducing a suitably designed heat barrier at the corresponding part of the discharge tube. The low condenser temperature is maintained by a heat sink attached to the discharge tube. The evaporator is initially provided with 50 mg of Cd^{114} which, for the operating conditions shown in Table I, lasts for about 50 hours. At this time the polarity of the tube is reversed. The roles of the original evaporator and condenser are also reversed by changing the heat transfer means at these points.

Table I shows the physical characteristics of the laser. With the cadmium evaporator cool, the arc drop is about 2600 volts. Optimum cadmium density is obtained when the arc drop is reduced to approximately 4/5 of this value. The helium pressure is not very critical, and

* Trade mark.

Table I—Characteristics of Sealed-Off He^3-Cd^{114} Laser

Bore Diameter	1.6 mm
Bore Length	41 cm
Current	50 mA
Voltage	2200 Volts
He Pressure	7 Torr
Tube Volume	230 cm^3
Active Volume	0.8 cm^3
Reservoir Temperature	~ 270°C
50 mg	Cd^{114}

corresponds to approximately a pressure diameter product of 10 Torr mm, in agreement with results published elsewhere.[1]

Table II gives the laser performance characteristics. The gain was measured by insertion of a calibrated attenuator in the cavity. The efficiency at 4416 Å is quite similar to that of a He-Ne laser. The efficiency in the ultraviolet at 3250 Å is better by at least one order of magnitude than that of the other gas lasers in this wavelength range.[2] The ratio of UV to violet output is somewhat higher than previously reported[3] for a laser using natural cadmium.

Life testing of the He-Cd laser has been carried out to about 550 hours, corresponding to 11 polarity reversals with no change in output. There is a reduction in the helium pressure of about 0.15 Torr/100 hours. A calculation of the helium leakage rate through the hot part

Table II—Performance of Sealed-Off He^3-Cd^{114} Laser

		4416 Å		3250 Å	
P_{OUT}		25 mW multimode 12 mW TEM_{00}		6 mW multimode	
GAIN		.33 m^{-1}		.15 m^{-1}	
EFFICIENCY		~ 2×10^{-4}		5×10^{-5}	
MIRRORS		Radius of Curvature	Transm. %	Radius of Curvature	Transm. %
		m		m	
	Multimode	1	0.02	1	1
		2	2	1	1
	TEM_{00}	∞	0.02		
		1	2		

[1] W. T. Silfvast, *Appl. Phys. Letters*, Vol. 13, p. 169 (1968).

[2] J. R. Fendley, Jr., "Continuous UV Lasers," *IEEE Jour. Quantum Elec.*, Vol. QE-4, p. 627, Oct. 1968.

[3] J. P. Goldsborough, "Continuous Laser Oscillation at 3250 Å in Cadmium Ion," *IEEE Jour. Quantum Elec.*, Vol. QE-5, p. 133, 1969.

of the Pyrex envelope (enclosing the ceramic liner) yields a value very close to this, using published data[4] of the helium diffusion rate at 273°C.

From the observed cadmium transport rate of 1 mg/hour, the cadmium ion density may be computed from Equation (1). Using the experimental value of $E = 50$ volts/cm and taking the value[5] of $\mu_+ = 1.4 \times 10^4$ cm/sec/V/cm for cesium (due to lack of data for cadmium) in helium (at 1 Torr and 0°C), Equation (1) yields $N^+_{CD} = 8 \times 10^{11} cm^{-3}$. For cadmium vapor in the range 10^{-3} to 10^{-2} Torr, the measured transport rate corresponds to a degree of ionization on the order of 1%.

Although the long-time stability (> 1 msec) of the He-Cd laser is excellent, power output fluctuations of as much as 10 to 15% peak-to-peak may be observed in the frequency range 10 to 100 kHz. These fluctuations are only partially correlated with the noise in discharge current, and are most likely due to cadmium vapor density fluctuations. Further studies are called for to establish firmly the source of noise and to devise methods for its reduction.

Effects of a Transverse Magnetic Field

It has been shown[6] that a transverse magnetic field has a pronounced effect on the power output of He-Cd lasers. When the magnetic field is parallel to the electric wave vector, as determined by the Brewster window of the laser, only the central Zeeman component, which is polarized parallel to the magnetic field, contributes to lasing. Then there is no change in the output spectrum of the laser (determined by the Doppler broadened line), and the power output of the laser increases monotonically with the magnetic field. This increase is presumably caused by plasma effects. When the transverse magnetic field is perpendicular to the electric field vector, only the outer Zeeman components are observed. Figure 4 shows the output spectrum for different values of the magnetic field. For zero magnetic field, a line width of about 1 GHz is observed, as reported by Silfvast.[1] At 450 gauss the splitting becomes comparable to the Doppler width. A slight increase in laser output is observed at this point. At about 800 gauss the Zeeman components have separated, and the laser output is lower than for zero field, presumably due to the reduced gain.

At low power levels, where only one longitudinal mode per Zeeman

[4] V. O. Altemose, "Helium Diffusion through Glass," *Jour. Appl. Phys.*, Vol. 32, p. 1309, July 1961.

[5] C. F. Powell and L. Brata, "Mobility of Alkali Ions in Gases," *Proc. Roy. Soc. Lond.*, Vol. A-138, p. 117, Oct. 1, 1932.

[6] K. G. Hernqvist, *Jour. Appl. Phys.*, Vol. 40, Sept. 1969.

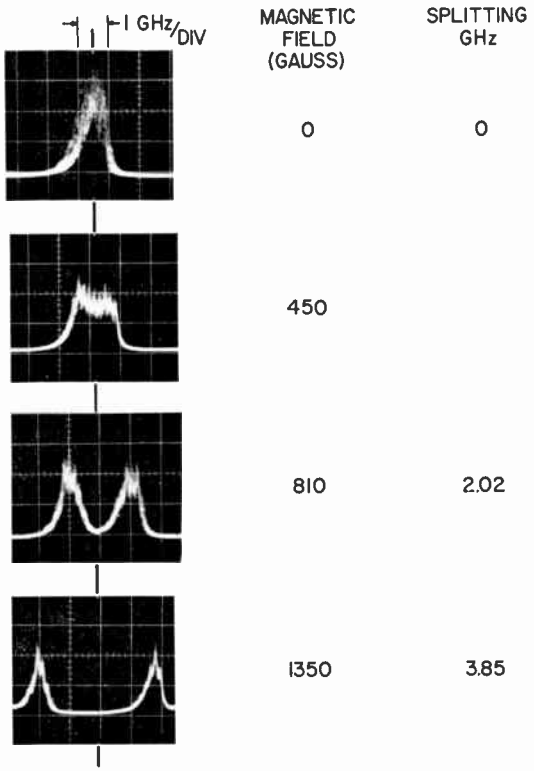


Fig. 4—Output spectra of He-Cd laser for transverse magnetic fields ($H \perp E_{Br}$).

component is operating, the separation of the outer $\Delta M = \pm 1$ Zeeman components was found to be ~ 3.4 MHz/gauss.

Conclusions

It has been demonstrated that simple sealed-off helium–cadmium lasers can be constructed. Their complexity is only slightly greater than that of helium–neon lasers.

The helium–cadmium complements the present family of gas lasers. It operates well at wavelengths in which existing lasers perform poorly and where photochemical processes are of great importance.

LOW-RADIATION-NOISE He-Ne LASER

By

K. G. Herngvist

RCA Laboratories
Princeton, N. J.

Summary—The construction and testing of a small (approximately 1 mW) He-Ne laser are described. This laser uses a potassium-activated alumina cold cathode and coaxial construction. Mirrors are attached to the laser tube outside the Brewster-angle windows. Measured power-output fluctuations are less than 0.2% in the frequency range up to 14 MHz. The laser should be ideal for information-handling applications where noise in this frequency range is objectionable.

Introduction

Power output fluctuations in the frequency range up to several tens of MHz are usually observed in gas lasers. Commercially available hot-cathode argon and helium–neon lasers exhibit fluctuations amounting to one or two percent. Cold-cathode lasers may exhibit even higher radiation noise, in some instances as high as 20%. Elaborate feedback circuits[1] or combination of rf and dc excitation[2] have been reported to reduce these fluctuations in He-Ne lasers. It is the purpose of this paper to describe the construction and performance of a simple dc-excited cold-cathode laser with inherently low power-output fluctuations.

Sources of Discharge Noise

The large power-output fluctuations observed in gas lasers are mostly due to variations in gas density, electron temperature, and plasma

[1] G. Forgo and M. J. O. Strutt, "Suppression of Discharge-Modulation Noise in a Gas Laser by Means of an External Modulator," *Electronics Letters*, Vol. 3, p. 547, Dec. 1967.

[2] S. A. Alyakishev, D. V. Gordeyev, E. P. Ostapchenko, and L. M. Pyatkova, "Radiation Noise of a He-Ne Laser at Various Methods of Excitation of the Active Medium," *Radio Eng. and Electronic Phys.*, Vol. 12, p. 1655, Oct. 1967.

density in the discharge column. Gas-density fluctuations are probably small in He-Ne lasers, and the gas density in the discharge region usually stabilizes after warm-up, a few minutes after ignition.

There are several classes of instabilities in the discharge, each of which is characteristic of the location of origin. One is cathode-emission fluctuations, especially prominent for cold cathodes, where the emission is due to ion bombardment. The ion bombardment leads to materials transport in the cathode region, changing of the emission, and corresponding re-arrangement of the discharge configuration. Other sources of instabilities are regions of impedance change of the discharge path, mainly at the entrance and at the exit of the bore. Noise-free discontinuous transitions can be designed, but may require grids or a specific pressure range.[3] The best solution appears to be a very gradual and symmetric transition.[4] Finally, striations in the main bore may result in radiation noise. These striations are limited to certain pressure and current ranges[5] that should, if possible, be avoided. In addition to these noise sources, a faulty anode design, giving rise to an anode drop, may result in instabilities.

Construction of Low-Radiation-Noise He-Ne Laser

The laser tube used in these studies was designed for an output power level of about 1 milliwatt at 6328 Å. The main bore has a diameter of 1 mm and a length of 10 cm. The tube is of coaxial design having a gas storage volume of about 300 cm^3. The cathode and the anode are symmetrically arranged with respect to the main bore, as seen in Figure 1. The conduction path from the hollow-cathode region to the entrance of the main bore is also coaxially symmetrical.

The main novel feature of the low-noise tube is the cold cathode. A detailed account of this cathode will be given in a future paper. This cathode, which will be termed the "ad-conductor" cathode uses a porous alumina coating deposited directly on glass. A tungsten wire is embedded in the coating, exposing no metal surface to the discharge. The aluminum oxide particles of the coating are 5-15 microns in diameter. Activation of the cathode is accomplished by introducing elemental potassium, which adheres to and completely penetrates the

---

[3] E. O. Johnson, J. Olmstead, and W. M. Webster, "The Tacitron, A Low Noise Thyratron Capable of Current Interruption by Grid Action," *Proc. IRE*, Vol. 42, p. 1350, Sept. 1954.

[4] K. G. Hernqvist and J. R. Fendley, Jr., "Construction of Long Life Argon Lasers," *IEEE Jour. of Quantum Electronics*, Vol. QE-3, p. 66, Feb. 1967.

[5] *Encyclopedia of Physics*, edited by S. Flugge, Vol. 22, pp. 138 and 139, Springer Verlag, Berlin, 1956.

coating. The potassium becomes adsorbed on the surface of each alumina particle and, by electron donor action,[6] makes surface conduction possible. As a result, the coating becomes electrically conductive. Potassium adheres more strongly to the aluminum oxide than to itself, assuring a uniform distribution of potassium in the cathode coating. Potassium leaving the cathode during operation will be ionized in the plasma outside the cathode surface and drawn back to the cathode by the process of cataphoresis. If the cathode is overheated at a certain area, potassium will leave, but will eventually return to assure an even coverage.

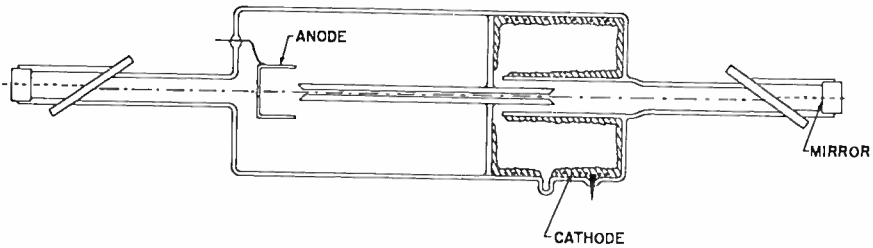


Fig. 1—Construction of low-radiation-noise He-Ne laser.

The electron emission of cold cathodes in the low-current glow-discharge mode is due to ion bombardment of the cathode. Cathode emission fluctuations result in changing the discharge configuration near the cathode. Due to the ion bombardment of the cathode, its emission characteristics change. Through the use of the continuously transportable potassium activator and a large cathode area, a state of uniform and stable emission is inherently approached in the cold cathode. If the current were to concentrate on a particular area of the cathode, that part of the cathode would be heated up and the potassium driven off. The resulting increase in the resistance of the coating here would then reduce the current. The tube shown in Figure 1 has a cathode emission area of about 150 cm^2.

The laser is filled with a mixture of 3 Torr He^3 and 0.6 Torr Ne. Mirrors are attached directly outside the permanently sealed Brewster-angle windows. The mirrors have a 60-cm radius of curvature and have transmissions of 0.04 and 0.9% at 6328 Å. Mirror separation is about 32 cm.

[6] J. D. Levine, "Absorption of Cesium on Insulators," NASA Report CR-349, Nat. Aeronautics and Space Admin., Washington, D. C., Dec. 1965.

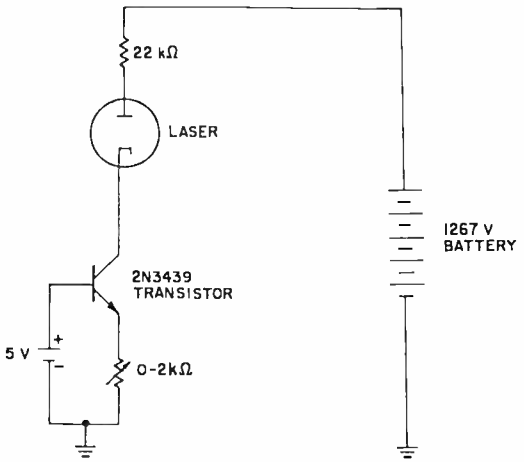


Fig. 2—Stabilizing circuit for He-Ne laser.

Performance Characteristics

The laser tube was operated in the current-stabilizing circuit shown in Figure 2. Series operation of the laser with the transistor and limiting resistor makes it possible to use a supply voltage only a few hundred volts higher than the tube drop.

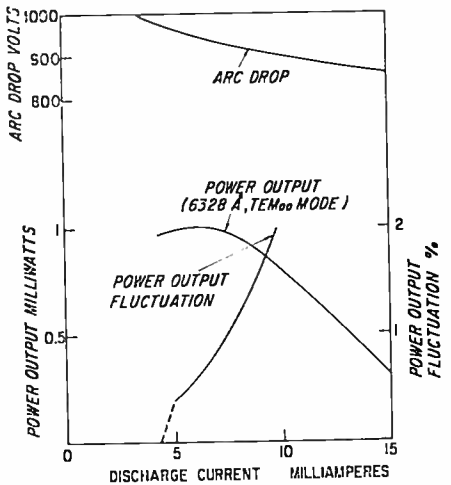


Fig. 3—Voltage power output and power-output fluctuation as functions of discharge current.

The light output was measured with a silicon cell. The signal generated by the cell was amplified and displayed on an oscilloscope, the bandwidth of the detection system being 14 MHz. Figure 3 shows voltage, laser output, and noise amplitude as functions of discharge

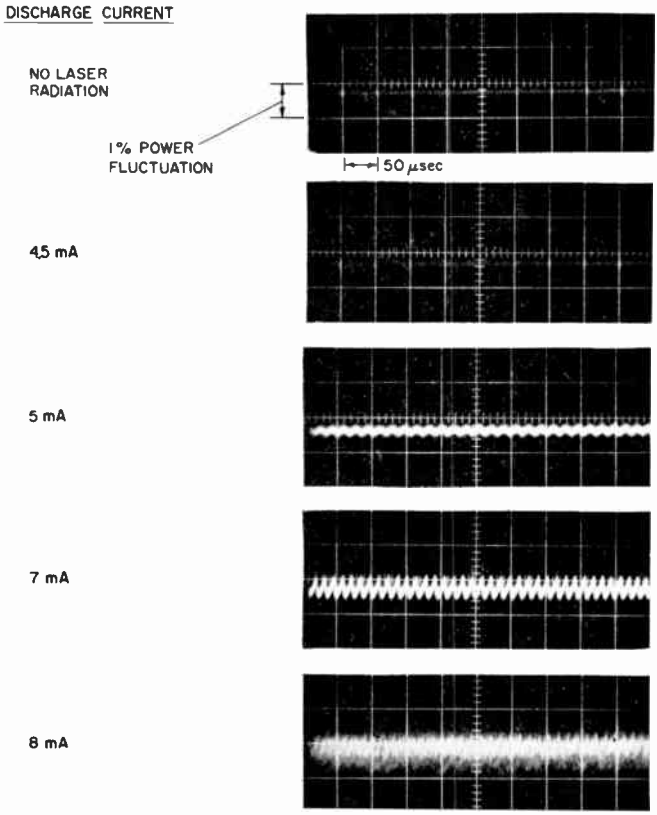


Fig. 4—Oscillograms showing power-output fluctuations at different discharge currents. One division corresponds to 1% power fluctuation.

current. Minimum noise occurs for currents equal to or less than the current corresponding to maximum output. Figure 4 shows the noise as displayed on the oscilloscope for different currents. At currents below 4.5 milliamperes, the power-output fluctuations are indistinguishable from the detector system noise and are less than 0.2%. A more or less coherent-type fluctuation sets in at 5 milliamperes and turns into a "hash"-like appearance at 8 milliamperes. The fluctuations increase further towards higher currents.

CONCLUSIONS

It has been demonstrated that simple low-radiation-noise He-Ne lasers may be constructed. Such lasers should be very effective in information-handling applications, where noise in the frequency region up to tens of megacycles are objectionable.

ACKNOWLEDGEMENTS

The author thanks D. C. Pultorak and G. L. Petrillo for technical assistance.

THE ACOUSTOELECTRIC EFFECTS AND THE ENERGY LOSSES BY HOT ELECTRONS

PART IV–FIELD AND TEMPERATURE DEPENDENCE OF ELECTRONIC TRANSPORT

By
A. Rose

RCA Laboratories
Princeton, N. J.

Summary—In this paper, which is the fourth in a five-part series, semi-classical expressions for spontaneous emission of phonons by energetic electrons are used to derive the temperature and field dependence of mobility and the mean energy of carriers versus applied field. A simple formalism is developed and applied to obtain the energy distribution of electrons around their mean energy. A comparison of theory and experiment on hot electrons shows that good agreement is obtained for energies near the band edge. At higher energies, more information about the properties of the band structure at the higher energies is needed before theory can be properly compared with experiment.

Preamble

THE PRESENT paper is the fourth in a series* on the interactions between electrons and phonons, and between electrons and acoustic waves. The major purpose of the series has been to trace the events that lead at high fields to hot electrons, impact ionization, and dielectric breakdown; and to trace the events in terms as simple and as physically graphic as possible. Certain concepts have emerged that were not evident at the outset, at least to this writer, and will be outlined here.

The rates of exchange of energy between electrons and a medium are characterized by a coupling constant β, which was previously defined as the fraction of the available coulomb field energy of the electron that can be transferred to the medium. It was shown that in

* Part I, "Small Signal Acoustoelectric Effects," *RCA Review*, Vol. 27, p. 98, March 1966; Part II, "Rates of Energy Loss by Energetic Electrons," *RCA Review*, Vol. 27, p. 600, Dec. 1966; and Part III, "Large Signal Acoustoelectric Effects," *RCA Review*, Vol. 28, p. 634, Dec. 1967.

the case of energy exchanges between electrons on the one hand and phonons or acoustic waves on the other, β is equal to the ratio of electrical to total energy of the corresponding acoustic wave. Further, when a fast electron excites resonances such as polar optical phonons, plasmons, or x-ray levels, β can also be written as $(K_L - K_H)/K_L$, where K_H and K_L are, respectively, the dielectric constants on the high and low sides of the resonance.

The coupling constant β was found to be valid for the classical acoustoelectric effect in which macroscopic acoustic waves are amplified by a drifting stream of electrons as well as for the quantum processes in which hot electrons emit phonons to a phonon-free or zero-temperature medium. The acoustoelectric effect is a process of *induced* emission, whereby the rate of loss of energy to the acoustic wave is proportional to the energy of the acoustic wave or, in quantum language, to the number of phonons in the wave.* On the other hand, the emission of phonons into a zero-temperature lattice by an energetic electron is conventionally called *spontaneous* emission.

The term spontaneous was Bohr's interpretation of Einstein's phrase "emission without excitation by external causes"† of a photon by an excited atom in a field-free surround. In quantum terminology, the rate of emission by an excited atom is proportional to $n_{Ph} + 1$ where n_{Ph} is the occupation number for photons of the emitted frequency and represents the induced part of the emission. It was conceptually reasonable from classical physics that the presence of real photons would indeed provoke or induce the atom to emit as well as to absorb. The term unity in $n_{Ph} + 1$, however, meant that even at zero temperature, in the absence of real photons in the field, the atom would still emit—apparently spontaneously, because there was nothing in the field to provoke it to emit.

It is for this reason that spontaneous emission has come to have a certain mystery attached to it. It is not uncommon to find some writers,[1] apparently in an attempt to avoid a logical vacuum, attributing spontaneous emission to the presence of zero-point vibrations. This association has two unfortunate consequences. First, it raises the question[1] of why zero-point vibrations, whose energy content is only a

* While the description of the energy of a coherent wave in terms of phonons may raise certain questions of meaning, it is well to bear in mind that the rate of amplification (as shown in Part II for $l > \lambda$) is the same whether the electrons amplify a coherent acoustic wave or a cloud of incoherent phonons having the same energy density.

† See, e.g., *The Conceptual Development of Quantum Mechanics*, by M. Jammer, McGraw-Hill Book Co., N. Y. (1966).

[1] L. I. Schiff, *Quantum Mechanics*, McGraw-Hill Book Co., Inc., New York, p. 388, 1949.

half quantum, should be equally as effective as real photons whose energy content is a whole quantum. Even more, it suggests that spontaneous emission is a purely quantum mechanical phenomenon that cannot be understood classically because zero-point vibrations have no classical counterpart.

In contrast to certain, at least historical, ambiguities attached to the concept of spontaneous emission in vacuum, the concept of spontaneous emission in a solid medium as described in Part II of this series has a simple, classical foundation. An electron in a solid polarizes or distorts the solid so that energy is transferred from the coulomb field of the electron to the surrounding solid medium. When the electron moves sufficiently fast, it leaves behind a trail of polarization whose energy content is equal to the average rate of emitting phonons (or other excitations). Spontaneous emission is no more mysterious than induced emission. In fact, either one implies the other. The fact that the medium can act on the electron, via the electric fields of its phonons, implies that the electron must be able to act upon (polarize) the medium and emit phonons. The electron and the medium are two systems. If we postulate action of one system (medium) on the other (electron), we are committed to the action of the second system (electron) on the first (medium). In brief, the interaction between electron and medium is a symmetric, two-sided connection that should be describable by a single coupling constant. The coupling constant β, which we have already outlined, does indeed characterize both spontaneous and induced emission of phonons by electrons in solids.

The major content of the present paper is the computation of the mean energy of electrons under the application of fields sufficient to raise that mean energy above the kT_l of the lattice.* These are commonly called hot electrons. The mean energy can be characterized by an equivalent temperature T_e by equating the mean energy to kT_e. However, this assignment does not ensure that the distribution of energies around the mean value satisfies the Boltzmann form $\exp(-E/kT_e)$. Such is not, in general, the case, and some estimates based on simplifying assumptions will be made to obtain a measure of the actual distributions.

Prior to the analysis of hot electrons, we outline a general method for obtaining the mobilities of electrons in thermal equilibrium with the lattice. The method represents a "back-door" approach to mobility. The usual approach computes by perturbation theory the scattering of

* For a treatment of hot electrons by perturbation theory, see *High Field Transport in Semiconductors*, by E. M. Conwell, Solid State Phys. Suppl., Academic Press, N. Y. (1967).

electrons by real phonons, that is, the *induced* emission and absorption of phonons. The present approach begins, instead, with the *spontaneous* emission of phonons and relates this to the induced emission via detailed balancing that states that the total electron–phonon interaction is $2n_{Ph}+1$ times the spontaneous interaction (here, n_{Ph} is the occupancy for phonons). We use this back-door approach for two reasons. First, we have already carried out a relatively simple semiclassical formalism in Part II for computing the spontaneous emission. The formalism is couched in real space. I do not know of a comparable formalism for induced emission. The second reason is that the back-door approach is a back door only by convention. As we have already pointed out in this preamble, induced and spontaneous emission are two aspects of the *same* physical interaction between electrons and phonons. We are free to choose either aspect as a means for computing the total interaction.

Temperature Dependence of Mobility in Semiconductors

Let dE/dt be the average rate at which an electron of given energy spontaneously emits energy to a given class of phonons. The actual emission is via a series of phonons randomly spaced in time. Hence, the average time required to emit one phonon is

$$\tau_e = \frac{\hbar\omega}{dE/dt} \tag{1}$$

From detailed balancing, we know that the total interaction between electrons and phonons occurs at the rate of $2n_{Ph}+1$ times that for spontaneous emission. Here, n_{Ph} is the occupancy for phonons of frequency ω,

$$n_{Ph} \equiv \left[\exp\left(\frac{\hbar\omega}{kT}\right)-1\right]^{-1} \doteq \frac{kT}{\hbar\omega} \text{ for } kT \gg \hbar\omega. \tag{2}$$

The average collision time for the total interaction is, then,

$$\tau_c = \frac{1}{2n_{Ph}+1}\,\tau_e$$

$$\doteq \frac{\hbar\omega}{2kT}\,\tau_e \text{ for } kT \gg \hbar\omega$$

$$= \frac{(\hbar\omega)^2/(2kT)}{(dE/dt)} \tag{3}$$

Now, the spontaneous emission (see Table I) is a strongly increasing function of ω, so that Equation (3) is dominated by the highest-frequency phonons. Each such collision is sufficient to randomize the momentum of the electron and to constitute one collision time for the purpose of computing mobility. We can, accordingly, use Equation (3) to write the mobility as

$$\mu \equiv \tau_e \frac{e}{m}$$

$$= \frac{e\,[\hbar\omega/(mv_t)]^2}{dE/dt} \tag{4}$$

where v_t is the mean thermal velocity of the electron and dE/dt is the spontaneous rate of energy loss by an electron having thermal energy.

We take the v_t dependence of dE/dt from Table I and compute the following temperature dependencies of mobility from Equation (4) for optical phonons for which $\omega =$ constant.

Polar Optical:

$$\mu \propto v_t^{-1} \propto T^{-1/2}, \text{ for } kT > \hbar\omega \tag{4a}$$

Non-Polar Optical:

$$\mu \propto v_t^{-3} \propto T^{-3/2}, \text{ for } kT > \hbar\omega \tag{4b}$$

For optical phonons and for $kT < \hbar\omega$, the dominant factor in the mobility is an exponential factor

$$\mu \propto \exp\left(-\frac{\hbar\omega}{kT}\right) \tag{5}$$

that is a measure of the fraction of electrons having enough energy

Table I—Time rates of energy loss dE/dt by electrons of velocity v. The middle column expresses these rates formally as $\beta(e^2\omega^2/Kv) \times$ geometric factor (derived in Part II), where β is the coupling constant and is set off by square brackets. The right-hand column gives the v-dependence of dE/dt after inserting the quantum constraints noted in the formal expressions.

Phonon and Coupling	Formal Expression for dE/dt	Final Expression for dE/dt in the Form Av^n
Polar Optical (Polarization Field)	$\left[\frac{\epsilon_o - \epsilon_\infty}{\epsilon_o}\right] \frac{e^2\omega^2}{\epsilon_\infty v} \ln\left(\frac{2mv^2}{\hbar\omega}\right)$	$\sim v^{-1}$
Non-Polar Optical (Deformation Potential)	$\frac{1}{2}\left[\frac{\pi K D^2}{\rho e^2 \omega^2 \lambda^2}\right] \frac{e^2\omega^2}{Kv}$ Note: $\frac{\lambda}{2\pi} = \frac{\hbar}{2mv}$	$\frac{D^2 m^2}{2\pi\rho\hbar^2} v$
Acoustic (Deformation Potential)	$\frac{1}{4}\left[\frac{B^2\omega^2 K}{4\pi e^2 \rho v_s^4}\right] \frac{e^2\omega^2}{Kv}$ Note: $\frac{\hbar\omega}{v_s} = 2mv$	$\frac{B^2 m^4}{\pi\rho\hbar^4} v^3$
Acoustic (piezoelectric)	$\frac{\pi}{4}\left[\frac{\epsilon_p^2}{KC}\right] \frac{e^2\omega^2}{Kv}$ Note: $\frac{\hbar\omega}{v_s} = 2mv$	$\frac{\pi\epsilon_p^2 e^2 m^2 v_s^2}{K^2 C\hbar^2} v$

Definitions for Table I

ϵ_o = low-frequency dielectric constant

ϵ_∞ = high-frequency (optical) dielectric constant

K = electronic part of dielectric constant

ϵ_p = piezoelectric constant

C = elastic modules (dynes/cm^2)

ρ = density (grams/cm^3)

v_s = phase velocity of sound

v = velocity of electrons

ω = angular frequency of phonon

B = deformation potential (electron volts in ergs/unit strain)

D = optical deformation potential (electron volts in ergs per centimeter relative shift of sublattices)

m = effective mass of electrons

($\hbar\omega$) to emit an optical phonon. Alternatively, the exponential factor is a measure of the occupancy of optical phonons that can be absorbed by the electrons.

In the case of acoustic phonons, the highest phonon frequency to be inserted in Equation (4) is given by the conservation of crystal momentum,

$$\frac{\hbar\omega}{v_s} \approx 2mv_t, \tag{6}$$

where v_s is the phase velocity of sound. Equation (6) inserted into Equation (4) yields

$$\mu \approx \frac{4ev_s^2}{dE/dt}. \tag{7}$$

We use Table I and Equation (7) to obtain
Acoustic (Piezoelectric Coupling)

$$\mu \propto v_t^{-1} \propto T^{-1/2}. \tag{8a}$$

Acoustic (Deformation Potential Coupling)

$$\mu \propto v_t^{-3} \propto T^{-3/2}. \tag{8b}$$

These mobilities and their temperature dependencies are based on a constant, isotropic effective mass for the electron. An anisotropic effective mass introduces some averaging process for the effective mass but does not alter the temperature dependence. An energy-dependent effective mass affects the temperature dependence of mobility in more or less obvious ways. That is, an effective mass that increases with increasing energy will cause the mobility to decrease more rapidly with increasing temperature.

Mobility in Metals

We treat the problem of mobility in metals separately from that in semiconductors for several reasons. The disparity between the mathematical level of a formal treatment for metals in terms of perturbation

theory[2] and of a real-space model[3] is even more pronounced than for semiconductors; the dielectric constant takes on a different form for metals; the temperature dependence is a more involved problem; and finally the real-space model gives an easy insight into the Fröhlich mechanism for superconductivity.

High Temperatures

We begin with temperatures greater than the Debye temperature so that Equation (3) gives the general form for the mobility collision time.

$$\tau_c = \frac{(\hbar\omega)^2/(2kT)}{dE/dt}. \tag{9}$$

Similarly dE/dt has the general form (see Table I)

$$\frac{dE}{dt} \approx \beta \frac{e^2\,\omega^2}{Kv_F}, \tag{10}$$

where v_F is the velocity of an electron at the Fermi surface. Combination of Equations (9) and (10) then yields the mobility

$$\mu \equiv \tau_c \frac{e}{m} = \frac{K\hbar^2 v_F}{2\beta emkT}. \tag{11}$$

In the case of semiconductors or ionic solids in particular, K is the electronic part of the dielectric constant. It enters into Equation (10) because we are interested in the force of the incident electron on the surrounding ions. This force is due to the vacuum coulomb field of the electron reduced by the electronic part of the dielectric constant. Similarly, in a metal K represents the shielding action of the ensemble of metal electrons on the force between any one electron and the surrounding ions. For distances greater than a Debye length this shielding can be written

---

[2] F. Bloch, "Über die Quantenmechanik der Elcktronen in Kristallgittern," *Zeitschrift fur Physik*, Vol. 52, p. 555, 1929; J. M. Ziman, *Electrons and Phonons*, Oxford Univ. Press, New York, p. 371, 1960.

[3] A. Rose and A. Rothwarf, Proc. IX Int. Conf. on Physics of Semicond., Moscow, Vol. 2, p. 694, 1968. Published by Nauka, Leningrad.

$$K = 1 + \left(\frac{L}{L_D}\right)^2 \doteq \left(\frac{L}{L_D}\right)^2. \tag{12}$$

We take for L the distance of closest approach between an electron and ion, and for this distance we take the uncertainty radius of the electron. Thus,

$$L \approx \frac{\hbar}{mv_F}, \tag{13}$$

and

$$L_D^2 = \frac{E_f}{4\pi n e^2}, \tag{14}$$

where n is the density of electrons in the metal and related to E_F by

$$n = \frac{8^{1/2}}{3\pi^2}\frac{(mE_F)^{3/2}}{\hbar_3} \doteq \frac{1}{\pi^2}\cdot\frac{(mE_F)^{3/2}}{\hbar^3}. \tag{15}$$

Combination of Equations (12)-(15) yields

$$K = \frac{4e^2}{2^{1/2}\pi\hbar v_F} \doteq \frac{e^2}{\hbar v_F}. \tag{16}$$

Insertion of Equation (16) into Equation (11) then gives

$$\mu \doteq \frac{eh}{2\beta mkT}. \tag{17}$$

The value of β is estimated to be about 0.5 and is based on the following argument. One definition of β was given in Part II as the ratio of electrical to total energy of the emitted phonons or of their equivalent macroscopic sound waves. For the jellium model of a metal, all of the energy of a sound wave is electrical, and β would be unity. But the jellium model of a metal is a plasma of ions immersed in a smoothed out sea of negative charge. A plasma has longitudinal vibrations but no transverse vibrations. The frequency of the latter

is zero, since no work is required to shear a plasma. In an actual metal, the frequencies of transverse and longitudinal vibrations are comparable. The energy for the transverse vibrations comes from an "elastic" force or energy not present in the jellium model. The elastic energy must be comparable to the electrical energy in order that the frequencies of transverse vibrations (which involve only the elastic energy) be comparable to the longitudinal frequencies (which involve the sum of elastic and electrical energies). Hence, the ratio of electrical to total energy is of order 0.5, and this is the value of β.

Equation (17), with $\beta = 0.5$, becomes

$$\mu \doteq \frac{e\hbar}{mkT}. \tag{18}$$

Note that the mobility collision time in Equation (18) is

$$\tau_c \doteq \frac{\hbar}{kT}. \tag{19}$$

This is simply the smallest value of τ_c that would be permitted by the uncertainty relation

$$\Delta E \, \Delta t \geq h \tag{20}$$

if we take $\Delta E = kT$. As suggestive as this interpretation is, it is not valid. According to Peierls[4], the ΔE of the uncertainty relation is of the order of E_F rather than kT. In any event, values of τ_c considerably smaller than that given by Equation (20) with $\Delta E = kT$ are encountered in polyvalent metals.

Equations (18) and (19) yield the experimental values of the monovalent metals within a factor of two. It is also true that the relatively complex expression for mobility derived by perturbation theory[2] reduces to Equations (18) and (19) in the jellium approximation for monovalent metals with $m = m_o$, the free electron mass.

Low Temperatures

The conductivity of metals at temperatures well below the Debye temperature has been observed and computed (Bloch theory) to vary

[4] R. E. Peierls, *Quantum Theory of Solids*, Clarendon Press, Oxford, England, p. 140, 1955.

as T^{-5}. We outline here the factors in the semiclassical argument that lead to the same temperature dependence.

At low temperatures, owing to the obvious energy constraints imposed by the Fermi distribution, the maximum phonon energy an electron can emit (or absorb) on the average is

$$\hbar\omega = kT$$

or

$$\lambda \propto T^{-1}. \tag{21}$$

From Equation (12), the screening factor will then vary as

$$K \propto \lambda^2 \propto T^{-2}, \tag{22}$$

since λ is now the distance of closest approach between the electron and the ions to which it can lose energy.

A further consequence of the low temperature is that the emission or absorption of a single phonon of momentum,

$$\hbar q = \frac{\hbar\omega}{v_s} \approx \frac{kT}{v_s}, \tag{23}$$

is no longer sufficient to completely scatter the momentum $\hbar k_F$ of a Fermi electron. The number of phonons required to turn the Fermi momentum through about 90° *if all of the phonons were added in the same sense* is, of course, k_F/q. Actually, the emission of phonons is a random process whose average contribution to the rotation of the Fermi momentum is zero. Hence, the rotation of the Fermi momentum must come from the rms deviation from the average. The rms deviation is proportional to the square root of the average number of phonons. Hence, $(k_F/q)^2$ phonons are needed on the average to randomize the momentum of a Fermi electron. The collision time of Equation (9) is, accordingly, increased by the factor

$$\left(\frac{k_F}{q}\right)^2 \propto \omega^{-2} \propto T^{-2} \tag{24}$$

The combined T dependencies of Equations (11), (22), and (24), then, lead to the appropriate T^{-5} at low temperatures.

Low-Temperature Pairing Energy

On the one hand, a Fermi electron at very low temperature cannot, by the exclusion principle, emit a high-energy phonon. On the other hand, it is clear from our semiclassical argument that the Fermi electron must still disturb the lattice as if it were in the act of emitting such a phonon. The "conflict" can be resolved if a second electron of equal and opposite momentum to that of the first electron cooperates with the first electron to absorb the latent energy in the trail of the first electron. This is the type of attractive energy first proposed by Fröhlich[5] as the source of the phenomenon of superconductivity.

The absorption by a second electron of phonon energy "virtually" emitted by a first electron is very much in the nature of a two-electron polaron. For example, the electron in the conventional one-electron polaron does not have sufficient energy to emit an optical phonon. Nevertheless, the electron disturbs the lattice as if it were about to emit such a phonon. The resolution here is that the electron re-absorbs its own phonons in a process known as virtual emission. The classical parallel is that of a slowly rolling ball on a rubber membrane. The leading edge of the ball does work on the membrane as if it were about to radiate vibrational energy to the membrane. At the same time, the trailing edge of the ball absorbs a comparable amount of energy from the previously stretched membrane. It is as if the trailing edge absorbs the energy emitted by the leading edge. While these two energies are only approximately equal in the macroscopic classical domain, they are identical in the quantum regime.

For the conventional one-electron polaron, the electron is moving slowly enough to re-absorb its own phonons. The Fermi electron in a metal, on the other hand, is moving too fast to absorb the phonons it "emits," and must make use of a second electron to form a "two-electron polaron."

We estimate the attractive energy for the two-electron problem by computing the depth of the energy trough left behind in the trail of the first electron.

The energy per unit *length* (see Equation (10)) in this trail is

$$\frac{dE}{dx} = \beta \frac{e^2\omega^2}{Kv_F^2}, \tag{25}$$

[5] H. Fröhlich, "Theory of the Superconducting State. I. The Ground State at the Absolute Zero of Temperature," *Phys. Rev.*, Vol. 79, p. 845, Sept. 1, 1950; *Rept. Prog. in Phys.*, Vol. 24, p. 1, 1961.

or, using Equation (16) for K,

$$\frac{dE}{dx} = \beta \frac{\hbar\omega^2}{v_F}.$$

The second electron then "samples" an element of length of the trough on the order of the uncertainty radius of the electron $\hbar/mv_F$. The attractive energy sensed by the second electron is then*

$$\begin{aligned} E &= \frac{\hbar}{mv_F} \frac{dE}{dx} \\ &= \beta \frac{\hbar^2\omega^2}{mv_F^2} \\ &= \frac{\beta}{2}\left(\frac{\hbar\omega}{E_F}\right)\hbar\omega. \end{aligned} \tag{26}$$

For reasonable values of the parameters, the attractive energy is equivalent to a few degrees Kelvin. If, on the other hand, dE/dx is estimated from the experimental values of resistivities of polyvalent metals, the interaction energy of Equation (26) ranges up to about 10°K, namely, the range of transition temperatures for many superconductors.

General Comment on Mobility

It is worth noting that the present semiclassical approach to mobility is particularly suited to estimating the transport properties of amorphous materials, in which there is a growing interest. The major factor in computing mobility is the rate of loss of energy, dE/dt, by an electron. The rate of loss of energy is largely insensitive to the periodicity of the medium, and should be equally valid for both crystalline and amorphous materials. The major uncertainty is in the choice of effective electron mass. To this extent one would expect that the *phonon contribution* to mobility would remain of the same order of magnitude in amorphous as in crystalline materials. This is born out in the case of metals, where the resistivity usually increases by only a

* A result of comparable magnitude was derived by K. S. Singwi by a more detailed argument, *Phys. Rev.*, Vol. 87, p. 1044 (1952).

factor of about two in going from the solid to the liquid state. Even this factor of two increase may be ascribed to the increase in elastic scattering due to the increased disorder. The temperature-dependent part of the resistivity, which is a measure of the phonon contribution, is substantially unchanged on melting.

It is also true (see Part II p. 621) that the rate of loss of energy by an electron in an ionic crystal is substantially the same whether it is computed conventionally by perturbation theory (electron–phonon interaction) or by treating the ions as a dense plasma and computing the rate of energy loss to plasma excitations. Periodicity, in brief, plays only a secondary role.

Mean Energies and Mobilities of Hot Carriers

In general, when an electric field is applied to a semiconductor, the average energy of the free carriers must be increased. The increase in energy is just sufficient to allow the average rate of energy input from field to free carriers to be passed on to the lattice via the emission of phonons. Loosely speaking, the temperature of the carriers is raised above the lattice temperature, thereby permitting a flow of energy from the carriers to the lattice. We say loosely speaking only because the energy distribution of the hot carriers is not, in general, a thermal distribution. We are concerned in this section with the mean energy of the hot carriers and discuss their distribution in energy separately below in the section on "Energy Distributions of Hot Carriers".

We make several approximations in order to simplify the calculations and to retain the emphasis on the physical processes taking place. These approximations still allow us to reproduce the first-order results already obtained in the literature by more detailed calculations.

The first approximation is that we neglect the energy spread of the carriers and treat them as if they all had the same energy. This is obviously a reasonable approximation for energy distributions that are sharply peaked at their mean energy.

The second approximation is that we confine the calculation to the dominant phonons in the spectrum of electron–phonon interaction. This means the shortest-wavelength phonons with which the electron can interact, namely, those with wavelengths comparable with the uncertainty radius of the electron. Since the interactions are strongly peaked at the shortest-wavelength phonons, the error involved is probably less than a factor of two.

Finally, we confine our treatment to "hot" as opposed to "warm" electrons. The mean energy of hot electrons is larger than about twice their thermal energy. The mean energy of warm electrons is only

somewhat greater than their thermal energy. The distinction between the two regimes is roughly given by the following consideration. In the absence of an applied field, there is a certain rate of exchange of energy between electrons and phonons, the net flow in thermal equilibrium being, of course, zero. At low applied fields, the mean energy of the electron is increased slightly so that the balance is tipped slightly in favor of more energy flowing from electrons to phonons than the reverse. If, for example, the rate of emission of energy by an electron were proportional to its energy, one would expect only a 1% increase in mean energy of electrons when I^2R losses incurred by the applied field were 1% of the thermal equilibrium rate of exchange of energy between electrons and phonons. By the same argument, hot electrons would be generated when the energy delivered by the applied field becomes comparable with the thermal equilibrium rate of exchange. The mean energy of electrons would then be about twice their thermal energy. Other energy dependencies of the rate of emission of energy by electrons would yield qualitatively similar results.

For hot electrons, then, we can equate the rate of energy input by the field to the spontaneous rate of energy loss by the electrons (see Fig. 1)

$$\mathcal{E}ev_d = \frac{\hbar\omega}{\tau_e} = Av^n. \tag{27}$$

Here, v_d is the drift velocity ($\mathcal{E}\mu$) of electrons, τ_e the net emission time for phonons of frequency ω, and Av^n is the rate of emitting energy by electrons of energy $mv^2/2$. The complete expressions for the various phonons and couplings are given in Table I. By way of reminder, n is unity for nonpolar optical phonons and for acoustic phonons with piezoelectric coupling. n is 3 for acoustic phonons with deformation-potential coupling. For polar optical phonons, n is minus unity (ignoring the logarithmic factor) so that, in the approximation of a single energy for all electrons, there is no stable value for the energy of hot electrons above the energy of optical phonons. This instability is the basis for intrinsic dielectric breakdown (by impact ionization) as well as the basis for the Gunn effect.

We make use of the relations

$$v_d = \mathcal{E}\mu = \mathcal{E}\frac{\tau_e e}{m}, \tag{28}$$

$$2mv = \begin{cases} \dfrac{\hbar\omega}{v_s} & \text{for acoustic phonons } (v_s = \text{constant}) \quad (29a) \\ \dfrac{\hbar}{\lambda} & \text{for optical phonons } (\omega = \text{constant}) \quad (29b) \end{cases}$$

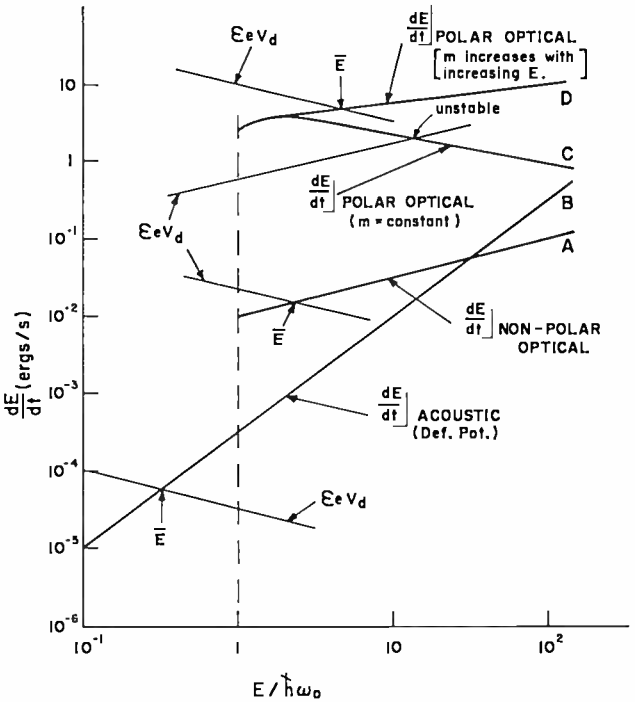


Fig. 1—Curves A, B, C, and D are representative rates of energy loss by electrons of energy E to the phonons indicated. Curves, labeled $\mathcal{E}cv_d$ are representative rates of energy gain by electrons of energy E from the applied field $\mathcal{E}$. The intersections marked $\bar{E}$ define the average energy of hot electrons.

$$\tau_c = \begin{cases} 2\dfrac{kT}{\hbar\omega}\tau_e & \text{for } 2kT > \hbar\omega \quad (30a) \\ \tau_e & \text{for } 2kT < \hbar\omega \quad (30b) \end{cases}$$

where kT $(= mv_t^2/2)$ is the energy of electrons in thermal equilibrium with the lattice. The condition $2kT > \hbar\omega$ ensures that the induced emission and absorption due to thermal phonons is dominant over spontaneous emission. If spontaneous emission is dominant, $\tau_c = \tau_e$.

Interaction with Acoustic Phonons $(2kT > \hbar\omega)$

Combination of the left-hand side of Equation (27) with Equations (28), (29a), and (30a) yields

$$v = \frac{v_d v_t}{2v_s}. \tag{31}$$

Equation (31) states that the random velocity v of hot electrons is proportional to their drift velocity. It also defines the threshold field for the onset of hot electrons, namely, that field for which the drift velocity v_d is somewhat in excess of twice the velocity of sound.

Combination of Equations (27) and (31) yields

$$v = \left(\frac{2e}{A}\frac{v_s}{v_t}\right)^{1/(n-1)} \mathcal{E}^{1/(n-1)} \tag{32}$$

From Equations (31) and (32) the following derived relations are also of interest.

$$v_d = 2\left(\frac{2e}{A}\right)^{1/(n-1)}\left(\frac{v_s}{v_t}\right)^{n/(n-1)} \mathcal{E}^{1/(n-1)} \tag{33}$$

$$\mu \equiv \frac{v_d}{\mathcal{E}} = 2\left(\frac{2e}{A}\right)^{1/(n-1)}\left(\frac{v_s}{v_t}\right)^{n/(n-1)} \mathcal{E}^{(2-n)1/(n-1)} \tag{34}$$

For acoustic phonons coupled via a deformation potential, $n = 3$ (see Table I) and Equations (32)-(34) yield, in summary form,

$$\begin{aligned} v &\propto \mathcal{E}^{1/2} \\ \mu &\propto \mathcal{E}^{-1/2} \\ v_d &\propto \mathcal{E}^{1/2} \\ l &\equiv \text{mean free path} = \text{constant}. \end{aligned} \tag{35}$$

Since the energy $(mv^2/2)$ of the hot electrons is proportional to $\mathcal{E}$ and since they begin to be heated when the drift velocity is about twice the velocity of sound, we can write

$$\text{mean energy of hot electrons} \equiv \bar{E} \approx \frac{\mathcal{E}}{\mathcal{E}_c} kT, \tag{36}$$

where $\mathcal{E}_c\mu \equiv 2v_s$. Equations (35) and (36) are in close agreement with Reference (6).

For acoustic phonons coupled via a piezoelectric field, $n = 1$. But for $n = 1$ all of the Equations (32)-(34) diverge. The meaning of this is that once the electrons begin to be heated, namely for fields greater than $2v_s/\mu$, there is no stable value for their energy at which the energy input from the field is balanced by energy loss to phonons. We can see this by noting, with the aid of Equation (31), that

$$\text{rate of energy input by field} = \mathcal{E}\, e\, v_d = \mathcal{E}\, e \frac{2v_s}{v_t} v,$$

$$\text{rate of energy loss to phonons} = Av.$$

Hence, when $2\mathcal{E}\, e\, (v_s/v_t) > A$, the electrons continue to receive more energy from the field than they can pass on to the phonons, and their mean energy continues to increase until some other type of phonon intervenes or until the assumptions leading to Equations (32)-(34) are no longer valid. The inequality $2\mathcal{E}\, e\, (v_s/v_t) > A$ occurs at the onset of hot electrons. This instability agrees with that analyzed by Laikhtman.[7]

Actually, energy loss to optical phonons will intervene to keep the energy of the hot electrons from diverging to large values. If the optical phonons did not intervene, the energy of the hot electrons would increase until Equation (30b) rather than Equation (30a) became valid. The condition for Equation (30b) is

$$\hbar\omega > 2kT,$$

or, since

$$\hbar\omega = 2mvv_s,$$

the energy of the hot electrons at the onset of Equation (30b) will then be

[6] W. Shockley, "Hot Electrons in Germanium and Ohm's Law," *Bell Syst. Tech. Jour.*, Vol. 30, p. 990, Oct. 1951.

[7] B. D. Laikhtman, "Deviations from Ohm's Law in Piezoelectric Semiconductors," *Soviet Physics-Solid State* (Translation), Vol. 6, p. 2573, May 1965.

$$\frac{mv^2}{2} > \left(\frac{kT}{mv_s^2/2}\right)\frac{kT}{4}. \tag{37}$$

At room temperature this energy is about 6 eV, and at liquid nitrogen about 0.3 eV.

Interaction with Acoustic Phonons ($\hbar\omega > 2kT$)

For $\hbar\omega > 2kT$, $\tau_c = \tau_\varepsilon$ and Equation (27) becomes

$$\frac{\mathcal{E}^2 e^2 \tau_c}{m} = \frac{\hbar\omega}{\tau_c} = Av^n. \tag{38}$$

From the right-hand side of Equation (38), together with Equation (29a), we get

$$\tau_c = \frac{\hbar\omega}{Av^n} = \frac{2m\,v_s}{A}\,v^{1-n}. \tag{39}$$

From the left-hand side of Equation (38) we get

$$v_d \equiv \frac{\mathcal{E} e \tau_c}{m} = \left(\frac{\hbar\omega}{m}\right)^{1/2} = (2v\,v_s)^{1/2}. \tag{40}$$

Combination of Equations (39) and (40) yields the relations

$$v = \left(\frac{2e^2 v_s}{A^2}\right)^{1/(2n-1)} \mathcal{E}^{2/(2n-1)} \tag{41}$$

$$\mu \equiv \frac{(2v\,v_s)^{1/2}}{\mathcal{E}} = \left(\frac{e}{A}\right)^{1/(2n-1)} (2v_s)^{n/(2n-1)}\,\mathcal{E}^{(2-2n)/(2n-1)} \tag{42}$$

Equations (40)-(42) applied to acoustic phonons with deformation potential coupling, for which $n = 3$, give in summary form

$$\begin{aligned} v &\propto \mathcal{E}^{2/5} \\ \mu &\propto \mathcal{E}^{-4/5} \\ v_d &\propto \mathcal{E}^{1/5} \\ l &\propto v\mu \propto \mathcal{E}^{-2/5} \\ \bar{E} &\propto v^2 \propto \mathcal{E}^{4/5} \end{aligned} \tag{43}$$

The same equations applied to acoustic phonons with piezoelectric coupling ($n = 1$) give

$$\begin{aligned} v &\propto \mathcal{E}^2 \\ \mu &\propto \text{constant} \\ v_d &\propto \mathcal{E} \\ l &\propto v\mu \propto \mathcal{E}^2 \\ \bar{E} &\propto v^2 \propto \mathcal{E}^4 \end{aligned} \tag{44}$$

Note that the field dependence of the energy of hot carriers applies to the range of energies for which Equation (30b) is valid, namely,

$$E > \left(\frac{kT}{mv_s^2/2} \right) \frac{kT}{4} .$$

Interaction With Optical Phonons

The rate of energy loss to optical phonons rises abruptly as the mean carrier energy approaches the energy of an optical phonon.

The maximum rate of loss of energy to polar optical phonons occurs at a carrier energy of about $1.5\hbar\omega$ and then declines slowly, so that at a carrier energy of 10 eV the rate of loss is about ⅓ that of the maximum (see Figure 1).

The rate of loss to nonpolar optical phonons continues to increase after the abrupt rise and at a rate proportional to the velocity of the hot carrier.

When the mobility is dominated by optical phonons, Ohm's law is satisfied at low fields, and the drift velocity increases linearly with field until it approaches the value $(\hbar\omega/m)^{1/2}$. At this point, the applied field is sufficient to accelerate the carrier to the energy of an optical phonon in a time equal to its low-field-mobility mean free time. As the field is increased beyond this point, the drift velocity remains constant at the value $(\hbar\omega/m)^{1/2}$, the energy of the carrier remains constant at a value $\approx \hbar\omega$, and the mean free time varies inversely with applied field. These statements are derivable from Equation (38), for which $\hbar\omega > kT$ and $\tau_e = \tau_c$:

$$v_d \equiv \frac{\mathcal{E}e\tau_c}{m} = \left(\frac{\hbar\omega}{m} \right)^{1/2},$$

$$\mu \equiv \frac{v_d}{\mathcal{E}} = \left(\frac{\hbar\omega}{m} \right)^{1/2} \mathcal{E}^{-1},$$

$$v \approx v_d,$$
$$l \propto \mathcal{E}^{-1},$$
$$\bar{E} \sim h\omega. \tag{45}$$

This pattern of behavior is known as "streaming", since the carrier is accelerated to an optical phonon energy in substantially a straight line, emits an optical phonon, and repeats the cycle. The pattern is maintained with increasing field until the time within which a carrier is accelerated to the energy of an optical phonon becomes shorter than the time for emitting an optical phonon. At this point, the mean energy of a carrier is increased beyond that of an optical phonon.

In the case of polar optical phonons, the energy of the carrier rapidly diverges with increasing field to energies that are limited by the advent of other sources of energy loss, such as impact ionization across the forbidden gap. This is the well known Fröhlich[8] model for intrinsic dielectric breakdown, as well as the model for the Gunn[9] instability. The critical field at which this instability sets in is obtained immediately by equating energy input to the maximum loss rate to optical phonons (see Table I):

$$\mathcal{E}_c' e v_d = \beta \frac{e^2\omega^2}{\epsilon_\infty v}. \tag{46}$$

Here, $$\beta = \frac{\epsilon_0 - \epsilon_\infty}{\epsilon_0} \text{ and } v \approx v_d = \left(\frac{h\omega}{m}\right)^{1/2},$$

so that Equation (46) becomes

$$\mathcal{E}_c' = \frac{\epsilon_0 - \epsilon_\infty}{\epsilon_0 \epsilon_\infty} \frac{e\omega m}{h}. \tag{47}$$

Some representative values for $\mathcal{E}_c$ are given in Table II together with approximate values for critical fields taken from impact ioniza-

[8] H. Fröhlich, "Theory of Electrical Breakdown in Ionic Crystals," *Proc. Roy. Soc. London*, Vol. A160, p. 230, 18 May, 1937. For more recent discussion see R. Stratton, *Progress in Dielectrics*, Heywood and Co., Ltd., London, Vol. 3, p. 235, 1961.

[9] J. B. Gunn, "Instabilities of Current in III-V Semiconductors," *IBM Jour. Res. & Develop.*, Vol. 8, p. 141, April 1964.

Table II

Material	E_g (ev)	ϵ_o	ϵ_∞	$\frac{\epsilon_o - \epsilon_\infty}{\epsilon_o \epsilon_\infty}$	$\frac{m}{m_o}$	$\hbar\omega$ (ev)	$\mathcal{E}_c$ (V/cm) Computed (Equation 47)	$\mathcal{E}_c$ (V/cm) Observed	Remarks
SiO_2	8	3.9	2.2	0.2	(≈1)	≈0.1	4.5×10^6	5×10^6	dielectric breakdown
NaCl	8	5.7	2.2	0.3	(≈1)	0.024	2×10^6	1.6×10^6	dielectric breakdown
CdTe	1.4	10.6	7.2	4×10^{-2}	0.1	0.02	2×10^4	10^4	Gunn effect
GaAs	1.4	12.5	11	10^{-2}	0.07	0.033	5×10^3	3×10^3	Gunn effect
InAs	0.36	14	11.6	1.5×10^{-2}	0.02	0.03	2×10^3	10^3	impact ionization
InSb	0.17	17	16	3.7×10^{-3}	0.015	0.025	3×10^2	4×10^2	impact ionization
CdS	2.5	9.3	5.2	8×10^{-2}	0.2	0.036	1.3×10^5	2×10^6	dielectric breakdown
ZnSe	2.6	8.1	5.9	4.5×10^{-2}	0.2	0.03	6×10^4	2×10^6	impact ionization
ZnO	3.3	8.5	3.7	0.15	0.3	0.07	7×10^4	4×10^6	impact ionization
GaP	2.2	10.2	8.5	2×10^{-2}	0.12	0.05	2.8×10^4	$\approx 10^6$	impact ionization
GaAs	1.4	12.5	11	10^{-2}	0.5	0.033	3.5×10^4	5×10^5	impact ionization
GaAs-GaP	2.0							10^6	impact ionization

Comments:

(1) It is not certain that dielectric breakdown in SiO_2 and NaCl is caused by impact ionization. The computed values only show that impact ionization is not expected to occur before the observed dielectric breakdown. The value $m/m_o = 1$ is chosen as a nominal value for computational purpose.

(2) For CdTe, GaAs, InAs and InSb, the concurrence of computed and experimental values for $\mathcal{E}_c$ is significant. Both the Gunn effect and impact ionization in small-bandgap materials emphasize the band properties near the band edge.

(3) For the remaining materials, the large discrepancy between observed and computed values of $\mathcal{E}_c$ is strong evidence that impact ionization is determined by the band properties at energies several volts above the band edge. Increased effective mass is one major contribution. Additional energy loss by intervalley scattering (e.g., in GaAs, GaP and GaAs-GaP) is a second.

References for Table II

SiO_2—A. M. Goodman, private communication; N. Klein and H. Gafni, "The Maximum Dielectric Strength of Thin Silicon Oxide Films," *IEEE Trans. on Electron Devices*, Vol. ED-13, p. 281, Feb. 1966.
NaCl—R. Cooper, "The Electric Breakdown of Alkali Halide Crystals," *Progress in Dielectrics*, Academic Press, Inc., New York, Vol. 5, p. 97, 1963; and R. Williams, "High Electric Fields in Sodium Chloride," *Jour. Phys. Chem. Solids*, Vol. 25, p. 853, 1964.

CdTe—M. R. Oliver, A. L. McWhorter, and A. G. Foyt, "Current Runaway and Avalanche Effects in n-CdTe," *Appl. Phys. Letters*, Vol. 11, p. 111, 15 Aug. 1967.

GaAs—Gunn Effect—State of the Art.
InAs—M. Steele and S. Tosima, "Electron-Hole Scattering in Solids Exhibiting Band-Gap Impact Ionization," *Japanese Jour. Appl. Phys.*, Vol. 2, p. 381, July 1963; J. W. Allen, M. Shyam, and G. L. Pearson, "Gunn Oscillations in Indium Arsenide," *Appl. Phys. Letters*, Vol. 9, p. 39, 1 July 1966.

InSb—M. Toda, "A Plasma Instability Induced by Electron-Hole Generation in Impact Ionization," *Jour. Appl. Phys.*, Vol. 37, p. 32, Jan. 1966; J. C. McGroddy and M. I. Nathan, *Jour. Phys. Soc. Japan*, Vol. 21, Suppl. 437, 1966.

CdS—R. Williams, "Dielectric Breakdown in Cadium Sulfide," *Phys. Rev.*, Vol. 125, p. 850, Feb. 1, 1962; A. Many, "High-Field Effects in Photoconducting Cadium Sulphide," *Jour. Phys. Chem. Solids*, Vol. 26, p. 575, 1965; K. W. Böer and K. Bogus, *Phys. Rev.*, Vol. 176, p. 899, 1968.

ZnSe—R. Williams, "Impact Ionization in ZnSe and Comparison with CdS," *Physics Letters*, Vol. 25A, p. 445, 25 Sept. 1967.

ZnO—R. Williams and A. Willis, "Electron Multiplication and Surface Charge on Zinc Oxide Single Crystals," *Jour. Appl. Phys.*, Vol. 39, p. 3731, July 1968.

GaP—H. G. White and R. A. Logan, "GaP Surface-Barrier Diodes," *Jour. Appl. Phys.*, Vol. 34, p. 1990, July 1963.

GaAs—H. Kressel and G. Kupsky, "The Effective Ionization Rate for Hot Carriers in GaAs," *Int. Jour. Electronics*, Vol. 20, p. 535, June 1966.

GaAs-GaP—R. Williams, "Impact Ionization and Charge Transport in GaAs:GaP 50% Alloy," *Jour. Appl. Phys.*, Vol. 39, p. 57. Jan. 1968.

tion, dielectric breakdown, or the Gunn instability. Factor of two agreement is confirmed in the upper half of Table II from critical fields of 10^2 to 10^6 volts/cm. A word of caution is, however, in order for NaCl, or the alkali-halides in general. While extensive work has been carried out in the last forty years on dielectric breakdown in the alkali-halides, and while the results appeared to bear out Fröhlich's model, it is highly likely that breakdown in these materials is caused by field emission from the electrodes together with heating due to large space-charge-limited currents rather than by impact ionization.[10] If the field emission could, however, be avoided, it is likely that the breakdown by the Fröhlich mechanism would take place at very nearly the same fields. In the smaller gap materials, the critical field is smaller and there is less likelihood of competitive processes such as field emission.

In the case of nonpolar optical phonons, we return to Equation (38) using the value $n = 1$ obtained from Table I. From the right-hand side of Equation (38) we get

$$v = \frac{\hbar\omega}{\tau_c A}, \tag{48}$$

and from the left-hand side,

$$\tau_c = \frac{1}{\mathcal{E}e}(\hbar\omega m)^{1/2}. \tag{49}$$

Equations (48) and (49) combine to yield

$$v = \frac{e}{A}\left(\frac{\hbar\omega}{m}\right)^{1/2}\mathcal{E},$$

$$\mu = \left(\frac{\hbar\omega}{m}\right)^{1/2}\mathcal{E}^{-1},$$

$$v_d = \left[\frac{\hbar\omega}{m}\right]^{1/2}$$

$$l = \text{constant}. \tag{50}$$

[10] R. Williams, "High Electric Fields in Sodium Chloride," *Jour. Phys. Chem. Solids*, Vol. 25, p. 853, 1964.

Here, the carrier energy increases as $\mathcal{E}^2$, while the drift velocity remains constant and the mobility and mean free time decrease as $\mathcal{E}^{-1}$. The energy of hot carriers can be written nominally in the form

$$\bar{E} \approx \left(\frac{\mathcal{E}}{\mathcal{E}_c}\right)^2 \hbar\omega, \tag{51}$$

where $\mathcal{E}_c$ is the field at which an electron is accelerated to the energy of an optical phonon in a time equal to the time required for an electron to emit an optical phonon. Thus,

$$\mathcal{E}_c \frac{e}{m} \tau_c = \left(\frac{2\hbar\omega}{m}\right)^{1/2}$$

or

$$\mathcal{E}_c = \frac{(2\hbar\omega m)^{1/2}}{e\tau_c} = \frac{2\hbar\omega}{el}. \tag{52}$$

If we take l to be about 10^{-6} cm (see discussion in the section on Rates of Impact Ionization) and $\hbar\omega = 0.06$ eV (values appropriate to silicon),

$$\mathcal{E}_c = 10^5 \text{ V/cm}$$

Significant impact ionization should then take place at fields sufficient to yield a mean energy of hot electrons of the order of half the energy of the forbidden gap. By Equation (51), this would mean

$$\mathcal{E} \approx 3 \times 10^5 \text{ volts/cm},$$

a value that compares favorably with experimental values around 2×10^5 volts/cm.[11]

Energy Distributions Of Hot Carriers

General Remarks

In the previous section the mean energies of hot carriers were obtained by simplifying assumptions that kept the physical processes in constant view. The results match those obtained by a variety of

[11] C. A. Lee, R. A. Logan, R. L. Batdorf, J. J. Kleimack, and W. Wiegmann, "Ionization Rates of Holes and Electrons in Silicon," *Phys. Rev.*, Vol. 134, p. A761, 4 May 1964.

authors by more detailed considerations—usually carried out in momentum space rather than in real space. We extend the analysis to obtain some insight into the distribution in energy about the mean energy,* again by using assumptions designed to keep the physical processes in evidence and the analysis compact. It must be remembered that there are still remarkably little, if any, direct data on any actual distributions, and that these are difficult to obtain.** The indirect estimates of energy distributions based, for example, on interpreting rates of impact ionization in back-biased junctions are tenuous indeed.

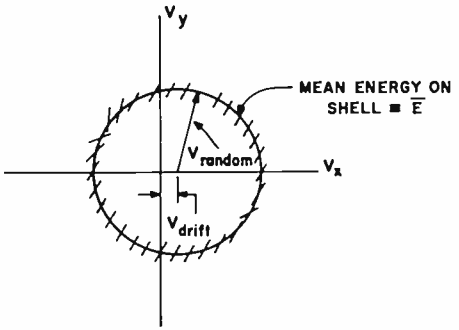


Fig. 2—A drifted Maxwellian distribution showing schematically the large ratio of random velocity v_r to drift velocity v_d.

Basis for Analytical Model

The problem to be analysed is sketched in Figure 2. We have a distribution of electrons in momentum space located predominantly on a roughly spherical shell that is displaced from the origin by the drift velocity. The random velocities are, in general, large compared with the drift velocity. A major exception is the case of streaming, discussed in the previous section under optical phonons, for relatively low-energy ($\approx \hbar\omega_{optical}$) electrons. A less pronounced departure is an elongation of the spherical distribution in the direction of drift. The major effect of these departures is to increase the mean drift velocity and hence the mean power input $\mathcal{E}ev_d$.

We wish to estimate the distribution of electrons in energy shells

* The emphasis in the present paper is on electron energies in excess of the optical phonon energy. For a treatment of energies below the optical phonon energy, see, e.g., R. S. Crandall, *Phys. Rev.*, Vol. 169, p. 585 (1968).

** See, e.g., E. D. Savoye and D. E. Anderson, *Jour. Appl. Phys.*, Vol. 38, p. 3245 (1969).

neighboring the shell of mean energy $\bar{E}$. For most problems of an applied nature, the distributions of electrons for energies greater than $\bar{E}$ is of major interest. The physical processes determining $\bar{E}$ are sketched in Figure 3. Here we show the dominant energy input, $\mathcal{E}ev_d = \mathcal{E}^2e^2\tau_e/m$, which forces electrons toward higher energy shells. A second curve gives the dominant energy loss $\hbar\omega/\tau_e$, which forces electrons towards lower energies. These two curves balance each other at the mean energy $\bar{E}$. The input curve is shown decreasing toward higher energies, while the loss curve is shown increasing towards higher energies. These slopes are based on the fact that in general,

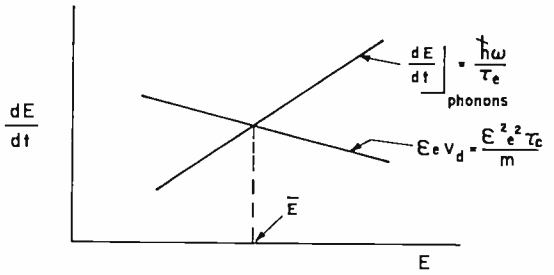


Fig. 3—Representative rates of energy gain and loss curves as a function of electron energy E. The intersection at $\bar{E}$ defines the mean energy of hot electrons.

τ_e and τ_e decrease toward higher energies. If all of the electrons were initially concentrated at $\bar{E}$, they would tend to spread out towards higher and lower energy shells by a kind of diffusion process resulting from the statistics of random emission and absorption of phonons. Our analytic model does in fact, calculate the energy distribution on the high side of $\bar{E}$ by equating the diffusive spreading force to the average loss rate $\hbar\omega/\tau_e$.[12]

Figure 3 is a plot in one-dimensional energy space, whereas the primary distribution in Figure 2 is in three-dimensional momentum space. The transformation would certainly be valid if the rate of jumping between opposite surfaces of the sphere were fast compared with the average drift under the action of either of the force curves. In this way an electron would tend to sample various parts of the sphere on the average as its drifts to or from the shell $\bar{E}$. Interaction with

[12] See also L. V. Keldysh, "Concerning the Theory of Impact Ionization in Semiconductors," *Soviet Phys. JETP* (Translation), Vol. 21, p. 1135, Dec. 1965; and T. Kurosawa, "Notes on the Theory of Hot Electrons in Semiconductors," *Jour. Phys. Soc. Japan*, Vol. 20, p. 937, June 1965 for more detailed treatments of diffusion in energy space.

acoustic phonons satisfies this condition easily, since the collisions are predominantly elastic (on the average, $2n_{ph} + 1$ collisions are incurred per net loss of one phonon). n_{ph} is the phonon occupancy, $kT/\hbar\omega \approx v_t/v_s$. In the case of nonpolar optical phonons, the condition is only marginally satisfied since $\tau_e = \tau_r$, that is, the rate of sampling different parts of the sphere is the same as the net rate of emission of phonons. This would tend to populate the sphere more heavily on the right-hand surface, that is, in the direction of the average drift. Such a distortion would tend to increase the average drift velocity and lead to poor saturation of the drift velocity in the hot-electron range. The fact that the drift velocity for n-type Ge saturates very well and for Si reasonably well (within a factor of two) is evidence that the distortion does not play a major role.

There is, however, one factor that we have ignored in transforming from three-dimensional momentum space to one-dimensional energy space. A particle in three-dimensional space will diffuse away from its point of origin with a velocity of the order of $v_r(l/r)$ where v_r is its random velocity, l its mean free path, and r its distance from the origin. This diffusive force stems from the geometrical fact that spherical shells at larger radii have larger volumes and, thereby, increase the probability that a randomly diffusing particle will gravitate toward larger radii. In the present case v_r, r and l refer to velocity, radius, and mean free path in velocity space. This means that, in the absence of spontaneous emission, an electron will drift toward larger energies. This drift should properly be added in Figure 3 to the term $\mathcal{E}ev_d$, forcing an electron to the right. However, as shown in the appendix, the correction is comparable with $\mathcal{E}ev_d$, which we already neglect on the high side of $\bar{E}$ since (as seen by inspection of Figure 3) it rapidly becomes small compared with $\hbar\omega/\tau_e$.

For the purpose of calculating the high-energy tail of electron energies, we have stripped our problem down to the simple balance between diffusive flow in energy space and the energy loss rate due to spontaneous phonon emission;

$$-D\frac{dn}{dE} = n\frac{\hbar\omega}{\tau_e}, \tag{53}$$

where n is the density of electrons in energy space. The diffusion constant D is, generally,

$$D = \langle v^2\tau_e \rangle. \tag{54}$$

where v is a random velocity and τ_c the mean free time between collisions. The random velocity in energy space is the energy change suffered by the electron moving stochastically with or against the field divided by the time between phonon collisions. Thus,

$$D = \left(\frac{\mathcal{E}el}{2\tau_c} \right)^2 \tau_c, \tag{55}$$

where l is the mean free path between collisions. The factor 2 comes from averaging the projection of l along the field. Equation (53) then becomes

$$-\frac{(\mathcal{E}el)^2}{4\tau_c}\frac{dn}{dE} = n\frac{\hbar\omega}{\tau_c}. \tag{56}$$

High-Energy Distribution for Interaction with Acoustic Phonons

We use Equation (56) together with the relations

$$\frac{1}{2} mv^2 = E,$$

$$\tau_c = \frac{2kT}{\hbar\omega}\tau_c,$$

$$\hbar\omega = 2mvv_s,$$

to obtain

$$n = n_0 \exp\left[\frac{-8mv_s^2E^2}{(\mathcal{E}el)^2kT} \right], \tag{57}$$

and the further relations*

$$l = v_t\tau_c,$$

$$\mu = \tau_c\frac{e}{m},$$

* Note: l is a constant for acoustic phonons; v_t is the thermal equilibrium velocity, and τ_c is the low-field collision time.

$$\mathcal{E}_0 = \frac{v_s}{\mu},$$

to obtain

$$n = n_0 \exp\left[-\left[\frac{E}{\frac{\mathcal{E}}{2\mathcal{E}_0} kT}\right]^2\right]. \tag{58}$$

But the mean energy of hot electrons for acoustic phonons coupled via a deformation potential is (see Equation (36))

$$\bar{E} = \frac{\mathcal{E}}{2\mathcal{E}_0} kT.$$

Hence, the final result is

$$n = n_0 \exp\left[-\left(\frac{E}{\bar{E}}\right)^2\right], \tag{59}$$

namely, a distribution more sharply peaked near $\bar{E}$ than would be given by the normal temperature distribution $\exp(-E/\bar{E})$.

The distribution obtained by Yamashita and Watanaba[13] by a more careful treatment has the same form as Equation (57).

Energy Distributions for Interaction with Nonpolar Optical Phonons

Again, we begin with the Equation (56) and, recognizing that $\tau_c = \tau_e$, we obtain at once

$$n = n_0 \exp\left[-\frac{4\hbar\omega E}{(\mathcal{E}el)^2}\right], \tag{60}$$

where, by way of reminder, the mean free path l is a constant. From Equations (51) and (52),

[13] J. Yamashita and M. Wantanabe, "On the Conductivity of Non-Polar Crystals in the Strong Electric Field, I," *Prog. Theor. Phys.*, Vol. 12, p. 443, Oct. 1954.

$$\bar{E} = \frac{(\mathcal{E}el)^2}{4\hbar\omega}, \tag{61}$$

and Equation (60) can then be written

$$n = n_o \exp\left(-\frac{E}{\bar{E}}\right). \tag{62}$$

In brief, the energy distribution for energies greater than $\bar{E}$ does, in this case, follow a normal temperature distribution. Again a more detailed analysis by Wolff[14] yields Equation (60), with a factor of 3 in place of 4 in the argument.

At fields for which $\mathcal{E}el < \hbar\omega$, $\bar{E} < \hbar\omega$ and Equation (61) loses its significance. At these low fields, the energy distribution of electrons is controlled mainly by statistical "leakage" through the optical phonon barrier. The probability that an electron will avoid emitting an optical phonon for n emission times is $\exp(-n)$. The energy acquired by such an electron will be

$$E = n(\mathcal{E}el). \tag{63}$$

Hence the energy distribution of electrons at these low fields is

$$n = n_o \exp\left(-\frac{E}{\mathcal{E}el}\right). \tag{64}$$

Multiple scattering processes have been ignored in writing Equation (63). These will tend to increase the "leakage" past the optical phonon barrier.

Energy Distribution for Interaction with Polar Optical Phonons

It was meaningful to compute a mean energy and an energy distribution for hot electrons interacting with acoustic phonons and with nonpolar optical phonons, because the rates of energy loss to these phonons increase with increasing electron energy. In brief, the distribution is stable (see Figure 1). In contrast, the rate of energy loss to polar optical phonons decreases with increasing electron energy for energies above that of one or two optical phonons. One can expect

[14] P. A. Wolff, "Theory of Electron Multiplication in Silicon and Germanium," *Phys. Rev.*, Vol. 95, p. 1415, Sept. 15, 1954.

certain instabilities to set in as the applied field is increased. And, indeed, if all of the electrons had the same energy, the instability would set in quite abruptly (see Figure 1) when the applied field delivered more energy to the electrons than the maximum rate at which they could lose energy to the phonons. This has been the model for intrinsic dielectric breakdown of ionic insulators—a phenomenon that is known to occur abruptly as the electric field is increased. While the mechanism for breakdown in high-bandgap insulators, particularly in alkali halides for which the model was designed, is still in some doubt, it is nevertheless true that the penetration of the optical phonon barrier does occur abruptly for materials such as GaAs, GaP, and CdTe, as observed in the Gunn effect.

For applied fields up to within a factor of two of the breakdown field, the mean energy of electrons should not significantly exceed that of an optical phonon, namely, about 0.05 eV. For fields at and above breakdown, the energy of the electrons should diverge toward infinity or until interrupted by a new loss mechanism. The latter is commonly impact ionization across the forbidden gap (as in the case of dielectric breakdown) or transfer to a higher mass band as in GaAs. In any event, significant hot-electron energies occur within a narrow range of electric fields near breakdown. There have been attempts[15] to describe analytically the mean energy or "temperature" of electrons in this range. Such analyses must be of questionable value, owing both to the approximations that are normally made to get a tractable analysis and to the difficulty of experimentally confirming any particular analytic results.

The dominant feature of the hot-electron energy is its rapid divergence from a few hundreths of an electron volt toward infinity within about a factor of two in electric field. What we present here is an approximate description of that divergence as derived from Equation (56). The fact that there is even a finite range of field that can be discussed analytically depends on the electrons having a finite energy spread as opposed to being located all at the same energy. The latter would lead to a step-function divergence as a function of field, as mentioned above. Hence, the following simplified analysis yields a form for the energy distribution.

We use an approximate form for the energy gain and loss curves as shown in Figure 4. The loss curve rises abruptly from zero to its maximum value at $E = \hbar\omega$. For higher energies it remains constant.

---

[15] R. Stratton, "The Influence of Interelectronic Collisions on Conduction and Breakdown in Polar Crystals," *Proc. Roy. Soc. London*, Vol. 246A, p. 406, 19 Aug. 1958.

The actual curve has a broad maximum and a slow decrease at higher energies. We take also the energy input curve to be substantially constant as a function of electron energy $\bar{E}$. Both these assumptions mean that $\tau_e = \tau_c = \text{constant}$. Equation (56) then becomes

$$-\frac{(\mathcal{E}el)^2}{4\tau_c}\frac{dn}{dE} = n\left[\frac{\hbar\omega}{\tau_e} - \frac{\mathcal{E}^2e^2\tau_c}{m}\right]. \tag{65}$$

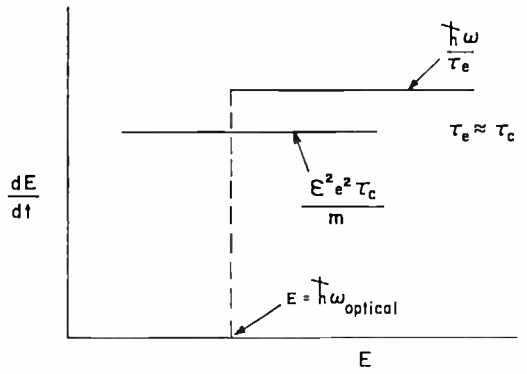


Fig. 4—Schematic approximation for rates of energy gain and loss for polar optical phonons for an electric field approaching breakdown.

For $l = v\tau_c$, $\tau_c = \text{constant}$, and $mv^2/2 = E$, Equation (65) is readily integrable in the form

$$n = n_o\left(\frac{E}{E_o}\right)^{-2\left[(m\hbar\omega/(\mathcal{E}^2e^2\tau_c^2)^2\right)-1]}$$

or

$$n = n_o\left(\frac{E}{E_o}\right)^{-2\left\{\left[(\hbar\omega/\tau_c)\ \mathcal{E}^2e^2v_d\right]-1\right\}} \tag{66}$$

n_o is the density of electrons at energy $E_o = \hbar\omega$. Equation (66) states immediately that when the rate of energy input $\mathcal{E}^2e^2v_d$ is equal to the rate of loss $\hbar\omega/\tau_c$, the electron energy diverges to infinity. This is the criterion for breakdown. If we take as an approximate measure for the average energy ($\bar{E}$) the value of E at which $n = n_o/2$, we can plot

$\bar{E}$ as a function of applied field $\mathcal{E}$. The result in Figure 5 shows that even at a field as high as half the breakdown field, the average energy of the electrons is only $\sqrt{2}\,\hbar\omega$. The calculation of Hilsum[16] based on Stratton's expression for electron temperature shows, if anything, an even more abrupt divergence at the breakdown field. That is, the field at which the electrons achieve an energy of $\sqrt{2}\hbar\omega$ lies even closer to the breakdown field than the value of ½ we calculate here.

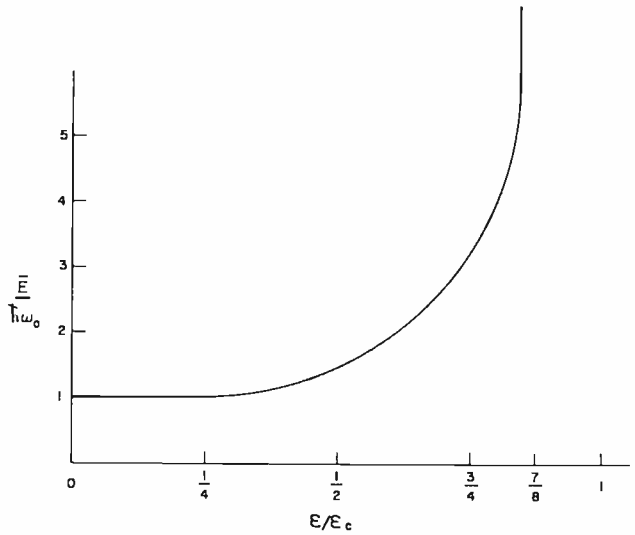


Fig. 5—An estimate of electron mean energy versus applied field for electrons interacting with polar optical phonons. $\mathcal{E}_c$ is the nominal electric field for breakdown and is given by $\mathcal{E}_c = (\epsilon_\infty - \epsilon_s)\, cm\omega/(\epsilon_\infty \epsilon_s \hbar)$. $\hbar\omega_o$ is the energy of the longitudinal optical phonon.

Our Equation (66) does not, of course, yield anything approaching a temperature distribution of energies. The departure from such a distribution is even more marked when one recognizes that even the distribution of Equation (66) must be cut off at electron energies such that $\mathcal{E}ev_d > \hbar\omega/\tau_\epsilon$. This is shown in Figure 1 on curve C where the two curves have their second crossing. These high-energy electrons continually gain energy from the field until they encounter a new energy-loss process.

Equation (66) does show that for larger bandgap energies, the dependence of the rate of impact ionization on electric field becomes

[16] C. Hilsum, "Transferred Electron Amplifiers and Oscillators," *Proc. I.R.E.*, Vol. 50, p. 185, Feb. 1962.

progressively steeper. For impact ionization, the value of E in Equation (66) is chosen to be comparable with the forbidden gap energy.

Rate Of Impact Ionization

An analysis of hot electrons should define the field at which significant impact ionization takes place or, in more detail, the rate of impact ionization as a function of field. The field at which significant impact ionization occurs is relatively easy to estimate. It is that field at which the mean energy of the hot carriers is of the order of a half or a quarter of the energy needed for impact ionization. The spread in energy of the actual distribution around this mean energy will then provide a significant fraction of carriers with energies sufficient to cause impact ionization. Equations (36) and (61) relate the mean energy to the electric field and readily yield estimates of the required fields, provided, of course, that the mean free path for phonon emission is known. These equations apply to interaction with acoustic phonons and with nonpolar optical phonons, respectively. For polar optical phonons, the critical field is given by Equation (47).

Nonpolar Optical Phonons

The rate of impact ionization as a function of electric field should be readily derivable from any of the energy distributions given by Equations (58), (60), (64), or (66). That is, one inserts into these equations a value for E (the energy at which significant impact ionization takes place) and a value for l (the mean free path for phonon emission) and proceeds to calculate n/n_o, a measure of the fraction of electrons with energies greater than E, as a function of electric field. Alternatively, the same procedure can be used to compute a value for l knowing experimentally the rate of impact ionization as a function of field. This is, indeed, what has been done for impact ionization in back-biased p-n silicon junctions. References to this literature are to be found, for example, in the key paper by C. A. Lee et al[11] on ionization measurements in such junctions. Also cited in this paper is the long controversy as to whether the energy distribution used to compute the rate of ionization should be of the form exp $(-A/\mathcal{E}^2)$ shown in Equation (60) and associated with Wolff[14] or of the form exp $(-A/\mathcal{E})$ shown in Equation (64) and associated with Shockley.[17] Finally, the paper cites the resolution of the controversy by the more detailed calcu-

[17] W. Shockley, "Problems Related to p-n Junctions in Silicon," *Solid-State Electronics*, Vol. 2, No. 1 p. 35, 1961.

lations by Baraff,[18] resulting in a mixture of these two forms. In sum, there appears to be general agreement that experimental data should be interpreted in terms of Baraff's universal curves in order to define a proper value for the mean free path for phonon emission.

The major point we wish to make here is that the difference between the models of Baraff and his predecessors, Wolff and Shockley, is small compared with the difference between all three authors and the actual situation. The reason for this statement is that all of the authors have assumed a constant mean free path, l, independent of energy. This is a proper assumption if the effective mass also is a constant. In fact, however, the effective mass is expected to vary in silicon from about 0.3 near the band edge to a value approaching unity at energies of a volt or more above the band edge. The precise variation is not important here. Since the mean free path depends inversely on the *square* of the effective mass,* the mean free path will decrease by almost a factor of ten in the energy range of interest. All of the calculations ignore this variation of mean free path by a *factor of ten* and, at the same time, produce results that are used to deduce a mean free path from experimental data supposedly accurate to about *ten per cent.* This mismatch between the models on which the several analyses have been based and the actual state of affairs, while bad enough, is not quite so serious as the factor of ten would suggest. All of the analyses tend to emphasize the mean free path at energies near the ionization energy rather than near the band edge. This comes about because the average energy $\bar{E}$ depends on the local, that is, high-energy mean free path, at least to the extent that the energy distribution is strongly peaked near $\bar{E}$.

Consistent with the approximate nature of the models for hot-electron energy distributions, we suggest here a simple and approximate method for interpreting experimental data. We note, for example that the data on impact ionization (see Lee et al[11]) cover only a factor of two in electric field. Over this factor of two, the rate of ionization by electrons increases by a factor of 100. The dominant fact is the steep rise of the rate of ionization with field, and not the particular shape of the rise curve. The latter is a tenuous hook on which to hang the validity of a theory. The question we ask, then, is what value of mean free path will yield a 100-fold increase in rate of ionization for a twofold increase in field at the field in question? In terms of Equation (60), and using the more accurate factor of 3 in the exponent,

[18] G. A. Baraff, "Distribution Functions and Ionization Rates for Hot Electrons in Semiconductors," *Phys. Rev.*, Vol. 128, p. 2507, Dec. 15, 1962.

* This dependence stems from the dependence of the coupling constant β on effective mass via the relation $\hbar/\lambda = 2\ mvv_s$.

$$\frac{\exp\left[-\dfrac{3\hbar\omega E_i}{(2\mathcal{E}el)^2}\right]}{\exp\left[-\dfrac{3\hbar\omega E_i}{(\mathcal{E}el)^2}\right]} = 10^2,$$

or

$$2.2\,\frac{\hbar\omega E_i}{(\mathcal{E}el)^2} = 4.6.$$

The relative importance of the several parameters is clearly assessable in this expression. The right-hand side is ln (100) and measures the steepness of the ionization curve. Using the values taken from Lee et al, $\hbar\omega = 0.06$ eV, $E_i = 1.6$ eV, and $\mathcal{E} = 2 \times 10^5$ V/cm, we compute $l = 110$ Å.

A similar computation using Equation (64) is expressed in the form

$$\frac{\exp\left(-\dfrac{E_i}{2\mathcal{E}el}\right)}{\exp\left(-\dfrac{E_i}{\mathcal{E}el}\right)} = 10^2,$$

and yields a value of 90 Å for l. These values are to be compared with 70 Å computed by Lee et al using Baraff's curves. It is worth noting that Equations (60) and (64) do not lead to widely different values of l.

Another simple estimate of l can be made by equating the average energy of electrons to, say, half the ionization energy at the highest fields for which measurements could be made before breakdown. This estimate favors the local high energy value of l.

$$\frac{(\mathcal{E}el)^2}{4\hbar\omega} = \frac{1.6}{2}.$$

For $\mathcal{E} = 4 \times 10^5$ V/cm, $l = 110$ Å.

One note to be kept in mind is that a mean free path of 100 Å yields an emission time of less than 10^{-14} sec, i.e., less than the ω^{-1} of the emitted phonons, and raises questions about the validity of perturbation theory.

Polar Optical Phonons

The field dependence of the rate of impact ionization in the case of electron interaction with *polar* optical phonons should be experimentally too difficult to define for materials with forbidden gaps of a volt or more, owing to the runaway character of the hot electrons. From Equation (66), the field required to produce an average energy of electrons of about a volt is within a few percent of the breakdown field. It would be difficult, indeed, to ensure such uniform ionization over a specimen cross section. These remarks are based on the assumption of a constant effective mass.

It is true that the field dependence of the rate of ionization has been successfully observed for InSb[19, 20] and analysed by Dumke[21]. The forbidden gap, however, is only 0.2 volt, and is sufficiently close to the energy of an optical phonon to permit statistical leakage of electrons past the optical phonon barrier at fields well removed from breakdown. Even here the rate of ionization increases at about the tenth power of the electric field. McGroddy and Nathan[20] computed a mean free path for phonon emission of 1 micron using Baraff's model for *nonpolar* optical phonons. Dumke computed substantially the same values using his model for *polar* optical phonons. Such universal agreement can only be read as the insensitivity of this parameter to "modeling." The dominant factor, as we have pointed out, for computing the mean free path for phonon emission is the electric field at which significant impact ionization takes place. The mean free path is also readily derivable from the material parameters, as shown in Part II of this series. There, for example, the mean free path for optical phonon emission was given as

$$l = r_{II} \frac{\epsilon_0 \epsilon_\infty}{\epsilon_0 - \epsilon_\infty} \frac{m_0}{m},$$

where r_0 = radius of the Bohr orbit (0.5Å), $\epsilon_0\epsilon_\infty/(\epsilon_0-\epsilon_\infty) \approx 2 \times 10^2$ and $m_0/m \approx 50$ for InSb. With these parameters l is again computed to be about 10^4 Å. For the alkali halides, at the other extreme, l is only a few angstroms.

---

[19] M. Toda, "A Plasma Instability Induced by Electron-Hole Generation in Impact Ionization," *Jour. Appl. Phys.*, Vol. 37, p. 32, Jan. 1966.

[20] J. C. McGroddy and M. I. Nathan, *Jour. Phys. Soc. Japan*, Vol. 21, Suppl. 437, 1966.

[21] W. P. Dumke, "Theory of Avalanche Breakdown in InSb and InAs," *Phys. Rev.*, Vol. 167, p. 783, 15 March, 1968.

While we have emphasized the highly precipitous and runaway character of the mean energy of hot electrons versus applied field in polar materials, it should be kept in mind that these arguments are based on a reasonably constant effective mass. If, on the other hand, the effective mass increases by a factor of three or more in the range of a few volts above the band edge, the loss curve dE/dt should increase (rather than decrease) with increasing energy (see Figure 1, Curve D). Under these conditions, an increasing electric field will lead to stable distributions of hot electrons whose mean energy versus field can readily be estimated in terms of the slope of the dE/dt curve. Alternatively, if at higher energies a new phonon loss mechanism is introduced, such as inter-valley scattering, the runaway character can be arrested and a stable high-energy-electron distribution maintained. Either or both these mechanisms can account, as Williams[22] has pointed out, for the relatively high fields required to impact ionize in CdS, ZnSe, ZnO, GaAs and GaP (see lower half of Table II).

Dielectric Breakdown

Dielectric breakdown by impact ionization will occur when an electron on the average creates one successor by impact ionization during its lifetime. The condition can be expressed approximately by

$$\frac{L_D}{L_I}\frac{n(E_I)}{n_o} = 1,$$

where L_D is the drift length of an electron during its lifetime, L_I the distance, in the field direction, an electron of energy E_I travels between impact ionizations, and $n(E_I)/n_o$ is the fraction of electrons having an energy E_I at which significant impact ionization takes place. E_I is somewhat larger than the energetic threshold for ionization. Since $n(E_I)/n_o$ is an exponential function of electric field, the breakdown field strength will depend only weakly (logarithmically) on the lifetime of an electron. Morever since $n(E_I)/n_o$ is such a steep function of electric field, one can conclude that the mean energy $\bar{E}$ of electrons will be in the order of half or more of E_I. This value of $\bar{E}$ then, defines an electric field (see Equations (36), (47), and (61) that should be within a factor of two of the breakdown field. A closer examination of breakdown field strength is hardly warranted, unless the fine details of band

[22] R. Williams, "High Electric Fields in II-VI Materials," *Applied Optics*, Supplement 3, p. 15, 1969.

structure in the neighborhood of E_I are taken into account. Even if such information were available, it is doubtful that the additional insight gained would be worth the computational effort.

Appendix

The rate of diffusion to larger radii in velocity space is $v_r(l_v/r)$, where

$$v_r = \frac{\mathcal{E}e\tau_c/m}{\tau_c} = \frac{\mathcal{E}e}{m},$$

$$l_v = \frac{\mathcal{E}e\tau_c}{m},$$

$$r = v.$$

Thus,

$$\frac{dv}{dt} = v_r \frac{l_v}{r} = \frac{\mathcal{E}^2e^2\tau_c}{m^2v},$$

But

$$\frac{dE}{dt} = mv\frac{dv}{dt}$$

$$= mv\frac{\mathcal{E}^2e^2\tau_c}{m^2v}$$

$$= \frac{\mathcal{E}^2e^2\tau_c}{m} = \mathcal{E}^2e\mu = \mathcal{E}ev_d.$$

IONOSPHERIC PHASE DISTORTION AND FARADAY ROTATION OF RADIO WAVES

By

T. Murakami and G. S. Wickizer

RCA Missile and Surface Radar Division
Moorestown, N.J

Summary—The propagation effects to be expected in wideband transmission of radio waves through the ionosphere are examined briefly and these effects are related to the condition or state of the ionosphere. The pertinent characteristics of the ionosphere are given and the procedure used to calculate the phase distortion and Faraday rotation caused by the ionosphere on a radio wave is illustrated. It is shown that although the phase-path delay and group-path delay vary inversely as the square of the signal frequency, the corresponding phase shifts vary inversely as frequency. The distortion produced by the phase-path and group-path delays are opposite in sign, resulting in a net phase distortion that is predominately parabolic.

Introduction

THERE ARE THREE principal windows through which we on earth may view objects beyond the earth's atmosphere. The first of these is the optical window, which has long been used; the others are the infrared and radio windows, which have become useful with the development of devices sensitive to signals in these frequency ranges. In radar and radio, we are concerned with the radio window that exists from a wavelength of about 20 meters to a few millimeters. The long-wavelength or low-frequency limit is determined by the earth's ionosphere, which reflects waves that are below some critical frequency. Absorption of radio waves by the lower atmosphere sets the high-frequency or short-wavelength limit of radio waves that we can receive from extraterrestrial sources.

Continued improvements in radar and radio astronomy equipment have refined the system performance to a point where environmental

factors have become limitations in achievable system accuracy. Errors caused by the troposphere and ionosphere are especially significant where wide-bandwidth signals are received and/or wide-aperture antenna systems are used. In the frequency ranges to be considered, tropospheric effects are usually negligible at elevation angles greater than about 10 degrees above the horizon, so their effects will not be treated here.

To limit the scope of such a large subject, this brief paper discusses only the refractive properties of the ionosphere with respect to radio waves, with emphasis on phase-distortion effects on the signal. A review of the characteristics of the ionosphere pertinent to ionospheric refraction is first given. This is followed by simplifications of the Appleton–Hartree equations giving the general formulas governing the behavior of the ionosphere. These formulas, in conjunction with ionospheric data, are then applied to the case of radio wave propagation.

Consideration is given to the phase distortion introduced by the ionosphere on a 500-MHz-bandwidth signal centered at 3350 MHz. Calculation is made of the phase-path length and group-path delay distortions over the bandwidth of interest. These calculations are dependent upon the particular ray path and on the electron distribution in the ionosphere. The details of the method of calculation accounting for these factors are given with the final results. An estimate is also given for the Faraday rotation of the plane of polarization of the signal.

Characteristics of the Ionosphere

As seen in the next section, nearly all of the effects of the ionosphere on electromagnetic waves vary directly as the square of the wavelength. These effects are proportional to the integrated electron density along the propagation path of the wave, and can be expressed as the total electron content of a one-meter-square column.

The data presented in this paper are based on vertical soundings made at Fort Monmouth, N. J.[1] during 1960, when sunspot numbers (R) were near present (July 1968) values. The original data were analyzed and processed by ITSA (Institute for Telecommunication Sciences and Aeronomy), and a tabulation on 35-mm microfilm for the year 1960 was obtained from their offices in Boulder, Colorado. Estimates of shinf (electron summation to infinite height) are used to serve

[1] Institute for Telecommunication Sciences and Aeronomy, *Ionospheric Predictions*, (Monthly, 3 months in advance) U.S. Government Printing Office. (Processed data for 1960 obtained from ITSA, Boulder, Colorado).

as the basis for estimates of electron content of transmission paths at other elevation angles. Thus the values of electron content given here apply to paths that extend through the ionosphere and would apply to radar targets that are more than 500 nautical miles above the earth, or to extraterrestrial signal sources.

Sunspot Cycle

An imporant factor to be considered in describing the state of the ionosphere is the position of the time period of interest with respect to the eleven-year sunspot cycle. A graph of the predicted sunspot activity is shown in Figure 1 for the period beginning in January 1967 and extending through February 1971. In this graph, the sunspot activity is measured in terms of the Wolf number R, given by[2]

$$R = k(10g + s) \tag{1}$$

where g = number of groups of sunspots,

s = number of observable individual spots,

k = correction factor for equipment and observer characteristics.

Sunspot activity has been recorded for a period of about 200 years, and shows a definite periodicity of approximately 11 years. Sunspot numbers are available on a daily basis, as well as monthly and yearly averages. In Figure 1, the mean activity of the previous 12 cycles is shown, along with curves of the mean activity plus 10% and of mean plus 20%. Points deduced from current data observed by ESSA (Environmental Sciences Services Administration) are also indicated in Figure 1. For the first part of the cycle, the present activity is seen to be between 10 to 20% above the mean for the past 12 cycles. The latest readings (posted 3 months in advance) show the activity to be close to the mean curve. From these observations, the 1968 maximum sunspot activity will correspond to a sunspot number of about 105. This is about one half the maximum of 201 observed in 1957.

Seasonal and Diurnal Variations of Electron Density

Figure 2 shows the seasonal and diurnal variations of the total electron content based on vertical soundings made at Fort Monmouth,

[2] Kenneth Davies, *Ionospheric Radio Propagation*, National Bureau of Standards Monograph 80, U.S. Government Printing Office.

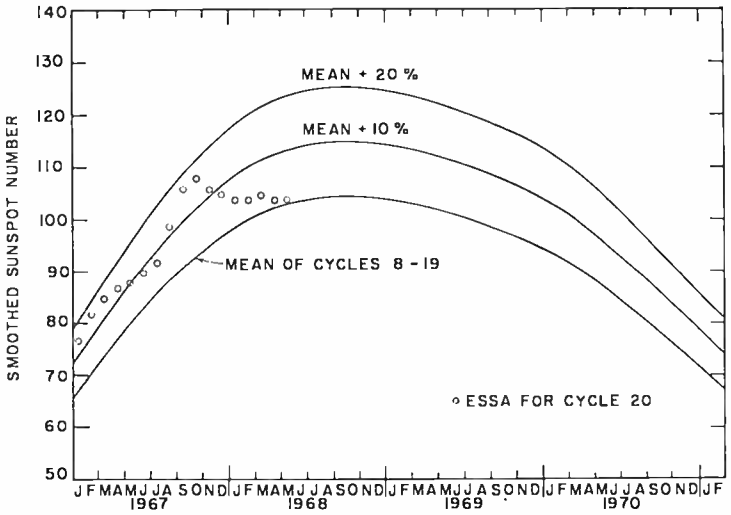


Fig. 1—Predicted sunspot numbers.

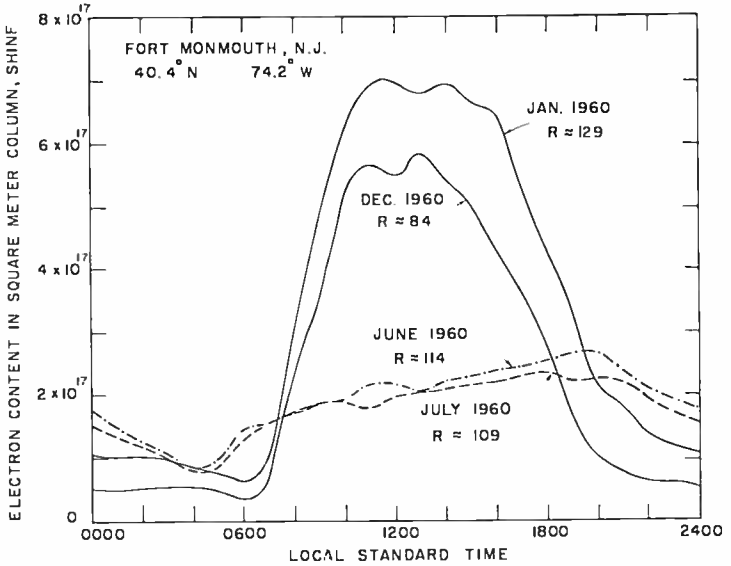


Fig. 2—Electron content in square-meter column, summation to H_{ISF}, (shinf).

N.J. during 1960. These are direct plots of the quantity shinf taken from the ITSA tabulation and represent monthly averages.

Most striking are the high values during mid-day in winter; also the low values during the night and the diurnal minimum that occurs at 0600 hours in winter. The total electron content during the summer

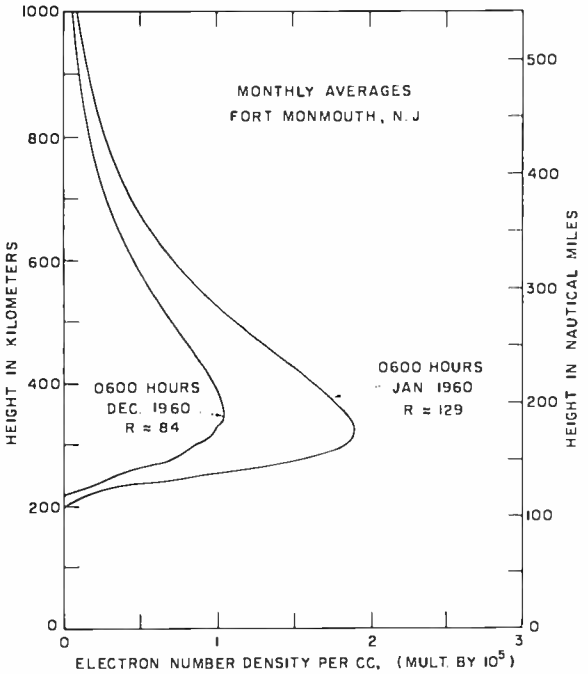


Fig. 3—Ionospheric electron-density profiles for winter diurnal minimum.

is seen to be much less with less relative variation. This means that the greatest distortion of radio waves caused by the ionosphere will occur during the middle of the day in winter. The least distortion will be found about one hour before sunrise, also in winter. From Figure 2 it is seen that the total electron content at the diurnal minimum in summer is about 60% more than the winter minimum and occurs earlier (at 0400 hours) due to the earlier sunrise. From the relative amplitudes of the curves of Figure 2, the degree of expected ionospheric distortion of radio waves can be deduced.

The electron number–density profiles in Figures 3 through 7 are intended more to illustrate the original data than to provide a direct

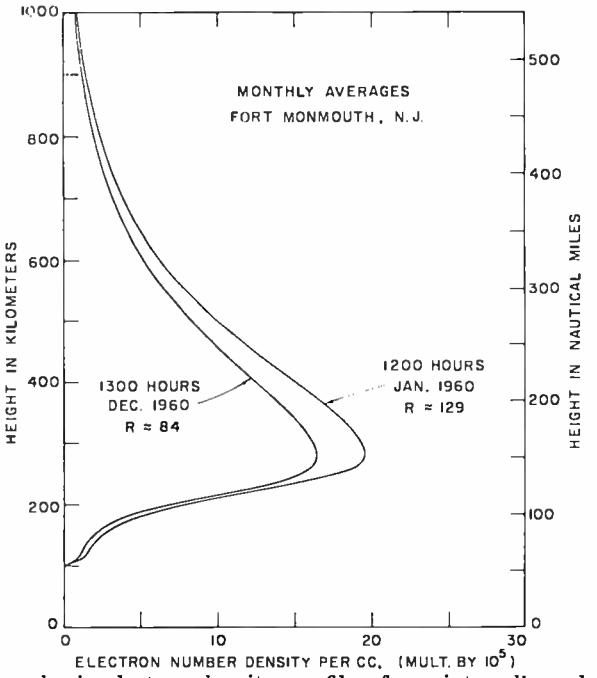


Fig. 4—Ionospheric electron-density profiles for winter diurnal maximum.

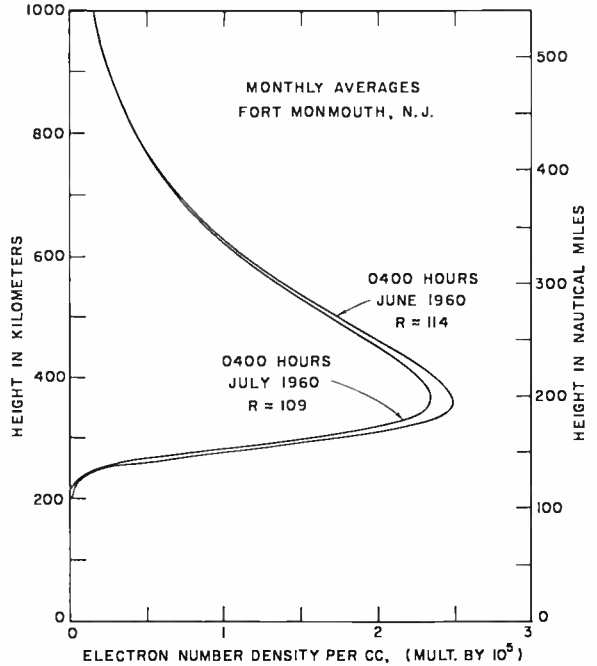


Fig. 5—Ionospheric electron-density profiles for summer diurnal minimum.

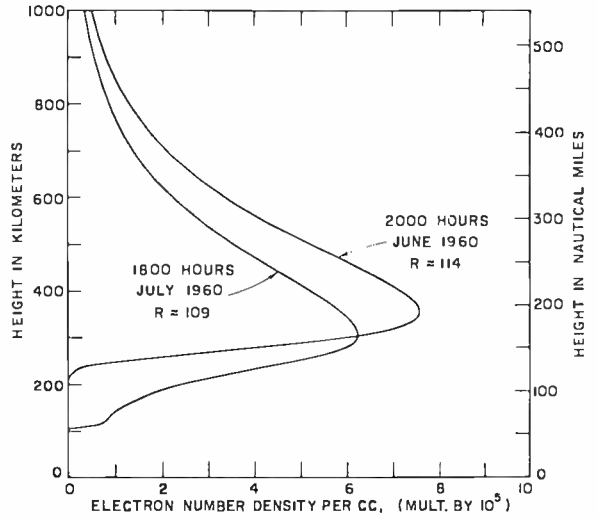


Fig. 6—Ionospheric electron-density profiles for summer diurnal maximum.

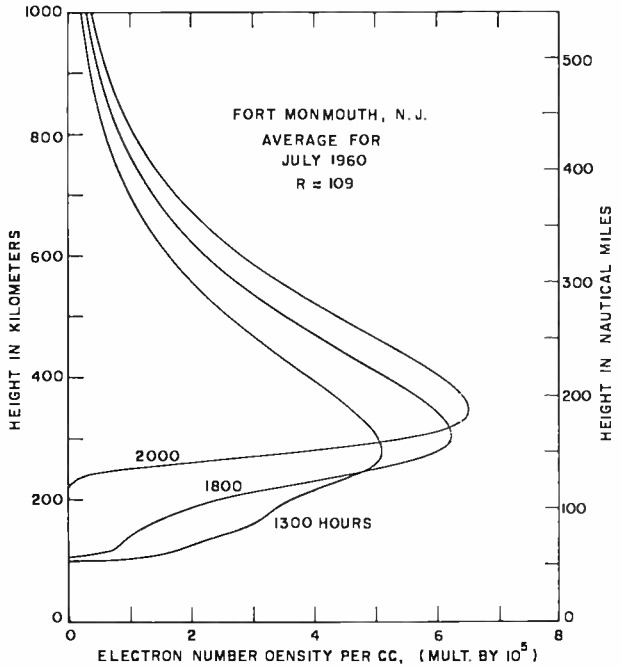


Fig. 7—Ionospheric electron-density profiles.

comparison of diurnal and seasonal variations. Some of the profiles will be used later to compute the total electron-density profiles along a given elevation angle. There are some points of interest illustrated by these curves:

(a) In Figure 3, the strong influence of sunspot number on the winter diurnal minimum electron density;
(b) In figure 4, the presence of the E-layer at 100 km height at midday; the order-of-magnitude increase in maximum electron density; the smaller influence of sunspot number on the winter diurnal maximum;
(c) In Figure 5, the small difference in the electron-density profiles for June and July 1960 at 0400 hours.
(d) In Figure 7, the decay of the E-layer and redistribution of electrons to greater heights between 1300 and 2000 hours.

Day-to-Day Variations

The curves of Figure 2 are monthly averages for the quantity shinf, which is not estimated for individual electron-content profiles. However, the quantity shmax (summation to height of maximum number density) is derived for each hourly profile and therefore provides a measure of the day-to-day variability of the ionospheric electron content. Values for shmax have been plotted in Figure 8 for the condition of winter diurnal maximum and minimum electron content for the month of December 1960 (taken at Fort Monmouth, N.J.).

Aside from the slow trend to lower values, possibly attributable to declining sunspot numbers, points of interest in Figure 8 are daily maximum-to-minimum ratios, and the variability of diurnal minimum values. It appears that the diurnal maximum-to-minimum ratio is at least an order of magnitude, and the day-to-day variation of the minimum electron content is about a factor of two above and below an average value.

Irregularities

The quiet atmosphere is not perfectly homogeneous, but contains horizontal gradients and irregularities. The presence of such irregularities becomes apparent under the dynamic conditions of satellite tracking, where the transmission path is sweeping horizontally through the ionosphere at velocities up to several hundred kilometers per second. Based on observations,[3] variations in total electron content are

[3] C. G. Little and R. S. Lawrence, "The Use of Polarization Fading of Satellite Signals to Study the Electron Content and Irregularities in the Ionosphere," *Jour. Research NBS*, Vol. 64D, No. 4, p. 335, July-Aug. 1960.

about ±1% of the smoothed average over a 400-km horizontal distance at a height of 300 km.

All descriptions of the ionosphere given above are based on average normal conditions, i.e., "quiet" or undisturbed days. This definition applies to a majority of days, the remainder being disturbed from various effects of solar flares in varying degrees.

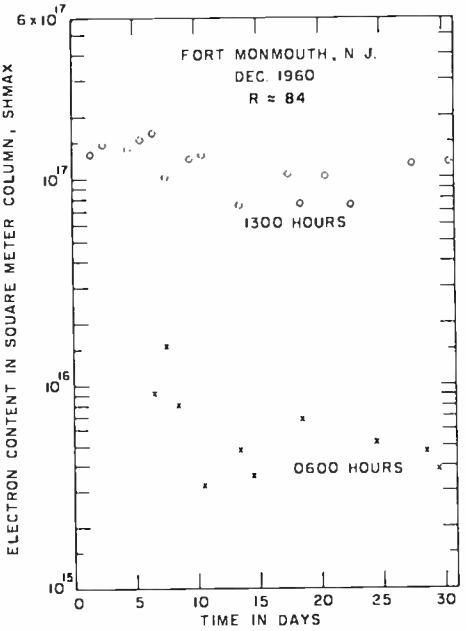


Fig. 8—Electron content in square-meter column, summation to H_{Max} (shmax).

In summary, measurements indicate that more than an order-of-magnitude difference can be expected between diurnal extremes of electron content in winter, at a sunspot number near 100. Minimum electron content appears to be more variable than maximum content on a day-to-day basis. The diurnal variation is much less in summer than in winter and the minimum values are higher. We would expect like variations in the radio-wave propagation, since these effects follow the electron-content behavior.

Simplifications of the Appleton–Hartree Equation

At frequencies that are large compared to the ionospheric penetration frequency, the exact expression for the refractive index can be

expressed in approximate form. The Appleton–Hartree equation for the complex refractive index of the atmosphere is given by[4]

$$n^2 = 1 - \frac{X}{1 - iZ - \frac{Y_T^2}{2(1 - X - iZ)} \pm \sqrt{\frac{Y_T^4}{4(1 - X - iZ)^2} + Y_L^2}} \quad (2)$$

where the following notation has been used:

c = free-space velocity of light,

e = charge on an electron

$f = \omega/(2\pi)$ = frequency of radio wave,

H = geomagnetic field intensity,

m = mass of an electron,

N = number density of free electrons,

ϵ_0 = electric permittivity of free space,

μ_0 = magnetic permittivity of free space,

$\lambda = 2\pi c/\omega$ = free-space wavelength of radio wave,

θ = angle between wave normal and geomagnetic field,

ν = frequency of collision of free electrons with heavy particles,

μ = real part of the refractive index,

$\omega_H = \mu_0 eH/m$ = gyro frequency of an electron,

$\omega_L = \omega_H \cos\theta$,

$\omega_N = [Ne^2/(\epsilon_0 m)]^{1/2}$ = angular plasma frequency,

$\omega_T = \omega_H \sin\theta$,

$X = \omega_N^2/\omega^2$,

$Y = \omega_H/\omega$,

$Y_L = \omega_L/\omega$, (longitudinal component)

$Y_T = \omega_T/\omega$, (transverse component)

$Z = \nu/\omega$.

In the upper regions of the ionosphere, the collision frequency ν is sufficiently small compared with the radio frequencies greater than

[4] J. A. Ratcliffe, *The Magneto-Ionic Theory and Its Applications to the Ionosphere*, Cambridge University Press, Cambridge, England; 1959.

about 1 MHz, so that we may set $Z = \nu/\omega = 0$. The square of the refractive index n^2 becomes a real quantity, μ^2, given by

$$\mu^2 = 1 - \frac{X}{1 - \frac{Y_T^2}{2(1-X)} \pm \sqrt{\frac{Y_T^4}{4(1-X)^2} + Y_L^2}}. \tag{3}$$

This quantity is seen to be double valued with the + sign referring to the so-called ordinary wave and the − sign to the extraordinary wave. For radio frequencies large compared to the gyro frequency, the propagation can be divided into two classes. For

$$\frac{Y_T^4}{(1-X)^2} \gg 4Y_L^2, \tag{4}$$

the propagation is said to be quasi-transverse (QT), since the longitudinal component Y_L can be ignored. In this case, the positive sign in Equation (3) gives a value of μ that tends towards the case where the magnetic field is absent, whereas the negative sign results in a μ that is dependent on the transverse magnetic field.

When

$$\frac{Y_T^4}{(1-X)^2} \ll 4Y_L^2, \tag{5}$$

the propagation is said to be quasi-longitudinal (QL). The refractive index in this case is dependent on the longitudinal component of the magnetic field for both the plus and minus signs in Equation (3). In the QL case, the refractive index is increased for the ordinary wave and reduced for the extraordinary wave when the magnetic field is present relative to the field-free condition.

Approximate Expression for Ionospheric Refractive Index

Under conditions where the magnetic field can be ignored, $Y_T = Y_L = 0$; so that Equation (3) reduces to

$$\mu^2 = 1 - X. \tag{6}$$

For the frequencies under consideration ($\omega/2\pi > 20$ MHz) X will be small compared to unity, so that

$$\mu = (1 - X)^{1/2} \cong 1 - \frac{X}{2}. \tag{7}$$

Denoting $1 - \mu$ by $\Delta\mu$ and using the defined quantity for X,

$$\Delta\mu = \frac{bN}{\omega^2}, \tag{8}$$

where

$$b = \frac{e^2}{2\epsilon_0 m} = 1.6 \times 10^3 \text{ (mks)}.$$

Equation (8) shows that under the conditions of no magnetic field in the ionosphere and no absorption, the deviation of the refractive index from unity is proportional to the electron density, and inversely proportional to the square of the operating frequency.

Ionospheric Distortion Characteristics

Two forms of ionospheric distortion of wide-band signals, other than absorption effects, that are investigated in this paper are (a) departure of phase from a linear shift across the utilized bandwidth and (b) the Faraday rotation of the plane of polarization of the signal. The net phase change across the frequency band of interest is due to two effects. These are the reduction in phase-path length, expressed in units of length, and the group-velocity delay, which is in terms of units of time. Both of these changes can be converted to electrical phase at the appropriate frequency from which the net phase shift across the specified bandwidth is derived.

Change in Phase-Path Length

To determine the effect of the ionosphere upon the phase-path length of the radio wave as it passes through the ionosphere, the effective phase-path length P, given by

$$P = \int_0^s \mu dl, \tag{9}$$

is used,[5] where the integral is taken along the ray path to s. The incremental change, Δl, in the phase-path length due to the ionosphere is then

$$\Delta l = \int_0^s (\mu - 1)\, dl = -\int_0^s \Delta\mu dl = -\frac{b}{\omega^2}\int_0^s N dl \quad \text{meters.} \tag{10}$$

The effect of the ionosphere electrons is to reduce the phase-path length relative to free space directly as the electron content of the ray path and inversely as the square of the signal frequency.

Group-Path Delay

Since the phase velocity v in the ionosphere given by

$$v = \frac{c}{\mu} = c\left[1 - \frac{Ne^2}{m\epsilon_0\omega^2}\right]^{1/2} \tag{11}$$

is a function of signal frequency, the ionosphere is said to be dispersive. Because of this, a pulse of radio energy centered at a given angular frequency ω travels in the ionosphere with its group velocity different from its phase velocity at a particular angular frequency ω.

The group refractive index μ' is defined as[6]

$$\mu' = \frac{d(\mu\omega)}{d\omega} = \mu + \omega\frac{d\mu}{d\omega}. \tag{12}$$

Using this quantity, the equivalent group-path length P', will be

$$P' = \int_0^s \mu' dl. \tag{13}$$

For small values of X, as in Equation (7), the value of μ' becomes

[5] R. S. Lawrence, C. G. Little, and J. S. A. Chivers, "A Survey of Ionospheric Effects Upon Earth-Space Radio Propagation," *Proc. IEEE*, Vol. 52, p. 4, Jan. 1964.

[6] Reference (2), p. 92.

$$\mu' \cong 1 + \frac{X}{2} \cong \frac{1}{\mu}. \tag{14}$$

Denoting by $\Delta l'$ the incremental change in the group-path length P' due to the presence of the ionosphere, we have

$$\Delta l' = \int_0^s (\mu' - 1)\, dl \cong \int_0^s (1 - \mu)\, dl = -\Delta l, \tag{15}$$

where Δl is defined by Equation (10). It is noted that the effect of the ionosphere is to increase the group-path length of the signal, relative to free space.

Polarization Rotation

In the presence of a magnetic field, the ionosphere behaves as a bi-refringent medium and splits the incident radio wave into the ordinary and extraordinary components. These waves independently suffer no change in polarization as they propagate through the ionosphere; however, since they have different phase velocities, the plane of polarization of the resultant wave rotates.

If we let λ_+ and λ_- correspond to the wavelengths of the two waves and dl the increment of the path length, the rotation of the two components will be

$$\begin{aligned} d\Omega_+ &= \frac{2\pi dl}{\lambda_+} \\ d\Omega_- &= \frac{2\pi dl}{\lambda_-}. \end{aligned} \tag{16}$$

These phase rotations are shown in Figure 9, which indicates the relative magnitudes and direction of rotation. From Figure 9 it is seen that

$$d\Omega = \frac{d\Omega_+ - d\Omega_-}{2}. \tag{17}$$

Use of Equation (16) and the relation

$$\lambda = \frac{v}{f} = \frac{c}{\mu f} = \frac{2\pi c}{\omega\mu} \tag{18}$$

gives the relation

$$d\Omega = \frac{\omega}{2c}(\mu_+ - \mu_-)dl\,. \tag{19}$$

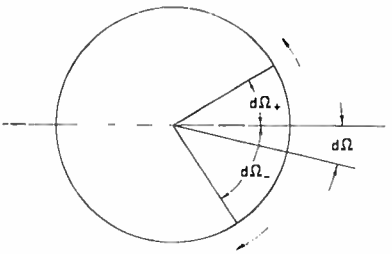


Fig. 9—Rotation of ordinary and extraordinary wave vectors.

If the approximation for the refractive index

$$\mu \cong 1 - \frac{X/2}{1 \pm Y_L} \tag{20}$$

is substituted for the two component waves, $\mu_+ - \mu_-$ can be approximated by XY_L in the high-frequency case; thus

$$d\Omega = \frac{\omega}{2c} XY_L dl = \frac{e^3\mu_o HN\cos\theta dl}{2cm^2\omega^2\epsilon_o}, \tag{21}$$

where the quantities defined in Equation (2) have been used. Substituting the quantities

$$\epsilon_o = 8.855 \times 10^{-12} \text{ farad/m},$$
$$\mu_o = 1.257 \times 10^{-6} \text{ henry/m},$$
$$m = 9.109 \times 10^{-31} \text{ kilogram},$$
$$e = 1.602 \times 10^{-19},$$
$$c = 3 \times 10^8 \text{ m/sec}$$

and integrating, results in the Faraday rotation

$$\Omega = \frac{0.0297}{f^2} \int_0^s NH \cos \theta dl \,. \tag{22}$$

This polarization rotation is inversely proportional to the square of the signal frequency and directly proportional to the longitudinal component of the magnetic field. Equations (10), (15), and (22) are standard formulas and are derived here only for completeness.

Table I—Order of Magnitude of Ionospheric Propagation Effects at 100 MHz

Phase-path length change	$\Delta l = -400$ meters,
Phase change	$\delta = -840$ radians,
Group-path delay	$\Delta T_o = 1.3 \times 10^{-6}$ sec,
Differential phase-path length	$\Delta l_+ - \Delta l_- = 6.4$ meters,
Polarization rotation	$\Omega = 6.6$ radians,

Application of Ionospheric Data to Radio Waves

From the equations derived above and numerical estimates of the electron content given previously, the effect of the ionosphere on wideband radio transmissions can be studied. Table I shows the magnitude of the ionospheric effects at a frequency of 100 MHz.[7] The values given in this table are based on the total ionospheric contribution for zenithal quasi-longitudinal propagation through the ionosphere where

$$\int_0^s N dl = 10^{17}/\text{meter-square column},$$

and $\omega_L = 5 \times 10^6$ radians/sec. For ionospheric conditions other than those used in Table I, proportional changes in

$$\int_0^s N dl \text{ and } \omega_L$$

[7] Reference (5), p. 8.

can be used to estimate the magnitude of the effects at 100 MHz. At frequencies other than 100 MHz, the frequency-dependent relations indicated in Equations (10), (15) and (22) are used.

The particular problem to be considered involves a radio signal of 500-MHz bandwidth centered at 3350 MHz. Computations are made of the phase departure from linearity in the frequency band for various elevation angles of the propagation path. Also, an estimate is made of the Faraday rotation of the plane of polarization.

Relative Phase Delay Due to Phase Path

Since the phase change, δ, corresponding to a phase-path length of Δl is given by

$$\delta = \omega t = \frac{2\pi}{\lambda} \Delta l , \tag{23}$$

the phase change varies inversely as the first power of the frequency. At 100 MHz the phase change is seen from Table I to be −840 radians. Direct extrapolation from 100 MHz to 3350 MHz leads to a value of −25.07 radians or −1437 degrees phase change compared to free space. The calculated relative phase change with respect to 3350 MHz has been plotted in Figure 10 for the frequency range from 3000 to 3700 MHz. The equation used for the calculation of the incremental phase delay $\Delta\phi_p$ is given by

$$\Delta\phi_p = \delta f_o \left(\frac{1}{f} - \frac{1}{f_c} \right), \tag{24}$$

with $\delta = -840$ radians, $f_o = 100$ MHz, and $f_c = 3350$ MHz.

Relative Group-Phase Delay

As seen from Equations (10) and (15), the group-path delay also varies inversely as the square of the operating frequency and directly as the total electron content of the signal path. Thus the delay at an angular frequency ω can be expressed as

$$\Delta T = \left(\frac{\omega_o}{\omega} \right)^2 \Delta T_o , \tag{25}$$

where ΔT_o is the group-path delay at angular frequency ω_o. If we

reference the group delay to the delay at the mid-band angular frequency ω_c, we can write the differential delay $\Delta\tau$ as

$$\Delta\tau = \Delta T - \Delta T_c = \left[\left(\frac{\omega_o}{\omega}\right)^2 - \left(\frac{\omega_o}{\omega_c}\right)^2\right]\Delta T_o . \tag{26}$$

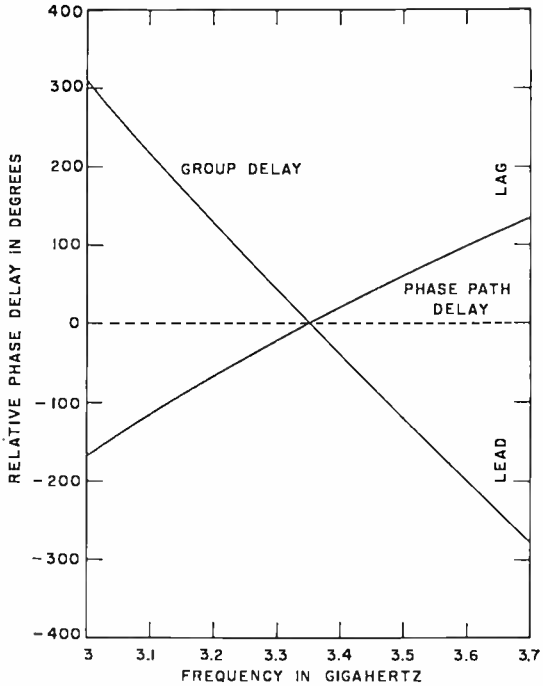


Fig. 10—Ionospheric phase delay per 10^{17} electrons in square-meter column.

The phase shift $\Delta\phi_g$ at the frequency ω associated with this delay will be $\omega\Delta\tau$ radians, or

$$\Delta\phi_g = \Delta T_o \omega_o^2 \left[\frac{1}{\omega} - \frac{\omega}{\omega_c^2}\right]. \tag{27}$$

From Table I, ΔT_o is seen to be 1.3×10^{-6} second at 100 MHz. Using these quantities in Equation (27), the relative group-phase delay has been calculated and plotted in Figure 10. It is noted that the phase shift is positive (more delay) for frequencies below the center fre-

quency and vice versa. This is just opposite to the relative phase-shift characteristic obtained for the phase-path change.

Relative Net Phase Delay

The two curves shown in Figure 10, for the relative phase and group delay as a function of frequency, not only have slopes of opposite sign but also exhibit nonlinearities that are opposite. Thus both the slope magnitude and the nonlinearity for the relative net phase delay are less than those exhibited by the curves shown in Figure 10 for the individual phase-path and group-delay phase shifts. By adding Equations (24) and (27), the relative net phase delay $\Delta\phi$ can be written as

$$\Delta\phi = \Delta\phi_p + \Delta\phi_g = \frac{\omega_o}{\omega}(\Delta T_o\omega_o + \delta) - \omega\frac{\Delta T_o\omega_o^2}{\omega_c^2} - \frac{\delta\omega_o}{\omega_c} \text{ radians,} \quad (28)$$

where $\Delta T = 1.3 \times 10^{-6}$ second at $f_o = 100$ MHz,

$\delta = -840$ radians at $f_o = 100$ MHz,

$f_o = 3350$ MHz for the case considered here.

For the above constants, the relative net phase shift in degrees becomes

$$\Delta\phi = \frac{c_o}{f} + c_1 f + c_2, \quad (29)$$

where f is expressed in GHz and

$c_o = -132.84550,$

$c_1 = -417.01938,$

$c_2 = 1436.6703.$

Equation (29) has been plotted in Figure 11 over a frequency range of plus and minus 350 MHz about the center frequency of 3350 MHz.

To determine the phase departure from linearity of the overall characteristic, a best-fit linear phase curve was subtracted from the relative net phase given by Equation (29). Thus the phase departure from linearity $\Delta\phi_l$ is found to be

$$\Delta\phi_l = (c_o/f) + c_1 f + c_2 - c_3(3.350 - f), \quad (30)$$

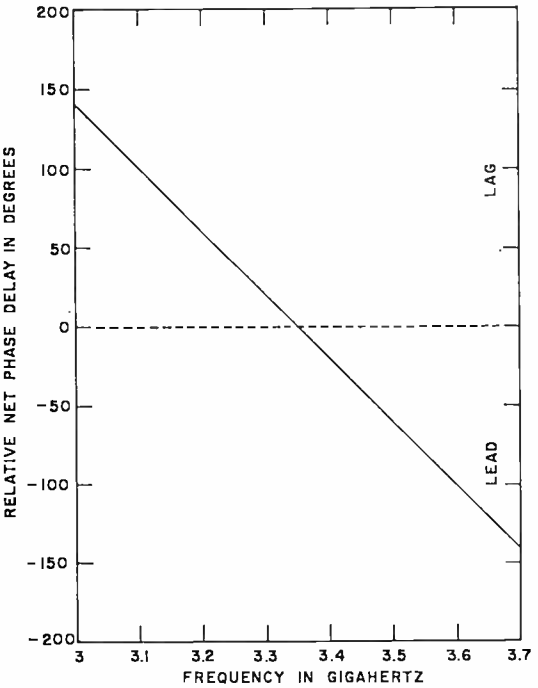


Fig. 11—Net relative ionospheric phase delay per 10^{17} electrons in square-meter column.

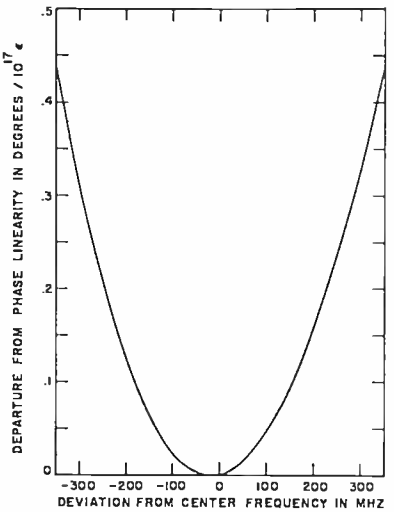


Fig. 12—Phase-shift nonlinearity in transmission path due to ionosphere.

where $c_3 = 405.05128$ and the other constants and the variable f are as defined for Equation (29). The phase-shift nonlinearity is plotted in Figure 12 for a symmetrical frequency departure from the mid-band frequency of 3350 MHz. This phase distortion is seen to be parabolic in shape and therefore corresponds to FM-type distortion. Since the above calculation was based on the phase distortion per 10^{17} electrons in a one-meter-square column, the total electron content in a given path must be first calculated before the total phase nonlinearity can be determined.

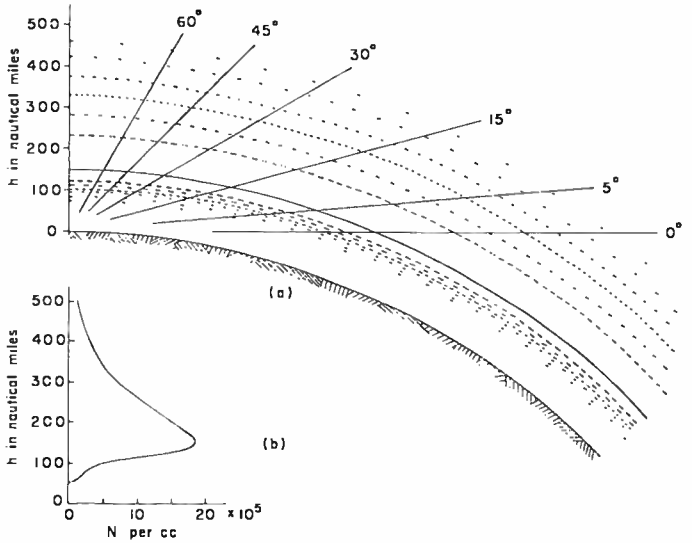


Fig. 13—Signal transmission paths.

Electron Content in Transmission Paths

Plots of shinf such as those shown in Figures 2-7 supply data for one elevation angle, 90°, or the zenith path. Figure 13(a) shows a sketch of the signal transmission path through the ionosphere for various elevation angles. The electron-density profile for the winter diurnal maximum (shown in detail in Figure 4 previously) is shown in Figure 13(b) to aid in illustrating the nature of the ray path. The distribution of electrons with height has been translated into coded contours indicated in the larger sketch. Each contour is 66% of the next higher contour, with the maximum density contour indicated by the solid-line contour at $h = 150$ nautical miles. The zenith path is seen to be the

minimum length, with the path length through the ionosphere increasing rapidly below elevation angles of 30°.

To determine the electron-density profiles along the various elevation angles shown, the distance along the ray path (R_1 in Figure 14) must be found for a given height h. From the law of cosines,

$$R_1 = [(R_o + h)^2 - R_o^2 \cos^2 \epsilon]^{1/2} - R_o \sin \epsilon . \tag{31}$$

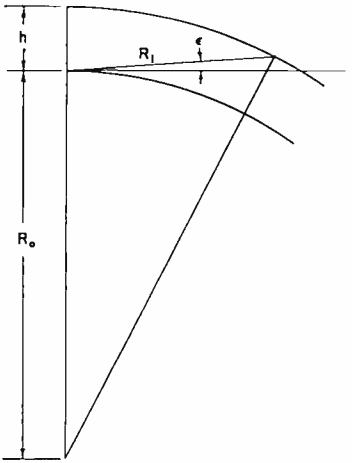


Fig. 14—Diagram showing relation between path length, height, and elevation angle.

With R_1 known, the electron-density profile along the ray path can be evaluated. This profile corresponds to the intersection of the ray path with the density contours shown in Figure 13(a). The contours for 30° and 60° elevation angles are shown in Figure 15.

The total electron content of transmission paths at a particular elevation angle was obtained by numerical integration of the electron-density profile for that angle. Figure 16 shows the resultant electron content per meter-square column for different elevation angles based on the average profile for January 1960 at 1200 hours (Figure 4). Assuming that other electron-density profiles can be represented by the January 1960 electron-density model, the curves of Figure 16 can be used for these cases by use of the proper multiplying factor.

For radio sources at or above 500 nautical miles above the earth, Table II gives the ratio, $\mathcal{R}$, of the total electron content at a given elevation angle to that at zenith. Using this factor $\mathcal{R}$ and the

values of shinf as indicated in Figure 2, the total electron content at a given elevation angle can be found for the diurnal and seasonal extremes. For example, from Figure 2, a value of 6.2×10^{17} electrons per meter-square column to infinite height was chosen to be repre-

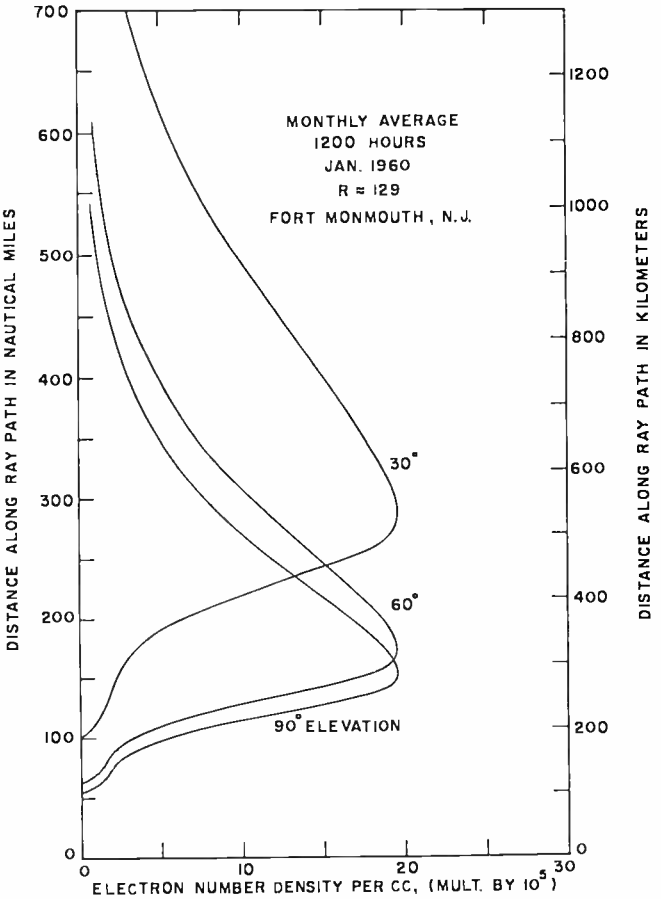


Fig. 15—Ionospheric electron-density profiles for different elevation angles.

sentative of the winter diurnal maximum, corresponding to a sunspot number R of about 100. Using Table II, it is seen that an elevation angle of 70°, $\mathcal{R} = 1.0604$ so that the total electron content at 70° was 1.0604 times that at zenith, thus the total content at 70° elevation is $6.2 \times 10^{17} \times 1.0604 = 6.57 \times 10^{17}$ per meter-square column. Thus at an elevation of 70° for the winter diurnal maximum condition, the factor multiplying the right-hand side of Equation (30) to give the

Table II—Ratio of Total Electron Content at Given Elevation Angle to That at Zenith, $\mathcal{R}$

Elevation Angle (Deg.)	$\mathcal{R}$
90	1.0000
80	1.0225
70	1.0604
60	1.1230
50	1.2584
40	1.4286
30	1.7282
20	2.1387
15	2.3991
10	2.7153
5	2.9605
0	3.0738

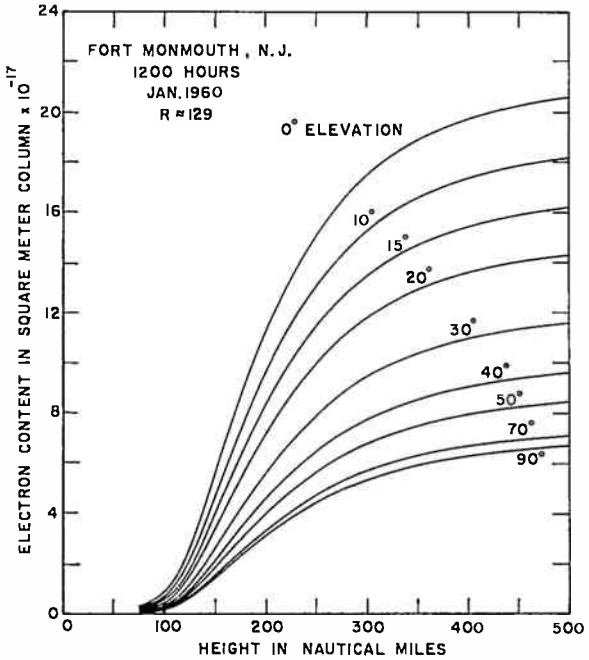


Fig. 16—Electron content in square-meter column for different elevation angles.

Table III—Values Used for Zenith Electron Content, E_z

	E_z (electrons/m^2 column)
Winter Maximum, 1200 hours	6.2×10^{17}
Winter Minimum, 0600 hours	0.5×10^{17}
Summer Maximum, 1900 hours	2.4×10^{17}
Summer Minimum, 0400 hours	0.8×10^{17}

total phase departure from linearity is 6.57. Similarly, the total phase nonlinearity can be determined for other conditions.

Phase Nonlinearity Due to Ionosphere

The phase distortion produced by the ionosphere is calculated using the values given in Table III for the total electron content at zenith. The factors multiplying the phase-nonlinearity curve given in Figure 12 are tabulated in Table IV. These factors are simply, $\mathcal{R}E_z$, where the values of $\mathcal{R}$ are listed in Table II. The values are for one-way propagation of the radio wave; for the radar case, these values would be multiplied by a factor of 2.

Table IV—Phase-Nonlinearity Multiplying Factors, K_N (one-way transmission)

Elevation Angle of Ray Path	Winter Maximum 1200 hrs.	Winter Minimum 0600 hrs.	Summer Maximum 1900 hrs.	Summer Minimum 0400 hrs.
90°	6.200	0.500	2.400	0.800
80°	6.340	0.511	2.453	0.818
70°	6.574	0.529	2.539	0.846
60°	6.963	0.561	2.695	0.898
50°	7.802	0.626	3.010	1.002
40°	8.857	0.707	3.408	1.131
30°	10.715	0.851	4.109	1.361
20°	13.260	1.037	5.040	1.658
15°	14.874	1.160	5.671	1.856
10°	16.835	1.280	6.302	2.048
5°	18.355	1.384	6.850	2.214
0°	19.058	1.421	7.073	2.274

Faraday Rotation

To calculate the Faraday rotation given by Equation (22), the magnetic-field factor $H \cos \theta$ must be evaluated; values for the electron content have been already obtained. The intensity of the earth's magnetic field H varies with the height above the earth, and also with the longitude and latitude. The angle θ depends on the direction of the earth's magnetic field and the particular direction of the radio wave, which can be distinguished at the earth receiver site as the azimuth and elevation angles.

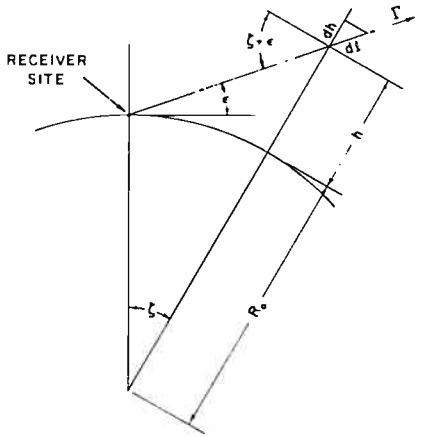


Fig. 17—Ray-path geometry.

For convenience, Equation (22) can be written in terms of a differential length of height above the earth instead of an element of ray-path length; thus

$$\Omega = \frac{0.0297}{f^2} \int_0^{h_t} H \cos \theta \csc (\zeta + \epsilon) N dh \,, \tag{32}$$

where dl in Equation (22) has been replaced by $dh \csc(\zeta + \epsilon)$, the relation being evident from Figure 17.

Since the magnetic-field factor is a stable parameter at a given receiver site, mean values can be assigned for specific signal source heights with corresponding variations for different values of azimuth and elevation angles. Therefore, Equation (32) can be expressed in the form

$$\Omega = \frac{0.0297}{f^2} \Psi(h_m,\Gamma) \int_0^{h_t} N dh \,, \tag{33}$$

where $\Psi(h_m,\Gamma)$ is the mean value of the magnetic-field factor $H \cos \theta \csc (\zeta + \epsilon)$ along the signal path Γ, and is evaluated at some intermediate value of height between 0 and h_t. A typical set of curves for Ψ is shown in Figure 18.[8] The data for these curves were taken at NRL (Naval Research Laboratories) Chesapeake Bay Annex.

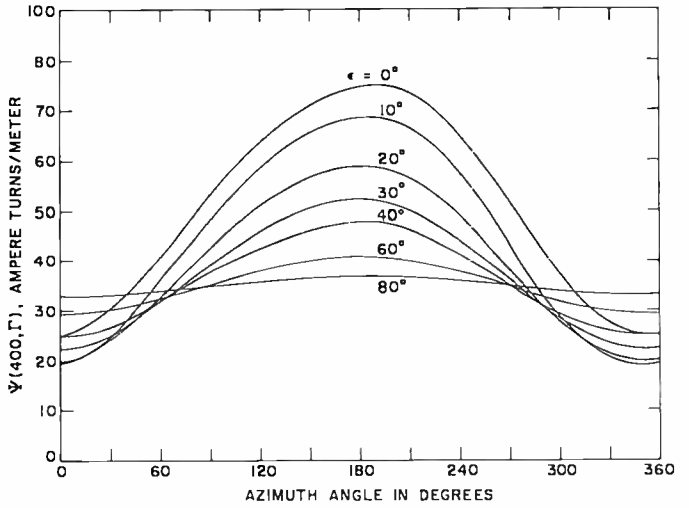


Fig. 18—The Ψ function for altitude of 400 kilometers.

Calculated Values of Faraday Rotation

The estimated values of the Faraday rotation for one-way signal transmission have been calculated using Equation (33) in conjunction with the curves of Figure 18 and the values of the integrated electron density given in Table III for various seasons and times. For the particular case being considered, $f = 3.350 \times 10^9$ Hz, so that

$$\Omega = 1.5163 \times 10^{-2} \Psi(400,\Gamma) E_z \,, \tag{34}$$

where E_z are the values in Table III. The mean height of the path

[8] J. M. Goodman, "Prediction of Faraday Rotation Angles at VHF and UHF," N.R.L. Report 6234, April 28. 1965, AD 615822.

through the ionosphere is normally assumed to be 400 kilometers for targets beyond the ionosphere. Curves showing the one-way Faraday rotation at the Chesapeake Bay site are given in Figures 19 and 20 for the winter and summer extremes.

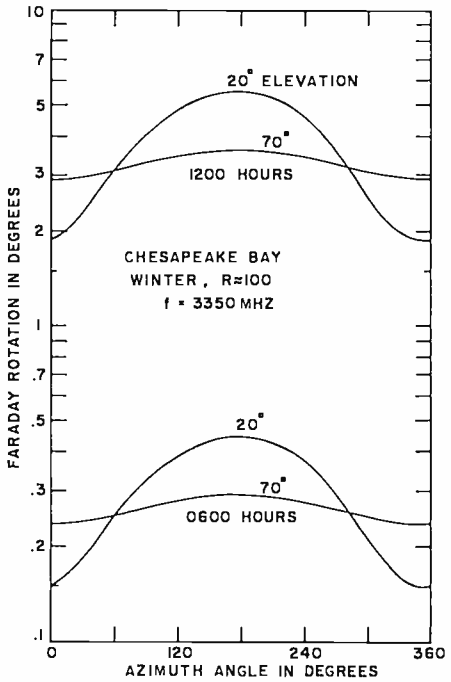


Fig. 19—Faraday rotation for one-way signal transmission through ionosphere.

In general terms, Faraday rotation is largest looking along the earth's magnetic field toward the south at low elevation angles. For the particular conditions chosen, the maximum rotation is about 5.5 degrees at winter diurnal maximum for elevation angles above 20 degrees. The rotation at the diurnal minimum in winter under the same conditions is less than 0.45 degree. In summer, the Faraday rotation for the extreme conditions lies between the winter extremes.

Conclusions

In this paper, the general procedure used to calculate the phase distortion caused by the ionosphere has been illustrated. Although the phase-path delay and the group-path delay vary inversely as the square

of the frequency, the corresponding phase shifts vary inversely as the frequency. The distortion produced by the phase-path and group-path delays are opposite in sign and the relative net phase distortion is predominantly parabolic in nature, which produces FM-type distortion if the signal used has a significant bandwidth. At a frequency of 3350

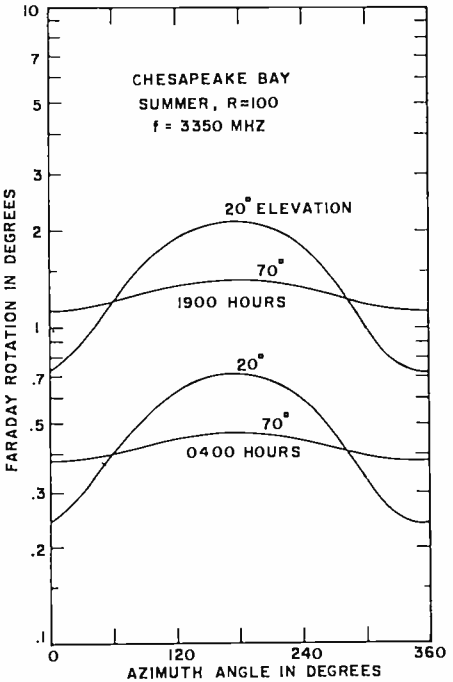


Fig. 20—Faraday rotation for one-way signal transmission through ionosphere.

MHz, the calculated Faraday rotation is seen to be relatively small and its effect can be minimized by proper choice of signal observation time.

Since the electron distribution in the ionosphere is statistical in nature, the characteristics of the ionosphere at a given time can only be estimated to be within a certain variance. Thus the quantities calculated here are only an estimate of the relative magnitude of ionospheric effects and cannot be used directly to compensate a signal-processing system for perturbations caused by the ionosphere. If the ionospheric measurements and data processing could be done rapidly enough, one might conceive of a signal-reception system that would be able to compensate for ionospheric effects in real time.

SPECTRAL ANALYSIS OF TURBULENT WAKES*

By

D. A. deWolf

RCA Laboratories
Princeton, New Jersey

Abstract

Summary—The power spectrum of backscattered radiation from an underdense turbulent re-entry wake illuminated by a train of short radar pulses can be significantly affected by pulse shape and by exponential decay of the mean electron density, as well as by the finite extent of each pulse (which limits the exposure time of any coherent wake target under observation). Expressions for the power spectrum and its broadening by these effects are derived and discussed.

Introduction

IT IS WELL KNOWN that the turbulent underdense wake of an object re-entering the atmosphere at velocities on the order of kilometers per second will reflect impinging radar signals appreciably. Any analysis of the reflected signals will attempt to distinguish at the onset between radar and wake effects. This is complicated by the fact that the dielectric properties of the wake vary irregularly in space and in time. An attempt to analyze these properties is best made by creating a situation in which the statistics are as close to being stationary as possible. A rapid train of short radar pulses serves this purpose, and the power spectrum of a sufficient number of reflected pulses is usually the starting point for wake analysis. Unfortunately the shape and strength of these pulses, as well as other features, all enter into the power spectrum and are entangled with those wake properties we wish to investigate. Most of these can be removed in standard fashion. The purpose of this work is to investigate certain distortions of the wake-power spectrum due to the limited extent and shape of the radar pulses, and also to investigate

* This work was supported by Government Contract No. DA-01-021-AMC-13095(Z).

the effect of exponential strength decay of the wake along its main axis upon the power spectrum.

A brief derivation of the parametric dependence of the backscattered pulse amplitude upon wake and transmitted-pulse properties is given in the first section. F. Lane[1] derived an expression for a monochromatic signal, carefully evaluating Doppler effects and discarding other unimportant relativistic effects. The present derivation incorporates the main, well-known result for weak scattering—Equation (5)—into an expression valid for short radar pulses; it enables us to work out the necessary statistics in the next section, carefully considering the required physical assumptions and their regions of validity. In Section 3 we show that pulses of limited extent broaden the turbulent velocity spectrum, particularly at low frequencies, and radically differently for square pulses than for Gaussian ones. Section 4 yields the broadening when there is exponential electron density decay; this also differs considerably according to whether the radar pulses are square or not.

1. The Electric Field Backscattered from a Range Cell

The re-entry situation is more complicated than it seems at first glance. A rapidly moving Re-entry Vehicle (RV) leaves behind it a wake of turbulent ionized hot gas that cools off, slows down, and dissipates its energy into the surrounding air. Since it is most likely that the statistics of this turbulent gas become stationary when they are interpreted as pertaining to the behavior in time of the gas properties at a fixed distance behind the RV, spectral analysis is performed upon the contents of a "range cell." At this stage a range cell may be thought of, simply, as a small section of wake with its center at a fixed distance behind the RV. The gas then streams through a range cell with the mean velocity $-\mathbf{U}(z)$ in the $\hat{\mathbf{z}}$-direction (carets denote unit vectors). Complications that will be avoided at first are that the radar is at a finite (although large), and thus changing, distance from the wake, and that the velocity of the gas with respect to the radar differs strongly from $\mathbf{U}(z)$. Here, it is assumed that radar pulses are incident in a direction $\hat{\mathbf{k}}$ making an angle θ with the axial wake direction $\hat{\mathbf{z}}$. The electric field $E_o(\mathbf{r},t)$ due to a radar pulse, the center of which has been emitted at $t = 0$, is assumed to be

$$E_o(\mathbf{r},t) = A_o(\hat{\mathbf{k}} \cdot \mathbf{r} - ct) \exp\{i(\mathbf{k}_o \cdot \mathbf{r} - \omega_o t)\}, \qquad (1)$$

[1] F. Lane, Paper given at AIAA 5th Aerospace Science Meeting, New York, New York, January 23-26, 1967.

where $k = \omega/c$ is an arbitrary wavenumber, $\mathbf{k} = k\hat{\mathbf{k}}$, the subscript 0 pertains to the carrier frequency, and A_0 is the (real) amplitude envelope of the pulse. In order to satisfy Maxwell's equations,

$$A_0(\hat{\mathbf{k}} \cdot \mathbf{r} - ct) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \tilde{A}_0(\omega) \exp\{i(\mathbf{k} \cdot \mathbf{r} - \omega t)\}, \tag{2}$$

where $\tilde{A}_0(\omega)$ is the weighting factor of the elementary plane waves determining the bandwidth of the pulses. By the same token $\tilde{A}_0(\omega) \exp\{i\mathbf{k} \cdot \mathbf{r}\}$ is the inverse Fourier transform of A_0 considered as a function of t.

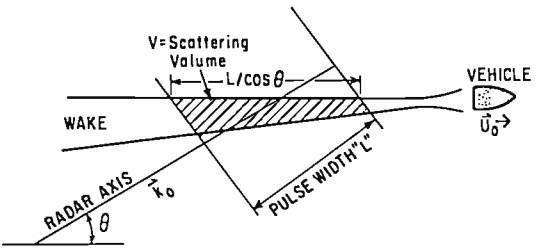


Fig.1—Backscatter situation.

The procedure we intend to follow for solving the backscatter problem is simple. We Fourier analyze $E_0(\mathbf{r},t)$ into its frequency components $\tilde{E}_0(\mathbf{r},\omega)$, and solve the harmonic Fourier-component Maxwell's equations for the total field $\tilde{E}(\mathbf{r},\omega)$ such that removal of the wake would lead to solutions $\tilde{E}(\mathbf{r},\omega) = \tilde{E}_0(\mathbf{r},\omega)$. Finally, we synthesize the resulting $\tilde{E}(\mathbf{r},\omega)$ into the scattered pulse $E(\mathbf{r},t)$.

$\tilde{E}_0(\mathbf{r},\omega)$ is obtained by applying Equation (2) to the Fourier transform of the right-hand side of Equation (1):

$$\tilde{E}_0(\mathbf{r},\omega) = \tilde{A}_0(\omega - \omega_0) \exp\{i\mathbf{k} \cdot \mathbf{r}\}. \tag{3}$$

Upon eliminating the magnetic field, the Fourier-transformed Maxwell's equations yield a vector wave equation

$$\begin{aligned} &\nabla \times \nabla \times \tilde{E}(\mathbf{r},\omega) - k^2 \epsilon(\mathbf{r},\omega) \tilde{E}(\mathbf{r},\omega) = 0 \\ &\lim_{\epsilon \to 1} \tilde{E}(\mathbf{r},\omega) = \tilde{E}_0(\mathbf{r},\omega), \end{aligned} \tag{4}$$

where $\epsilon(\mathbf{r},\omega)$ is the relative dielectric permittivity of the gas reacting to a harmonic disturbance proportional to $\exp(-i\omega t)$. For weak scatterers, $\epsilon(\mathbf{r},\omega)$ is close to unity and the Born approximation may be utilized under a number of further conditions discussed by Salpeter and Treiman[2]. For backscatter, i.e., for the asymptotic scattered field $\tilde{E}_b(-\bar{\mathbf{k}},\omega)$ at $\mathbf{r} = \lim_{r\to\infty} -r\bar{\mathbf{k}}$, we find $\tilde{E}_b(-\bar{\mathbf{k}},\omega)$ is polarized in the same direction as $\tilde{E}_o(\mathbf{r},\omega)$ (therefore no vector notation is required in this approximation) and is the following integral of the gas quantity $q(\mathbf{r},\omega) \equiv k^2[1-\epsilon(\mathbf{r},\omega)]$:

$$\tilde{E}_b(-\bar{\mathbf{k}},\omega) = \lim_{r\to\infty} \frac{e^{ikr}}{r} \frac{1}{4\pi} \int d^3r_1 q(\mathbf{r}_1,\omega) \exp\{i\mathbf{k}\cdot\mathbf{r}_1\}\tilde{E}_o(\mathbf{r}_1,\omega). \quad (5)$$

We now form the inverse Fourier transform of the right-hand side of Equation (5) at finite r. This field we name $E_b(\mathbf{r},t)$ and its limit as $r\to\infty$ is the desired solution. We shall retain r finite, but at a location very distant from the wake compared to the dimensions of the effective scattering part of the wake, and we assume that the local quantity $rE_b(r,t)\exp\{-ik_o r\}$ does not differ appreciable from its asymptotic value of $r\to\infty$. Inserting Equation (3) into Equation (5) and performing the inverse Fourier transform, we obtain

$$E_b(r,t) = \frac{1}{4\pi r} \exp\{i(k_o r - \omega_o t)\} \times$$

$$\int d^3r_1 \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega q(\mathbf{r}_1,\omega) \bar{A}_o(\omega-\omega_o) \exp\{-i(\omega-\omega_o)(t - 2\bar{\mathbf{k}}\cdot\mathbf{r}_1\}. \quad (6)$$

Upon assuming that the variation of $q(\mathbf{r}_1,\omega)$ within the effective bandwidth of $\bar{A}_o(\omega-\omega_o)$ is negligible (this is certainly true for an electron plasma where $q(\mathbf{r})$ is a constant times the electron density $n_e(\mathbf{r})$, one can perform the $d\omega$ integration, utilizing Equation (2), to obtain

$$E_b(r,t) = \frac{\exp\{-i\omega_o t'\}}{4\pi r} \int d^3r_1 q(\mathbf{r}_1) \exp\{2i\mathbf{k}_o\cdot\mathbf{r}_1\} A_o(2k\cdot\mathbf{r}_1 - ct'),$$

$$t' \equiv t - \frac{r}{c}, \quad (7)$$

[2] F. F. Salpeter, and S. B. Treiman, "Backscatter of Electromagnetic Radiation from a Turbulent Plasma," *Jour. Geophys. Res.* Vol. 69, pp. 869-881, 1964.

where t' is the time interval lapsed after the emitted pulse has reached the location corresponding to a distance r from its source (specifically the location of the RV). This field is now normalized to the field incident upon the RV by taking large but finite r and multiplying by $(4\pi)^{1/2}r/A_o$, so that the absolute square of the thus obtained $E_b(t)$ corresponds to a radar cross section (RCS);

$$E_b(t) = (4\pi)^{-1/2} \exp\{-i\omega_o t\} \int d^3r_1 q(\mathbf{r}_1) \exp\{2i\mathbf{k}_o \cdot \mathbf{r}_1\} \frac{A_o(2\hat{\mathbf{k}} \cdot \mathbf{r}_1 - ct)}{A_o(0)}. \tag{8}$$

Since $A_o(0)$ corresponds to the center of the pulse, Equation (8) expresses the fact that at given time $t = 2\hat{\mathbf{k}} \cdot \mathbf{r}_o/c$ the radar return is strongest from points at location $\mathbf{r}_o$ in the wake and other locations contribute with reduced strength $A_o(2\hat{\mathbf{k}} \cdot \mathbf{R}_1)$, where $\mathbf{R}_1 = \mathbf{r}_1 - \mathbf{r}_o$ is the relative distance from $\mathbf{r}_o$. Thus returns are from an effectively finite region of the wake around $\mathbf{r}_o$ determined by the pulse shape; this region is the so-called range cell.

2. Autocorrelation and Power Spectrum

Equation (8) is useful only for media that can be described by a $q(\mathbf{r}_1)$ suitably connected to the physical properties. For a turbulent plasma there are basic difficulties; $q(\mathbf{r}_1)$ is proportional to the electron density, which is a function of time. Therefore the basic description of the plasma cannot be used in our analysis since $q(\mathbf{r}_1)$ must be independent of time. Strictly speaking, the proper $q(\mathbf{r}_1,\omega)$ of the plasma should be computed, but since that would invoke knowledge of the behavior in time at all locations—which we do not have—such a computation is not feasible. Fortunately, the fact that we work with a rapid succession of short pulses each of which traverses the wake in a time quite short compared to relaxation times of the density enables us to make a very good approximation by utilizing Equation (8) with the classical expressions for $q(\mathbf{r}_1)$ as if the wake is stationary during the very short pulse-transition time $\delta t = 2L/c$, where L is the effective length of a range cell. For microsecond pulses, $\delta t \sim 10^{-6}$ sec, and for typical fluid (we shall use the term "fluid" to describe bulk behavior of the gaseous wake matter) velocities less than 10^3 m/sec, the displacements in δt are quite negligible compared to a UHF wavelength. We consider Equation (8) for the m-th of N pulses emitted by the radar (time between emission of two successive pulses is ΔT). For the first pulse time t is defined as $t = 2\hat{\mathbf{k}} \cdot \mathbf{r}_o$ as before, corresponding to a fixed

distance r_o such that $t = 2\hat{\mathbf{k}} \cdot \mathbf{r}_o$. For the m-th pulse we have time $t + m\Delta T$ instead of t. The RV moves with a velocity $\mathbf{U}_o$ with respect to the radar as a result of which the vector $\mathbf{r}_1$ of the first pulse becomes $\mathbf{r}_1 + \mathbf{U}_o m\Delta T$ in the m-th pulse in order to retain the origin at the radar. With these simple adjustments the m-th pulse E_m can be written[3]
$E_m = (4\pi)^{-1/2} \exp\{-i\omega_o t\} \exp\{-i(\omega_o - 2\mathbf{k}_o \cdot \mathbf{U}_o) m\Delta T\} \times$

$$\int d^3 r_1 q_m(\mathbf{r}_1 \exp\{2i\mathbf{k}_o \cdot \mathbf{r}_1\} \frac{A_o(2\hat{\mathbf{k}} \cdot \mathbf{R}_1)}{A_o(0)}. \tag{9}$$

The approximation has been made that $\mathbf{k}_o$ does not change direction in N consecutive pulses (each of which encounters a specific configuration $q_m(\mathbf{r}_1)$ in the range cell). The vector $\mathbf{R}_1 = \mathbf{r}_1 - \mathbf{r}_o$ indicates location with respect to the "center" of the range cell.

Now the autocorrelation of the field with respect to time is formed as a (discrete) function of $\Delta t = m\Delta T$: $\mathcal{R}_m \equiv N^{-1} \sum_{n=1}^{N} E_{n+m} E_n^*$, thus,

$$\mathcal{R}_m = \frac{\exp[-i(\omega_o - 2\mathbf{k}_o \cdot \mathbf{U}_o) m\Delta T]}{4\pi A_o^2(0)} \times$$

$$\int d^3 R_1 \int d^3 R_2 \frac{1}{N} \sum_{n=1}^{N} q_{n+m}(\mathbf{R}_1) q_n^*(\mathbf{R}_2)$$

$$\exp\{2i\mathbf{k}_o \cdot (\mathbf{R}_1 - \mathbf{R}_2)\} A_o(2\hat{\mathbf{k}} \cdot \mathbf{R}_1) A_o(2\hat{\mathbf{k}} \cdot \mathbf{R}_2) \tag{10}$$

It has been assumed that $\mathcal{R}_m$ does not depend on any particular sequence of N pulses, nor the magnitude of N (provided N is not small), i.e., that the statistics in forming the mean sum over N pulses are stationary. In particular we write $q_n(\mathbf{R}) = \bar{q}(\mathbf{R})[1 + \Delta_n(\mathbf{R})]$ where $\Delta_n = \bar{q}_n/\bar{q}$ is a stochastic discrete function of n for a given coordinate $\mathbf{R}$. A source of difficulty in continuing is that although it is assumed that $\bar{q}(\mathbf{R})$ and $<|\Delta(\mathbf{R})|^2>$ are both stationary in the RV rest frame, only the latter quantity is independent of location $\mathbf{R}$ in all cases we consider. In order to make assumptions about the autocorrelation

$$N^{-1} \sum_{n=1}^{N} \Delta_{n+m}(\mathbf{R}_1) \Delta_n^*(\mathbf{R}_2) \equiv <\Delta_{n+m}(\mathbf{R}_1) \Delta_n^*(\mathbf{R}_2)>, \text{ we}$$

[3] C. T. Tai, "Present Views on Electrodynamics of Moving Media," *Radio Sci.* Vol. 2 (New Series), pp. 245-248, 1967.

must account for the fact that the origin of coordinate $\mathbf{R}_1$ has moved by $-m\mathbf{U}\Delta T$ compared to the origin of $\mathbf{R}_2$. The actual spatial separation between these two points is $\mathbf{R}_1 + m\mathbf{U}\Delta T - \mathbf{R}_2$. Furthermore, the point $\mathbf{R}_1 + m\mathbf{U}\Delta T$ is being observed a time $m\Delta T$ later than $\mathbf{R}_2$. Consequently, the central assumption for homogeneous turbulence takes the following form:

$$<\Delta_{n+m}(\mathbf{R}_1)\Delta_n^*(\mathbf{R}_2)> = <|\Delta|^2> C(\mathbf{R}_1 - \mathbf{R}_2 + m\mathbf{U}\Delta T; m\Delta T), \quad (11)$$

where $C(\Delta\mathbf{R};\Delta t)$ is the four-dimensional space–time autocorrelation of the dielectric quantity $\Delta(\mathbf{R},t)$. Under the assumption that Equation (11) is a slowly varying discrete function of m, we rewrite $m\Delta T = \Delta t$ and assume that $C(\mathbf{R}_1 - \mathbf{R}_2 + m\mathbf{U}\Delta T; m\Delta T)$ is a good discrete sampling of the continuous function $C(\mathbf{R}_1 - \mathbf{R}_2 + \mathbf{U}\Delta t; \Delta t)$ so that Equation (10) for the fluctuating part* becomes

$$\mathcal{R}(\Delta t) = \frac{<|\Delta|^2> \exp\{-i(\omega_o - 2\mathbf{k}_o \cdot \mathbf{U}_o)\Delta t\}}{4\pi A_o^2(0)} \times \int d^3R_1 \int d^3R_2 \bar{q}(\mathbf{R}_1)\bar{q}^*(\mathbf{R}_2) C(\mathbf{R}_1 - \mathbf{R}_2 + \mathbf{U}\Delta t;\Delta t) \exp\{2i\mathbf{k}_o \cdot (\mathbf{R}_1 - \mathbf{R}_2)\} A_o(2\hat{\mathbf{k}} \cdot \mathbf{R}_1) A_o(2\hat{\mathbf{k}} \cdot \mathbf{R}_2) \quad (12)$$

Let us define new coordinates $\mathbf{R} = \mathbf{R}_2$, $\Delta\mathbf{R} = \mathbf{R}_1 - \mathbf{R}_2 + \mathbf{U}\Delta t$. Autocorrelation C depends only upon $\Delta\mathbf{R}$, and the effective integration volume over element $d^3\Delta R$ is only several correlation lengths l wide in any dimension. Since we assume that the range cell under consideration is centered sufficiently far away from the RV in order that $\mathbf{R}_1 + \mathbf{U}\Delta t$ will lie in the wake, even for $\Delta t \sim m\Delta T$, we may ignore boundary effects in the $\Delta\mathbf{R}$ integration and perform it as if it were unrestricted. Since l is assumed to be so small that neither $\bar{q}$ nor A_o change appreciably over a length $\sim l$, we ignore the $\Delta\mathbf{R}$ dependence of these quantities. Thus:

$$\mathcal{R}(\Delta t) = \frac{<|\Delta|^2>}{4\pi A_o^2(0)} \exp\{-i(\omega_o - 2\mathbf{k}_o \cdot \mathbf{v})\Delta t\} \int d^3\Delta r C(\Delta\mathbf{R};\Delta t) \exp\{2i\mathbf{k}_o \cdot \Delta\mathbf{R}\} \times \int d^3R \bar{q}(\mathbf{R} - \mathbf{U}\Delta t)\bar{q}^*(\mathbf{R}) A_o[2\hat{\mathbf{k}} \cdot (\mathbf{R} - \mathbf{U}\Delta t)] A_o(2\hat{\mathbf{k}} \cdot \mathbf{R}) \quad (13)$$

* The term due to $<\bar{q}(\mathbf{R}_1)\bar{q}^*(\mathbf{R}_2)>$ is diffraction-like, and since it contributes only to the dc value, it is of no interest in spectral analysis.

where $\mathbf{v} = \mathbf{U}_o - \mathbf{U}$ is the actual average fluid velocity with respect to the radar ($v \ll U_o$), and where the $d^3\Delta r$ integral is just the spatial spectrum $\Phi_3(2\mathbf{k}_o; \Delta t)$, which is a correlation function with respect to time and a Fourier spectral decomposition of the spatial correlation. Thus:

$$\mathcal{R}(\Delta t) = \frac{<|\Delta|^2>}{4\pi} \exp\{-i(\omega_o - 2\mathbf{k}_o \cdot \mathbf{v})\Delta t\}\Phi_3(2\mathbf{k}_o;\ \Delta t)\, g(\Delta t)$$

$$g(\Delta t) = \frac{1}{A_o^2(0)} \int d^3R \bar{q}(\mathbf{R} - \mathbf{U}\Delta t)\, \bar{q}^*(\mathbf{R}) A_o[2\hat{\mathbf{k}} \cdot (\mathbf{R} - \mathbf{U}\Delta t)] A_o(2\hat{\mathbf{k}} \cdot \mathbf{R}). \tag{14}$$

Upon setting $\Delta t = 0$ and choosing q constant, the above expression reduces to the well-known weak-scattering expression for the radar cross section $\mathcal{R}(0)$ of a homogeneous turbulent gas. Upon multiplying $\mathcal{R}(\Delta t)$ by $\exp(i\omega\Delta t)$ and integrating over all Δt we obtain the power spectrum $S(\omega)$, formally of the form

$$S(\omega) = \frac{<|\Delta|^2>}{4\pi} \int_{-\infty}^{\infty} d\omega_1 G(\omega_1)\Phi_4(2\mathbf{k}_o; \tilde{\omega} - \omega_1),$$

$$\tilde{\omega} \equiv \omega - \omega_o + 2\mathbf{k}_o \cdot \mathbf{v}, \tag{15}$$

where $G(\omega)$ and $\Phi_4(2\mathbf{k}_o; \omega)$ are the Fourier transforms with respect to Δt of $g(\Delta t)$ and $\Phi_3(2\mathbf{k}_o; \Delta t)$.

The power spectrum $S(\omega)$ is determined by the convolution of $\Phi_4(2\mathbf{k}_o; \omega)$ with $G(\omega)$, Doppler-shifted by $\omega_o - 2\mathbf{k}_o \cdot \mathbf{v}$. For a cylindrical wake with radius r_ω we can utilize $\hat{\mathbf{k}} \cdot \mathbf{U} = U\cos\theta$, and then simplify Equation (15) to read

$$S(\omega) = \frac{<|\Delta|^2>}{4\pi} \frac{\pi r_{\omega}^2}{U} \int_{-\infty}^{\infty} d\omega_1 G(\omega_1/U)\Phi_4(2\mathbf{k}_o; \tilde{\omega} - \omega_1), \tag{16}$$

where $G(\omega/U)$ is the power spectrum with respect to wave number ω/U of the product $\bar{q}(z)A_o(2z\cos\theta)/A_o$ of mean q and pulse shape. It is the convolving factor $G(\omega/U)$ that distorts the actual turbulence spectrum $\Phi_4(2\mathbf{k}_o; \omega)$ in the measurement process.

Favre,[4] among others, has argued in favor of a simple Gaussian form of Φ_4. This form, sufficient for our purposes, is

$$\Phi_4(2\mathbf{k}_0;\omega) = \Phi_3(2\mathbf{k}_0;0)\,\frac{\exp\left[-\omega^2/8k_0^2<(\Delta v)^2>\right]}{\left[2k_0^2<(\Delta v)^2>/\pi\right]^{1/2}}, \qquad (17)$$

and this spectral shape is characterized by one parameter: the sigma width $\sigma_\phi^2 = 4k_0^2 <(\Delta v)^2>$. It is the purpose of this work to investigate how this parameter can be retrieved from the broadened spectrum $S(\omega)$. A measure of the broadening is the variance width σ_s^2 defined as

$$\sigma_s^2 = \frac{M_2(S)}{M_0(S)} - \left[\frac{M_1(S)}{M_0(S)}\right]^2 = -\left[\frac{\partial^2}{\partial t^2}\ln \mathcal{R}(t)\right]_{t=0}, \qquad (18)$$

where M_0, M_1, M_2 are the integrals of $S(\omega)$, $\omega S(\omega)$, $\omega^2 S(\omega)$, respectively, from $\omega = -\infty$ to $\omega = +\infty$. This width is the same as the sigma width for a Gaussian. Other characterizations of the spectral width are possible, of course, but of no specific advantage above this one for which an addition-of-width theorem, $\sigma_s^2 = \sigma_\phi^2 + \sigma_G^2$, holds whenever each of these widths can be defined.

3. Finite-Exposure-Time Broadening

When $\bar{q}$ does not vary, the spectral broadening is due to the fact that elements of the dielectric fluid move through the range cell of effective length L at velocity U and are seen only for a time L/U. The expected broadening is therefore of the order of U/L. In order to consider this in detail we should examine the Fourier transform $G(\omega)$ of $g(\Delta t)$ as given in Equation (14). If we assume that U is axially directed we see that $g(\Delta t)$ is really an even function of $U\Delta t$ and thus can consider $K = \omega/u$ and $z = U\Delta t$ as a pair of conjugate Fourier-transform variables since $ikz = i\omega\Delta t$. Therefore we rename $U\Delta t = z$ and consider $G(K)$ the Fourier transform of $g(z)$, an even function of z:

$$G(K) = \left|\bar{q}\int_{-z_0}^{\infty} dz\,\frac{A_0(2z\cos\theta)}{A_0(0)}\,e^{iKz}\right|^2, \qquad (19)$$

[4] A. J. Favre, "Review on Space-Time Correlations in Turbulent Fluids," *Jour. Appl. Mech.*, pp. 241-257, June 1965.

where $z = -z_o$ is the leading edge of the wake ($\vec{q} = 0$ at $z < z_o$). For far-wake range cells ($z_o \gg L$) we can replace $-z_o$ by $-\infty$ without serious error in Equation (19) and thus we require the Fourier transform of $A_o(2z)$; $\theta \neq 0$ results are easily reconstructed from $\theta = 0$ results and therefore not considered explicitly. In determining the variance width it should be stressed that the range of K-values that are meaningful in G is finite. In practice $g(z)$ is obtained for a finite stretch of time T (defining a finite extent UT of variable z) and consequently $G(K)$ is undistorted for $K < 2\pi(UT)^{-1}$. If the actual $G(K)$ has higher wave-number components, these will be aliased down and contaminate the lower values. We will assume that is not the case. For a set of N discrete values E_m the actual processing involves a finite discrete analog of Fourier transformation; $G(K)$ consists of N values $K_m = \omega_m/U$ with $\omega_m = 2\pi m/N\Delta T$, and is periodic modulo $\Omega = 2\pi/\Delta T$. In practice, we have unambiguous values of $G(K)$ only for $|K| \leqslant \Omega/U$ and the moments $M_o(G)$, $M_1(G)$ and $M_2(G)$ are really obtained by integration of K between finite bounds $\pm\,\Omega/U$. When G falls off rapidly to zero within $|K| \leqslant \Omega/U$ no appreciable error at values of $|K|$ close to Ω/U due to an aliasing function multiplying G, will arise.

We now apply Equation (18) to Equation (19) with the above provisions for replacing $-z_o$ by $-\infty$ and for restricting K to $|K| \leqslant \Omega/U$. From the form of Equation (19) it then follows that $M_1(G) = 0$. Since $A_o(\pm\infty) = 0$, we find

$$M_o(G) = \int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} dz_2 A_o(2z_1)\, A_o(2z_2)\, \frac{2\sin\left[(z_1 - z_2)\,\Omega/U\right]}{(z_1 - z_2)}\,,$$

$$M_2(G) = U^2 \int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} dz_2 A_o'(2z_1)\, A_o'(2z_2)\, \frac{2\sin\left[(z_1 - z_2)\,\Omega/U\right]}{(z_1 - z_2)}\,, \tag{20}$$

where $A_o'(2z) = \partial A_o(2z)/\partial z$. An important point to be made for further simplification is that Ω/U can be quite large in practice. Typically, radars can emit microsecond or shorter pulses at rates above 500 pulses per second, and U will be of the order of 5 km/sec, so that $\Omega/U > 1$ m^{-1}. If $A_o(2z_1)$ and $A_o'(2z_1)$ vary slowly for Δz_1 several times U/Ω (i.e., over several meters) no great error will be made in

replacing $2(z_1 - z_2)^{-1} \sin [z_1 - z_2)\Omega/U]$ by $\pi\delta(z_1 - z_2)$. This will be true in general at the center of a pulse, i.e., at $z_1 = 0$, but as the pulse edges are approached, the variation of $A_o(2z)$, and perhaps of $A_o'(2z)$ too, may increase to the extent that this is no longer true. There are two simple cases for which further progress is possible, and they illustrate the extremes very well.

(i) Gaussian Pulses

One type of often occurring radar pulse is essentially Gaussian $A_o(z) = A_o \exp(-z^2/2L^2)$ with the signal level vanishing under the noise level at $L_o > L$. For a one-microsecond pulse, $L \sim 150$ meters and L_o can be several standard deviations (L) long. The relative change in $A_o(2z)$ is $|\partial \ln A_o(2z)/\partial z| = 4z/L^2$ per unit length, so we wish to compare L_o/L to $\Omega L/U$. For the above discussed typical radar parameters, $\Omega L/U > 100$ and $L_o/L \ll 100$; consequently $M_o(G)$ can be simplified as above. Similarly $|\partial \ln A'(2z)/\partial z| \ll \Omega/U$ except for a small region around $z = 0$ where $A_o'(2z)$ is close to zero. Consequently,

$$\sigma_G^2 \approx \frac{U^2 \int dz[A_o'(2z)]^2}{\int dz[A_o(2z)]^2} \approx 2U^2L^{-2}, \tag{21}$$

where the integration is from $-L_o/4$ to $+L_o/4$ (not essentially different from the $-\infty$ to $+\infty$ integration). Other, related, pulse shapes with the same variance width L of $A_o(z)$ will yield larger values of σ_G, according to the Fourier-transform uncertainty principle (Papoulis[5]).

(ii) Square Pulses

Very short pulses with good signal-to-noise ratio can have steep rise and fall edges. Following $A_o(2z)$ for growing $|z|$ we note that initially $A_o(2z)$ and $A_o'(2z)$ vary slowly, but at a certain point $|z| = L/4$ (assuming symmetric pulses) there is an abrupt increase in $A_o'(2z)$ beyond which $A_o(2z)$ decreases most rapidly into the noise-level. We will represent such pulses by idealized square pulses, and the criterion governing real pulses is that there is a point $|z| = L/4$ at which the transition $|\partial \ln A_o(z)/\partial z| \ll \Omega/U$ to $|\partial \ln A_o(z)/\partial z| \gg \Omega/U$ occurs within several times U/Ω. This latter distance being in

[5] A. Papoulis, *The Fourier Integral and its Applications*, McGraw-Hill, Inc., New York, New York, 1962.

the order of 1 meter, we see that pulses as short as 0.2 microsecond ($L/2 \sim 30$ m) can fall in this category. In this case we set $A_o(2z) = A_o[H(z + L/4) - H(z - L/4)]$ where $H(z)$ is the Heaviside unit step function. If $\Omega L/U \gg 1$, we then find that Equation (20) reduces to

$$\sigma_G^2 \approx \frac{8U\Omega}{\pi L}\left[1 - \frac{\sin(\Omega L/2U)}{(\Omega L/2U)}\right] \approx \frac{8U\Omega}{\pi L}. \tag{22}$$

We note that the square pulse σ_G^2 is larger by a factor $\sim \Omega L/U$ than the Gaussian σ_G^2 and that it depends furthermore on the pulse repetition rate Ω. The occurrence of steep rise and fall edges induces *additional* spectral broadening above the well-known exposure-width broadening, Equation (21). In both cases, however, we note that σ_G is frequency independent, whereas $\sigma_\phi \propto k$. Thus this broadening will be more appreciable at low than at high radar frequencies. Of course, actual radar pulses may belong to neither extreme category, in which case more complicated broadening—presumably with values of σ_G^2 intermediate between those of Equations (21) and (22)—of the power spectrum will occur.

4. Modifications Due to Exponential Decay of q

We have seen that q is proportional to the average electron density n in an ionized wake. Electrons are created close to the vehicle and the ensuing ionized gas streams back and loses its free electrons and ions through various processes of recombination. Under a number of common re-entry conditions the ensuing electron-density decay in time can be described as exponential; i.e., in the fluid rest frame;

$$\bar{n}(t) = \bar{n}(0)\, e^{-\beta t}. \tag{23}$$

Because fluid elements are moving in the body rest frame, we rewrite this as

$$\bar{n}(Z + Ut, t) = \bar{n}(Z, t)\, e^{-\beta t} = \bar{n}(Z, t)\, e^{-(\beta/U)Ut}. \tag{24}$$

If the statistics of n are stationary in the body frame, then $\bar{n}(Z, t) = \bar{n}(Z)$, and the temporal decay in Equation (23) dictated by β is experienced as a spatial decay governed by $d = U/\beta$. Consequently, we will now set

$$\bar{q}(Ut) = \bar{q}(0) \exp(-Ut/d),\ d \equiv U/\beta \tag{25}$$

to obtain

$$G(K) = \left| \bar{q}(0) \int_{-z_0}^{\infty} dz \frac{A_0(2z \cos \theta)}{A_0(0)} \exp(-z/d + iKz) \right|^2 \quad (26)$$

as the appropriate generalization of Equation (19) to the case of exponential density decay.

Instead of Equation (20) we now find expressions for M_0 and M_2 that differ solely by the replacements $A_0(2z) \exp(-z/d)$ for $A_0(2z)$ in M_0 and $[A_0'(2z) - A_0(2z)/d] \exp(-z/d)$ for $A_0'(2z)$ in M_2. Considering Gaussian and square pulses again we obtain the following results for each category under similar assumptions as before:

(i) Gaussian Pulses, Exponential Decay

$$\sigma_G^2 = U^2 \frac{\int dz[A_0'(2z) - A_0(2z)/d]^2 \exp(-2z/d)}{\int dz[A_0(2z)]^2 \exp(-2z/d)} \approx 2U^2L^{-2} \quad (27)$$

The integrations are from $z = -L_0/4$ to $z = +L_0/4$, but for range cells deep enough in the wake, the error made in taking them from $z = -\infty$ to $z = +\infty$ is negligible and the value $\sigma_G^2 = 2U^2L^{-2}$ arises as a good approximation. Thus Gaussian pulses are broadened as if there were no electron-density decay across the range cell.

(ii) Square Pulses, Exponential Decay

In this case we again take $A_0(2z)$ as the difference of Heaviside functions at $z = \pm L/4$ and, upon integrating in Equation (20) carefully, we find that only the term containing $A_0'(2z_1)\, A_0'(2z_2)$ in the integrand contributes. The result is

$$\sigma_{ti}^2 \approx \frac{4\Omega U}{\pi d \sinh(L/2d)} \left[\cosh(L/2d) - \frac{\sin(\Omega L/2U)}{(\Omega L/2U)} \right], \quad (28)$$

which reduces to Equation (22) when $L/2d \to 0$ as a result of allowing $d \to \infty$. We gain insight into this broadening width by allowing $L/2d$ to be large (yet $L/2d \ll \Omega L/2U$ in order that Equation (28) remain a valid approximation). In that case it follows that $\sigma_G^2 \approx 4\Omega U/\pi d$ and it

is clear that the densitp decay plays an important role in broadening the spectrum.

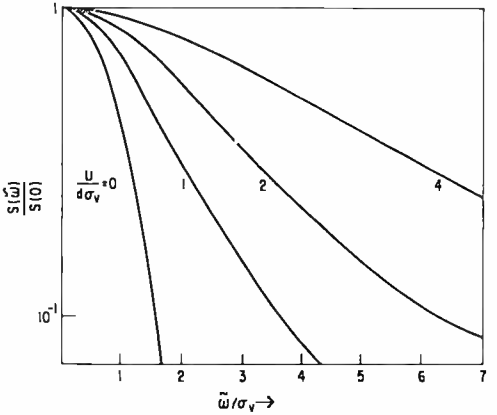


Fig. 2—Gaussian wake spectrum broadened by strong exponential electron-density decay in the range cell.

An alternative way of establishing these results is to compute $G(K)$ as in Equation (26) but with $z_o = -\infty$. For Gaussian pulses, $G(K)$ will be independent of the length d, and for square pulses an expression is obtained that will convolve with $\Phi(2\mathbf{k}_o;\omega)$ in the limit $L/2d \to \infty$ to give

$$S(\omega) \propto \mathrm{Im}\left\{\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} dx\,\frac{\exp(-x^2)}{x - [\omega/\sigma_v + iU/d\sigma_v]}\right\}, \tag{29}$$

$$\sigma_v^2 \equiv 8k_o^2 <(\Delta v)^2>$$

where the imaginary (as well as the real) part of the braced integral is the so-called plasma dispersion function, which is tabulated (Fried and Conte[6]) for varying ω/σ_v and $U/d\sigma_v$. The broadening with increasing $U/d\sigma_v$ is demonstrated in Figure 2, which displays a number of normalized spectra $S(\omega)/S(0)$ governed by Equation (28). Since $L_o/2d \gg 1$ for validity, the $U/d\sigma_v = 0$ curve holds only for $U = 0$.

ACKNOWLEDGMENT

The author is grateful to R. S. Ruffine and S. Stein for suggesting this problem as an extension of their preliminary investigation of square-pulse broadening by frozen-flow turbulence in the wake, and for ensuing discussions.

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

BISTATIC CLUTTER IN A MOVING RECEIVER SYSTEM

By

E. G. McCall

RCA Missile and Surface Radar Div.
Moorestown, N.J

Summary—This paper discusses the characteristics of ground clutter as seen by a moving receiver in a bistatic radar system. A method of analyzing this problem and predicting performance is developed. The method is illustrated by example using parameters from a representative system. In the example chosen, computer calculations were made and plotted to show the distribution of the received clutter amplitude versus frequency. Variations of this distribution with time are plotted to show that this must also be considered.

The results are analyzed and interpreted to gain insight into the phenomenon so that the method can be applied to similar situations. A method of simplifying the calculations is also presented and the results are compared with the detailed calculations.

Introduction

UNDER certain conditions the ability of a radar to detect targets is limited by unwanted returns from ground or sea clutter. To maximize the signal-to-clutter power ratio by appropriate signal processing and to predict the resulting performance, it is necessary to know the characteristics of the clutter returns competing with the signal. These characteristics depend not only on the clutter itself but also on the type of radar.

Analyses of clutter characteristics over the years have paralleled the growing sophistication of radar systems. The earliest studies (for noncoherent monostatic radars) were concerned with the total clutter power within the range-angle resolution cell containing the target. With the development of moving target indicators (MTI) and other coherent methods, the distribution of the clutter in velocity became

Note: The editors would like to express their appreciation to Dr. S. M. Sherman for his assistance in the review and editing of this paper in the author's absence.

an additional factor of importance (or in the case of cw radar, the sole factor). Both surface-based and airborne situations have been analyzed. For radar systems in which the transmitter and receiver are not co-located, bistatic characteristics of clutter have also received attention. Reference (1) considers the case for a monostatic system in an aircraft and is applicable to either a cw or a pulse doppler application. Reference (2) considers both the monostatic airborne case and the bistatic case, but is concerned only with the received frequencies and does not present any clutter data.

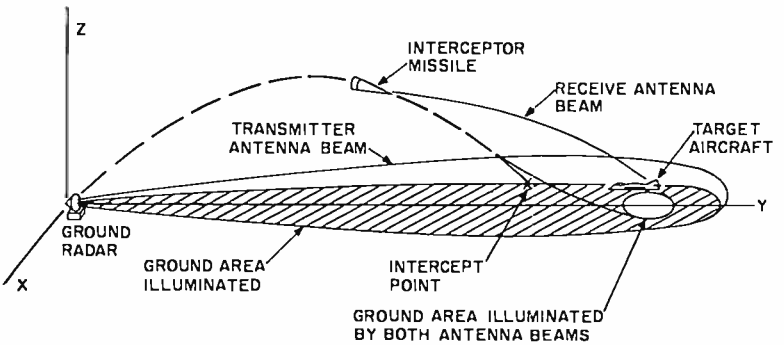


Fig. 1—Radar–missile–target geometry showing illuminated clutter areas.

The problem analyzed in this paper goes beyond those reported on previously. The type of radar system to which it applies can be characterized as follows:

(a) The transmitter is at a fixed point.

(b) The transmitted signal is a coherent pulse burst.

(c) The receiver is on a moving platform.

Representative of this type of system is a semi-active surface-to-air missile system. The transmitter is surface-based. The target is an aircraft. The receiver is in the front section of the missile approaching the target. The transmitted pulses are reflected from the target and received by the receiver in the missile, together with clutter returns. The situation is illustrated in Figure 1.

Clutter produces errors in missile guidance commands, and these errors are most critical during the last few seconds (the terminal phase) of missile flight. Therefore this study concentrates on the

[1] J. L. Farrell and R. L. Taylor, "Doppler Radar Clutter," *IEEE Trans. on Aerospace and Navigational Electronics*, p. 162 Sept. 1964.

[2] H. A. Crowder, "Ground Clutter Isodops for Coherent Bistatic Radar," *IRE National Convention Record*, Vol. 7, Part 5, p. 88-94 (1959).

clutter reflection characteristics during the terminal phase. The problem is especially important when the missile is diving on a target that is very close to the ground, for then large clutter returns compete with the target. To minimize the clutter returns, the system uses high resolution in both range and doppler.

In this paper, a method of analyzing this problem and predicting performance is developed. The method is illustrated with the results of calculations for a representative system. In the example chosen, computer calculations were made and plotted to show the distribution of received clutter amplitude versus frequency. The process was repeated for a sequence of time intervals to show how the clutter spectrum changes with time. Although the numerical results will naturally vary with the parameters chosen, they reveal certain characteristic features that give insight into the phenomenon, so that the method can be applied to similar situations.

A method of simplifying the calculations is also presented and the results are compared with the detailed calculations.

Some methods of reducing the adverse effects of clutter are suggested, but these are not followed through in detail, since the general principles are already well known. The main purpose of the paper is to define the clutter problem and to present the method of analysis.

Amplitude Calculations

In the geometry of the problem, the radar transmitting antenna is located at the center of a right-handed XYZ coordinate system. The target is incoming radially, at any azimuth, and at a low altitude. The missile is assumed to fly a parabolic trajectory, (parabolic in the YZ plane and also in the XY plane) and with a constant velocity. At specific points along the missile flight path, clutter calculations are made (see Figure 2). At each point, the necessary angles and ranges necessary for clutter calculations are computed from the geometry.

Table 1 is a list of the symbols used in the figures and throughout this paper.

The basis for comparison of the amount of clutter power received at any time is its ratio to the desired target signal received. The desired target signal power received by the bistatic receiver is

$$S = \frac{P_T G_{RT}}{4\pi R_T^2} \frac{\sigma_T}{4\pi} \frac{G_{MT}\lambda^2}{4\pi R_{TM}^2}, \tag{1}$$

where P_T = peak transmitter power,

G_{RT} = ground radar antenna gain in the direction of the target,

R_T = distance between the radar and the target, in meters,

σ_T = radar cross section of the target, in square meters,

G_{MT} = missile antenna gain in the direction of the target,

R_{TM} = distance between the target and the missile, in meters,

λ = ground radar transmission wavelength, in meters.

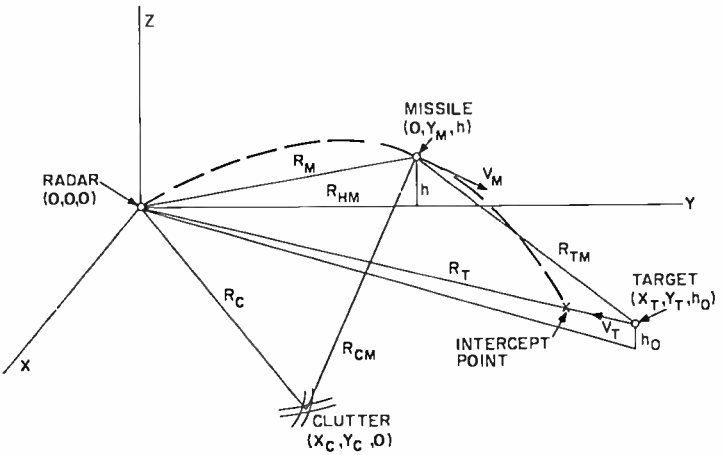


Fig. 2—Radar–missile–target geometry parameters at the time of clutter calculations.

The clutter power received by the bistatic receiver is reflected from an extended area, so that the antenna gain and the range parameters are not constant over the area. Therefore, the clutter power must be computed as a summation of the power reflected from many small elements of area. If the areas are small enough, an average antenna gain and range can be used for each element of area without any appreciable error. The clutter power received is then $\sum_{i=1}^{n} C_i$, where

$$C_i = \frac{P_T G_{RC_i}}{4\pi R_{C_i}^2} \frac{\sigma_{C_i}\sigma_o}{4\pi} \frac{G_{MC_i}\lambda^2}{4\pi R_{MC_i}^2}. \qquad (2)$$

where the symbols are the same as in Equation (1), except that the subscript C (clutter) replaces subscript T (target).

The antenna gains, range, and clutter area vary with the clutter patch under consideration and are a function of i in the summation. A method of segmenting the clutter area will be developed that is

consistent with the method by which the receiver processes the return signals. The parameter σ_o is the clutter reflection coefficient.

From the radar tracking data for the target and the missile, the radar can command the position of a receiver range gate in the missile to be coincident with the target signal. This range gate can then be made to track the target signal and thereby eliminate a large amount of the ground clutter. With a range-gated system, only reflections from ground areas that are at the same bistatic range as the target can arrive at the same time as the target return and be passed by the range gate. To locate these ground areas mathematically, the geometry of the model must be examined. The radar transmitter and the missile receiver are fixed points at the particular time at which the clutter analysis is to be made. The locus of all points at the same bistatic range as the target describes an ellipsoid, with the radar and the missile as the foci. The points on the ground where this ellipsoid intersects the ground are the locations of the clutter-producing areas. As the low-altitude intercepts are at short ranges, the earth can be considered to be flat. The intersection of a plane and an ellipsoid is an ellipse. Therefore the clutter-producing ground areas are located in an elliptical band around the radar with the transmitter as one focus of the ellipse. The width of the band is determined by the width of the range gate.

In a high pulse-repetition-frequency (prf) system there are other ground areas that contribute to the clutter power. These areas are at bistatic ranges that are both greater and smaller than the target bistatic range, the difference being equal to the range equivalent of a multiple of the prf period. These areas are also ellipses that have the radar as a focus. Figure 3 shows a family of these ellipses for a particular target-missile geometry.

The equation of the clutter area ellipses is of the form

$$\frac{X^2}{a^2} + \frac{(Y-b)^2}{c^2} = 1, \tag{3}$$

where

$$a = \frac{R_E^2 - R_M^2}{2[R_E^2 - R_{HM}^2]^{\frac{1}{2}}}$$

$$b = \frac{R_{HM}[R_E^2 - R_M^2]}{2[R_E^2 - R_{HM}^2]}$$

$$c = \frac{R_E[R_E^2 - R_M^2]}{2[R_E^2 - R_{HM}^2]}$$

and

$$R_{HM} = Y_M$$

$$R_E = R_T + R_{TM} \pm K \frac{3 \times 10^8}{\text{prf}} \begin{bmatrix} -0 \\ +300\,\tau \end{bmatrix} \text{ (in meters)}$$

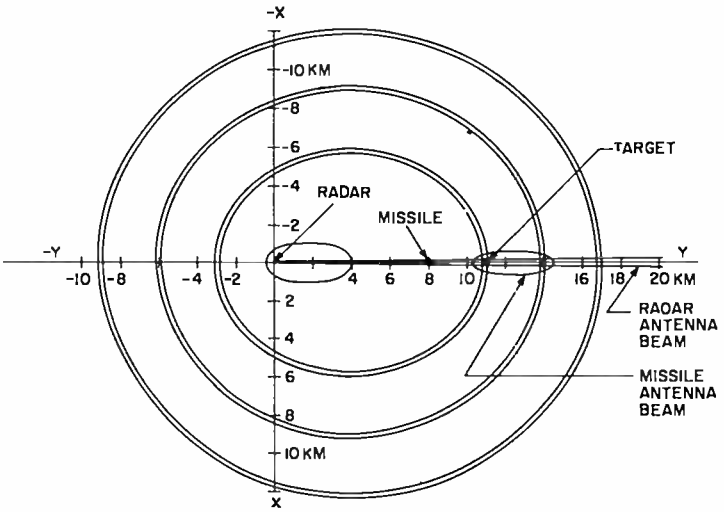


Fig. 3—Bistatic ground clutter areas.

The distances R are defined in Figure 2. R_E is the bistatic target range, plus or minus a multiple of the range equivalent of the prf period and with tolerance equivalent to the range gate width. The factor K in the expression for R_E is an integral number that determines the ellipse number. When $K = 0$, the ellipse is at the bistatic target range and is the primary clutter ellipse. Negative values for K produce the smaller ellipses, while the positive values produce the larger ellipses. The limit on negative values for K is that $R_E > R_M$; when $R_E = R_M$, the ellipsoid collapses to a line between the radar transmitter and the receiver. The limit on the positive values of K has been set at +2 for most calculations, for it has been found that the clutter power contributions from larger ellipses was negligible. The factor

$\begin{bmatrix} -0 \\ +300\tau \end{bmatrix}$ determines the width of the clutter band on the ground when the ranges are in meters. The range gate width, τ, is measured in microseconds.

Area Subdivision by Isodop

Now that the location and the shape of the ground clutter areas are known, they still must be divided up into elements of area for the summation. It can be done in the same manner that the radar signal processor will discriminate against the clutter, i.e., by doppler filtering. This method will also give a graph of how the clutter power is distributed with doppler frequency. The equation for the doppler frequency of clutter returns for the patch of clutter area along the missile velocity vector is

$$f_{cd} = \frac{V_M}{c} f_o = \frac{V_M}{\lambda}, \tag{4}$$

where V_M is the velocity of the missile (meters/sec), f_o is the transmitter carrier frequency, and c is the speed of light. This formula is correct because only the receiver is in motion, while both the transmitter and reflecting area are stationary. The doppler frequency of clutter areas off the missile velocity vector can be found by

$$f_{cd} = \frac{V_M}{\lambda} \cos \gamma, \tag{5}$$

where γ is the angle between the missile velocity vector and the line of sight to the clutter area of interest.

From Equation (5) it can be seen that for a fixed angle γ, there is a cone-shaped figure around the velocity vector on which the clutter doppler frequency is constant. Points on the ground where this cone intersects the ground plane have a constant doppler frequency. The lines connecting these points of constant doppler frequency are called isodops. The isodops are geometrical figures that are dependent upon the dive angle of the missile, ζ. This is the angle between the horizontal and the missile velocity vector. The shapes of the isodops are found from the relative magnitudes of the angles ζ and γ.

$$\begin{aligned} &\text{For } \gamma < \zeta, \text{ the isodops are ellipses;} \\ &\text{for } \gamma = \zeta, \text{ the isodops are parabolas;} \\ &\text{for } \gamma > \zeta, \text{ the isodops are hyperbolas.} \end{aligned} \tag{6}$$

When the isodops are symmetrical with respect to the axes, the equation is again of the form

$$\frac{X^2}{a^2} + \frac{(Y-b)^2}{c^2} = 1 \text{ where} \tag{7}$$

$$a = \frac{h \sin \gamma}{[\sin^2 \zeta - \sin^2 \gamma]^{\frac{1}{2}}} \tag{8}$$

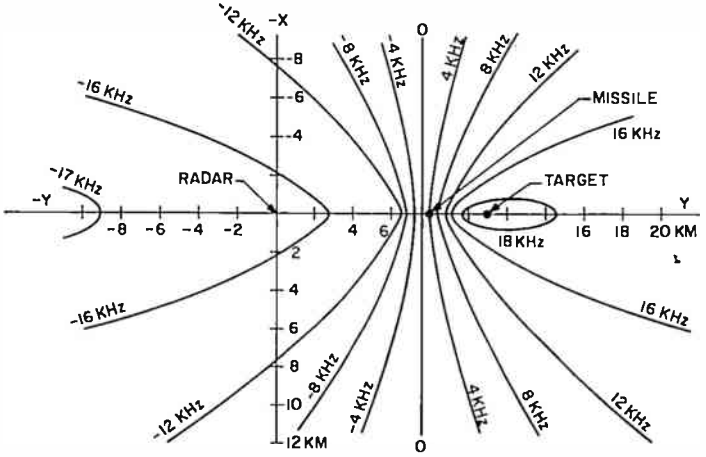


Fig. 4—Bistatic ground isodops.

$$c = \frac{h \cos \gamma \sin \gamma}{\sin^2 \zeta - \sin^2 \gamma} \tag{9}$$

$$b = R_{MH} + \frac{h \sin \zeta \cos \zeta}{\sin^2 \zeta - \sin^2 \gamma} \tag{10}$$

where R_{HM} is the distance between the radar and the missile vertical projection, i.e., Y_M, and h is the vertical projection, i.e., Z_M.

Equation (7) is the equation of either an ellipse or a hyperbola. The sign of the first term is determined by the magnitude of a (Equation (8)), which is dependent upon the relative values of γ and ζ. When $\gamma > \zeta$, the value of a^2 is negative. The first term then becomes negative and Equation (7) is the equation of a hyperbola. Equation (7) cannot be used for the parabola, $\gamma = \zeta$, and a more general form of the equation is required. Figure 4 shows a family of isodops for a particular target and missile geometry.

When the target and the interceptor missile are not both in the YZ plane, the equation for either the clutter ellipses or the isodops becomes more general. If the geometry is chosen so that the interceptor missile is in the YZ plane, the target aircraft will be out of the YZ plane. The equation for the clutter ellipses will remain the same, and the equation for the isodops will be the general equation for a conic section in which the axes are not parallel to the coordinate axes.

$$Ax^2 + By^2 + Cxy + Dx + Ey + F = 0, \tag{11}$$

where

$$A = \cos^2 \gamma - \sin^2 \epsilon \cos^2 \zeta,$$
$$B = \cos^2 \gamma - \cos^2 \epsilon \cos^2 \zeta,$$
$$C = -2 \cos \epsilon \sin \epsilon \cos^2 \zeta,$$
$$D = 2 \sin \epsilon \cos \zeta \, [R_{MH} \cos \epsilon \cos \zeta - h \sin \zeta],$$
$$E = -2 \, [R_{MH} \cos^2 \gamma - \cos \epsilon \cos \zeta \, (R_{MH} \cos \epsilon \cos \zeta - h \sin \zeta)],$$
$$F = \cos^2 \gamma \, (R_{MH}^2 + h^2) - (R_{MH} \cos \epsilon \cos \zeta - h \sin \zeta)^2,$$

and ϵ is the horizontal attack angle, the angle between the projection of the missile velocity vector on the X-Y plane and the Y axis.

The relationship between the angles ζ and γ remains as given in Equation (6).

In dividing up the clutter areas into elements of area, the intersections of the isodops and the clutter ellipses are used. The doppler resolution of the system is used as the isodop interval spacing. The clutter power received from the area between any two of these isodops is the power that would be collected in a band-pass filter equal in width to the doppler resolution. Figure 5 shows how this area is determined. A simultaneous solution of the center line of the clutter ellipse and the isodop of interest gives one end-point of the clutter area. The intersection of the clutter ellipse and the next isodop gives the other end-point of the clutter area. The width of the area is taken to be the range equivalent of the range gate in a bistatic system. This area is an approximation, but it is quite reasonable.

In any doppler frequency band, there are areas on several clutter ellipses that contribute power. To find the total clutter power in any band, the contribution from each of these areas must be computed and summed. A computer program is needed to make the required calcu-

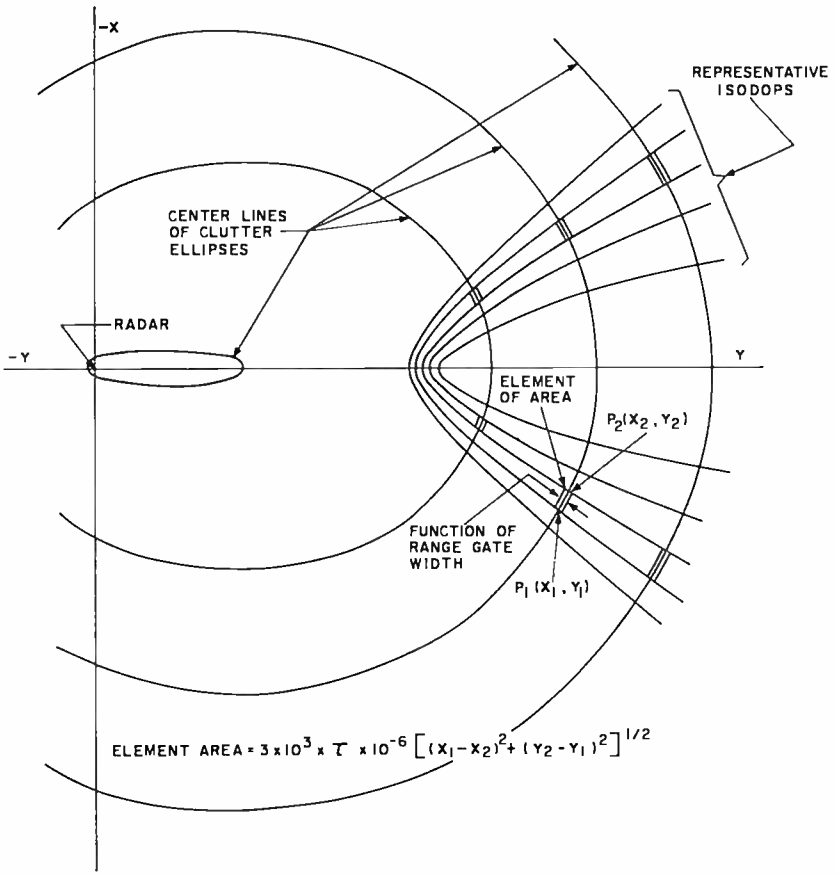


Fig. 5—Clutter areas at a common frequency.

lations and summations. The computer makes the following computation:

$$C = \underbrace{\sum_{0}^{\pm f_{cd\,\max}}}_{\text{doppler freq.}} \underbrace{\sum_{-K\,\max}^{+2}}_{\substack{\text{clutter} \\ \text{ellipse}}} \left[\frac{P_T \lambda^2 \sigma_o}{4\pi^3} \right] \left[\frac{G_{RC} G_{MC} \sigma_C}{R_C^2 R_{CM}^2} \right], \qquad (12)$$

where G_{RC} is the ground antenna gain in the direction of the clutter, G_{MC} is the missile antenna gain in the direction of the clutter, and the distances R_C and R_{CM} are defined in Figure 2. The terms in the right-hand bracket are recomputed for each change in the summation process.

The computer starts with the smallest clutter ellipse and the intersection of the zero doppler line and the first positive doppler line. There are two doppler areas within this doppler frequency band, one on each side of the ellipse. The antenna gains, the area, and the ranges are computed for each of these areas and the resulting clutter power is computed and summed. Then the corresponding clutter areas on the next larger clutter ellipse are computed, together with the antenna gains and ranges. The clutter power from these areas are then computed and added to the sum. This process continues to the second clutter ellipse outside the ellipse that is at the target range. These are the only ones that contribute significantly to the total clutter power. The next step is to increase the frequency and compute the clutter power in the next doppler resolution band. The process is repeated for all clutter ellipses. The doppler frequency continues to be increased in steps until it reaches the maximum doppler frequency. Then the computer returns to zero doppler and moves in the negative doppler direction in steps until it reaches the maximum negative doppler frequency. The sum of the clutter power in each band is printed out for inspection and recorded on magnetic tape for automatic plotting.

Typical Example

Computations based upon a typical example have been made using the procedure just described. The system parameters, assumed antenna patterns, and the missile flight path are given in Table I. Figure 6 is the clutter doppler frequency spectrum plot for a particular position of the target and missile along their flight paths. The clutter power is shown in dB relative to the target signal power. The target signal power (normalized to unity) is shown at its doppler frequency for the clutter case plotted. This plot was made by an automatic curve plotter from the computed data that was stored on the magnetic tape.

The computer program required to calculate the clutter doppler spectrum is complicated, because the solution of the equations for the intersections of the clutter ellipses and the isodops requires the solution of a fourth-order equation and selection of the valid roots. This indicates that the computations can be made for the most general case where the ground radar, the missile, and the target are not in the same vertical plane.

Interpretation of Results

Now, what information does the plot in Figure 6 contain? First, consider the broad spectrum of the transmitted and received power.

Table I—Assumed Constants and System Parameters

Target Velocity, $V_T = 500$ meters/second

Missile Velocity, $V_M = 1000$ meters/second

Target Height, $h_o = Z_T = 100$ meters

Range Gate Width, $\tau = 0.2$ microsecond

Pulse Repetition Frequency, prf = 50,000 and 60,000

Transmitted Wavelength, $\lambda = 5.45 \times 10^{-2}$ meter

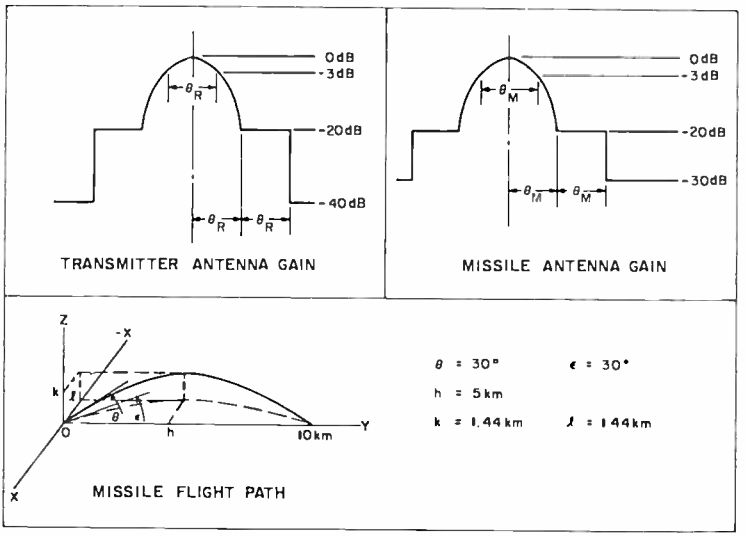
Radar Antenna Beamwidth, $\theta_R = 1.33$ degrees

Missile Antenna Beamwidth, $\theta_M = 14.3$ degrees

Doppler Filter Bandwidth, $B = 100$ Hz

Clutter Reflection Coefficient, $\sigma_o = -20$ dB

Equivalent Target Cross Section $\sigma_T = 0.5$ square meter



An infinite train of transmitted pulses has a line-frequency spectrum. The amplitudes of the lines follow the $(\sin x)/x$ distribution, where x is a function of the width of the pulses. The spacing between the lines is equal to the prf. If the pulse train has a finite length, the spectral lines have width and shape and the level does not drop to zero between the lines. In order to separate the effects on the spectrum in this analysis, the pulse train is assumed to be infinite so that the transmitted power is in the form of a line spectrum. The received power from a moving target is also a line spectrum that is shifted in frequency by an amount proportional to the missile and target velocities. The re-

ceived clutter power is also shifted in frequency due to the velocity of the missile. The amount of the frequency shift is therefore less than for an incoming target. Also the shift is not a fixed amount due to the extended nature of the clutter areas. If we refer to Figure 6 again, we see that it shows only the part of the target and clutter frequency spectrum around the central position of the transmitted line spectrum. If the plot were extended to the left and to the right, a similar spectral distribution would appear about each pulse-repetition frequency line.

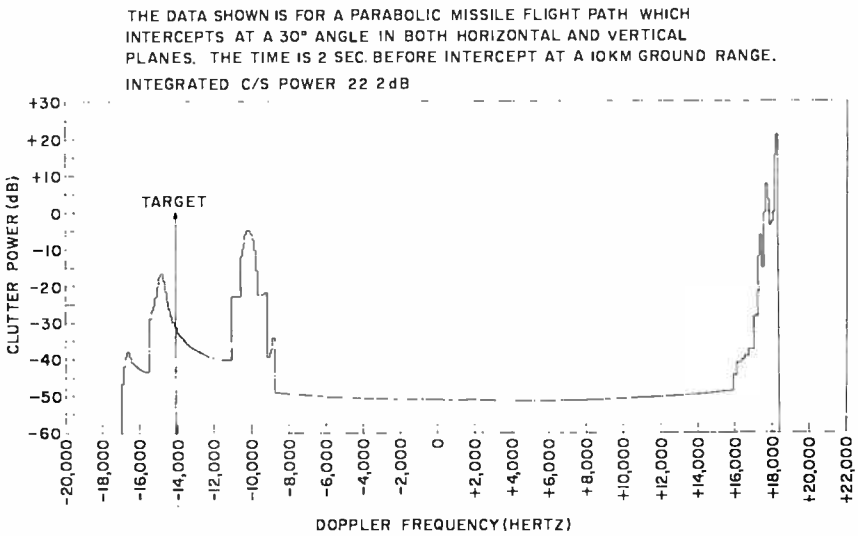


Fig. 6—Clutter power versus doppler frequency.

Figure 6 shows a peak of clutter power that is 21 dB higher than the target-signal amplitude. If the clutter power in all the 100-Hz intervals is added together, the level is 22.2 dB higher than the target-signal power. Thus, without frequency discrimination, it would be impossible to detect the target. The clutter level at the target-signal frequency is 32 dB below the target. Therefore, with frequency discrimination, the target can be detected. The peak of clutter power near the target-signal leads one to wonder if the relative frequency positions between target-signal and clutter peak remain separated enough so that the target can always be detected. This separation has been evaluated by computing the clutter frequency distribution as a function of time to intercept. A series of points were chosen along the missile trajectory at time positions ¼ second apart. At each of these points, a clutter frequency distribution similar to Figure 6 was

plotted. These plots showed a peak of clutter moving from the maximum positive clutter doppler frequency to the maximum negative doppler frequency as time progresses. Thus, for the system parameters and missile path chosen, there will be a time when the clutter peak occurs in the frequency range of the target filter. At this time it would be impossible to get tracking data on the target (a remedy is suggested later).

To understand the movement of the clutter peak, plots of the clutter ellipses and the isodops must be examined. Another computer program was developed to compute the positions of these lines and to plot the figures on an automatic curve plotter. A typical plot is shown in Figure 7. The radar is at the intersection of the X and Y axes. The projection of the missile on the XY plane is at a range of 8 km while the target is at 11 km range. One clutter ellipse is shown at a closer range than the target and two are at longer ranges. The half-power beam width of the ground antenna is shown in the direction of the target. Also, the missile antenna beam is pointing at the target and its half-power beam width is shown. The maximum clutter power is returned from the clutter area at the same range as the target. This clutter area is in the main beam of both the ground radar and the missile antenna. Other clutter areas that contribute significant amounts of clutter power are those that are also illuminated by the ground radar and that may or may not be within the antenna beam width of the missile antenna. There is an area on the next larger clutter ellipse that is in the main beam of both antennas, but the longer range makes the clutter amplitude lower, although still significant. The clutter area on the smaller ellipse, which is illuminated by the ground radar and is in the side-lobe region of the receiver antenna, is also significant. The shorter range makes the clutter contribution significant, although not as high as the primary clutter area. The peaks of clutter all come from the areas on each clutter ellipse that are illuminated by the main beam of the ground-radar antenna. The approximate doppler frequency can be determined from the position of these significant clutter areas relative to the isodops. With the exception of the smallest clutter ellipse, the significant clutter areas are all near the higher doppler frequency. The significant area of the smallest clutter ellipse can be seen to be at a lower doppler frequency. A comparison of the location of these frequencies to the doppler spectrum plot show that these areas are the sources of the peaks of clutter.

The effect of the decreasing range between target and missile as the missile approaches the target is to reduce the size of any ellipses inside the primary ellipse (the one at the bistatic target range). As the

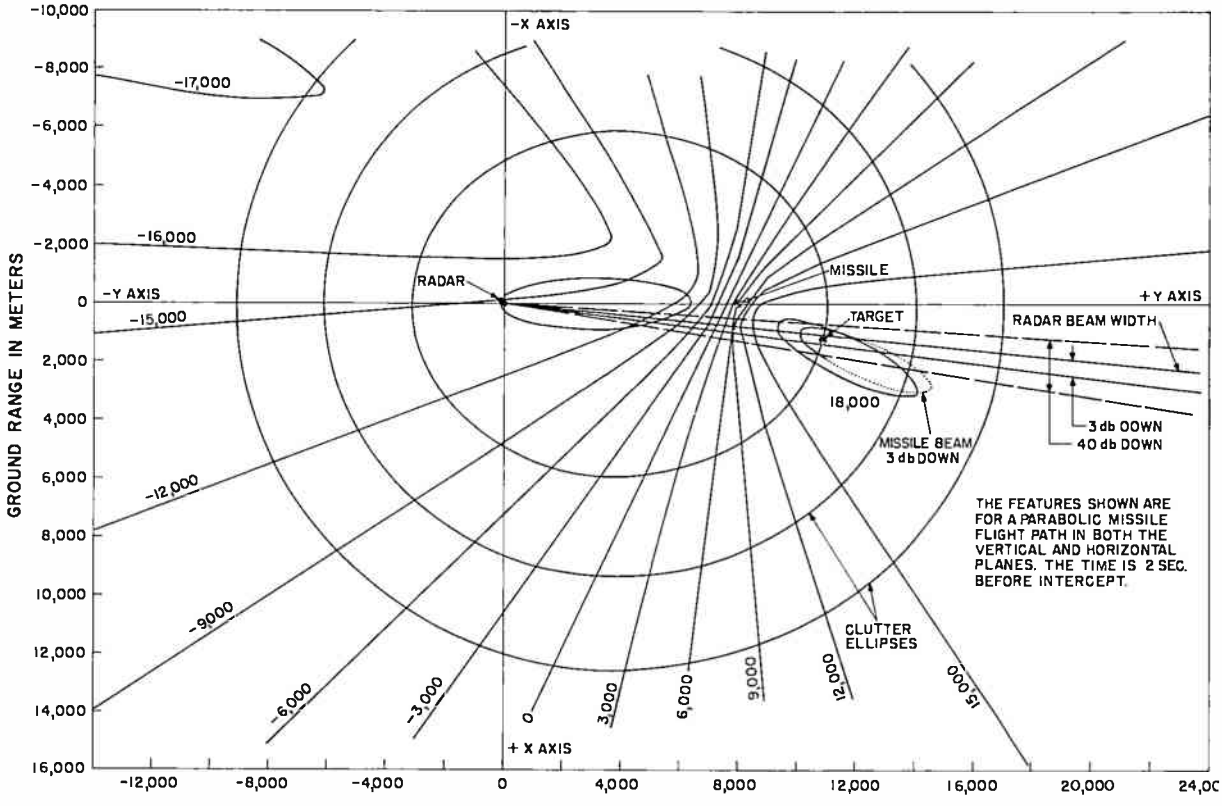


Fig. 7—Bistatic clutter and isodops.

ellipses are reduced in size, the significant clutter area of the smallest ellipse moves toward the more negative doppler lines. Therefore, the minimum peak in the clutter doppler spectrum moves toward the negative end of the spectrum. This continues until there are no ellipses smaller than the primary ellipse that contains clutter areas that have

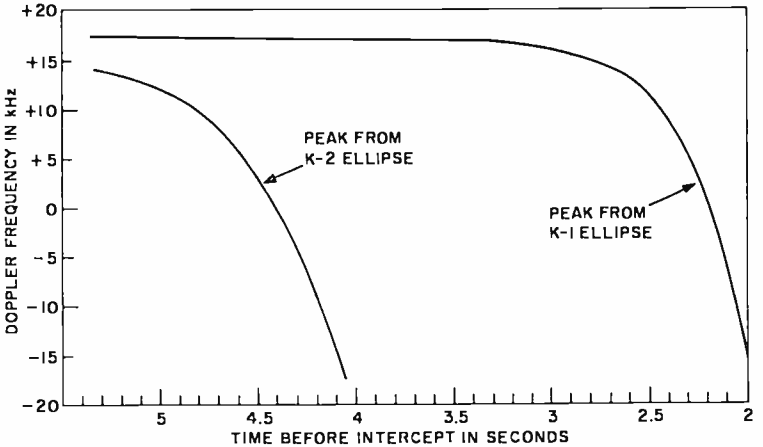


Fig. 8—Doppler frequency of clutter peaks due to close-in clutter ellipses.

the same bistatic range as the target. The number of peaks is equal to the number of clutter ellipses, and the number of peaks that move from positive doppler to negative doppler is equal to the number of clutter ellipses that are inside the primary clutter ellipse.

The rate of change in the frequency of the negative peak of clutter is of interest, since it moves quite rapidly. Its frequency has been plotted in Figure 8 for a particular set of system parameters. This figure shows the frequency change for the two smaller ellipses. The time at which the peak would remain in any target filter is seen to be very short.

Effect of Clutter and Methods of Signal/Clutter Improvement

Clutter in the same filter as the target results in errors in the output. The measure of the error depends upon the type of information channel and the method of demodulation being used. The method by which the clutter power enters the target filter can be either direct or indirect. In the direct method, the received clutter power has a por-

tion of the spectral distribution at the target frequency. An indirect method is one in which, through some frequency mixing process, the peaks of the clutter power mix to a sum or difference output at the target frequency. This can occur in an amplifier or some other device in the data-transmission and processing system.

Clutter power can also enter the target filter if the pulse train is finite in length. A pulse train of finite length has some power distributed between the spectral peaks in the received spectrum. This power contributes to the clutter power that enters the target filter. Also, a target filter of realizable design will have a non-zero response at frequencies on each side of its pass band. The response of the target filter at the peaks of clutter will also determine the amount of clutter power that enters the target filter.

Now what can be done to improve the target-signal/clutter-power (S/C) ratio in a case of the type just presented? There are at least five alternatives:

(1) Increase the system prf so that the spectral frequency lines are farther apart. This keeps the received target spectral lines spaced between the maximum clutter doppler and the minimum clutter doppler from the next higher clutter line. This method is acceptable if the missile and target velocities are known accurately enough and the required prf is not too high.

(2) Jump the prf at an appropriate time. A scheme for anticipation or recognition could be used to determine the time for interference and a time to jump prf to a new value. This would prevent a peak of clutter from entering the target filter although some low level of clutter could always be in the target filter.

(3) Have a clutter tracking filter. This filter would be adjusted to remove the major peaks in the clutter spectral distribution. It can be seen that the peaks occupy a rather narrow range of frequencies near the maximum possible doppler frequency. In the example, a 3-kHz filter following the range gate at some suitable i-f frequency could remove the major clutter peaks without sacrificing target detection. For the target doppler frequency to be in this region, it would have to be flying a course orthogonal to the missile velocity vector, and this is a most unlikely situation.

(4) Accept the clutter condition. In view of the fact that the peak of clutter that moves with frequency would remain at the same frequency as the target for such a short time, the resulting

errors could be negligible or not worth the complications necessary to take other action. With a minimum of effort, perhaps the bad data points (when a clutter peak entered the target filter) could be eliminated or ignored. It is quite possible that such a system could be operated satisfactorily if the design considers the clutter model correctly.

(5) Some improvement in the S/C ratio can be obtained by adjusting the flight path of the missile. In this method the missile would fly a "hook shot" flight path. This means that the missile would have an out-of-plane component of flight in addition to the "up and out and dive", in-plane, flight path. This kind of flight path reduces the total C/S amplitude during the early part of the flight because the significant areas of clutter in the clutter ellipses, which are at longer ranges than the bistatic target range, are in the side-lobe region of the receiver antenna pattern. This reduces the clutter power contribution from these areas during this portion of the flight path. It does not affect the peak in the clutter doppler spectrum that moves in frequency as the missile moves along its trajectory.

Simplified Method of Calculation

After making detailed calculations of the clutter spectral distribution and determining the causes of the peaks, a simplified approach can be justified to obtain a frequency distribution. Some calculations are required, but the detailed summation is not always necessary. It has been recognized that the clutter areas within the main beam of the ground radar are the significant areas and result in the peaks in the clutter frequency distribution. Therefore a simplified clutter frequency distribution can be constructed by computing the clutter power return from the significant areas as shown in Figure 9 and the associated doppler frequency for each of these areas. These values are the peaks in the clutter frequency distribution.

The relative amplitude of these clutter peaks can be found from the ratio of Equations (1) and (2).

$$\frac{C}{S} = \frac{\dfrac{P_T G_{RC} G_{MC} \lambda^2 \sigma_C \sigma_0}{(4\pi)^3 R_C^2 R_{CM}^2}}{\dfrac{P_T G_{RT} G_{MT} \lambda^2 \sigma_T}{(4\pi)^3 R_T^2 R_{TM}^2}} = \frac{G_{MC} \sigma_C \sigma_0 R_T^2 R_{TM}^2}{G_{MT} \sigma_T R_C^2 R_{CM}^2} \tag{13}$$

(Note that $G_{RC} = G_{RT}$ for these clutter areas.) The ratio of the antenna gains in Equation (13) can be estimated from some general information about the problem and experience with other clutter diagrams. At the primary clutter area, where the bistatic target and clutter ranges are equal, the antenna gains, G_{MC} and G_{MT}, are equal. The significant clutter areas at closer ranges will have a gain ratio

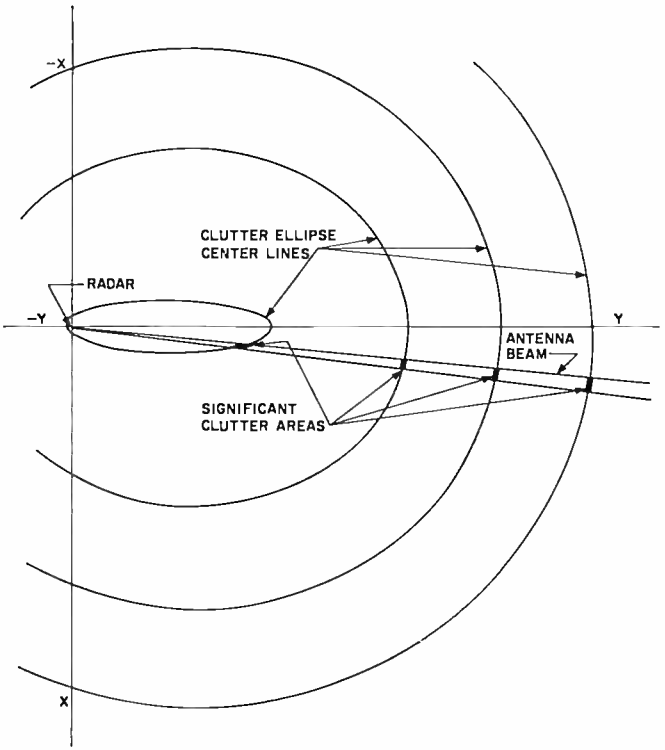


Fig. 9—Significant clutter areas for simplified calculations.

equal to the dB difference between missile antenna peak and side-lobe gains. The clutter areas at longer ranges can have a gain ratio close to either that of the primary area or that of the closer range clutter areas. If the missile is at a long range from the target, the antenna gains are approximately the same. At shorter missile-to-target ranges, the antenna gain ratio will be the same as for the short-range clutter areas.

The clutter area, σ_C, can be approximated by

$$\sigma_C = 300\,\tau\,R_C\theta_R \text{ square meters,} \qquad (14)$$

where τ is the range gate width in microseconds and θ_R is the radar antenna beam width in miliradians. The range in meters, can be computed or approximated from the system geometry.

Using Equations (13) and (14), assumptions about the ratio of the antenna gains and the values for σ_T and σ_0, and the ranges used in Equation (13), the ratio of C/S can be found for each of the significant clutter areas shown in Figure 9. These are the peaks of clutter in the clutter frequency distribution plot. The minimum level is approximately determined by the following:

$$\text{minimum clutter level} = (\text{peak clutter level}) - [(\text{transmitter antenna peak gain} - \text{average side lobe level}) + (\text{receiver antenna peak gain} - \text{average sidelobe level})] \tag{15}$$

For the parameters used earlier this is $24.7 - [40 + 30] = -45.3$ dB C/S ratio for the minimum clutter level.

The doppler frequency of each clutter area of interest can be found from Equation (5):

$$f_{cd} = (V_m/\lambda) \cos \gamma,$$

where the angle γ is the angle between the missile velocity vector and the line of sight between the missile seeker and the clutter area of interest. The value of cos γ can be found from the following equations and reference to Figure 10.

$$\cos \gamma = \frac{R_{CM}^2 + R_1^2 - R_2^2}{2R_{CM}R_1}, \text{ for } \zeta > 0 \tag{16}$$

$$R_1 = [X_V^2 + (Y_V - Y_M)^2 + h^2]^{\frac{1}{2}}, \tag{17}$$

$$R_2 = [(X_V - X_C)^2 + (Y_V - Y_C)^2]^{\frac{1}{2}}, \tag{18}$$

$$R_{CM} = [X_C^2 + (Y_C - Y_M)^2 + h^2]^{\frac{1}{2}}, \tag{19}$$

$$X_V = h\left(\frac{\sin \epsilon}{\tan \zeta}\right), \tag{20}$$

$$Y_V = Y_M + h\left(\frac{\cos \epsilon}{\tan \zeta}\right), \tag{21}$$

$$X_C = R_C \sin \rho, \tag{22}$$

$$Y_C = R_C \cos \rho, \tag{23}$$

$$R_C = \frac{c\,(1-e^2)}{1-e\,\cos\,\rho}, \tag{24}$$

$$e = \left[1 - \frac{a^2}{c^2}\right]^{\frac{1}{2}}. \tag{25}$$

The values for a and c can be found from the clutter ellipse relationships given in Equation (3).

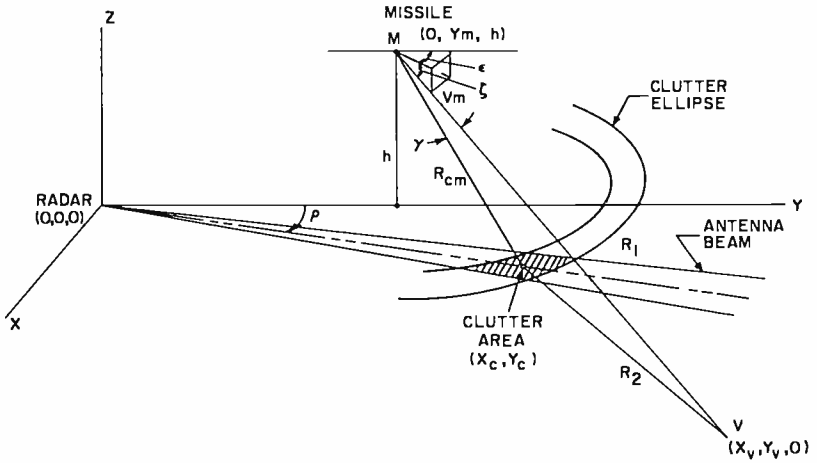


Fig. 10—Clutter area doppler frequency.

The clutter doppler frequency can then be found for each of the significant clutter areas by solving Equation (5), using Equations (16) through (25), with the system parameters suitable for each area. The width of each clutter peak is only 100 to 300 hertz based upon the detailed calculations made for similar problems. If a more accurate estimate is required, the doppler frequency of the significant areas could be calculated at each side of the area rather than finding it for the center only. The results of a plot of the simplified calculations are shown in Figure 11. This should be compared with the curve of Figure 6 (two seconds before target intercept).

Adjusting the Calculations to a Specific Terrain

The calculations made thus far were all based upon a constant σ_0 over the entire ground area. Without any other information, this is a reasonable approximation to the actual conditions. If a ground map of

the anticipated intercept area is available, appropriate adjustments to the frequency distribution can be made. To do this, one of the clutter-isodop plots would have to be superimposed over the terrain map. A realistic σ_0 could be assigned to general areas depending upon the make-up of the terrain. Changes in the clutter frequency distribution plotted for a constant σ_0 can then be made by an adjustment based

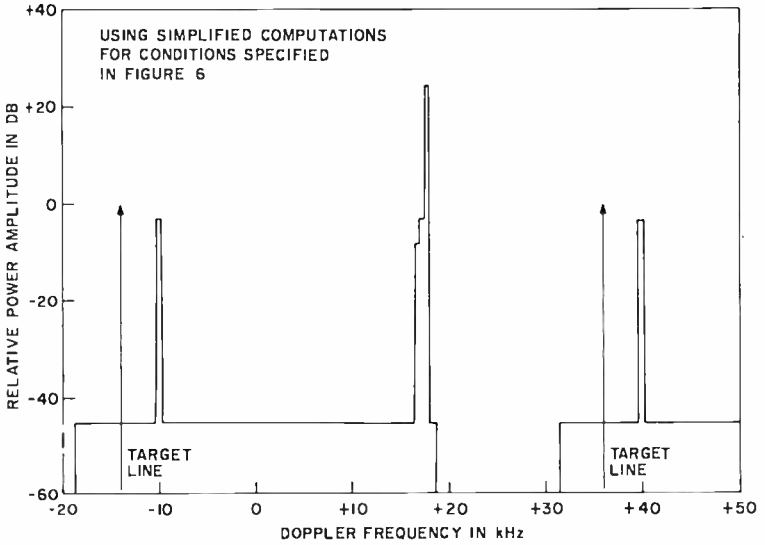


Fig. 11—Clutter amplitude versus doppler frequency.

upon the σ_0 values assigned on the map. The adjustments would be made in the frequency range indicated by the isodops. This could reduce the amplitude of the computed peaks, add peaks at other doppler frequencies, or cause an adjustment in the minimum level.

Conclusions

Through the solution of some detailed geometrical relations, it is possible to compute the amplitude versus frequency distribution of received clutter in a given bistatic radar system. Furthermore, the detailed calculations can be reduced to some simple computations with quite acceptable results. It is possible to apply the results of calculations to a particular ground area to make use of values of σ_0 that are representative of a realistic environment. The calculations show that there are definite peaks in the clutter spectrum and not all remain fixed in frequency as time passes and the receiver moves toward the target. There are several ways in which the clutter power can interfere with

the received target signal and result in target measurement errors. Fortunately there are several things that can be done to minimize these effects:

(1) range gating to reduce the amplitude of the clutter power,
(2) frequency discrimination (doppler filtering) to separate the received target and clutter power,
(3) optimize the missile flight path to further reduce the received clutter amplitude,
(4) change the prf at an optimum time to prevent a peak in the clutter spectrum from entering the target filter,
(5) use a clutter tracking filter to remove the peaks of clutter.

Through an understanding of the sources of the clutter returns, the anticipated doppler frequencies, and several methods for reduction of the clutter, it is possible to have guided missile systems operate at low altitudes in the presence of ground clutter.

Acknowledgments

The author is indebted to many of his associates at the RCA Missile and Surface Radar Division for encouragement, suggestions and interest in this study. Special credit should be given to H. C. Cox who wrote the computer programs to obtain the clutter data used in this analysis.

MINIATURE MICROSTRIP CIRCULATORS USING HIGH-DIELECTRIC-CONSTANT SUBSTRATES

By

B. Hershenov* and R. L. Ernst†

Summary—A broad-band miniature circulator using an $\varepsilon_r = 83$ dielectric as a substrate for quarter-wavelength impedance transformers has been tested. Good performance over the frequency band of approximately 2.8 to 3.7 GHz has been measured.

MICROSTRIP circulators designed with distributed circuit elements in the form of a central resonant disk coupled through quarter-wavelength impedance transformers have been described by Hershenov.[1] While this approach results in very compact circulators at X-band, the size of the circulators becomes very large at lower frequencies. At 2 GHz, for example, the diameter of an all-garnet circulator would be approximately 2 inches, which is considerably larger than most other substrates used in a microwave integrated-circuit system.

At L- and S-bands (1–4 GHz) the major contributions to the overall microstrip circulator size are the quarter-wave transformer lengths between the ferrimagnetic disk and the 50-ohm connectors or transmission lines. The transformers are necessary when 20-dB isolation bandwidths greater than 8–12% are required. The transformer lengths can be reduced by using temperature-compensated, high-dielectric-constant materials.

Figure 1 shows a circulator (that operates below resonance) built for operation at 3.5 GHz. The circulator consists of a polycrystalline garnet puck cemented into the annulus of a temperature-compensated, high-dielectric-constant ceramic.* Both the garnet and dielectric are 0.019 inch thick with relative permitivities of 14.5 and 83, respectively.

* RCA Laboratories, Princeton, N. J.

† Defense Advanced Communication Laboratory, Princeton, N. J.

[1] B. Hershenov, "X-Band Microstrip Circulators," *Proc. IEEE*, Vol. 54, p. 2022, Dec. 1966.

* American Lava Corporation designation N750-T96.

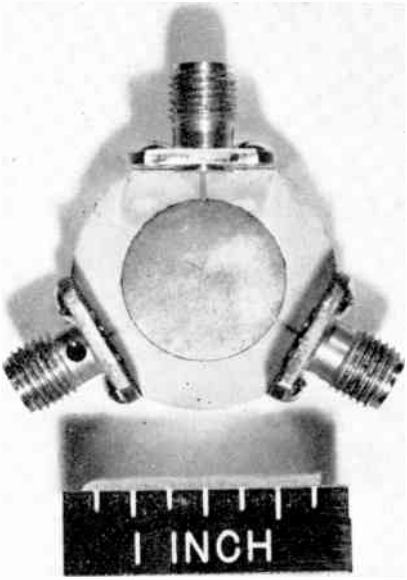


Fig. 1—S-band microstrip circulator.

The ceramic relative permitivity has a negative temperature coefficient of 750 ± 120 ppm/°C. (Ceramics with relative dielectric constants ranging between 50 and 70 are available with significantly lower temperature coefficients.) The copper on both surfaces of the garnet and ceramic is deposited by evaporation after a chrome flash. The disk conductor and transformer lines are defined by standard photolithographic techniques.

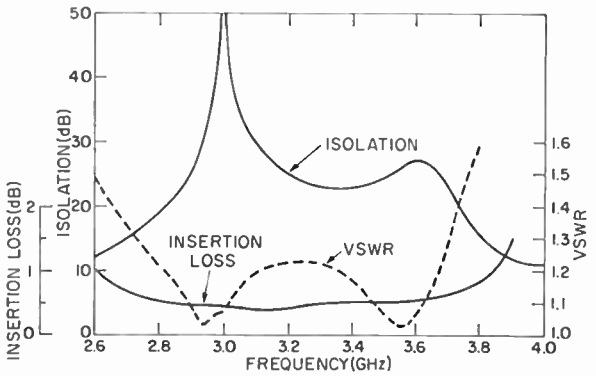


Fig. 2—Microstrip circulator characteristics.

Figure 2 is a plot of preliminary results obtained with the circulator using a permanent magnet 0.6 inch in diameter and 0.15 inch thick. The 20-dB isolation bandwidth is 28% with the insertion loss between 0.4 and 0.65 dB. It is apparent that the large difference in garnet and ceramic dielectric constants does not affect the circulator performance. Empirical variations of the present impedance-transformer values may yield further performance improvement. (A fired glass frit for bonding the puck in the ceramic annulus would be a major improvement over the present high-temperature epoxy technique.)

Table I—Size Comparison for Different Substrate Materials

Substrate	Resonant Disk Diameter (inches)	Impedance Transformer (inches)	Total Diameter (inches)
Garnet Puck in Alumina ($\epsilon_r = 9.6$)	0.572	0.320	1.212
All Garnet	0.572	0.270	1.112
Garnet Puck in N 750-96 Ceramic ($\epsilon_r = 83$)	0.572	0.106	0.784

Table I illustrates the size reduction achieved with the high-dielectric-constant ceramic. Further size reduction is possible for all three examples of Table I by designing for above-resonance operation. A 30% reduction of the puck diameter can be achieved, but with a reduced isolation bandwidth. It may be possible to use the high-dielectric-constant ceramic to reduce the size of the ferrimagnetic puck. The outer portion of the ferrimagnetic disk contributes little to the nonreciprocal action of the circulator, and it could be replaced by the ceramic with its appreciably higher relative dielectric constant. This approach would require good bonding techniques between the ferrimagnetic and ceramic, but would reduce the resonant-disk diameter as well as the quarter-wave transformer lines. A smaller ferrimagnetic disk would be used and the copper conducting disk would cover it and overlap a portion of the ceramic.

Finally, higher-dielectric-constant ceramics can be used for larger size reductions, but they will have a larger percentage dielectric-constant change with temperature. Their use will be dictated by circuit requirements and environmental conditions.

ACKNOWLEDGMENT

The authors wish to acknowledge the assistance of H. L. Davis and A. Mikelsons in expediting the fabrication and testing of the circulator.

RCA Technical Papers

Second Quarter, 1969

Any request for copies of papers listed herein should be addressed to the publication to which credited.

"Comparison Between Measured Critical Currents and a Model of Transiently Stabilized Nb_3Sn Magnets," E. R. Schrader, *Jour. Appl. Phys.* (April) 1969

"Effects of Oxidation on Electrical Characteristics of Silicon-on-Sapphire Films," E. C. Ross and Coauthor, *Jour. Appl. Phys.* (April) 1969

"The Engineer Is a Loner," R. F. Ficchi, *The Electronic Engineer* (April) 1969

"An Evaporated Heterojunction Diode Strain Sensor," R. M. Moore and C. J. Busanovich, *Proc. IEEE* (Letters) (April) 1969

Discussion on "FM Bandwidth as a Function of Distortion and Modulation Index," I. T. Bucher, *Trans. GCOM* (Correspondence) (April) 1969

"Microwave Conductivity of Granular Superconducting Aluminum Films," R. V. D'Aiello and Coauthor, *Jour. Appl. Phys.* (April) 1969

"Properties of Efficient Silicon-Compensated $AlGa_{1-x}As$ Electroluminescent Diodes," H. Kressel, F. Z. Hawrylo, and N. Almeleh, *Jour. Appl. Phys.* (April) 1969

"Radiation Resistance of Al_2O_3 MOS Devices," K. H. Zaininger and A. S. Waxman, *Trans. IEEE GED* (April) 1969

"Stacked High-Power Avalanche-Diode Oscillators," S. G. Liu, *Proc. IEEE* (Letters) (April) 1969

"Steady-State Flux Jumping in Superconducting Niobium Titanium Tubes in Superimposed ac and dc Magnetic Fields," D. A. Gandolfo and Coauthors, *Jour. Appl. Phys.* (April) 1969

"Thin-Film Josephson Junctions Using Getter-Sputtered Niobium," J. E. Nordman, *Jour. Appl. Phys.* (April) 1969

"Threshold Gate Approximations Based on Chow Parameters," R. O. Winder, *Trans. IEEE GC* (Short Notes) (April) 1969

"Designing Lumped Elements Into Microwave Amplifiers," M. Caulton and W. E. Poole, *Electronics* (April 14) 1969

"Sonic Film Unit Is a Sound Bet," R. Shahbender, *Electronics* (April 14) 1969

"Cathodochromism in Photochromic Materials," Z. J. Kiss, and W. Phillips, *Phys. Rev.* (15 April) 1969

"Approximate Solutions for a Coupled Pair of Microstrip Lines in Microwave Integrated Circuits," A. Schwarzmann, *Microwave Jour.* (May) 1969

"An Automatic System for the Control of Multiple Drone Aircraft," B. E. Keiser and P. Z. Peebles, Jr., *Trans. IEEE GAES* (May) 1969

"Barrier Height Diminution in Schottky Diodes Due to Electrostatic Screening," S. S. Perlman, *Trans. IEEE GED* (May) 1969

"Coupled Microstrip-Line Varactor Doublers," E. F. Belohoubek and A. Rosen, *Trans. GMTT* (Correspondence) (May) 1969

"Designing MOS Systems for Radiation Environments," A. G. Holmes-Siedle and K. H. Zaininger, *Solid State Tech.* (May) .. 1969

"Emission of Visible Radiation from Extended Plasmas in Silicon Diodes During Second Breakdown," D. J. Dumin, *Trans. IEEE GED* (May) 1969

"Induced dc Negative Resistance in Avalanche Diodes," A. S. Clorfeine and R. D. Hughes, *Proc. IEEE* (Letters) (May) 1969

"Multipath Fading in FSK Communication Links—An Experimental Investigation," G. P. Richards, W. T. Bisignani, and S. H. Roth, *Trans. GAES* (May) 1969
"Photo Erasable Dark Trace Cathode-Ray Storage Tube," W. Phillips and Z. J. Kiss, *Proc. IEEE* (Letters) (November) 1968
"Author's Reply," W. Phillips and Z. J. Kiss, *Proc. IEEE* (Letters) (May) 1969
"Role of Optical Flux and of Current Density in Gradual Degradation of GaAs Injection Lasers," N. E. Byer, *IEEE Jour. Quantum Electronics* (May) 1969
"Solid Solutions in the Pseudobinary (III-V)-(II-VI) Systems and Their Optical Energy Gaps," W. M. Yim, *Jour. Appl. Phys.* (May) 1969
"Solving the Equation for the Maximum a Posteriori Estimate," A. Guida, *Proc. IEEE* (Letters) (May) 1969
"Studies of Electron Bombardment Damage in GaAs by Thermally Stimulated Conductivity Measurements," H. Schade, *Jour. Appl. Phys.* (May) 1969
"Wide Tuning Range S-Band Maser," D. J. Miller and G. G. Weidner, *Proc. IEEE* (Letters) (May) 1969
"Loop Gain Measurements Made with Paired Amplifiers," K. Feher, *Electronics* (Designer's Casebook) (May 26) 1969
"Current-Voltage Characteristics of Silver-n-Type GaP Schottky Barriers," C. R. Wronski, *RCA Review* (June) 1969
"The Design and Operating Characteristics of a 1200-Element Ferroelectric-Electroluminescent Display," G. W. Taylor, *Trans. IEEE GED* (June) 1969
"Digit-by-Digit Transcendental-Function Computation," R. J. Linhardt and H. S. Miiller, *RCA Review* (June) 1969
"Double Injection Electroluminescence in Anthracene," J. Dresner, *RCA Review* (June) 1969
"The Effect of Barrier Recombination on Production of Hot Electrons in a Metal by Forward Bias Injection in a Schottky Diode," R. Williams, *RCA Review* (June) 1969
"High-Frequency Behavior of Microstrip Transmission Lines," L. S. Napoli and J. J. Hughes, *RCA Review* (June) 1969
Comment on "Local-Mode Absorption and Defects in Compensated Silicon-Doped Gallium Arsenide," H. Kressel, H. Nelson, and F. Z. Hawrylo, *Jour. Appl. Phys.* (Communications) (June) .. 1969
"Single-Frequency Argon Laser," L. Gorog and F. W. Spong, *RCA Review* (June) 1969
"Switching and Storage Characteristics of MIS Memory Transistors," J. T. Wallmark and J. H. Scott, *RCA Review* (June) 1969
"Theory of the Switching Behavior of MIS Memory Transistors," E. C. Ross and J. T. Wallmark, *RCA Review* (June) 1969
"Transmission and Reflection Group Delay of Butterworth, Chebychev, and Elliptic Filters," C. M. Kudsia and N. K. M. Chitre, *RCA Review* (June) 1969
"Vapor Pressure Data for the Solid and Liquid Elements," R. E. Honig and D. A. Kramer, *RCA Review* (June) 1969
"Origin of Photocarriers in Phthalocyanines," S. E. Harrison, *Jour. Chem. Phys.* (1 June) 1969
"Zeeman Effect of Cr^{3+} Fluorescence in $SrTiO_3$," W. J. Burke and R. J. Pressley, *Phys. Rev.* (10 June) 1969

Patents Issued to RCA Inventors

Second Quarter 1969

April

H. Ball and D. N. McLaughlin, *Frequency and Phase Error Detection Means for Synchronization Systems*, 3,441,342.
W. H. Barkow, *Torsional Light Deflector*, 3,436,139.
J. B. Beck, *Deflection Circuit with Bidirectional Trace and Retrace Switches*, 3,441,791.
G. P. Benedict, *Two Phase Clock Pulse Generator Employing Delay Line Having Input-Output Means and Characteristic Impedance Termination Means at Each End*, 3,441,751.
A. Blicher, *Insulated-Gate Field-Effect Transistor with Critical Bulk Characteristics for Use as an Oscillator Component*, 3,439,236.
C. E. Bush, *Tape Reel or Similar Article*, D213,699.
J. D. Callaghan, *Antenna for a Radio or Television Set*, D213,968.
J. L. Dailey, *Modulating or Q-Switching a Laser*, 3,437,951.
A. G. F. Dingwall, *Method of Fabricating Semiconductor Devices*, 3,437,533.
M. B. Finkelstein, *Electronic Splicing Control System*, 3,441,666.
L. J. French, *Electro-Optical Memory*, 3,440,620.
J. J. Gibson, *Logic Circuits Employing Field-Effect Transistors*, 3,439,185.
C. E. Goltsos and A. Amato, *Temperature Controlled Circuit Boards*, 3,440,407.
G. L. Grundmann, *Electron Beam Deflection and Low-Voltage Supply Circuit*, 3,436,591.
P. Haferl, *Black Level Setting Circuit for Color Subcarrier Modulator*, 3,437,745.
O. P. Hart, *Field-Effect Oscillator Circuit with Frequency Control*, 3,436,681.
G. V. Jacoby, *Self-Synchronizing Readout with Low Frequency Compensation*, 3,441,921.
E. E. Janson and L. P. Thomas, *Threshold Control for Sync Separator Noise Protection Circuit and for AGC Stage*, 3,441,669.
M. F. Kaminsky and G. Spector, *Card Processing Apparatus*, 3,440,409.
Z. J. Kiss, *Method of Preparing Luminescent Materials*, 3,438,881.
J. A. Konkel and F. A. Joy, *Hue-Adjust Circuit for a Color Television Receiver*, 3,436,470.
T. G. Marshall, *Ladder Network Filters Having a Negative Resistance To Compensate for Lossy Reactive Components in the Filter*, 3,439,291.
J. A. McDonald and P. C. Wilmarth, *Stabilization of Television Deflection Circuits*, 3,441,790.
J. C. Miller and C. M. Wine, *Coordinate Converter System*, 3,439,317.
A. K. Rapp, *Driver Sense Circuit Arrangement*, 3,440,4444.
H. D. G. Scheffer, *Wire Handling Apparatus*, 3,438,403.
H. D. G. Scheffer, *Wire Stringing Machine*, 3,438,405.
K. Siwko, *Inter-Stage Coupling Circuit for Neutralizing Internal Feedback in Transistor Amplifiers*, 3,441,865.
H. Weinstein, *Stress-Wave Thin-Film Memory*, 3,440,625.
R. E. Werner, *Bidirectional IGFET with Symmetrical Linear Resistance with Specific Substrate Voltage Control*, 3,441,748.

C. F. Wheatley, Jr., *Television Deflection System Including AFC Circuit with Regenerative Phase Detector*, 3,441,673.
R. O. Winder, *Parity Circuits Employing Threshold Gates*, 3,439,328.
T. A. Zimmerman and T. W. Burrus, *Blanking Circuits for Television Receivers*, 3,436,475.

May

J. J. Amodei, *Thermal Feedback for Stabilization of Differential Amplifier Unbalance*, 3,445,777.
M. Artzt and M. M. Sowiak, *Character Generator*, 3,444,319.
J. Avins, *Automatic Frequency Control Apparatus Especially Suitable for Integrated Circuit Fabrication*, 3,444,477.
H. R. Beelitz, *Differential Amplifier*, 3,445,780.
J. D. Callaghan, *Antenna*, 3,445,854.
B. R. Czorney and E. F. Cave, *PN-Junction Semiconductor with Polycrystalline Layer on One Region*, 3,443,175.
T. J. Desmond, L. S. Greenburg, and H. Weisberg, *High Speed Controlled Rectifiers with Deep Level Dopants*, 3,445,735.
A. G. F. Dingwall and D. K. Wilde, *Thermoelectric Device Having a Graphite Member Between Thermoelement and Refactory Hot Strap*, 3,442,718.
R. A. Dischert and N. L. Hobson, *Coordinated Sensitivity and Amplification Control System*, 3,445,590.
F. C. Easter, *Current Limiting Voltage Regulator*, 3,445,751.
E. C. Fox, *Signal Translating Circuit Providing Signal-Controlled Time Delay*, 3,444,396.
H. J. Gerritsen, *Beat Frequency Holograms*, 3,444,316.
Z. J. Kiss, *Luminescent Materials and Apparatus for Generating Coherent Radiation*, 3,447,097.
A. J. Leidich, *Direct Coupled Amplifier with Feedback for DC Error Correction*, 3,444,476.
A. J. Leidich, *Phase Splitting Circuit for a Direct Coupled Push-Pull Amplifier*, 3,445,776.
S. W. Liddle and C. E. Bush, *Endless-Loop Tape Cartridge*, 3,443,767.
S. W. Liddle and C. G. Hawkins, *Endless Tape Cartridge*, 3,443,768.
A. J. Mortimer, *Switching Circuits*, 3,443,188.
R. G. Olden, *Electrophotographic Method*, 3,442,645.
R. F. Sanford, *Half Wavelength Monopole Antenna with Spaced Loading Coils*, 3,445,849.
R. A. Santilli, *Gain Control Biasing Circuits for Field-Effect Transistors*, 3,443,240.
E. R. Schrader, *Superconducting Ribbon*, 3,443,021.
E. D. Simshauser, *Printer with Print Bars Supported by Parallelogram Linkage Arrangement*, 3,444,975.
R. D. Thompson, *Color Cathode-Ray Tube with Radiation-Emitting Strip-Like Indexing Areas Having Serrated Edges*, 3,443,139.
J. A. Tourtellot, *Phonograph with Circuit Board Mounted on the Motorboard Beneath the Turntable*, 3,444,333.
R. H-Y Yen, *Memory Accessing System*, 3,445,818.

June

R. W. Ahrons, *Cryoelectric Memories*, 3,452,333.
K. W. Angel, *Transmit-Receive Switch*, 3,452,299.
W. H. Barkow and D. Brasen, *Light Deflecting Apparatus*, 3,447,853.
F. Cohen, *Semiconductor Device Having Increased Resistance to Second Breakdown*, 3,448,354.
J. W. Coleman, *Electron Gun Suitable for Electron Microscope*, 3,452,241.
W. F. W. Dietz, *Electron Beam Deflection Circuit*, 3,449,623.
W. F. W. Dietz, *Electron Beam Deflection and High-Voltage Generation Circuit*, 3,452,244.
A. G. F. Dingwall, *Thermoelectric Device With Graphite Elements*, 3,451,858.
J. J. Gibson and J. R. Burns, *Logic Circuits Employing Complementary Pairs of Field-Effect Transistors*, 3,449,594.
H. E. Haynes and K. C. Hudson, *Magneto-Optic Display System*, 3,448,211.
G. H. Heilmeier and L. A. Zanoni, *Method for Preparing a Ferroelectric Body and Devices*, 3,449,824.

W. Henn, *Stepping Switch Employing Chain of Logic Gates Having Means for Locking a Gate in a Given State*, 3,450,897.
K. G. Hernqvist and J. D. Levine, *Method of Operating a Thermionic Converter*, 3,452,224.
N. W. Hursh, *Protection Circuit*, 3,450,935.
M. B. Knight, *Television Scanning and Power Supply System*, 3,450,936.
S. Lynn and W. C. Ries, *Capacitor Anode Holder*, 3,449,650.
H. I. Moss and W. P. Stoller, *Method for Pressing Particulate Material*, 3,448,184.
R. C. Owens, *Television Kinescope Voltage Cable Assembly*, 3,448,323.
S. S. Perlman and R. S. Silver, *Ferroelectric Insulated-Gate Field-Effect Device*, 3,450,966.
D. H. Pritchard, *Cathode-Ray Tube*, 3,452,233.
D. L. Ross, *Luminescent Material and Laser Apparatus Utilizing Said Material*, 3,451,009.
H. D. G. Scheffer, *Wire Handling Apparatus*, 3,448,777.
J. H. Scott, Jr., and J. A. Olmstead, *Remote Cutoff Junction-Gate Field-Effect Transistor*, 3,449,647.
J. A. Vallee, *Conversion from Self-Clocking Code to NRZ Code*, 3,448,445.
J. A. Vallee, *Conversion from Self-Clocking to NRZ Code*, 3,452,348.
J. P. Watson, *Electromagnetic Clutch Having Eccentrically Mounted Poles*, 3,451,516.
C. F. Wheatley, Jr., *Television Deflection Circuits*, 3,449,622.

AUTHORS

ALVIN S. CLORFEINE received his Bachelor's degree in electric engineering from the College of the City of New York in 1959. He obtained his M.S. in electrical engineering the following year from Carnegie Institute of Technology (now Carnegie-Mellon University) and his Ph.D. in E.E. from the same school in 1966. Since joining RCA Laboratories in 1961, he has been engaged in research on semiconducting and superconducting devices and circuits for microwave frequencies. Specific areas of research have included tunnel diodes, superconducting films, and avalanche diodes. His work with superconductors resulted in an "Industrial Research-100" award.

Dr. Clorfeine is a member of the IEEE, Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.

DAVID A. DE WOLF obtained the doctorandus degree in theoretical physics at the University of Amsterdam in 1959, specializing in many-body techniques in quantum statistics. His work since June 1962, at which time he joined RCA Laboratories, has been concerned with electromagnetic wave propagation problems; and a number of contributions to the theory of wave propagation in turbulent air led to the degree of D. Tech. awarded by the Technological University of Eindhoven in 1968. Originally a member of the Plasma And Space Applied Physics affiliated laboratory, he transferred to the Communications Research Laboratory in 1964. He has worked on propagation problems arising in nuclear-disturbed atmospheres, re-entry physics situations, turbulent air at the earth's surface, planetary atmospheres, as well as on radiation problems such as VLF radiation into magnetoplasma, and has contributed publications in each of these areas.

Dr. de Wolf is a member of the Dutch Physical Society, a senior member of IEEE, and an elected member of U.S. National Commission 2 of the International Scientific Radio Union (U.R.S.I.).

ROBERT L. ERNST received the BEE degree with honor from Manhattan College in 1961 and the MSEE degree from the Polytechnic Institute of Brooklyn in 1965. In 1961 Mr. Ernst joined the Western Union Telegraph Company where he was active in the design and development of microwave components. In 1963, he joined the Defense Advanced Communications Laboratory of RCA. Here he has developed many solid-state microwave circuits such as resistive and reactive frequency converters, varactor diode frequency multipliers, transistor oscillator-frequency multipliers, voltage tunable transistor oscillators, and bulk-effect oscillators. He has done extensive theoretical work in the fields of interference reduction techniques and parametric amplifiers. Currently, he is working with microwave transistor amplifiers, both large signal and small signal, and reduced-size microstrip ferrite circulators. Mr. Ernst is a Member of the IEEE, Eta Kappa Nu, and Tau Beta Pi.

JAMES R. FENDLEY, JR. received the B.A. degree from St. Mary's University, San Antonio, Texas in 1952 and the M.S.E. degree from the University of Pennsylvania, Philadelphia, in 1969. He joined RCA in 1955, working in large power tube and thermionic energy converter development at Lancaser, Pa. From 1961 until March 1969 he worked at RCA Laboratories, first on energy conversion and during recent years, on gas lasers. Presently he is with Special Power Devices Engineering in Lancaster, continuing gas laser development work.

ISTVAN GOROG received the B.Sc (1961), M.Sc (1962) and Ph.D. (1964) degrees in Electrical Engineering from the University of California at Berkeley. His dissertation was a study of the near ultraviolet and visible light emission from a "dynamic pinch." Since September of 1964, Dr. Gorog has been a Member of the Technical Staff of the RCA Laboratories, David Sarnoff Research Center, where his main areas of concern have been quantum electronics and electro-optical systems. During 1968 he was on leave of absence from RCA as a National Science Foundation Post-Doctoral Fellow at the Laboratori Nazionali di Frascati and the European Space Research Institute in Frascati, Italy, where he was working on the problems of production of high temperature plasmas by laser irradiation of solid targets and scattering of laser radiation by collective plasma fluctuations.

Dr. Gorog is a member of the American Physical Society and of Eta Kappa Nu.

KARL G. HERNQVIST graduated in Electrical Engineering at the Royal Institute of Technology, Stockholm, Sweden, in 1945. He received the Licentiate of Technology degree in 1951 and the Doctor of Technology degree in 1959. From 1946 to 1952 he was employed by the Research Institute of National Defense in Stockholm, working in the field of microwave electronics. From 1948 to 1949, Dr. Hernqvist was a trainee of the American-Scandinavian Foundation at RCA Laboratories, to which he returned in 1952; he is presently working on gas lasers. Dr. Hernqvist is a Member of the Institute of Electrical and Electronics Engineers and Sigma Xi.

BERNARD HERSHENOV received his B.S. degree in Physics in 1950, his M.S. degree in Mathematics in 1952, and his Ph.D. in Electrical Engineering in 1959, all from the University of Michigan. From 1951-1952 he worked for the University of Michigan Dental Materials Laboratory studying the physical properties of dentin. From 1952-1959 he was employed as a research assistant, research associate, and finally as an associate research engineer with the University of Michigan Research Institute. During this period he worked on domain wall resonance in ferrites, high-power traveling wave tubes, and crossed-field devices. From 1959-1960 he worked on high-power unimoded magnetrons for the G.E. Company. He joined the Microwave Research Laboratory of RCA Laboratories in March 1960 and in January 1968 became head of the Microwave Integrated Circuits Group. Since joining RCA, he has worked on crossed-field amplifiers, ferrite devices, and magnetic semiconductors. During the period 1964-1966 he also served as coadjutant in the Mathematics Department of University College, Rutgers University, at New Brunswick, New Jersey.

Dr. Hershenov is a member of Phi Kappa Phi, American Physical Society, IEEE and Sigma Xi.

RAYMOND J. IKOLA received the B.S.E. and M.S.E. degrees in electrical engineering from the University of Michigan in 1962 and 1963, respectively. In June of 1963 he joined RCA Laboratories. Since joining RCA, he has investigated the use of varactor diodes for millimeter wave applications and studied certain magnetoacoustic effects in microwave ferrites. He has also worked on the avalanche diode as a microwave oscillator and amplifier. The academic year 1965-1966 was spent in full-time graduate study under a David Sarnoff Fellowship, during which time he completed the course requirements for the doctorate in electrophysics at the Polytechnic Institute of Brooklyn.

Mr. Ikola is a member of Eta Kappa Nu, Sigma Xi and the IEEE.

E. G. MCCALL received his B.E.E. degree from Clemson College in 1948 and his M.S.E.E. from the University of Pennsylvania in 1963. He was first employed by RCA in the Broadcast Transmitter Division where he assisted in the development and design of the first UHF television transmitters. During a three-year period from 1952 to 1955 he worked for Allen B. DuMont Labs. and Air Associates, Inc., where he designed UHF transmitters and a Navy communications transmitter. In 1955 he returned to RCA as an engineer and as an engineering leader in the Airborne Systems Division until 1963. His responsibilities included system evaluation and optimization for both airborne radar and television systems. In 1963 Mr. McCall transferred to the Missile and Surface Radar Division as a systems engineer, where he has been assigned to such projects as the Terrier Missile System evaluation, an IR Flame Ranging System, SAM-D, and the ASMS program. Mr. McCall is presently on leave from the Company.

He is a member of Tau Beta Pi, Phi Kappa Phi, and the IEEE.

T. MURAKAMI received the B.S. degree in E.E. from Swarthmore College in 1944, and the M.S. degree from the Moore School of Electrical Engineering, University of Pennsylvania in 1947. From 1944 to 1946 he was an assistant and research associate in the Department of Electrical Engineering at Swarthmore College. From 1946 to 1961 he was associated with the Advanced Development Section of the RCA Home Instruments Division, Camden, N. J., working on radio-frequency circuit development and analyzing various radio and television problems. Since January 1961 he has been with the Advanced Techniques Development group of the RCA Missile and Surface Radar Division at Moorestown, N. J. Mr. Murakami is a Senior member of the Institute of Electrical and Electronics Engineers and a member of Sigma Xi and Sigma Tau.

LOUIS S. NAPOLI received his B.S. in 1959 and the M.S. in 1961 in Electrical Engineering, both from Rutgers University. He has pursued further studies in electrophysics at the Polytechnic Institute of Brooklyn. He joined the technical staff of RCA Laboratories in June, 1959, and has engaged in research in phase-locked oscillators as a microwave logic device. He has specialized in research relating to microwave phenomena in gaseous plasmas, solid-state microwave devices, and is now engaged in research on solid-state microwave integrated circuits. His work in collaboration with Dr. George Swartz on amplification at 24 Gc by the interaction of an electron beam with a cesium plasma was cited by Industrial Research Magazine as one of the 100 most important achievements in 1963.

Mr. Napoli is a member of Sigma Xi, the Institute of Electrical and Electronics Engineers, Tau Beta Pi, and Eta Kappa Nu.

ALBERT ROSE received the A.B. degree from Cornell University in 1931 and the Ph.D. degree in Physics in 1935. From 1931 to 1934 he was a teaching assistant at Cornell University. From 1935-1942 he was with the RCA Tube Division in Harrison, N. J., and since 1942 he has been associated with RCA Laboratories at Princeton, N. J. From 1955 to 1957 he directed the research at Laboratories RCA, Ltd., in Zurich, Switzerland. Dr. Rose is a Fellow of the Institute of Electrical and Electronics Engineers and of the American Physical Society.

C. SUN received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan in 1958 and the M.S. degree and Ph.D. degree in electrical engineering from Cornell University in 1962 and 1965, respectively. During 1958-1960, he was a ground radar maintenance officer in the Chinese Air Force. Since September, 1964, he has been with the RCA Microwave Applied Research Laboratories in Princeton. He has worked on low-noise traveling-wave tubes, development of varactors, wide-band and high-power varactor multipliers, and optical demodulators. Dr. Sun is a member of Sigma Xi.

GILBERT S. WICKIZER received the B.S. degree in electrical engineering from the Pennsylvania State University in 1926, and after graduation was employed by RCA in the transoceanic communication facilities at Rocky Point and Riverhead, New York. From 1927 until 1961, he was with RCA at Riverhead, N. Y., engaged in research and development on a wide range of subjects in the general field of radio reception. Receiving equipment and field-strength measuring techniques were developed for propagation studies as the useful spectrum of frequencies expanded progressively through the HF, VHF, and UHF bands. In 1961, Mr. Wickizer transferred to the Missile and Surface Radar Division of RCA at Moorestown, N. J., where he pursued propagation studies in connection with radar operation until his retirement in 1968.

He is a member of Eta Kappa Nu and a Fellow of the IEEE.