RCA REVIEW

THE ROLE OF SURFACE STATES IN THE PHOTOELECTRONIC PROPERTIES OF INSULATING CADMIUM SULFIDE CRYSTALS

By

Peter Mark

RCA Laboratories
Princeton, N. J.

Summary—*The manner in which certain surface states affect the photoelectronic properties of insulating CdS crystals is reviewed. Two types of surface states are considered—chemisorbed ions, which can be called extrinsic surface states because they are a class of impurity, and intrinsic surface states, which arise solely because of the termination of the lattice by a surface. The model for photo-induced chemisorption on an insulator surface is presented and the chemisorption kinetics are derived. Supporting experimental results for the photo-induced chemisorption of oxygen and carbon monoxide on insulating CdS crystals are reviewed. Next, the experiments are reviewed that demonstrate that the intrinsic surface states function as traps for photogenerated carriers. From this and other experimental evidence, the phenomenological model for the intrinsic surface states is constructed. The paper concludes with a discussion of the physical origin of the intrinsic surface states. The theory for intrinsic surface states of ionic crystals, due to J. D. Levine, is reviewed; according to this theory, the surface states follow directly from the fact that the Madelung energy at a surface lattice ion site is always less than the Madelung energy for the same site deep in the bulk. It is shown that the properties of the intrinsic surface states derived from physical arguments coincide with those of the intrinsic surface states deduced from the experiments.*

I. Introduction

SURFACE states are those electronic levels of a bounded crystalline solid that are associated with the presence of a surface that terminates the periodic lattice. They can be divided into two fundamental categories—extrinsic or intrinsic—depending on their origin.

Extrinsic surface states arise from the presence on the surface of any sort of "impurity," such as adsorbed ions, oxide, or other layers of chemical origin, or because the surface has been mechanically damaged. Since matter generally interacts through the agency of a surface, it is not surprising that extrinsic surface states, as defined, play a major role in the chemical and physical properties of matter, some of which are reviewed below.

Intrinsic surface states exist solely because the surface interrupts the periodic lattice, and they are present even on an uncontaminated

surface. Intrinsic surface states derive from the properties of a periodic lattice, and consequently differ according to the type of interaction (binding) giving rise to the cohesive energy of the lattice. Tamm surface states,[1] as interpreted by Fowler[2] and Goodwin,[3] are associated with metallic binding, since they are computed from a plane-wave approximation with a periodic potential. Shockley,[4] Goodwin,[5] and several other workers[6,7] have considered the surface states arising from a tight-binding approximation with the lattice viewed as a collection of weakly coupled atoms. This approach is appropriate for covalent binding. A model for intrinsic surface states of ionic crystals has recently been proposed[8,9] that considers the long-range electrostatic interactions of the lattice ions via the coulombic Madelung potential as the dominant contribution to the cohesive energy. Although these three types of surface states differ significantly in their details, they have one underlying feature in common that they also share with the extrinsic surface states; all can give rise to localized electronic states that are associated with the surface.

The purpose of this paper is to review the experimental evidence for the observation of a particular extrinsic surface state—that due to chemisorbed ions—and of intrinsic surface states on insulating crystals of the quasi-ionic II-VI compound CdS,[10–13] and also to review the physical models for these two types of surface states.[8,9] Both the extrinsic and the intrinsic surface states could be detected on the same crystal by studying the behavior of the photocurrent caused by greater-

[1] I. Tamm, "Über Eine Mögliche Art der Elektronenbindung an Kristalloberflächen," *Physik. Zeitschr. Sowjetunion*, Vol. 1, p. 733, 1932.

[2] R. H. Fowler, "Notes on Some Electronic Properties of Conductors and Insulators," *Proc. Roy. Soc.*, Vol. A141, p. 56, July 1933.

[3] E. T. Goodwin, "Electronic States at the Surfaces of Crystals. I. The Approximation of Nearly Free Electrons," *Proc. Cam. Phil. Soc.*, Vol. 35, p. 205, April 1939.

[4] W. Shockley, "On the Surface States Associated with a Periodic Potential," *Phys. Rev.*, Vol. 56, p. 317, Aug. 15, 1939.

[5] E. T. Goodwin, "Electronic States at the Surfaces of Crystals. II. The Approximation of Tight Binding: Finite Linear Chain of Atoms," *Proc. Cam. Phil. Soc.*, Vol. 35, p. 221, April 1939; "Electronic States at the Surfaces of Crystals. III. The Approximation of Tight Binding: Further Extensions," *Proc. Cam. Phil. Soc.*, Vol. 35, p. 232, April 1939.

[6] J. Bardeen, "Surface States and Rectification at a Metal Semiconductor Contact," *Phys. Rev.*, Vol. 71, p. 717, May 15, 1947.

[7] *Solid Surfaces*, ed. by H. C. Gatos, North-Holland Publishing Co., Amsterdam, 1964.

[8] J. D. Levine and P. Mark, "Theory of Intrinsic Surface States on Ionic Crystals," and "Observation of Intrinsic Surface States of Insulating CdS Crystals," papers BG-13 and BG-14, *Bull. Amer. Phys. Soc.*, Vol. 10, p. 1093 (1965).

[9] J. D. Levine and P. Mark, "Theory and Observation of Intrinsic Surface States on Ionic Crystals," *Phys. Rev.*, Mar. 11, 1966 (to be published).

[10] P. Mark, "Photo-Induced Chemisorption on Insulating CdS Crystals," *Jour. Phys. Chem. Solids*, Vol. 25, p. 911, Sept. 1964.

than-band-gap light and flowing parallel to the large-area surface of platelet crystals with this surface exposed to various ambients.

The extrinsic surface states (the chemisorbed ions) can act as recombination centers[11,12] for photocarriers from the bulk, and can be observed whenever the surface is exposed to a chemically inert acceptor adsorbate while the surface is illuminated with greater-than-band-gap light. The model for such photo-induced chemisorption on insulators is developed in Section II. Section III deals with a description of the CdS crystals and experimental measuring techniques.[10-12] The experimental results for oxygen and carbon monoxide chemisorbed on insulating CdS crystals, which illustrate this model, are presented in Section IV.

The intrinsic surface states were detected when the surface was freed of chemisorbed ions. These states function as traps for the photocarriers (electrons and holes) from the bulk.[9,11] The experimental evidence for the existence of intrinsic surface states, together with a phenomenological model based on this evidence, is reviewed in Section V. Section VI is devoted to a discussion of the interaction between the intrinsic surface states and an adsorbate as it affects the chemisorption mechanism. Finally, the underlying physics of intrinsic surface states on ionic crystals[9] is briefly reviewed in Section VII and it is shown that the physical model and the phenomenological model have enough features in common that the latter can be interpreted self-consistently in terms of the former.

All the equations are written in practical units.[14]

II. Models for Chemisorption

The process of chemisorption is the formation of ions on a surface by the exchange of electrons between physically adsorbed species and the adjacent bulk. A particularly significant manifestation of this process occurs when the adsorbate acts as an acceptor for electrons from the conduction band of a wide-band-gap n-type semiconductor. Then, the surface charge of the ions is compensated by the immobile space charge of the ionized donors in a Schottky-type depletion layer. [The analogous situation involving holes from the valence band will

---

[11] P. Mark, "The Role of Chemisorption in Current Flow in Insulating CdS Crystals," *Jour. Phys. Chem. Solids*, Vol. 26, p. 959, June 1965.

[12] P. Mark, "Chemisorption and Trapping on Insulator Surfaces," *Trans. N. Y. Acad. Sci.*, Vol. 27, p. 946, 1965.

[13] P. Mark, "The Energy of Desorption of Photochemisorbed Oxygen and Nitrous Oxide on Insulating CdS Crystals," *Jour. Phys. Chem. Solids*, Vol. 26, p. 1767 (1965).

[14] W. Shockley, *Electrons and Holes in Semiconductors*, pp. 211-213, D. van Nostrand and Co., Princeton, N. J., 1950.

not be discussed.] The importance of this depletion-type chemisorption (hereafter referred to simply as "chemisorption") in many physical and chemical processes involving wide-band-gap semiconductor surfaces has been recognized for some time. Among the physical processes is the photoconductivity[15] of granular ZnO as utilized in the electrophotographic "Electrofax" process,[16] while the most important chemical consequence of chemisorbed ions is heterogeneous catalysis; many metal oxides that are extrinsic wide-band-gap (> 2 eV) semiconductors act as heterogeneous catalysts.[17]

It will be shown that chemisorption on semiconductor surfaces is spontaneous; the ions are formed when the surface is exposed to an adsorbate. On the other hand, chemisorption on insulator surfaces is not spontaneous; it is possible to control the chemisorption process on insulator surfaces by physical means, such as illumination, that offer certain advantages for examining the chemisorption mechanism and may also have useful practical application. The model for photoinduced chemisorption on insulators, that is, depletion-type chemisorption occurring only during illumination with greater-than-band-gap light, is an extension of the model for spontaneous chemisorption on a wide-band-gap semiconductor.[15] For clarity, the latter is first briefly reviewed.

Model for Spontaneous Chemisorption on Wide-Band-Gap Semiconductors

The process by which chemisorbed ions are formed on the surface of an n-type, extrinsic semiconductor by the capture of electrons from the bulk by an acceptor adsorbate is illustrated in Figure 1. Each picture is a plot of a portion of the electronic energy level structure versus the distance from the geometric surface, and shows the upper and lower portions, respectively, of the allowed occupation levels forming the valence and conduction bands, as well as the energy gap between them, which is devoid of occupation states when the lattice is perfectly periodic. The discrete occupation level, E_d, introduced into the gap by donors, and the free energy (i.e., the chemical potential) of the electrons, commonly called the Fermi energy or level E_f, are also indi-

---

[15] D. A. Melnick, "Zinc Oxide Photoconduction, an Oxygen Adsorption Process," *Jour. Chem. Phys.*, Vol. 26, p. 1136, 1957.

[16] J. A. Amick, "A Review of 'Electrofax' Behavior," *RCA Review*, Vol. 20, p. 753, Dec. 1959.

[17] G. Parravano and M. Boudart, "Chemisorption and Catalysis on Oxide Semiconductors," *Advanc. Catalys.*, Vol. 7, p. 213, 1955; F. S. Stone, "Chemisorption and Catalysis on Metallic Oxides," *Advanc. Catalys.*, Vol. 13, p. 1, 1962.

cated. An example would be ZnO, for which $\Delta E_g = 3.2$ eV and $\Delta E_d = 0.04$ eV. The donors are excess interstitial Zn^+ ions.[18]

The chemisorbate-free surface is illustrated in Figure 1(a). The bands are drawn flat to the surface, signifying that there is no net charge on the surface with respect to the adjacent volume. On exposure

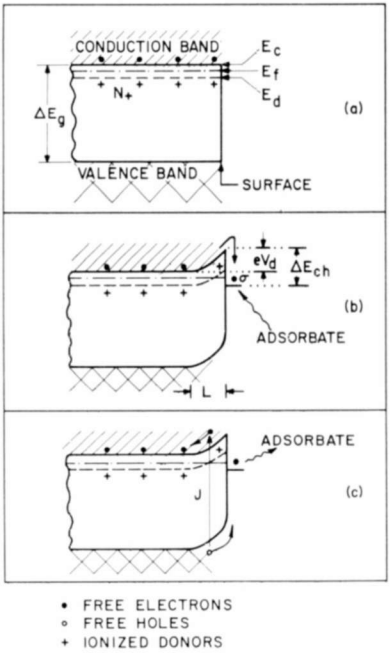


Fig. 1—Model for spontaneous chemisorption on the surface of a wide-band-gap extrinsic semiconductor. (a) Energy-band configuration of an adsorbate-free surface; E_c is the bottom of the conduction band, E_f is the equilibrium Fermi level, E_d is the donor level, ΔE_g is the band gap, N_+ is the ionized donor concentration. (b) Surface after chemisorption showing the Schottky-type surface barrier; V_d is the diffusion potential, ΔE_{ch} is the energy of desorption, L is the thickness of the depletion layer, N_+ is again the ionized donor concentration in the depletion layer, σ is the concentration of chemisorbed ions. (c) Photodesorption when $t_p < \tau_p$; J represents the optical generation of free electron–hole pairs.

to an electronegative (acceptor) adsorbate (such as oxygen), some of the electrons near the surface are spontaneously trapped by physically adsorbed oxygen to form negative ions on the surface. These ions are electrostatically compensated by the positive space charge of the immobile ionized donors distributed in the adjacent bulk and a Schottky-type surface barrier is formed. This situation is depicted in Figure 1(b).

[18] S. R. Morrison, "Surface Barrier Effects in Adsorption, Illustrated by Zinc Oxide," *Advanc. Catalys.*, Vol. 7, p. 259, 1955.

The rate-limiting step during chemisorption is considered to be the diffusion of the electrons through the Schottky barrier field.[15,18] Chemisorption proceeds until the thermal release rate of the electrons from the surface ions back to the conduction band just balance their arrival rate. However, there is a maximum concentration σ_m of adsorbed ions that are electrically compensated in an adjacent depletion layer;[19]

$$\sigma_m = \left(\frac{2K\epsilon_0 N_+ V_d}{e} \right)^{1/2} \tag{1}$$

where K is the dielectric constant, ϵ_0 is the electric permittivity of free space (8.85×10^{-14} farad/cm), N_+ is the concentration (cm^{-3}) of ionized donors in the depletion layer, V_d is the barrier diffusion potential, and e is the magnitude of the electronic charge (1.6×10^{-19} coulomb). One finds that for some extreme values, $K = 10$, $V_d = 1$ volt and $N_+ = 10^{19}$ cm^{-3}, the upper limit to σ_m is 10^{13} cm^{-3}, which corresponds to only 10^{-2} of a monolayer.[19] Usually, for less extreme values for N_+ and V_d, σ_m will be much smaller.

The adsorbed ions can be desorbed either by heating the system, which enhances the thermal release of electrons from the ions, or by the production of holes (minority carriers) with the aid of light absorbed in the valence band.[15,20] The latter case, known as photodesorption, is illustrated in Figure 1(c). The photoholes drift to the surface in the barrier field where they neutralize the adsorbed ions, and the photoelectrons in the conduction band compensate the positive space charge of the ionized donors. Thus, photodesorption tends to restore the initial flat band configuration. This mechanism has accounted for the surface photovoltage[21] and the photoconductivity[15] of ZnO and for the operation of the Electrofax process.[16]

Photodesorption requires that an appreciable fraction of the photoholes reach the surface within the interval of the hole (minority carrier) recombination lifetime τ_p. The holes drift to the surface with the transit time

---

[19] P. B. Weisz, "Electronic Barrier Layer Phenomena in Chemisorption and Catalysis," *Jour. Chem. Phys.*, Vol. 20, p. 1483, Sept. 1952; "Effects of Electronic Charge Transfer between Adsorbate and Solid on Chemisorption and Catalysis," *Jour. Chem. Phys.*, Vol. 21, p. 1531, Sept. 1953.

[20] D. B. Medved, "Photoconductivity and Chemisorption Kinetics in Sintered Zinc Oxide Semiconductor," *Jour. Phys. Chem. Solids*, Vol. 20, p. 255, Nos. 3/4, 1961.

[21] R. J. Collins and D. G. Thomas, "Photoconduction and Surface Effects with Zinc Oxide Crystals," *Phys. Rev.*, Vol. 112, p. 388, Oct. 15, 1958.

$$t_p = \frac{L^2}{\mu_p V_d}, \tag{2}$$

where L is the barrier thickness and μ_p is the hole mobility. Thus, for photodesorption to occur,[9]

$$t_p < \tau_p. \tag{3}$$

This inequality can be formulated as a condition on N_+. The relation between L, V_d, and N_+ in a Schottky-type barrier is given by[22]

$$L^2 = \frac{2K\epsilon_0 V_d}{eN_+}. \tag{4}$$

Combining Equations (2) and (4) yields

$$t_p = \frac{2K\epsilon_0}{e\mu_p N_+}, \tag{5}$$

which shows that the transit time is inversely proportional to N_+. Equation (5) and inequality (3) give a relation among material parameters that must be satisfied for photodesorption to occur.[9] This relation plays a very important role in the insulator problem.

Model for Photo-Induced Chemisorption on Insulators

Insulators differ from wide-band-gap semiconductors in that the thermal equilibrium carrier concentration n in the conduction band is many decades smaller. Thus, the equilibrium Fermi level, which is a measure of n, is further removed from the conduction-band edge. This situation is depicted in Figure 2(a). Chemisorption does not occur spontaneously because of the paucity of free electrons. However, if a sufficient number of electrons can be photogenerated near the surface by light of energy greater than the band gap, photo-induced chemisorption ought to occur. The mechanism is illustrated in Figure 2(b), (c), and (d).

The change in the electronic occupation of the bands on illumination in the fundamental absorption band is shown in Figure 2(b). Electrons are excited across the band gap, leaving holes of concentration p in the valence band. For an n-type photoconductor, $n \gg p$, since

---

[22] For example, see A. van der Ziel, *Solid State Physical Electronics*, pp. 241-253, Prentice-Hall, Englewood Cliffs, N. J., 1957.

the recombination lifetime (τ_n) of the electrons in the conduction band is much greater than the recombination lifetime of the holes (τ_p) in the valence band.[23] Preservation of charge neutrality requires that a large fraction of the holes become quickly localized in the recombination centers R through which recombination with electrons in the

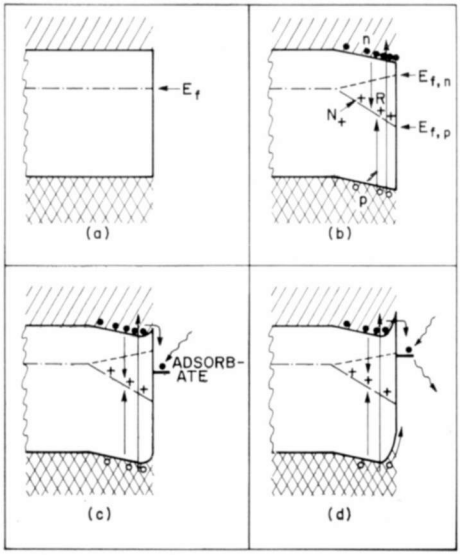


Fig. 2—Model for photo-induced chemisorption on an insulator surface. (a) Unilluminated adsorbate-free surface. (b) Adsorbate-free surface under illumination with light of greater-than-band-gap energy. $E_{f,n}$ and $E_{f,p}$ are the steady-state quasi-Fermi levels for electrons and holes, respectively. The bulk recombination transitions are indicated by the downward and upward arrows between the conduction band and the valence band, respectively, and the recombination centers R. N_+ is the concentration of photoholes in the recombination centers. (c) Initial response on admission of an adsorbate showing the appearance of a surface barrier. (d) Steady state when $t_p < \tau_p$; surface coverage of chemisorbed ions is limited by minority carrier desorption.

conduction band occurs. Although this is only a steady-state situation, one can assign a quasi-Fermi level to n and to p as if these were equilibrium concentrations.[23] These are also shown in Figure 2(b).

Concentration gradients in n and p are created by the nonuniform optical absorption, and the free carriers diffuse away from the surface. The diffusion is considered to be minority-carrier controlled;[24] the

[23] A. Rose, "Recombination Processes in Insulators and Semiconductors," *Phys. Rev.*, Vol. 97, p. 322, Jan. 15, 1955; *Concepts in Photoconductivity and Allied Problems*, Chap. 3, John Wiley and Sons, New York, 1963.

[24] P. Mark, "Ambipolar Diffusion of Free Carriers in Insulating CdS Crystals," *Phys. Rev.*, Vol. 137, p. A203, 1965.

electrons diffuse further per unit time than the holes, and this net charge separation produces an electrostatic field—the Dember field[25]—that reduces the diffusion of the electrons to that of the holes. The energy bands have a constant slope between the illuminated surface and the unexcited bulk because the Dember field is constant; its spatial extent is several ambipolar diffusion lengths.[24] A detailed description of this energy-band configuration is given in the Appendix.

The illuminated insulator surface bears a strong resemblance to the semiconductor surface in thermal equilibrium. A conducting channel has been prepared near the surface, and one expects the carriers in the channel to participate in chemisorption. The main differences are that the insulator surface region is bipolar (holes as well as electrons are present) and that the free carrier concentrations can be varied over several decades by changing the illumination level. On exposing the illuminated insulator surface to an electronegative adsorbate, the initial response, shown in Figure 2(c), is the formation of negative ions on the surface in analogy to the semiconductor case. A Schottky-type barrier is formed by the electrostatic field between the negative ions on the surface and a compensating portion of the photogenerated holes in the recombination centers in the adjacent bulk.

An ultimate steady-state coverage of chemisorbed ions is reached when the chemisorption transitions are balanced either by thermal release of electrons from the surface ions back to the conduction band or by photodesorption. Which of these desorption mechanisms set the steady-state coverage of adsorbed ions depends on whether or not the inequality (3) is satisfied. If this condition is not satisfied ($t_p > \tau_p$), thermal desorption limits the coverage to the values of σ_m given by Equation (1), where N_+ is now the concentration of photoholes localized in the bulk recombination centers and can be proportional to the illumination level.[23] Since the electrons of the surface ions remain coupled to the conduction band, the chemisorbed ions act as traps for the photoelectrons and thus ought not to influence the magnitude of the steady-state photocurrent flowing parallel to the surface in question.[23] However, if the inequality (3) is satisfied ($t_p < \tau_p$), an appreciable fraction of the holes reach the surface and steady-state coverage of chemisorbed ions is set by photodesorption to a value that is always less than that given by Equation (1). Thus, a steady state is envisioned, illustrated in Figure 2(d), involving the simultaneous flow

---

[25] H. Dember, "Über Eine Photoelektromotorische Kraft in Kuperoxydul-Kristallen," *Physik. Zeitschr.*, Vol. 32, p. 554, 15 July, 1931. See also T. S. Moss, *Optical Properties of Semiconductors*, Chap. 4, pp. 61-63, Butterworth's, London, 1961.

of charge (electron and holes) to the surface and of adsorbate gas to and from the surface. If the adsorbate gas is removed (either by pumping or by displacement with an inert gas), the holes arriving at the surface will desorb the adsorbate and tend to restore the original adsorbate-free surface.

When $t_p < \tau_p$, the adsorbate ions offer an additional recombination path for the photogenerated carriers, and the steady-state photocurrent flowing parallel to the surface should be reduced.[10-12] Further predictions can be made. The process should be reversible, even at temperatures where the thermal release rate of electrons from the surface ions is negligible, provided that no chemical changes have occurred. Only molecules (or atoms) with an appreciable electron affinity should participate in chemisorption on an n-type photoconductor surface. No changes in the adsorbed ion concentration should occur in the dark at temperatures where the thermal desorption rate is negligible.[10,13]

III. Material Specifications and Experimental Procedure

To illustrate the model for photo-induced chemisorption on insulators, electrical measurements were made with the representative insulator cadmium sulfide (CdS) which has a band gap[26] of 2.4 eV. This material was chosen for several reasons. First, despite the fact that CdS crystallizes in the wurtzite structure,[9] indicating a considerable covalency in the binding, it is nevertheless legitimate to treat this II-VI compound as essentially ionic, and to account for the covalent contribution to the binding by considering fractional electronic charges at the sites of the lattice ions.[9,27] The essentially ionic nature of CdS will be invoked again in Section VII, where the origin of the intrinsic surface states is discussed. Second, techniques for preparing large single-crystal specimens with predictable electrical and optical properties are well established.* In this investigation, the crystals were grown from the vapor phase and were in the form of platelets from 20 to 50 microns thick, with surface areas ranging from 0.05 to 0.2 cm.^{2} The crystals grow naturally with their large-area surfaces parallel to the $(11\bar{2}0)$ plane, which is normal to the crystallographic a-axis and parallel to the optic c-axis.[9] The room-temperature dark resistivity of the crystals exceeded 10^{10} ohm cm. A third reason for using CdS is that there exists a vast literature that contains much of the peripheral

---

[26] R. H. Bube, *Photoconductivity of Solids*, Chap. 7, p. 233, John Wiley & Sons, New York, 1960.

[27] J. L. Birman, "Electronic Energy Bands in ZnS: Potential in Zinc-blende and Würtzite," *Phys. Rev.*, Vol. 109, p. 810, Feb. 1, 1958.

* See Reference (26), Chap. 4.

information necessary for interpreting the present results—information such as the photoconductivity excitation spectrum, energy transport mechanisms, and carrier lifetimes and mobilities for a great variety of impurity and defect structure conditions.† Finally, electroding is no problem; barrier-free (ohmic) contacts can be prepared with indium or indium–gallium–tin alloys.[28]

The measuring technique is simple. The sample is mounted in a chamber either in vacuum or in a selected ambient. The chamber has a window through which the crystal can be illuminated and has provisions for making electrical contacts and for controlling the temperature of the crystal. The current parallel to the exposed surface is measured as a function of various parameters such as temperature, ambient surrounding the crystal, applied voltage, and illumination (intensity and wavelength). The measurements were performed with those CdS crystals that exhibited the following two photoconductivity properties: (1) photoconductivity only under optical excitation in the fundamental absorption band and (2) low photosensitivity. The first condition assures that the width of the conducting channel parallel to the illuminated surface is set by free-carrier diffusion, which makes possible the direct determination of this width as well as the lifetime, τ_p, of the holes.[24] The second condition means selection of crystals with the smallest electron lifetime, τ_n. According to Rose's model for sensitization,[23] photoconductors with small τ_n values tend to have correspondingly large τ_p values, a condition that is advantageous for observing photodesorption. Most measurements were made with the natural surfaces of the crystals. Chemical etching had no consistent effect on the results. The only treatment prior to each run was the removal of physically adsorbed moisture by heating to 150°C in a flowing dry atmosphere.

IV. Detection of Extrinsic Surface States—Chemisorbed Ions

Experimental Observations

The photocurrent under constant illumination with light of greater-than-band-gap energy ($\lambda_{peak} = 0.4\ \mu$, bandwidth $= 0.1\ \mu$, photon flux density $= 1.4 \times 10^{15}$ cm^{-2} sec^{-1}) depends on the ambient to which the crystal is exposed. The effect is illustrated in Figure 3. Curve (a)

† See bibliography of Reference (26).

[28] R. W. Smith, "Properties of Ohmic Contacts to Cadmium Sulfide Single Crystals," *Phys. Rev.*, Vol. 97, p. 1525, March 15, 1955.

shows the response of the current at 20°C when dry nitrogen, flowing at 1 atm is used to displace an original ambient of dry oxygen, also flowing at 1 atm. The nitrogen is admitted at $t = 0$. The current rises by about a factor of six. An identical result is obtained on evacuation, or on displacement with other inert gases (e.g., the rare gases); only the partial pressure of the adsorbate gas is important. If the flow of nitrogen is interrupted before the new steady state is reached, the

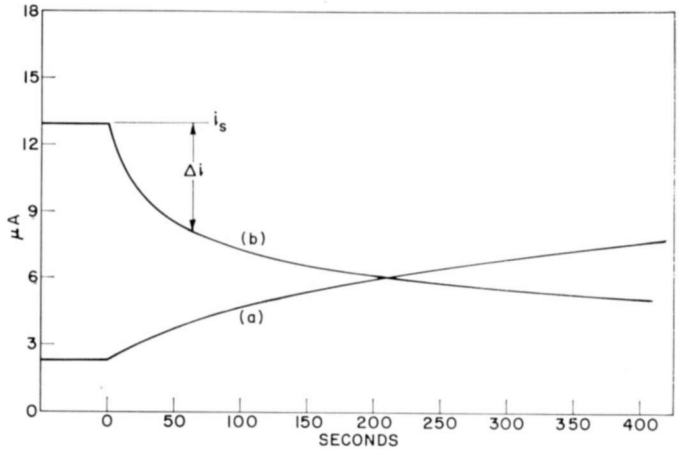


Fig. 3—Reversible photo-induced chemisorption of oxygen on an insulating and insensitive CdS crystal at 18°C. (a) Change in the photocurrent flowing parallel to the illuminated surface on photodesorption. (b) Change in the photocurrent on chemisorption [from *Trans. N. Y. Acad. Sci.*, Vol. 27, p. 946 (1965)].

current increase is arrested until the flow of nitrogen is resumed. After the new and larger steady-state current is reached and the crystal is again exposed to oxygen, the photocurrent decreases to its original value, as shown by curve (b). A photo-Hall-effect measurement verified that the change in the current is caused principally by a change in n and not in the electron mobility.[10] This reversible behavior is predicted by the model for the case where $t_p < \tau_p$. Similar results are obtained with oxygen up to 170°C, with nitrous oxide over the temperature interval 20°C to 170°C, and with iodine vapor and carbon monoxide at room temperature. Measurements with carbon monoxide at elevated temperatures reveal significant departures, and are discussed at the end of this section.

Chemisorption Kinetics

Curve (b) in Figure 3 represents the change in the photocurrent under steady illumination during the uptake of oxygen on the surface

by chemisorption. It is possible to express this behavior analytically by the following formalism, which is an adaptation of Melnick's theory[15] for spontaneous chemisorption on semiconductors. The rate-limiting step for chemisorption is considered to be the diffusion of electrons to the surface through the surface barrier.[11,15,18,20] Thus, the net rate of accumulation of chemisorbed ions is

$$\frac{d\sigma}{dt} = \sigma_0 c_e v n \exp\left[\frac{-eV_d\sigma}{kT}\right] - \gamma\sigma. \tag{6}$$

Here σ is the concentration of chemisorbed ions, σ_0 is the concentration of physically adsorbed (neutral) species, c_e is the capture cross section of the physisorbed species for an electron from the bulk, v is the thermal velocity of the electrons in the conduction band, and n is the concentration of conduction electrons. The term $\gamma\sigma$ represents photodesorption, so that $\gamma = c_p v_p p$, where c_p is the capture cross section of the chemisorbed ions for free holes, v_p is the thermal velocity of the holes in the valence band, and p is the free-hole concentration in the valence band. The diffusion potential V_d depends on σ through Equation (1):

$$V_d = \frac{e\sigma^2}{2K\epsilon_0 N_+}. \tag{7}$$

The solution of this problem is an expression for the current as a function of time during chemisorption. This dependence is obtained with several assumptions, the first of which is that the desorption term can be neglected during the initial stages of uptake. The consequent solution is valid over only the initial uptake interval. This assumption is physically reasonable during the initial stages of uptake since, for CdS, $c_e v n \sigma_0 \exp[-eV_d/(kT)] \gg c_p v_p p\sigma$ because V_d is small, $\sigma_0 \gg \sigma$, $n \gg p$, and $v > v_p$ (it is assumed that $c_e \simeq c_p$). Thus, the differential equation reduces to

$$\frac{d\sigma}{dt} = an \exp[-b(N_+)\sigma^2], \tag{8}$$

where $a = \sigma_0 c_e v$ and

$$b = \frac{e}{2K\epsilon_0 N_+ kT}. \tag{9}$$

In order to convert Equation (8) into an equation for the current, a second assumption is made—trapping of electrons in the bulk can be neglected, so that every chemisorbed ion represents a photoelectron removed from the conduction band. Thus

$$V|\Delta n| = S|\Delta\sigma|, \tag{10}$$

where V and S are, respectively, the volume of the photoconducting channel and the area exposed to the adsorbate. The incremental quantities are

$$\Delta n = n_s - n(t), \tag{11}$$

and

$$\Delta\sigma = \sigma(t) - \sigma_s, \tag{12}$$

where n_s and σ_s are the values of n and σ prior to chemisorption. Equations (10), (11), and (12) yield the following relation between n and σ:

$$\sigma(t) = \sigma_m - \delta n(t), \tag{13}$$

where $\delta = V/S$ and $\sigma_m = \sigma_s + \delta n_s$. The final connection between σ and the photocurrent, i, is provided by the relation

$$i = F\mu_n e\mathcal{E}n \tag{14}$$

where F is the area of the conducting channel normal to current flow, μ_n is the electron mobility (assumed constant) and $\mathcal{E}$ is the applied field driving the current.

The foregoing equations can be combined to give the following integral, which expresses the time dependence of the current during chemisorption;

$$\int_{i}^{i_s} \{\exp[b\alpha i(\alpha i - 2\sigma_m) + b\sigma_m{}^2]\} \frac{di}{i} = \left(\frac{a}{\alpha}\right) t \tag{15}$$

where $\alpha = \delta/(F\mu_n e\mathcal{E})$ and i_s is the photocurrent prior to chemisorption. Although this integral cannot be expressed analytically, it can be simplified by noting that i is a decreasing function of time. Then, if $\alpha i_s < 2\sigma_m$, the expression $-2b\sigma_m\alpha i$ will be a good approximation for

the first term in the exponential and will be increasingly better as i decreases with increasing time. Now $\alpha i_s = \delta n_s$ so that $\alpha i_s - 2\sigma_m = -(\delta n_s + 2\sigma_s)$. Thus, in order for $\alpha i_s < 2\sigma_m$, $\sigma_s \gg \delta n_s$ or $\sigma_m \simeq \sigma_s$. If it is also assumed that N_+ (and hence b) does not vary during chemisorption, Equation (15) leads to the following relation for the photocurrent under steady illumination during chemisorption of an adsorbate:

$$\Delta i(t) = i_s - i(t) = \lambda \ln \left[\frac{t}{t_0} + 1 \right] \tag{16}$$

when terms in $(\Delta i/i_s)^2$ are small. This is the well known Elovich equation,[29] familiar from other adsorption experiments.[30] The parameters λ and t_0 are given by

$$\lambda = \frac{N_+ K \epsilon_0 k T \mu \mathcal{E}}{\sigma_m \delta}, \tag{17}$$

and

$$t_0 = \left[\frac{a}{2b\delta\alpha i_s \sigma_m} \right] \exp\,(b\sigma_m)^2. \tag{18}$$

If $N_+ \propto J$, the photon flux density (photons/sec cm^2) of the illumination, $\lambda \propto J\mathcal{E}$ and $t_0 \propto \exp\,(\text{constant}/J)$. A plot of Equation (16) as Δi versus log t should show positive curvature when $t < t_0$, while as t increases beyond t_0, Δi should approach asymptotically a straight line of slope λ intersecting the time axis at t_0. Equation (16) does not hold for all t. As soon as minority carrier desorption becomes the rate-limiting step, a negative curvature to the new steady-state current should appear.

A test of the model is provided by plotting the uptake curve (b) of Figure 3 as Δi versus log t. This is shown by the curve labeled "100%" in Figure 4(a). The open circles represent the experimental data, and the straight line approximated by the solid dots is obtained by an empirical selection of a value for t_0. The curve labeled "40%" was obtained by repeating the uptake experiment with the illumination reduced to 40% of maximum. Similarly, the two curves of Figure 4(b)

---

[29] S. Yu. Elovich and J. Zhabrois, *Gaz. Fiz. Khim. SSSR*, Vol. 13, p. 1761, 1939.

[30] P. T. Landsberg, "On the Logarithmic Rate Law in Chemisorption and Oxidation," *Jour. Chem. Phys.*, Vol. 23, p. 1079, June 1955; "A Note on Logarithmic Rate Laws," *Jour. Appl. Phys.*, Vol. 33, p. 2251, July 1962.

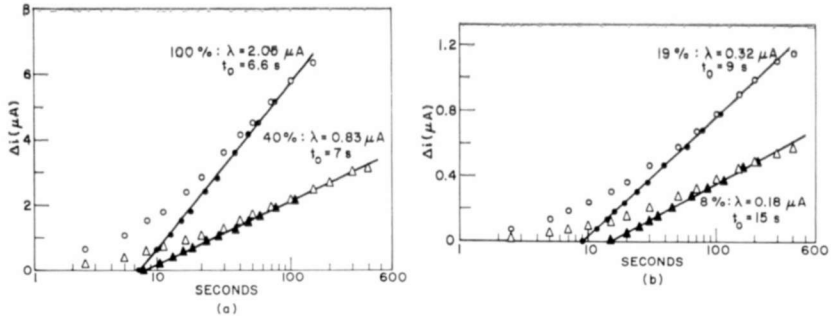


Fig. 4—Elovich-type plot [Equation (16)] of the change in the photocurrent on chemisorption for four illumination levels. (a) $J = 100\%$, 40% (b) $J = 19\%$, 8%. The values of λ, corrected for the Naperian base, and t_0 corresponding to each intensity are listed next to the curves. Maximum illumination flux density: 1.4×10^{15} photons/cm^2 sec. [from *Trans. N. Y. Acad. Sci.*, Vol. 27, p. 946 (1965)].

show the uptake curves at 19% and 8% of maximum illumination. In each case, a straight line of slope λ can be constructed from the experimental points by an empirical choice of t_0. The numerical values of λ (corrected to the Naperian base) and t_0 are given for each curve. Each curve is terminated just after it departs from linearity toward the new steady state.

Both λ and t_0 depend on J according to the condition $N_+ \propto J$. This is shown by the curves of Figure 5. In Figure 5(a), λ is plotted against J on a logarithmic field and yields a straight line of unit slope, and Figure 5(b) shows that a plot of log t_0 versus $(1/J)$ gives a straight line. The proportionality between λ and $\mathcal{E}$ was established by repeating the uptake experiment at different voltages but at the same illumination.

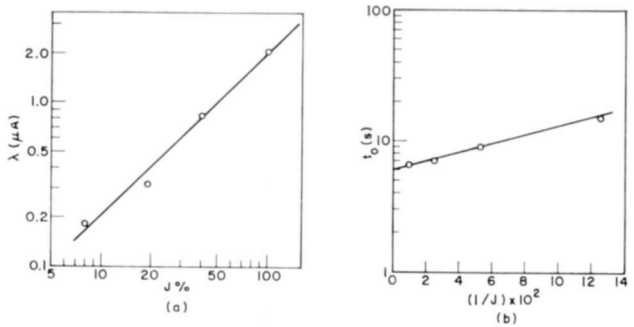


Fig. 5—(a) Plot of λ versus J showing that $\lambda \propto J$. (b) Plot of log t_0 versus $(1/J)$ showing that $t_0 \propto \exp$ (constant/J) [from *Trans. N. Y. Acad. Sci.*, Vol. 27, p. 946 (1965)].

Comments on the Scope of the Analysis

It is clear that this particular CdS crystal demonstrates the main features of the model for photo-induced chemisorption using oxygen as the adsorbate. But this does not mean that all similar materials should show the same behavior, for several restrictive assumptions have been made in the development of the model that may not be valid for other materials. It is well to review some of these restrictions and to see how they might affect the chemisorption kinetics.

One basic assumption is the inequality expressed in (3), $t_p < \tau_p$. If this inequality is not satisfied, the chemisorbed ions act as traps for majority photocarriers rather than as recombination centers for the

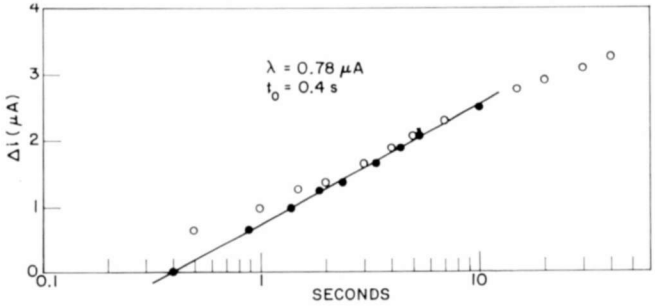


Fig. 6—Elovich-type plot of the photo-induced chemisorption of CO on CdS at 18°C.

photocarriers. This means that the chemisorbed ions do not strongly affect the steady-state photocurrent, but influence mainly the response time, which is the speed with which the photocurrent responds to a sudden change in the illumination level. However, a transient change in the photocurrent may appear during chemisorption. Violation of inequality (3) also means that the chemisorbed ions cannot be photodesorbed by greater-than-band-gap light. The surface can then be freed of chemisorbed ions by heating, by bombardment with energetic particles, or possibly optically by absorption directly by the chemisorbed ions. The various assumptions made in the derivation of Equation (16) are certainly not universally true. The relaxation of any of these restrictions will produce distortions in the chemisorption kinetics as expressed by Equation (16) that may or may not be accounted for within the general framework of the present analysis.

Finally, a very important restriction is that the adsorbate be chemically inert. The photo-chemisorption of carbon monoxide serves as an example of what can occur when this is not the case. Figure 6 shows the Elovich-type kinetics for the photocurrent under constant illumi-

nation during the uptake of CO on CdS at 18°C. The change in the steady-state photocurrent is completely reversible with partial pressure of the adsorbate at this temperature. However, a strong departure from this behavior occurs above 120°C. A recorder trace of the variation of the photocurrent on exposing the crystal to CO at 168°C is shown in Figure 7(a). The curve shows a definite minimum followed

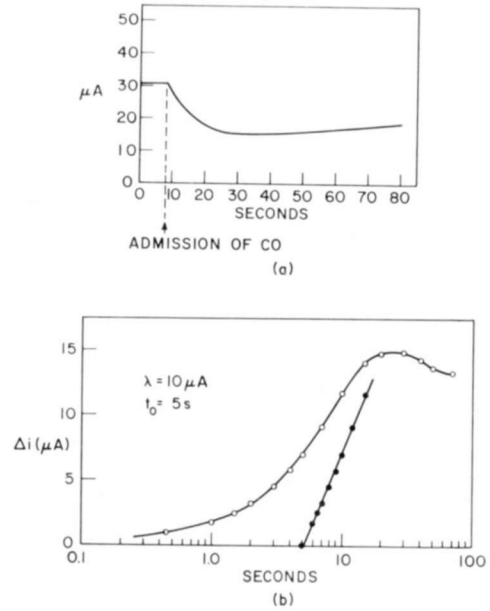


Fig. 7—(a) Change in the photocurrent in CdS on exposure to CO at 168°C. (b) Same curve on an Elovich-type plot.

by a partial recovery to a new steady-state photocurrent that is still less than the original value. The change in the steady-state photocurrent is not reversible at this temperature. Figure 7(b) shows the uptake kinetic of Figure 7(a) plotted as an Elovich-type curve. For the first 10 seconds after exposure, the response conforms with the foregoing analysis but thereafter, a marked departure occurs. This behavior is very similar to the reversals observed by Garner, Grey, and Stone[31] in the chemisorption of CO on semiconducting Cu_2O films. These workers found normal Elovich-type chemisorption at 20°C and

[31] W. E. Garner, T. J. Gray, and F. S. Stone, "The Oxidation of Copper and the Reactions of Hydrogen and Carbon Monoxide with Copper Oxide," *Proc. Roy. Soc.* (London), Vol. A197, p. 294, June 1949.

spontaneous reversals in the uptake curve, very much like those of Figure 7, at 200°C. The high-temperature result with Cu_2O was interpreted by postulating that carbonate ions ($CO_3^=$) are formed on the surface of the Cu_2O. It is possible that a similar reaction takes place on the sulfide surface through the formation of carbonyl sulfide (COS).[32] But this conjecture and similar speculations about the surface chemistry of CdS are outside the scope of this review. These results for CO chemisorbed on CdS are mentioned here primarily as an illustration of what can be expected when some of the assumptions leading to the derivation of Equation (16) are not fulfilled. Further examples are discussed at the end of Section VI.

V. Surface Trapping and Intrinsic Surface States

Another effect that is not predicted by the foregoing model relates to the localization of charge on the surface, not as chemisorbed ions, but as electrons and holes.[11] When the surface is freed of chemisorbed ions, either by photodesorption or by heating as described in the previous section, a different class of surface states becomes apparent—states that function as traps for free carriers of both signs rather than as recombination centers. This information is derived from the behavior of the photoconductivity response time and the thermally stimulated current, the magnitudes of which are closely related to the extent of carrier trapping.

Response Time

The response time of a photoconductor is equal to the majority carrier (electrons, in this case) recombination lifetime τ_n in the absence of trapping and, when traps are present, becomes larger than τ_n by a factor approximately equal to the ratio of the trapped-to-free charge, which is always greater than unity.[23] The effect is schematically shown in Figure 8. Since the response time is the speed with which the photocurrent responds to a change in the illumination, the effect of trapping is to cause the change in the photocurrent to lag behind an abrupt change in the illumination, either when it is diminished or when it is intensified.[23] The former case, giving rise to photoconductivity decay, is illustrated by Figures 8(a) through 8(c). The decay is rapid (Figure 8(a)) when there are few traps and more gradual when there are many traps (Figure 8(b)). Figure 8(c) depicts the temperature dependence of the trap-dominated decay and elucidates the physics of the process. All other things being equal, the

[32] R. J. Ferm, "The Chemistry of Carbonyl Sulfide," *Chem. Revs.*, Vol. 57, p. 621, 1957.

greater the temperature the more rapid the decay. The reason for this is that the trapped electrons are situated in localized levels below the conduction band and rely on thermal energy for transition into this band. The remainder of Figure 8 is a schematic illustration of the effect of trapping on the photocurrent growth following a sudden

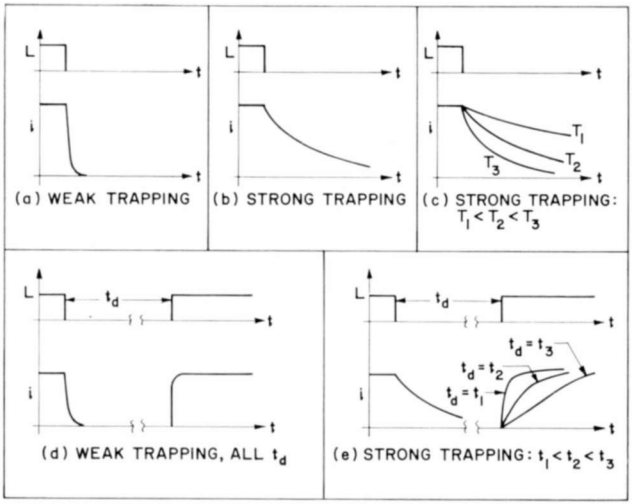


Fig. 8—Schematic behavior of the photocurrent response for strong and weak trapping. L represents the illumination and i represents the photocurrent. (a) Decay for weak trapping. (b) Decay for strong trapping. (c) Behavior of decay for strong trapping for different temperatures: $T_1 < T_2 < T_3$. (d) Growth for weak trapping. (e) Growth for strong trapping showing the effect of varying periods of darkness, t_d, between successive illuminations: $t_1 < t_2 < t_3$.

intensification of the illumination. Here the prior illumination history is important since the lag in the photocurrent growth depends on the number of *unoccupied* traps. Thus, if there are very few traps that empty rapidly after illumination, the subsequent growth on re-illumination will be fast and independent of the elapsed time interval t_d between the two successive illuminations (Figure 8(d)). On the other hand, if there are many traps, the growth will depend on t_d. Specifically, the growth becomes more extended as t_d is increased (Figure 8(e)).

Figures 9 through 12 show that the nature of the ambient affects the response time in a manner consistent with an ambient sensitive trapping model. Figure 9 shows several photoconductivity decay curves from an insulating, insensitive CdS crystal.[10] The curves were obtained

by suddenly terminating the illumination (at $t = 0$). Curve (a) is the decay at room temperature in 1 atmosphere of dry air (passed through a dry-ice-in-acetone moisture trap); curve (b) is the decay at 3×10^{-5} torr, all other parameters being unaltered from curve (a) (note that the steady-state photocurrent at $t = 0$ is greater in vacuum than in

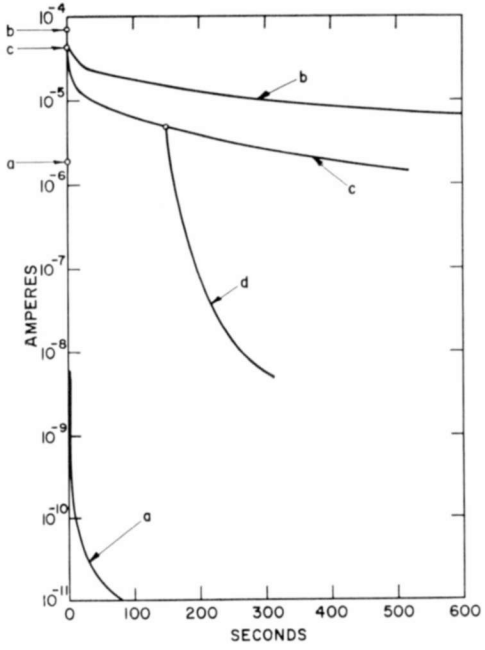


Fig. 9—Photocurrent decay in CdS at 18°C for various surface conditions. (a) 1 atmosphere dry air. (b) 3×10^{-5} torr. (c) 3×10^{-3} torr. (d) Effect of admitting dry air at 1 atmosphere during decay that commenced at 3×10^{-3} torr [from *Jour. Phys. Chem. Sol.*, Vol. 25, p. 911 (1964)].

air, as it should be); curve (c) is the decay at 3×10^{-3} torr; and curve (d) is obtained by suddenly exposing the crystal to dry air at 1 atmosphere during a decay that commenced at 3×10^{-3} torr.

Figure 10 shows a series of oscilloscope traces of the photocurrent decay designed to show the short time response of the decay. The upper trace corresponds to curve (b) of Figure 9 and the lowest trace corresponds to curve (a). Note the different calibration of the time scales. The negative curvature of the bottom trace (and also of the middle trace) is the closing profile of the mechanical shutter; the decay time is less than 1 millisecond. The middle trace is particularly interesting. It shows the decay at 3×10^{-5} torr, the same pressure as the upper trace, but at an illumination level of 0.027% of that used to obtain the

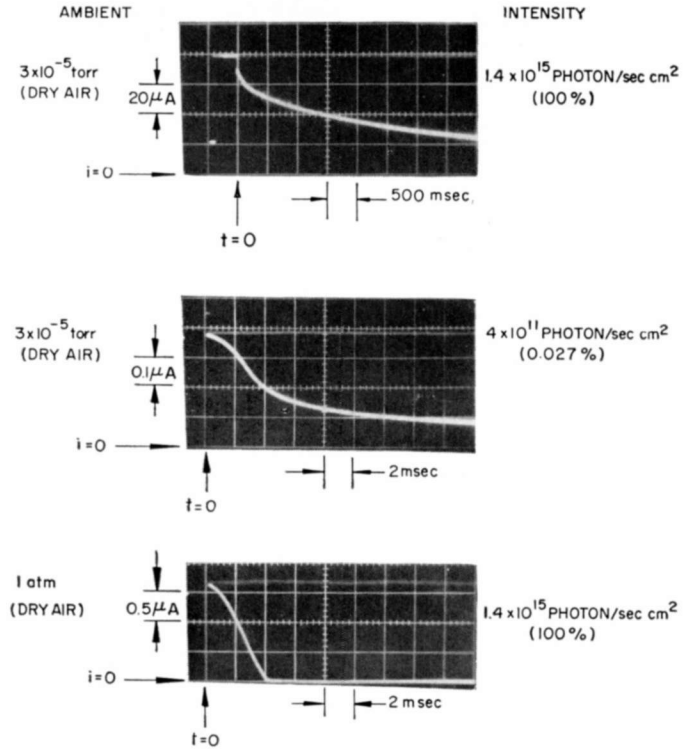


Fig. 10—Behavior of the photocurrent decay in CdS at 18°C for various surface conditions and illumination levels.

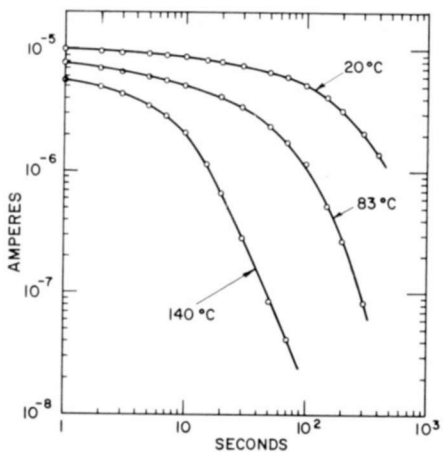


Fig. 11—Temperature dependence of the photocurrent decay in CdS for an adsorbate-free surface. Measurements were recorded at 1 atmosphere of dry nitrogen.

upper trace. Notice that the decay is *faster* at the *lower* illumination. The importance of this unusual effect is treated later.

The effect of temperature on the photocurrent decay of a CdS crystal with a chemisorbate-free surface is shown in Figure 11. These curves were measured in flowing dry nitrogen at 1 atmosphere rather than in vacuum. The curves are plotted on a logarithmic field to compress the time scale.

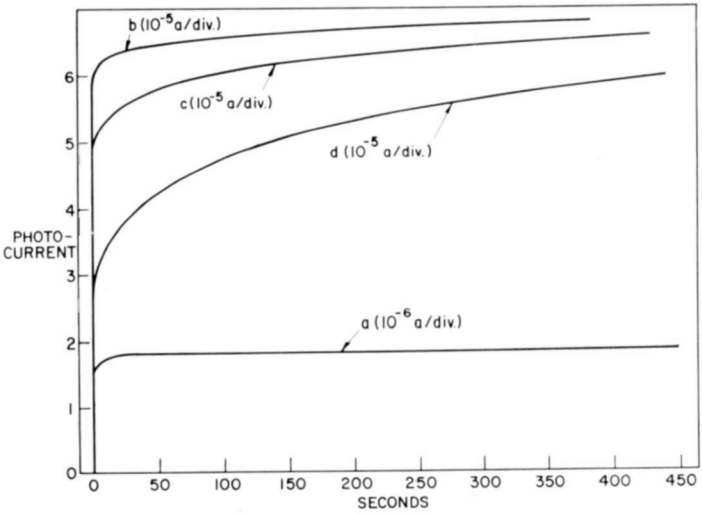


Fig. 12—Photoconductivity growth in CdS at 18°C for various surface conditions. (a) Growth in dry air at 1 atmosphere for all values of t_d between 1 and 30 minutes. (b), (c) and (d) represent the growth at 3×10^{-5} torr for $t_d = 1$, 5, and 30 minutes, respectively [from *Jour. Phys. Chem. Sol.*, Vol. 25, p. 911 (1964)].

As a final example of the response-time effect, Figure 12 shows the influence of the ambient on the photocurrent growth at room temperature.[10] Curve (a) represents the growth in 1 atmosphere of dry air for all values of t_d between 1 and 30 minutes, while curves (b), (c), and (d) represent the growth at 3×10^{-5} torr for the intervals $t_d = 1$, 5, and 30 minutes, respectively.

The following conclusions can be drawn from the response-time measurements. The response time is fast when the specimen is in contact with an adsorbate atmosphere and slow when the ambient is free of an adsorbate constituent. The slow response time exhibits all the features usually ascribed to trapping. The response time is rapidly accelerated when an adsorbate is admitted to a surface previously freed of chemisorbed ions.

Thermally Stimulated Current

Further evidence that trapping is strongly affected by surface conditions is furnished by the behavior of the thermally stimulated current (TSC). This is the current in excess of the dark current that is observed when a pre-excited photoconductor is heated in darkness to thermally release trapped carriers.* The integrated area between the

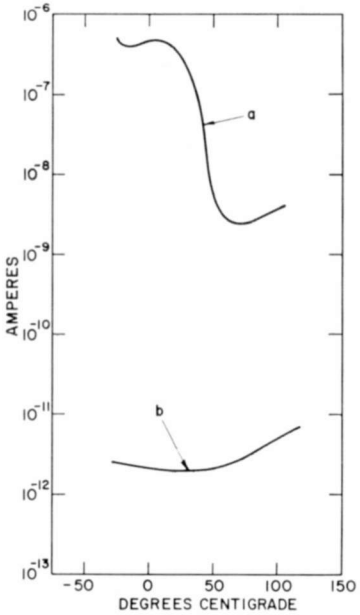


Fig. 13—Thermally stimulated current in CdS for two different surface conditions. (a) Adsorbate-free surface in nitrogen. (b) Surface with oxygen chemisorbed on it. Heating rate: 0.5°C/sec.

TSC and the dark current is proportional to the amount of trapped charge released during the measurement provided the heating rate is constant; the proportionality factor is the photoconductivity gain. The TSC measurements are consistent with the response-time measurements; the amount of trapped charge released with the surface devoid of chemisorbed ions is several decades in excess of that released with oxygen chemisorbed on the surface. This can be seen from Figure 13, which shows two TSC curves obtained from a CdS crystal. Curve (a) was recorded in a nitrogen atmosphere after photodesorption and represents the TSC for the adsorbate-free surface, while curve (b), which was measured in oxygen, is the TSC with oxygen chemisorbed

* See Reference (26), Chap. 9, pp. 292-299.

on the same surface. The increase in the current at the high-temperature end is the dark current, which also depends on surface conditions. For a complete discussion of the dark-current effect, the reader is referred to Reference (13).

The reduction of these data using the photoconductivity lifetime leads to the following conclusion: 10^{14} trapped electrons are released in the absence of oxygen while only 10^{8} trapped electrons are released when oxygen is present.[9] The TSC measurement shows directly that an acceptor adsorbate strongly affects the amount of trapped charge that the crystal can store. It is unlikely that this ambient sensitive trapping is the result of the diffusion of impurities into or from the bulk. The rapid rate at which the response time changes during chemisorption (curve (d) of Figure 9) would then require an unrealistically large diffusion coefficient of 10^{-10} cm^2/sec at room temperature for impurities. [This value is obtained from the relation $D = \lambda^2/\tau$, where D is the diffusion coefficient, λ is the diffusion length, and τ is the diffusion time. Curve (d) of Figure 9 shows that the current drops by one decade in 20 seconds on the admission of an adsorbate. Thus, if the trapping effect is due to the adsorbate diffusing into the photoexcited surface layer, which is several microns thick,[24] $D \approx (3 \times 10^{-4})^2/20 \approx 5 \times 10^{-9}$ cm^2/sec.] Evidently, the trapping observed with no adsorbate is a true surface phenomenon in that the trapped electrons are situated on the surface. But when the amount of released charge is normalized to the geometric surface area of the crystal (0.14 cm^2) one finds that 7×10^{14} electrons/cm^2 are trapped on the adsorbate-free surface, a number corresponding to nearly one electron per surface lattice site. [By contrast, if it is assumed that the trapped charge released with oxygen on the surface originates in the bulk, the concentration one obtains is 7×10^{12} cm^{-3}, a very small value. This figure is obtained by expressing the number 10^{8} as a volume density. The volume of the conducting channel is: (surface area) $\times$ (ambipolar diffusion length) $= 1.4 \times 10^{-1}$ cm^2 $\times$ 10^{-4} cm $= 1.4 \times 10^{-5}$ cm^3].

Phenomenological Model for Intrinsic Surface States

The following phenomenological model was proposed to account for the unusually large surface coverage of trapped charge in the absence of an adsorbate.[8,9,11] The electrons trapped on an adsorbate-free surface are not compensated in a depletion layer, as are the chemisorbed ions, but rather by a nearly equal number of holes, which are also trapped on the surface, so as to leave only a negligible net surface charge relative to the adjacent bulk. In other words, the trapped electrons are *surface compensated* by trapped holes, and the occupation of the

traps is no longer subject to the condition expressed by Equation (1). It is doubtful whether electron-hole pairs can be trapped on the surface at a concentration approaching nearly one per surface ion pair without appreciable direct recombination. However, an electron released from a surface trap into the bulk cannot recombine until a compensating hole is released from the surface into the bulk. Thus the lifetime of an electron released from a surface trap can become greater than the lifetime of a photogenerated electron. Using the photoconductivity lifetime in the reduction of the TSC then tends to give values of trap concentrations that are too large. But, to understand the behavior of the response time with the illumination level for the chemisorbate-free surface (discussed at the end of this section), it is still necessary to accommodate a trapped electron concentration on the surface exceeding the limitation imposed by Equation (1). Since both electrons and holes are considered trapped on the surface, there must be surface states for both carrier types. The fact that these surface states manifest themselves when no adsorbate is present is taken as an indication that they are intrinsic surface states. Figure 14(a), which depicts the energy-band scheme near the surface in thermal equilibrium, shows how the intrinsic surface states are disposed relative to the bulk bands.[9] There are two distinct sets of surface states, possibly distributed slightly in energy. One set is situated near the valence band and is occupied by electrons so as to function as traps for holes. The other set lies near the conduction band. It is unoccupied by electrons and serves as electron traps. The physical origin of the intrinsic surface states is outlined in Section VII.

In order to see some of the consequences of this model, it is instructive to regard in some detail the way in which the occupation of the intrinsic surface states is altered when the surface is illuminated with light of greater-than-band-gap energy.[9] What happens under such excitation is illustrated in Figures 14(b) and 14(c).[9] Optically generated electrons and holes appear in the conduction and valence bands, respectively. For the present case of an n-type photoconductor, charge neutrality is maintained by the recombination centers, which localize most of the photoholes very rapidly. If one assumes that the capture cross sections of the surface traps for their respective carrier types are about equal, then the initial trapping rate of electrons exceeds that of holes ($n \gg p$) and the electron traps fill more rapidly at the outset than the hole traps. At this point, the trapped electrons are compensated almost entirely by holes localized in the recombination centers in the adjacent bulk, and a Schottky-type barrier begins to form that retards the migration of electrons to the surface while assisting the

motion of holes to the surface. If condition (3) is violated, that is, if $\tau_p < t_p$, virtually no holes reach the surface, the electrons in the surface traps will remain compensated in the depletion layer, and the Schottky barrier grows until access to the electron traps is pinched off by the diffusion potential. Thus, little evidence of surface trapping can be expected when $\tau_p < t_p$, since Equation (1) limits the accumulation of bulk compensated surface charge to about 10^{-4} of a monolayer. This

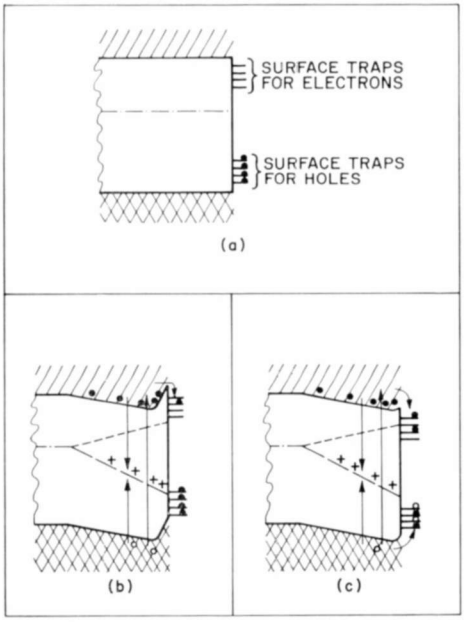


Fig. 14—Model for the intrinsic surface states of an adsorbate-free surface as deduced from the experiments. (a) Surface at thermal equilibrium showing two distinct sets of surface traps; the levels near the conduction band are unoccupied and serve as electron traps, while the states near the valence band are occupied and serve as traps for holes. (b) Steady state with surface under illumination for the case $t_p > \tau_p$; the electrons in the surface traps are compensated in the depletion layer. (c) Steady state with the surface illuminated for the case $t_p < \tau_p$; the electron and hole surface traps are now equally occupied.

situation is depicted in Figure 14(b). Alternatively, if $\tau_p > t_p$, most of the photoholes will penetrate to the surface and the restriction imposed by Equation (1) on the occupation of surface electron traps will no longer hold. Under these circumstances the surface compensation referred to earlier becomes effective, and the occupation of the surface traps is greatly increased, as shown schematically in Figure 14(c).

The relationship between the occupation of the surface traps and the relative magnitude of τ_p and t_p serves to explain the unusual de-

pendence of the response time on the illumination level observed in Figure 10, namely, that the response time becomes slower as the illumination increases.[9] There is no way to account for this behavior on the basis of bulk transitions in a homogeneous photoconductor in which the photoconductivity is trap dominated.[23,33] However, an explanation does follow directly from the present model of surface-compensated trapping.[9] From Equation (5) and the fact that N_+ is proportional to the illumination level, J, it is clear that $t_p \propto (1/J)$. Thus, if $t_p > \tau_p$ at low J, there should be little evidence of trapping, and one should see a relatively fast response. But if t_p becomes less than τ_p as J increases, strong trapping should occur with a correspondingly slower response. This is precisely what is observed.[9]

VI. Interaction between the Intrinsic Surface States and an Adsorbate

The interaction between the charge trapped in the intrinsic surface states and the adsorbate is responsible for the strong and rapid acceleration of the photoconductivity decay when an adsorbate is admitted to a previously adsorbate-free surface. It will be shown that this effect is caused by the action of the adsorbate molecules as recombination centers for electrons and holes trapped in the intrinsic surface states.

Consider the energy-band diagrams of Figure 15. The steady state under strong illumination ($t_p < \tau_p$) in the absence of an adsorbate is illustrated in Figure 15(a). When the illumination is switched off, the photocurrent decay commences but it will be slow because of the action of the surface traps. The band picture and the schematic behavior of the photocurrent corresponding to this phase of the decay are shown in Figure 15(b). The effect of admitting an acceptor adsorbate at this point in the decay is shown in Figure 15(c). When an adsorbate molecule meets the surface, it is first converted into a negative ion at the site of a trapped electron. This means that the free energy of the chemisorbed ion must be lower than that of the trapped electron, a fact that is indicated on the figure by situating the level of the chemisorbed ion below that of the electron traps.[11,13] The transition leading to the formation of a negative adsorbate ions is labeled "1." The second step in the recombination process is the neutralization of the negative adsorbate ion by a neighboring trapped hole so that one is left with the original neutral adsorbate molecule but less one trapped electron–hole pair. The neutralization transition is labeled "2." In this way, each adsorbate molecule can cause many trapped

[33] A. Rose, "An Outline of Some Photoconductive Processes," *RCA Review*, Vol. 12, p. 362, Sept., Pt. I, 1951.

electrons and holes to recombine. Indeed, this must be so since prior to chemisorption the surface can accommodate about 10^{14} cm^{-2} electron–hole pairs, while after chemisorption the surface-ion concentration cannot exceed 10^{11} cm^{-2}. The chemisorbed ions must necessarily be able to migrate over the surface to achieve this recombination gain.

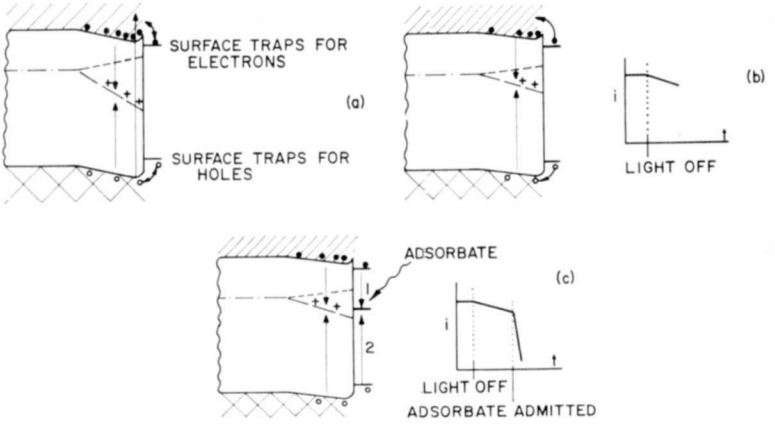


Fig. 15—Model explaining how trapped surface charge is recombined through the action of an adsorbate. (a) Surface under steady illumination showing the surface traps. (b) Initial stage of photocurrent decay with adsorbate-free surface. (c) Effect of admitting an adsorbate. Transition 1 represents the formation of negative ions and transition 2 represents the neutralization of the chemisorbed ions by the trapped holes. The schematic behavior of the photocurrent is also shown.

The motion of adsorbed ions on SiO_2 layers has been reported[34,35] with a mobility of the order of 10^{-3} cm^2/volt sec. This value is sufficient to account for the recombination gain but not large enough for an ionic surface current to make a measurable contribution to the photocurrent.[9] Thus, the adsorbate rapidly removes, by direct recombinations, the trapped electrons and holes that normally sustain the photocurrent decay, and a strong acceleration of the decay ensues. This is also shown schematically in Figure 15(c).

The mechanism for photo-induced chemisorption on an insulator, including the influence of intrinsic surface states, can now be sum-

[34] H. Statz and G. A. DeMars, "Electrical Conduction via Slow Surface States on Semiconductors," *Phys. Rev.*, Vol. 111, p. 169, July 1, 1958.

[35] Shockley, Hooper, Queisser, and Schroen, "Mobile Electric Charges on Insulating Oxides with Application to Oxide Covered Silicon p-n Junctions," *Surface Science*, Vol. 2, p. 277, 1964.

marized.[11] One begins with steady-state illumination and an adsorbate-free surface as depicted in Figure 15(a). The immediate response to the admission of an adsorbate is the rapid recombination of the electrons and holes in the surface traps according to the mechanism just described. This is shown in Figure 16a. In this way, virtually all the surface-compensated trapped charge can recombine via the adsorbate

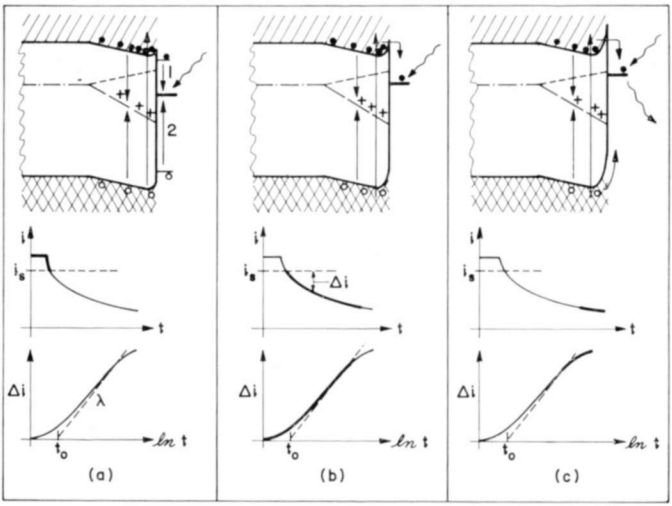


Fig. 16—Model for photo-induced chemisorption showing the influence of the intrinsic surface states. (a) Rapid combination of charge in the surface traps on initial exposure to an adsorbate. (b) Accumulation of bulk compensated surface charge in the form of chemisorbed ions. (c) Approach to the new steady state when photodesorption limits the accumulation of chemisorbed ions ($t_p < \tau_p$). The schematic behavior of the photocurrent is also shown.

without an accumulation of bulk compensated surface charge. The photocurrent may drop while the adsorbate is eradicating the trapped charge but one should not expect this initial decrease in the photocurrent to conform to the Elovich equation. The reason is that the Elovich equation is derived on the assumption that the rate-limiting step during chemisorption is the diffusion of electrons through a surface barrier. But an appreciable surface barrier may exist only after some bulk compensated charge has accumulated on the surface, and this may or may not happen during the recombination of the trapped surface charge. The present theory is too crude to take such details into account. However, there is evidence that the Elovich rate law (Equation (6)) is not applicable for arbitrarily short times after

exposure to an adsorbate. It is sometimes necessary to choose an initial current, i_s, (in Equation (15)) somewhat less than (but never greater than) the actual steady-state photocurrent prior to adsorption in order to construct an Elovich curve from the measured uptake curve. This possibility is shown in the schematic uptake curve of Figure 16(a).

After the removal by direct recombination of most of the trapped surface charge, the adsorbate begins to accept electrons from the conduction band and the accumulation of bulk compensated surface charge in the form of chemisorbed ions proceeds according to the Elovich rate law. This is shown in Figure 16(b). Eventually, photodesorption by minority carriers limits the accumulation of surface ions and the final approach to a new steady state proceeds more slowly, as shown in Figure 16(c). If the excitation is now removed, the photocurrent decay is very rapid because the chemisorbed ions prevent trapped charge from accumulating on the surface.[10,11]

VII. Physical Origin of the Intrinsic Surface States

The purpose of this section is to review briefly the physical origin of intrinsic surface states on ionic crystals and to underscore the several common features between the results of this physical model and the phenomenological model of Sections V and VI, which was derived from a self-consistent interpretation of the experiments. The physical arguments reviewed here are due to J. D. Levine.[8,9] The reader is referred to the original paper, "The Theory and Observation of Intrinsic Surface States on Ionic Crystals," for a more complete presentation.[9]

The basis of the physical argument is the classical model of ionic binding, which considers the long-range electrostatic forces among the lattice ions as the principle interaction contributing to the cohesive energy.[36,37] The origin of the bulk states from the separate atomic levels can be understood with the aid of Figure 17(a), which shows the dependence on the interionic distance (r) of the electronic energy in an ideal ionic lattice composed of cations (M^+) and anions (X^-). The right side of the diagram corresponds to infinite dilation of the lattice with no coulombic interactions. The figure shows the energy-level-term scheme of M relative to M^+ and of X^- relative to X. Thus, the depth of the M and X^- levels represent, respectively, the ionization

[36] F. Seitz, *Modern Theory of Solids*, p. 408 ff and p. 447 ff, McGraw-Hill Book Co., New York, 1940.

[37] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals*, p. 86, Dover Publications, New York, 1964.

potential, I, of M, and the electron affinity, A, of X. The fact that M lies below X^- indicates that the system is stable as neutral M and X atoms at infinite separation, and that the energy $I - A$ must be expended to transfer an electron from M to X to form an ion pair.

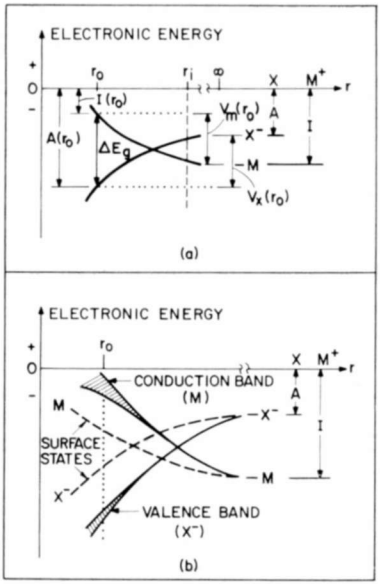


Fig. 17—(a) Energy-level scheme as a function of interionic distance for an ionic crystal illustrating how the stable ionic configuration arises from coulombic interactions among the lattice ions. (b) Origin of the intrinsic surface states on an ionic crystal from the reduced coulombic interaction of lattice ions situated on the surface.

Consider next a lattice composed of M and X atoms, with a large but finite interatomic distance, r_i. If the term scheme for the ionization potential of M and the electron affinity of X is drawn for r_i, one would find that the magnitude of I has become smaller and that the magnitude of A has become larger according to

$$-I(r_i) = -I(\infty) + V_M(r_i), \tag{19}$$

and

$$-A(r_i) = -A(\infty) - V_X(r_i), \tag{20}$$

where V_M and V_X are the coulombic interaction potentials at each M^+ and X^- lattice site, respectively, due to all the other lattice sites now considered occupied by M^+ and X^- ions. Note that Equation (19)

refers to taking an electron from a *neutral* M to vacuum, while Equation (20) refers to the removal of an electron from an X^- *ion* to vacuum. The potentials are

$$V_M(r) = \frac{c_M e}{r}, \tag{21}$$

and

$$V_X(r) = \frac{c_X e}{r}, \tag{22}$$

for a lattice of singly charged ions. Here c_M and c_X are the Madelung constants, which express the fact that in the computation of the coulombic interaction one must consider each lattice ion as surrounded by an infinite sea of other lattice ions. The magnitude of the Madelung constants depends only on the geometry of the lattice.

As one continues to reduce r, both I and A continue to change according to Equations (19) and (20). Eventually, the two levels cross, as shown in Figure 17. To the left of the crossover, where I is less than A, the ionic configuration is stable, and it becomes appropriate to call the lower level the X^- state and the upper level the M^+ state. The lattice stabilizes at the equilibrium interionic distance, r_0, when the repulsive quantum mechanical overlap forces balance the attractive coulombic interactions. One must now expend the energy $A(r_0) - I(r_0)$ to form a neutral M — X pair in the ionic lattice. This energy corresponds to the band gap of the lattice.

The intrinsic surface states arise quite naturally from this picture by considering the experience of a lattice ion on the surface of a finite lattice as it is contracted to its equilibrium configuration.[8,9] The Madelung constant for a test lattice site on the surface will always be less than that for a similar test site located deep in the bulk, because of the way the coulomb sums used to evaluate the Madelung constants are affected when the lattice is considered as a semi-infinite three-dimensional crystal with a surface rather than as a lattice of infinite extent in all dimensions.[9] The surface states are computed by inserting the reduced Madelung constant peculiar to the surface lattice sites into Equations (19) and (20). The surface states are shown in Figure 17(b) together with a more realistic representation of the bulk states, which now show broadening into bands to reflect the quantum mechanical interactions between neighboring lattice ions.

The properties of the intrinsic surface states derived from the

simple $M^+ - X^-$-type ionic lattice can now be enumerated.[9] (1) There is one surface state for each surface lattice ion. (2) For surfaces terminating in an equal number of positive and negative ions, there will be two different sets of surface states; one set is derived from the M^+ bulk band (the conduction band) and the other set originates with X^- bulk band (the valence band). The surface states are labeled accordingly in Figure 17(b). The concentrations of the M^+ and X^- surface states are equal, respectively, to the surface M^+ and X^- ion concentrations. (3) To a good approximation, the surface states are equidistant in energy from their respective bulk bands. The interionic distance for real ionic lattices lies almost always to the left of the crossover of the surface states. Thus, the X^- surface states are situated in the bottom half of the forbidden gap near the valence band, while the M^+ surface states lie in the upper half of the gap near the conduction band. (4) The equilibrium occupation of the surface states coincides with that of the bulk band from which the surface states originate. That is, the X^- surface states are occupied by electrons and function as traps for holes, while the M^+ surface states are unoccupied and serve as electron traps. Clearly, these surface states have all the features of the intrinsic surface states that were postulated to account for the experiments performed with the adsorbate-free surfaces of the CdS crystals.

Some of the properties of the intrinsic surface states are altered when more complicated surface configurations are considered,[9] e.g., when there is an unequal concentration of positive and negative ions terminating the lattice, or for high index surfaces where edges and corners must be included in the surface topography. However, the entire picture remains quite valid even when there is a considerable covalent contribution to the binding as there is with several oxides and sulfides (including CdS).[9] The accepted procedure in such cases is to account for the covalency with an effective ionic charge (at each lattice site) that is smaller than the ideal ionic charge.[9,27] In the case of CdS, the ideal ionic charge according to 100% ionic binding is 2. Including the covalent contribution to the binding, the effective ionic charge becomes[9,27] 0.5. These and other complications are treated in the original paper.[9]

It is appropriate to mention here one important consequence of the action of the intrinsic surface states as traps for electrons and holes from the bulk, which bears on the stability of the surface. Physically, the surface traps are the terminating lattice ions at the surface. Occupation of the traps by the appropriate carrier types reduces the effective charge of some of these surface ions relative to the lattice.

That is, the periodic charge pattern of the surface, which figures in the computation of the Madelung constant, will be perturbed so as to lower the Madelung constant. This effect can become appreciable when a large fraction of the surface traps have been filled with carriers from the bulk. The net effect is to make the surface less stable when it is illuminated than when it is in darkness.[9] This effect has recently been reported[38] with the observation that the evaporation rate of insulating CdS crystals at 700°C and 10^{-6} torr can be enhanced five-fold by illuminating the free surface with greater-than-band-gap light. Further, the rate-limiting step in the evaporation is the transfer of photogenerated electrons and holes to the surface and the quantum efficiency for the photo-enhanced evaporation is near unity. The interpretation given to this effect[38] is that the photogenerated electron and holes become trapped on the surface Cd and S ions, respectively, neutralize their charges, and reduce their binding to the underlying bulk; this interpretation is clearly consistent with the present model of intrinsic surface states.

Acknowledgments

The author wishes to thank A. Rose and S. Larach for their continuing interest during the course of this work, and particularly J. D. Levine, with whom it has been a singular pleasure to collaborate during the latter portion of this work.

Appendix—Energy-Band Configuration Near the Illuminated Surface of a Photoconducting Insulator

The energy-band configuration of the illuminated surface of a photoconducting insulator as shown in Figure 2(b) is obtained as follows (see Reference (24) for the complete development). Because of diffusion, the free-electron concentration decreases with distance x from the illuminated surface according to the relation

$$n(x) = n(0) \exp(-x/L_a). \tag{23}$$

Here $n(0)$ is the free-electron concentration at the illuminated surface and L_a is the ambipolar diffusion length; $L_a = (D_a\tau_n)^{1/2}$, where D_a is the ambipolar diffusion coefficient; $D_a = D_p[2\mu_n\tau_p/(\mu_n\tau_n + \mu_p\tau_p)]$;

[38] G. A. Somorjai and J. E. Lester, "Charge-Transfer-Controlled Vaporization of Cadmium Sulfide Single Crystals. I. Effect of Light on the Evaporation Rate of the (0001) Face," *Jour. Chem. Phys.*, Vol. 43, p. 1450, 1 Sept. 1965.

$D_p = \mu_p kT/e$. The energy difference between the conduction band edge and the quasi-Fermi level for electrons is $\Delta E_{f,n}(x)$ and is defined by the relation (for nondegenerate statistics)

$$n(x) = N_c \exp[-\Delta E_{f,n}(x)/(kT)], \tag{24}$$

where N_c is the effective density of states in the conduction band, k is the Boltzmann constant, and T is the absolute temperature. The position dependence of $\Delta E_{f,n}(x)$ is determined by Equations (23) and (24):

$$\Delta E_{f,n}(x) = \Delta E_{f,n}(0) + (kT/L_a)x. \tag{25}$$

Thus, $\Delta E_{f,n}(x)$ has a constant and positive slope with respect to the conduction-band edge. Similarly, it can be shown that the quasi-Fermi level for holes also has a constant slope with respect to the valence-band edge. The condition for this to be true is that at any x between the illuminated surface and the unexcited bulk, the ratio of n to p must be constant[24] and independent of x.

The Dember field $\mathcal{E}_D$ is obtained by equating the total current (field driven plus diffusion) normal to the illuminated surface to zero and solving for the field;

$$\mathcal{E}_D(x) = -\frac{kT}{e}\frac{\mu_n\tau_n - \mu_p\tau_p}{\mu_n\tau_n + \mu_p\tau_p}\frac{1}{n}\frac{dn}{dx}. \tag{26}$$

Upon substitution of Equation (23) into (26), one obtains

$$\mathcal{E}_D = \frac{kT}{eL_a}\frac{\mu_n\tau_n - \mu_p\tau_p}{\mu_n\tau_n + \mu_p\tau_p} = \text{constant}. \tag{27}$$

The slope dE/dx of the energy band is $e\mathcal{E}_D$ so that the energy bands have the constant slope equal to the slope of $\Delta E_{f,n}$ relative to the bottom of the conduction band multiplied by the factor $(\mu_n\tau_n - \mu_p\tau_p)/(\mu_n\tau_n + \mu_p\tau_p)$, which is always positive and less than unity when $\mu_n\tau_n > \mu_p\tau_p$.

The energy-band diagrams of the illuminated insulator surface have been drawn according to this development with $\mu_n\tau_n > \mu_p\tau_p$.

A SIMPLIFIED APPROACH TO THE ANALYSIS OF ELECTROMAGNETIC-WAVE PROPAGATION CHARACTERISTICS OF PLASMA-COATED SURFACES

By

L. W. Zelby*

RCA Defense Electronic Products
Camden, N. J.

Summary—*An electromagnetic system consisting of a plasma layer on a surface of arbitrary impedance is known to lead to a number of wave characteristics not usually obtained in the case of dielectrics. A system comprising a plasma slab on such an arbitrary impedance is completely analyzed using the impedance concept, Snell's law, and propagation vectors. This very simple and straightforward analysis readily leads to results usually obtained only by rather complicated mathematical manipulations.*

Introduction

THE response of systems including gaseous plasma to electromagnetic excitation has been the object of intensive study during recent years. The complexity of plasmas has led to the application of extremely elegant mathematical techniques to the various problems involving propagation of electromagnetic waves in or near plasmas. As a consequence, many of the calculations for the determination of various singularities and their loci in the complex propagation constant plane are extremely involved, whether the system is physically realizable or not.

The approach employed in this paper deviates from this trend and shows that a relatively simple analysis of some plasma systems can provide a substantial amount of knowledge about the type of waves that may exist there. This approach may provide groundwork for systematizing the problems of synthesis of systems in which plasma properties are significant relative to wave propagation. This work is concerned chiefly with the characteristics of plasma-coated surfaces, and is based on the use of wave vectors and wave impedances. Both wave vectors and wave impedances are eminently suited for a general approach—the wave vectors because of their covariant nature, and the wave impedances because of their relation to the physically significant

* Consultant from the Moore School of Electrical Engineering, University of Pennsylvania.

quantity, energy flow. Furthermore, the natural invariance of Maxwell's equations,[1] and the resulting topological character of the solutions,[1,2] allows a very general analysis in any convenient coordinate system. This work, then, represents an extension of previous attempts at unification and simplification of the solution of problems concerned with electromagnetic guided waves.[3,4]

The behavior of systems comprising plasma-coated surfaces is becoming increasingly important due in large measure to space exploration and to attempts at submillimeter-wave generation and transmission. The former involves communication with missiles or space vehicles through plasma sheaths, while the latter involves changes in the wave-guiding properties of systems in the presence of corona due to extremely high field intensities. In spite of their apparent similarity, properties of plasma-coated surfaces are not as well known as those of surfaces coated with dielectrics.[5-8]

Several relatively simple systems have already been analyzed using Green's functions and other, sometimes quite complicated, integral techniques. These systems represented plasma layers terminated by a short or open circuit in the presence[9-12] or absence[13-15] of magnetic

---

[1] E. J. Post, *Formal Structure of Electromagnetics*, North-Holland Publishing Co., Amsterdam, Netherlands, Chap. 3 (1962).

[2] J. D. Lawson, "Electromagnetic Wave Problems," *Electronic Radio and Engineer*, Vol. 36, No. 9, p. 332, Sept. 1959.

[3] L. W. Zelby, "Unified Approach to the Theory of Guided Waves," *International Jour. Elec. Eng. Ed.*, Vol. 3, No. 1, p. 39, May 1965.

[4] L. W. Zelby, "Surface Waves Along a Plasma-Air Boundary," *Proc. IEEE*, Vol. 51, No. 12, p. 1774, Dec. 1963.

[5] H. M. Barlow and A. L. Cullen, "Surface Waves," *Proc. I.E.E.*, Vol. 100, Pt. III, p. 329, Nov. 1953.

[6] J. Brown, "The Types of Waves Which May Exist Near a Guiding Surface," *Proc. I.E.E.*, Vol. 100, Pt. III, p. 363, Nov. 1953.

[7] A. E. Karbowiak, "Radiation and Guided Waves," *IRE Trans. on Antennas and Propagation*, Vol. AP-7, Supplement, p. S191, Dec. 1959.

[8] L. W. Zelby, "Propagation Modes on a Dielectric Coated Wire," *Jour. Franklin Inst.*, Vol. 274, No. 2, p. 85, Aug. 1962.

[9] S. R. Sheshadri, "Excitation of Surface Waves on a Perfectly Conducting Screen Covered with Anisotropic Plasma," *IRE Trans. on Microwave Theory and Techniques*, Vol. MTT-10, No. 5, p. 573, Nov. 1962.

[10] S. R. Sheshadri and W. F. Pickard, "Surface Waves on an Anisotropic Plasma Sheath," *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-12, No. 5, p. 529, Sept. 1964.

[11] S. Adachi and Y. Mushiake, "Surface Wave on a Perfectly Conducting Plane Covered with Magnetoplasma," *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-12, No. 5, p. 569, Sept. 1964.

[12] S. V. Zivanovic, H. M. Musal, Jr., and R. I. Primick, "Determination of Plasma Layer Properties From the Measured Electromagnetic Transmission Coefficient," *IEEE Trans. on Antennas and Propagation*, Vol. AP-12, No. 5, p. 618, Sept. 1964.

induction. Dispersion relations have been obtained, and eigenvalues for a specific set of parameters determined. Eigenvalues for one of the systems[14] were obtained in a very elementary manner,[4] and this method will be extended here to more general cases.

Besides having practical significance, the plasma layer is also of academic interest because the real part of the dielectric coefficient,

$$K_r = 1 - \frac{\dfrac{\omega_p^2}{\omega^2}}{\dfrac{1+\nu^2}{\omega^2}}, \tag{1a}$$

is always less than unity, whereas in dielectric media[16] it is greater than unity except at extremely high frequencies,*

$$(K_r)_d = 1 + \frac{\dfrac{\Omega_p^2}{\omega^2}}{\dfrac{\omega_0^2}{\omega^2} - \dfrac{1+\nu_M^2}{\omega^2}}. \tag{1b}$$

In the above equations $\omega_p^2 = ne^2/(m\epsilon_0)$ is the angular plasma frequency, n is electron density, e is electron charge, m is electron mass, ϵ_0 is permittivity of free space, Ω_p^2 is related to the volume density of oscillators whose resonant frequency is ω_0, ν_M accounts for quasi-frictional forces introduced by molecular collisions, and ν is the electron collision frequency for momentum transfer.

---

[13] W. V. T. Rusch, "Propagation Constants of Surface Waves on a Plasma-Clad Cylinder," *IRE Trans. on Antennas and Propagation*, Vol. AP-10, No. 2, p. 213, March 1962.

[14] T. Tamir and A. A. Oliner, "The Spectrum of Electromagnetic Waves Guided by a Plasma Layer," *Proc. IRE*, Vol. 51, p. 317, Feb. 1962.

[15] G. Meltz and R. A. Shore, "Leaky Waves Supported by Uniaxial Plasma Layers," *IEEE Trans. on Antennas and Propagation*, Vol. AP-13, No. 1, p. 94, Jan. 1965.

[16] J. A. Stratton, *Electromagnetic Theory*, p. 323, McGraw-Hill Book Company, New York, New York, 1941.

* Since[17] $\omega_{violet} < \omega_0$, the real part of the dielectric coefficient for dielectrics will be less than unity only for frequencies in excess of 10^6 gc, clearly out of the range of interest for microwave propagation.

[17] A. Sommerfeld, *Optics*, p. 91, Academic Press, Inc., New York, New York, 1954.

The character of the wave vectors (real, imaginary, or complex) determines the nature of the wave; the character of the impedances of the system determines the character of the corresponding wave vectors, thus determining the set of waves that can be supported by the system. In the specific case considered here, it is convenient to discuss the character of the vectors first and then that of the impedances. Rectangular geometry is used to simplify the functional relations. The generality of this approach (applicable to all cylindrical, orthogonal coordinate systems in which the scalar Helmholtz equation is separable) is justified.

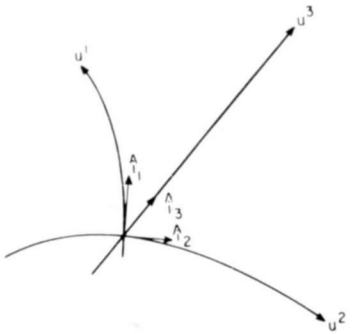


Fig. 1—Cylindrical, orthogonal coordinate system.

General Analysis

Consider the cylindrical, orthogonal coordinate system shown in Figure 1. In such a coordinate system, the scalar Helmholtz equation,

$$(\nabla^2 + k_0^2)\phi = 0, \tag{2}$$

is always separable[18] and can be reduced to

$$\frac{d^2U^{(i)}(u^{(i)})}{(du^{(i)})^2} + U^{(i)}(u^{(i)}) \sum_{j=1}^{3} \alpha_{ij}k_j^2 = 0, \qquad i = 1,2,3. \tag{3}$$

Here α_{ij} are the elements of the Stäckel matrix, which need not be selected uniquely; ∇^2 is the Laplacian; $k_0^2c^2 = \omega^2$ is the angular frequency; c is the speed of light in vacuum; and k_0 is the norm of the wave vector $\boldsymbol{k}_0$.

---

[18] P. Moon and D. E. Spencer, "Separability Conditions for the Laplace and Helmholtz Equations," *Jour. Franklin Inst.*, Vol. 253, No. 6, p. 585, June 1952.

The coefficients k_j^2, which are the components of the wave vector, are determined by the boundary conditions, and the analysis of their forms (real, imaginary, or complex) will yield information about the behavior of the system. This is seen more readily when the above equations are written out:[19]

$$\frac{d^2U^{(1)}}{(du^{(1)})^2} - U^{(1)}(u^{(1)})\,[k_2^2 - \alpha_{13}(u^{(1)})k_3^2] = 0 \tag{4a}$$

$$\frac{d^2U^{(2)}}{(du^{(2)})^2} - U^{(2)}(u^{(2)})\,[k_2^2 + \alpha_{23}(u^{(2)})k_3^2] = 0 \tag{4b}$$

$$\frac{d^2Z(z)}{dz^2} + Z(z)\,[k_0^2 + k_3^2] = 0 \tag{4c}$$

with $U^{(3)}(u^3) = Z(z)$. In a circular, cylindrical system, $r,\theta,z,\alpha_{23} = 0$, $\alpha_{13} = r^{-2}$; and $k_2^2 = \nu^2$. These conditions reduce the above equations to a more familiar form. Quite evidently, k_i^2 represent the respective eigenvalues, and can only be determined[20] to within $\pm k_i$.

The more familiar, but possibly less rigorous, approach would be the splitting of the Laplacian into $\nabla^2 = \nabla_t^2 + \partial^2/\partial z^2$ in Equation (2), so that it becomes

$$[\nabla_t^2 + (k_0^2 - k^2)]\phi = 0. \tag{5}$$

Here $-k^2$ represents the eigenvalue of $\partial^2/\partial z^2$, and $(k^2 - k_0^2) = -h_0^2$ represents the eigenvalue of ∇_t^2, which by comparison with Equation (4c) equals k_3^2.

The vectorial representation of these eigenvalues is shown in Figure 2, with $\boldsymbol{k}$ in the direction of the z-axis, and $\boldsymbol{h}_0$ transverse to it. In Figure 2b, the components of $\boldsymbol{h}_0$ resolved along the $u^{(1)}$ and $u^{(2)}$ axes are shown.

Since (1) the physical description of the field distribution depends largely on the topology of the system,[1,2] (2) the wave vector is covariant (its image corresponds to oriented surface elements and, consequently, transforms as coordinates of a plane), and (3) the

[19] P. Moon and D. E. Spencer, *Field Theory Handbook*, p. 12ff, Springer Verlag, Berlin, W. Germany, 1961.

[20] R. B. Adler, "Waves on Inhomogeneous Cylindrical Structures," *Proc. IRE*, Vol. 40, p. 339, March 1952.

Helmholtz equation is separable in the coordinate systems under consideration, it is evident that for simplicity, subsequent discussion can be based on cartesian coordinates without any loss of generality. Transformation to other cylindrical, orthogonal systems, and their associated eigenfunctions, can readily be accomplished.[18,21]

It must be remembered that in cylindrical systems, in the absence of perfectly conducting boundaries, both TE and TM modes are generally needed to satisfy boundary conditions, whereas in many rectangular systems either of the modes will satisfy them. This simply means

Fig. 2—Vectorial representation of eigenvalues.

that when transforming from one rectangular coordinate system to another, the wave impedances must be added in parallel.* The algebra becomes quite cumbersome, but this is the only complication. The additional condition needed to evaluate the constants is that the product of the TE and TM wave impedances in any direction equal the square of the characteristic impedance of the medium.†

The Wave Vectors

In a rectangular coordinate system, the wave vector may be re-

---

[21] P. Moon and D. E. Spencer, "Separability in a Class of Coordinate Systems," *Jour. Franklin Inst.*, Vol. 254, No. 3, p. 227, Sept. 1952.

* There is an equivalence between a curved surface with a constant impedance and a plane surface with an equivalent variable impedance that can be shown using the inhomogeneous boundary condition $\partial\phi/\partial n + A\phi = 0$ at the surface ($\partial/\partial n$ is the normal derivative, and A is related to the impedance). In some simple cases, i.e., linearly varying impedance[22] and oblique incidence on a plane, equivalence is practically obvious; but it has not been proven, either rigorously or in general. (H. N. Kritikos, private communication.)

[22] H. N. Kritikos, "Boundary Waves Along an Impedance Plane with a Linearly Varying Impedance," *IEEE Trans. on Antennas and Propagation*, Vol. AP-13, No. 4, p. 577, July 1965.

† See Reference 16, p. 355.

solved along the respective axes, i.e.

$$k_0 = \hat{\imath}k_x + \hat{\jmath}k_y + \hat{k}k,$$

so that

$$k_0^2 = k_x^2 + k_y^2 + k^2 = h_0^2 + k^2. \tag{7}$$

In a lossless system, the k's must be either real or pure imaginary.[20] In either case, however, k_0^2 will be real. With this in mind, it is sufficient to consider the wave characteristics on the basis of $k_0^2 = h_0^2 + k^2$, since an exactly analogous discussion can be then extended to $h_0^2 = k_x^2 + k_y^2$. When the surface under consideration is at x = constant,

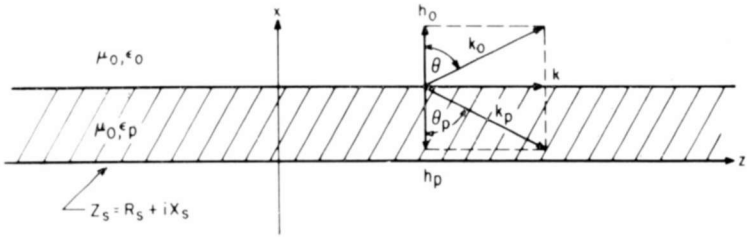


Fig. 3—Propagation vectors associated with a plasma slab on a surface of arbitrary impedance.

if $k_y^2 > 0$ and $k^2 > 0$, the propagation takes place at an angle to the z-axis. If all the k's are real, the wave is a homogeneous plane wave; if some of the k's are imaginary, in *any* orthogonal system, the wave is an inhomogeneous wave with the real and imaginary k's orthogonal to each other.[23,24]

Consider now the system of Figure 3, representing a plasma layer on an arbitrary impedance for which the following relations hold:

$$h_0^2 = k_0^2 - k^2 \tag{8a}$$

$$h^2 = k_p^2 - k^2 = n^2 k_0^2 - k^2 \tag{8b}$$

$$k = k_0 \sin \theta_0 = n k_0 \sin \theta \tag{8c}$$

where $n^2 = (\epsilon/\epsilon_0)$ is the index of refraction of the plasma. Equations

[23] F. J. Zucker, "Electromagnetic Boundary Waves," *Onde Superficiali C.I.M.E.*, Varenna, Italy, 4-13, Sept. 1961.

[24] F. J. Zucker, "The Guiding and Radiation of Surface Waves," *Proc. Symposium on Modern Advances in Microwave Techniques*, Polytechnic Institute of Brooklyn, Brooklyn, New York, p. 403, Nov. 1956.

(8a) and (8b) represent the wave equation (or "separability") in free space and plasma slab, respectively. Equation (8c) is Snell's law. Assuming, for simplicity, that the plasma is cold and collisionless,

$$n^2 = 1 - \frac{\omega_p^2}{\omega^2}. \tag{9}$$

Simultaneous solution of Equations (8) for k_0^2 and k^2 leads to

$$k_0^2(1 - n^2) = h_0^2 - h^2 \tag{10a}$$

$$k^2(1 - n^2) = n^2 h_0^2 - h^2, \tag{10b}$$

so that the phase velocity in the axial direction is

$$\frac{v^2}{c^2} = \frac{k_0^2}{k^2} = \frac{h_0^2 - h^2}{n^2 h_0^2 - h^2}, \tag{11}$$

which must be greater than zero. Two different cases may now be distinguished, $n^2 > 0$ and $n^2 < 0$.

Case of $n^2 > 0$

When $n^2 > 0$, two situations must be considered. The first is when both h_0^2 and h^2 are greater than zero, i.e., the field components are periodic in the transverse direction. Equation (11) can then be written

$$\frac{v^2}{c^2} = \left[1 - \frac{\omega_p^2 h_0^2}{\omega^2(h_0^2 - h^2)} \right]^{-1} \tag{12}$$

with the requirement that $n^2 h_0^2 > h^2$ for real k (see Equation 10(b)). Clearly, $v/c \geq 1$, as shown in Figure 4, curve 1. This situation represents a reflected wave in free space, and a wave-guide-type mode in the plasma. The cutoff condition for this situation occurs when $\omega_p h_0^2/[\omega^2(h_0^2 - h^2)] = 1$, which leads to the following relation:

$$0 < \theta_0 < \theta < \sin^{-1}\left(\frac{\omega}{\sqrt{\omega^2 - \omega_p^2}} \right) < \frac{\pi}{2} \tag{13}$$

with both angles (θ_0, θ) real. Furthermore, the condition $n^2 > 0$, which implies $(\omega_p^2/\omega^2) < 1$, must be changed to the more stringent require-

ment $(\omega_p^2/\omega^2) < 1 - (h^2/h_0^2)$, demonstrating the influence of the eigenvalues.

The second situation arises when $h_0^2 < 0$. In this case it can be said that the necessary and sufficient condition for $k_0^2 > 0$ and $k^2 > 0$ is $h^2 < 0$ and $|h| > |h_0|$. This follows directly from Equations (10).

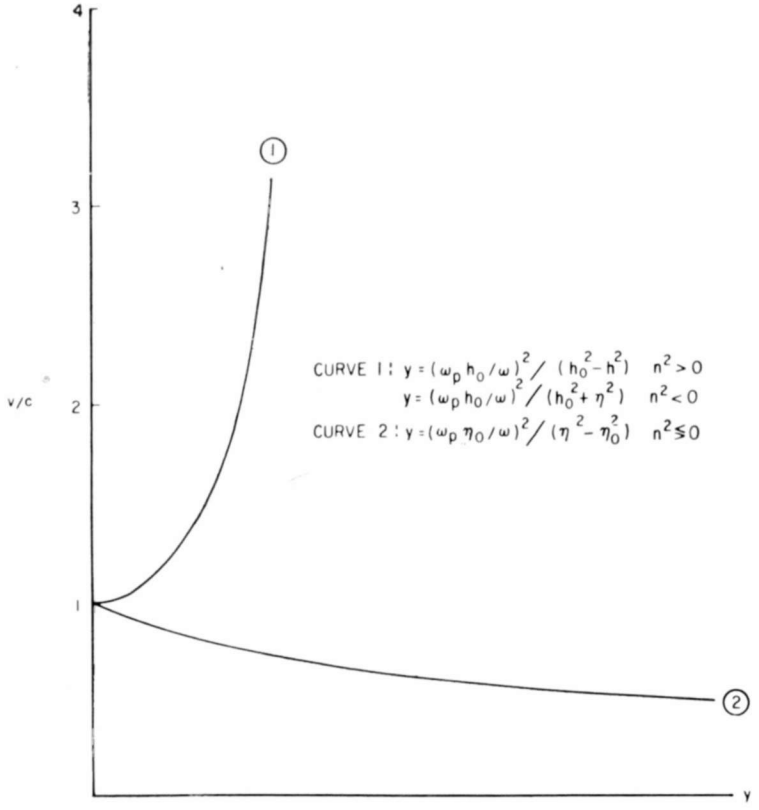


Fig. 4—v/c as a function of the parameter y.

The phase velocity becomes (with $-\eta^2 = h^2$ and $-\eta_0^2 = h_0^2$)

$$\frac{v^2}{c^2} = \left[1 + \frac{\omega_p^2 \eta_0^2}{\omega^2 (\eta^2 - \eta_0^2)} \right]^{-1} \tag{14}$$

and $v/c \lesseqgtr 1$, as shown by curve 2 in Figure 4. Both angles are now complex, with the real part equal to $\pi/2$ radians, i.e. $\theta_0 = (\pi/2) - i\xi_0$, and $\theta = (\pi/2) - i\xi$, and the imaginary parts satisfy

$$\cosh \xi_0 = n \cosh \xi, \tag{15}$$

which is obtained from Equation (8c). This situation represents a boundary-type wave. The interesting aspect of this wave, if it can be supported by the system,* is that $v < c$, even though the phase velocity in the plasma is always greater than c. It is well known[6,8,24] that for this type of boundary wave the plasma phase velocity always lies between the wave-front velocities of the media on either side of the boundary.

It should be pointed out that the situation $h_0^2 < 0$ and $h^2 > 0$ cannot arise in the case of a plasma layer, since this requires[8] that $n^2 > 1$.

Case of $n^2 < 0$

In this case, again, two situations can be distinguished. In one, θ_0 is real and θ complex, so that

$$\frac{v^2}{c^2} = \left[1 - \frac{\omega_p^2 h_0^2}{\omega^2 (h_0^2 + \eta^2)} \right]^{-1} \tag{16}$$

and $v/c \geq 1$ (see curve 1 of Figure 4). The layer now acts as a reflector, with the wave partly penetrating the plasma layer—a behavior similar to that of a waveguide below cutoff. The signal frequency must lie between the values given by the relation

$$1 < \frac{\omega_p^2}{\omega^2} < 1 + \frac{\eta^2}{h_0^2}. \tag{17}$$

The second situation represents a boundary-type wave satisfying conditions described previously by Equations (14) and (15).

Wave Impedances

A plane wave incident on a plasma slab can be conveniently resolved into components parallel and perpendicular to the surface. Thus, the impedance of the slab to the normal component of the incoming wave can be represented by[25]

$$Z_L' = Z_c' \frac{Z_s' \cos \gamma + i Z_c' \sin \gamma}{Z_c' \cos \gamma + i Z_s' \sin \gamma} \tag{18}$$

* This depends on the conditions under which the characteristic equation can be satisfied, i.e., on the relation between wave impedances, which will be discussed next.

[25] S. Ramo and J. R. Whinnery, *Fields and Waves in Modern Radio*, John Wiley and Sons, Inc., 2nd ed., p. 291 (1953).

with $\gamma = hd$, where d is the thickness of the slab and Z_s' is the surface impedance. In the case of E modes, i.e. when the electric field satisfies the relation $\boldsymbol{E} \cdot (\boldsymbol{k}_0 \times \boldsymbol{k}) = 0$, $Z_c' = \sqrt{\mu_0/\epsilon} \cos\theta$; in the case of H modes, $\boldsymbol{E} \times (\boldsymbol{k}_0 \times \boldsymbol{k}) = 0$, $Z_c' = \sqrt{\mu_0/\epsilon} \sec\theta$.[25] Writing out the two impedances explicitly, normalizing with respect to the impedance of free space (i.e., $Z = Z' \sqrt{\epsilon_0/\mu_0}$), and using the relation $\cos\theta = h/(nk_0)$,

$$Z_{LE} = \frac{Z_s \cos\gamma + \dfrac{i\gamma \sin\gamma}{n^2 k_0 d}}{\cos\gamma + iZ_s n^2 k_0 d (\sin\gamma/\gamma)} \tag{19a}$$

$$Z_{LH} = \frac{Z_s \cos\gamma + ik_0 d (\sin\gamma/\gamma)}{\cos\gamma + i\left(\dfrac{Z_s}{k_0 d}\right)\gamma \sin\gamma}. \tag{19b}$$

It is evident that when $\gamma = m\pi$, m being an integer, Z_L is equal to the surface impedance. On the other hand, if $\gamma = (2m+1)\pi/2$, Z_L is proportional to the surface admittance. Furthermore, the general form of the expression for Z_L remains unchanged whether γ is real or imaginary, because the terms $\cos\gamma$ and $(\gamma)^{\pm 1} \sin\gamma$ are real in either case.

The integral form of the reflected part of the solution of the inhomogeneous Holmholtz equation is of the form

$$\phi \propto \int_C R(k)\, g(x,y,k) \exp(ikz)\, dk.$$

C represents a suitable contour (in the case of a Fourier transform it is the real k axis); R is the reflection coefficient

$$R(k) = \frac{Z_L - Z_0}{Z_L + Z_0}; \tag{20}$$

Z_0 is the characteristic wave impedance of the medium above the slab and is of the same form as Z_c with ϵ replaced by ϵ_0 and θ by θ_0. The poles of $R(k)$ represent the isolated singularities of the integrand and thus lead to the determination of the type of waves that the system can support. The inhomogeneous (Leontovitch) boundary conditions are included in $R(k)$ with the proper expressions for the impedances.*

* See Reference (25), p. 298.

Furthermore, $R(k)$ has no branch-points[3,26] unless they are introduced through Z_s. (If Z_s represents the surface impedance, the only singularities of $R(k)$ are poles.)

Thus, the natural behavior of the system (free modes) can now be investigated by consideration of the poles of $R(k)$ in light of the character of the eigenvalues discussed previously.

Impedance of the Plasma-Coated Surface

In general, the impedance Z_L is complex whether or not Z_s is complex. However, if the layer is lossless, Z_L is purely reactive when Z_s is purely reactive. Substituting $Z_s = \pm iX_s$ in Equations (19),

$$Z_{LE} = i \frac{\dfrac{\gamma \sin \gamma}{n^2 k_0 d} \pm X_s \cos \gamma}{\cos \gamma \mp X_s n^2 k_0 d (\sin \gamma / \gamma)} \tag{21a}$$

$$Z_{LH} = i \frac{k_0 d (\sin \gamma / \gamma) \pm X_s \cos \gamma}{\cos \gamma \mp \left(\dfrac{X_s}{k_0 d} \right) \gamma \sin \gamma} . \tag{21b}$$

Z_L will be inductive (capacitive) when Z_s is inductive (capacitive) and γ lies in the range $m\pi < \gamma < \pi(2m+1)/2$ whether n^2 is less or more than 0; Z_L will be capacitive (inductive) otherwise (i.e., $\pi(2m-1)/2 < \gamma < m\pi$). Since an inductive impedance[15,24] is required for boundary waves with $(v/c) < 1$, the range of γ for a given Z_L can readily be determined.

Imaginary values of Z_L imply that Z_0 must also be pure imaginary if $R(k)$ is to have poles.* This, in turn, means that the angle of incidence, θ_0, is complex and the system will support an inhomogeneous wave with k real and h_0 imaginary. Consequently, in accordance with previous discussion, h must be imaginary and Equations (14) and (15) will hold whether $n^2 > 0$ or $n^2 < 0$. If we let $\gamma = \pm i\beta$, then for Z_{LE} to be inductive, it is necessary that $\coth \beta > X_s$ when $n^2 > 0$ and

[26] O. Aboul-Atta, *Investigation of Necessary and Sufficient Conditions for Uniqueness of Solution of a Class of Problems in E-M Theory*, Dissertation in Electrical Engineering, University of Pennsylvania, Philadelphia, Pa. (1965).

* Leaky waves,[23,24] connected with complex values of h even in the presence of lossless media, are not considered here. They may arise, however, as a result of curvature[5] introduced when the system is transformed. A simple example, transformation to circular cylindrical coordinates, demonstrates the point.

that $Z_s = iX_s$. Regardless of the value of β, Z_{LE} will be inductive when $n^2 < 0$, and Z_{LH} will be inductive for all n^2.

The cases for $Z_s = 0$ or $Z_s \to \infty$ have been studied[4,14] and the range of allowed values of n^2 and h for surface waves are well known. They can be very easily obtained from Equations (19) and (20).

When Z_s is either real or complex, the system becomes dissipative, and the character of the waves is changed because the eigenvalues become complex. The degree of change depends on the relative magnitudes of the real and imaginary parts of the eigenvalues. It is evident from Equations (19) and (20) that in all but the simplest cases, the roots of $R(k) = 0$ must be determined by a computer.

Conclusions

The simple analysis presented shows that in the case of a cold collisionless plasma layer, a waveguide-type mode exists in the layer when $\omega_p^2/\omega^2 < 1 - (h^2/h_0^2)$ (a condition more stringent than $\omega_p^2/\omega^2 < 1$) and that reflected waves propagate simultaneously in the space above the layer. When the system supports surface waves, their characteristics are somewhat different from those supported by dielectric layers. In the dielectric case, the phase velocity of the surface wave lies between the wave-front velocities of the media on either side of the boundary, while in the plasma case, the surface-wave phase velocity is always less than the speed of light, even though the plasma phase velocity may become infinite. Further, it has been shown that the nature of the waves, and the range of the corresponding eigenvalues, can be obtained rather easily for a plasma-coated surface by consideration of relations between the wave vectors (eigenvalues) and pertinent impedances, thus avoiding complex mathematical techniques. The generality of this approach has been discussed on the basis of the mathematical formalism of Maxwell's and Helmholtz' equations, and its validity was justified. Extension of this work by the application of coordinate transformations is suggested.

Acknowledgment

The encouragement of and profitable discussions with A. Boornard and J. Vollmer during considerations of various problems associated with this work, and during the preparation of the manuscript, are gratefully acknowledged.

CERENKOV INTERACTION IN A UNIFORMLY LOADED WAVEGUIDE*

BY

J. J. STEKERT

Summary—*An analysis and computer experiment are carried out on a simple, circularly symmetrical waveguide, homogeneously filled with a uniform dielectric, which is undergoing Cerekov interaction with a flowing density-modulated electron stream. The dispersion (ω,k) characteristics of the system reveal the presence of two complex root loci for each mode, one denoting evanescence and the other convective instability and gain associated with coupling of the slow space-charge wave to the waveguide mode. The latter loci extend from the region of synchronism for each mode to the immediate vicinity of the dispersion plane origin, thereby presenting an extended frequency region for gain. The gain of the convective instability, however, does peak in the region of mode synchronism. The peak gain is found to be approximately proportional to the cube root of the average beam current, as is the bandwidth at half-peak gain. Peak gain also maximizes as a function of the threshold parameter, βn. The Cerenkov threshold condition, $\beta n = 1$, and the characteristic radiation angle are found to hold for the system. It is concluded that aside from the possible implications of the extended gain region, the differences between this interaction and those of more conventional traveling-wave devices are not fundamental, but are reflections of the specific difference in the properties of the slowing circuits.*

INTRODUCTION

THE phenomenon known as Cerenkov radiation was established with the series of experiments conducted by P. H. Cerenkov from 1934 to 1938,[1] and the concurrent theory proposed by Frank and Tamm in 1937.[2] In 1947, V. L. Ginsburg[3] suggested that such an interaction might be useful as a basis for microwave generation, and subsequently, the search for new techniques has led numerous researchers in this direction.

Prior to 1955, Cerenkov radiation had been viewed predominantly as a single charge interaction, suitable for use in high-energy particle

* This work was partially supported by the United States Air Force, Rome Air Development Center, under Contract No. AF30(602)-3061.

[1] For a list of Cerenkov's publications, see J. V. Jelley, *Cherenkov Radiation and Its Applications*, p. 290, Pergamon Press, 1958.

[2] I. M. Frank and Ig. Tamm, "Coherent Visible Radiation from Fast Electrons Passing Through Matter," *Dokl. Akad. Nauk. SSSR*, London, Vol. 14, p. 109, 1937.

[3] V. L. Ginsburg, *Dokl. Akad. Nauk, SSSR*, Vol. 56, p. 253 (1947a).

detection. Soon after, such investigators as J. R. Pierce,[4] J. G. Linhart,[5] and G. Mourier[6] suggested that Cerenkov radiation and traveling-wave interactions by extended electron beams bore many similarities. More recently, several people have examined various aspects of such a relationship[7–10] and some have delineated differences that distinguish the two types of interaction.[11,12]

In this paper, we study the dispersion properties of a simple, circularly symmetrical, cylindrical system undergoing Cerenkov interaction with a flowing, modulated electron stream. A derivation and analysis of the dispersion equation is followed by a detailed series of computer solutions and their interpretation. As such, this may be regarded as a computer experiment on a theoretical system. While the system itself has no physical realization at present, it does relate to Cerenkov devices actually studied in the laboratory. Thus, it not only provides an insight into the properties of the interaction, but it also yields information about the performance capabilities of the laboratory devices.

Formulation of the Problem

Consider the system shown in Figure 1. It is composed of a circular, cylindrical, perfectly conducting waveguide of radius b and uniformly filled with a lossless, nonmagnetic, homogeneous dielectric. A current of uniform transverse cross section is constrained to move through the dielectric only in the longitudinal direction with speed u. Assume all field quantities to be proportional to

---

[4] J. R. Pierce, "Interaction of Moving Charges with Wave Circuits," *Jour. Appl. Phys.*, Vol. 26, p. 627, May 1955.

[5] J. G. Linhart, "Cerenkov Radiation of Electrons Moving Parallel to a Dielectric Boundary," *Jour. Appl. Phys.*, Vol. 26, p. 527, May 1955.

[6] G. Mourier, "L'effect Cerenkov et la generation d'ondes hertziennes," *L'Onde Electrique*, Vol. 38, p. 132, Feb. 1958.

[7] B. W. Hakki, "Cerenkov Radiation in Slow-Wave Guiding Structures," *Trans. IRE Elec. Dev.*, Vol. ED8, p. 497, Nov. 1961.

[8] C. H. Durney and R. W. Grow, *Cerenkov Microwave Radiation to an Infinite Dielectric Medium*, Technical Report ONR-2, Microwave Devices Laboratory, University of Utah, May 1, 1961.

[9] C. H. Durney and R. W. Grow, *Traveling-Wave Tubes with Nonperiodic Dielectric Circuits*, Technical Report ONR-7, Microwave Devices Laboratory, University of Utah, May 1, 1964.

[10] P. D. Coleman, "Cerenkov Radiation and Allied Phenomena," *Proc. Symposium on Quasi-Optics*, Vol. XIV, p. 199, Polytechnic Institute of Brooklyn, June 1964.

[11] I. Palócz and A. A. Oliner, "Leaky Space-Charge Waves I: Cerenkov Radiation," *Proc. IEEE*, Vol. 53, p. 24, Jan. 1965.

[12] J. J. Stekert, "Travelling Wave Characteristics of Cerenkov Interaction," *Proc. IEEE*, Vol. 53, p. 1237, Sept. 1965.

$$e^{i(\omega t - kz)},$$

where ω = angular frequency, k = longitudinal wave number, and where we have chosen the longitudinal direction to be the z axis. Since only longitudinal interactions of the beam are to be permitted, we will limit consideration to the TM modes of the system.

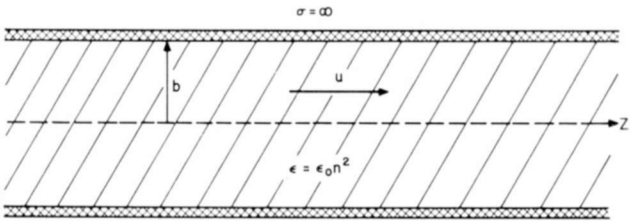


Fig. 1—The homogeneously loaded waveguide.

By straightforward manipulation of Maxwell's equations, the differential equation for the z component of the electric field inside the waveguide is obtained in terms of source current and charge density;

$$[\nabla_t^2 + (k_0^2 n^2 - k^2)]E_z = -i\frac{k}{\epsilon_0 n^2}\rho + i\omega\mu_0 j, \tag{1}$$

where

E_z = the z component of the electric field,

∇_t^2 = transverse Laplacian operator,

$k_0 = \dfrac{\omega}{c}$ = free space wave number,

$i = \sqrt{-1}$,

n = index of refraction of the dielectric,

ρ = source charge density,

j = source current density,

ϵ_0 = permittivity of free space,

μ_0 = permeability of free space.

To evaluate the source terms, it will be assumed that E_z, u, ρ and j

can be represented in the form

$$\alpha = \alpha_0 + \alpha_1 e^{i(\omega t - kz)}, \tag{2}$$

where

$$|\alpha_1| << |\alpha_0| \quad \text{and} \quad \alpha_0 = \text{constant}.$$

If we let $j = u\rho$ and utilize the equation of continuity with the Lorentz force equation, application of the small-signal assumption of Equation (2) allows a separation of the d-c terms from the smaller a-c terms. Further, neglecting small quantity terms of second order, we find that

$$j_1 = \rho_0 u_1 + \rho_1 u_0, \tag{3}$$

$$-kj_1 + \omega\rho_1 = 0, \tag{4}$$

$$\omega u_1 - u_0 k u_1 = -\frac{e}{im} E_z', \tag{5}$$

where

e = charge of the electron

m = mass of the electron

E_z' = a-c electric field on the moving electrons.

In Equation (5), the magnetic force term has been dropped because of the constraint to one-dimensional beam flow.

From Equations (3), (4), and (5), it is found that

$$\frac{-ik}{\epsilon_0 n^2}\rho_1 + i\omega\mu_0 j_1 = [k_e^2(\beta n)^2 - k^2]\left[\frac{k_{p0}^2}{(\beta n)^2(k_e - k)^2}\right] E_z', \tag{6}$$

where

$$k_e = \frac{\omega}{u_0}, \tag{7}$$

$$\beta = \frac{u_0}{c}, \tag{8}$$

$$k_{p0}^2 = \left(\frac{\omega_p}{c}\right)^2 = \frac{-e\rho_0}{\epsilon_0 m}\frac{1}{c^2}. \tag{9}$$

If E_z' is self-consistently assumed to be the same field as E_z of Equation (1), we can write

$$[\nabla_t^2 + h^2]\, E_z = 0, \tag{10}$$

where

$$\begin{aligned} h^2 &= [k_e^2(\beta n)^2 - k^2]\,[1 - K^2], \\ K^2 &= \frac{k_{p0}^2}{(\beta n)^2 (k_e - k)^2}. \end{aligned} \tag{11}$$

By choosing circular cylindrical coordinates, the field solutions of Equation (10) within the waveguide become straightforward. Considering only cylindrically symmetric solutions, they are

$$E_z = AJ_0(hr)\, e^{i(\omega t - kz)} \tag{12}$$

$$E_r = A \frac{ikh}{k_e^2(\beta n)^2 - k^2} J_1(hr)\, e^{i(\omega t - kz)} \tag{13}$$

$$H_\theta = A \frac{i\omega\epsilon_0 n^2 h}{k_e^2(\beta n)^2 - k^2} J_1(hr)\, e^{i(\omega t - kz)}. \tag{14}$$

where J_0, J_1 are Bessel functions.

Since Equation (10) is homogeneous, the arbitrary constant A cannot be determined without explicit source information or initial conditions. However, at the waveguide boundary, $E_z(b) = 0$ and, therefore,

$$J_0(hb) = 0. \tag{15}$$

The dispersion relation for the system then becomes

$$[k_e^2(\beta n)^2 - k^2]\left[1 - \frac{k_{p0}^2}{(\beta n)^2 (k_e - k)^2}\right] = \left(\frac{p_m}{b}\right)^2, \tag{16}$$

where p_m is the m^{th} zero of the Bessel function $J_0(X)$.

Analysis of the Dispersion Equation

Solutions of Equation (16) for real $k_e(\omega)$ and both real and complex $k = k_r + ik_i$ will reveal the wave propagation properties of the system. The loci of solutions for real k will give an indication of the

types of mode coupling present and delineate the regions of the dispersion plane where solutions exist for complex k, while the imaginary parts of the solutions for complex k will give specific information about the gain or evanescence that may be present in the wave propagation.

The dispersion plane (k_e,k_r) with its characteristic lines is shown in Figure 2. We will consider only the upper half plane, although the solutions will be symmetric about the origin. The lines $k_e = \pm k_r/(\beta n)$ are the asymptotes for the unperturbed wave-guide modes. They

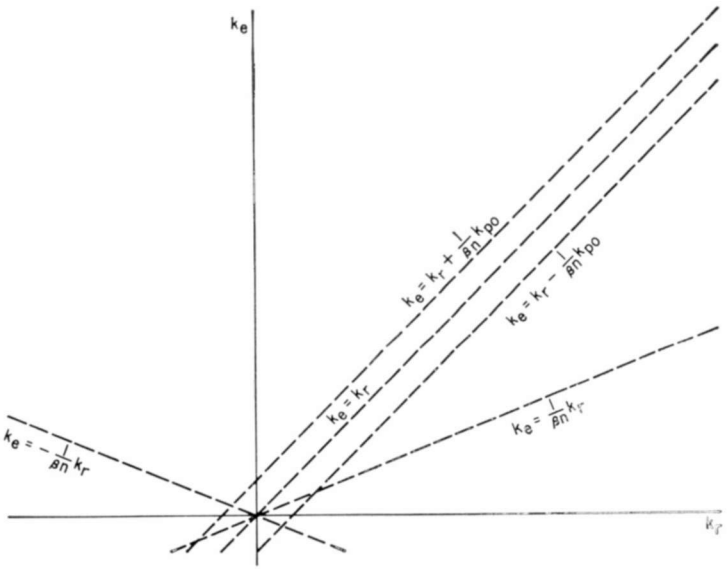


Fig. 2—The characteristic lines of the dispersion plane.

represent the dispersion properties of plane waves propagating in the $\pm z$ directions in a medium with an index of refraction n.

The lines $k_e = k_r \pm [k_{p0}/(\beta n)]$ are the asymptotes for the fast and slow space-charge modes of the uncoupled electron beam. Here they will be called the Hahn–Ramo lines,[13,14] and the region of the dispersion plane contained between them will be called the channel. Finally, the line $k_e = k_r$ is the beam synchronous line, along which a wave is in synchronism with the average velocity of the electron beam. Reference to these characteristic lines will simplify the analysis of Equation (16).

[13] W. C. Hahn, "Small Signal Theory of Velocity Modulated Electron Beams," *Gen. Elec. Rev.*, Vol. 42, p. 258, June 1939.

[14] S. Ramo, "Space Charge and Field Waves in an Electron Beam," *Phys. Rev.*, Vol. 56, p. 276, Aug. 1, 1939.

Regions for Real Solutions

Some information about the location of the dispersion solutions can be gained immediately by inspection of Equations (11) and (15). Inasmuch as the modified Bessel function $I_0(y) = J_0(iy)$ has no finite zeros, it is evident that for Equation (15) to hold, h must be real when k is real. Equation (11) shows that outside the channel this is the case only on the fast-wave side of the lines $k_e = \pm k_r/(\beta n)$, where

$$k_e^2(\beta n)^2 - k^2 > 0 \quad \text{and} \quad 1 - K^2 > 0,$$

and inside the channel only on the slow-wave side of the same lines, where

$$k_e^2(\beta n)^2 - k^2 < 0 \quad \text{and} \quad 1 - K^2 < 0.$$

Therefore, solutions for real k must be confined to these regions.

A glance at Figure 2 will show that the resulting configuration can be expected to qualitatively differ in accordance with whether βn is greater than or less than unity. When $\beta n > 1$, the main area of the channel extends into the fast-wave region, thus possibly causing a bifurcation of real solutions that may exist there.* When $\beta n < 1$, the channel lies almost entirely in the slow-wave region, thereby allowing the possibility of real solutions over its entire length, as well as over the entire fast-wave region, except near the origin.

Qualitative Form of the Solutions

For a more detailed picture of the resulting dispersion configuration, the behavior of Equation (16) can be investigated for real solutions in various parts of the dispersion plane of Figure 2. We will consider first the case when $\beta n > 1$. In the fast-wave region, far removed from the channel, $1 - K^2 \to 1$, and Equation (16) becomes

$$k_e^2(\beta n)^2 - k_r^2 = \left(\frac{p_m}{b}\right)^2. \tag{17}$$

This is the equation for unperturbed hyperbolic waveguide modes,[15] one corresponding to each p_m. Thus, for large $|k_r|$ in the regions far from the channel, the real solutions become asymptotic to the lines

* Fast- and slow-wave regions are defined here relative to the lines $k_e = \pm k_r/(\beta n)$.

[15] S. Ramo and J. R. Whinnery, *Fields and Waves in Modern Radio*, 2nd ed., John Wiley and Sons, 1953, p. 374ff.

$k_e = \pm k_r/(\beta n)$. As one moves in along the waveguide modes toward the channel from each side, and hence closer to the synchronous line, $(k_e - k_r)$ becomes smaller, as does $1 - K^2$. The quantity $k_e^2\,(\beta n)^2 - k_r^2$ must then increase to maintain the solution. Hence, the solution locus moves up along each side of the channel. As it proceeds along the channel and away from the waveguide asymptotic lines, $k_e^2\,(\beta n)^2 - k_r^2 \to k_e^2(\beta n)^2$. Thus, in the limit as $k_e \to \infty$, $1 - K^2 \to 0$, and the solutions asymptotically approach the Hahn–Ramo lines from each side of the channel. The resulting configuration for one mode is shown qualitatively in Figure 3a.

To investigate the real solutions in the allowed region near the origin, it is convenient to rewrite Equation (16) as a quartic equation in k.

$$k^4 - 2k_e k^3 - \left[(\beta n)^2 k_e^2 - k_e^2 + \left(\frac{k_{p0}}{\beta n}\right)^2 - \left(\frac{p_m}{b}\right)^2 \right] k^2 \tag{18}$$

$$- 2\left[\left(\frac{p_m}{b}\right)^2 k_e - (\beta n)^2 k_e^3 \right] k - \left[(\beta n)^2 k_e^4 - k_{p0}^2 k_e^2 - \left(\frac{p_m}{b}\right)^2 k_e^2 \right] = 0.$$

It is at once evident that a real solution exists at the origin. Substitution of either $k_e = 0$ or $k_r = 0$ shows that $k_e = k_r = 0$ is a double root. Furthermore when $k = 0$,

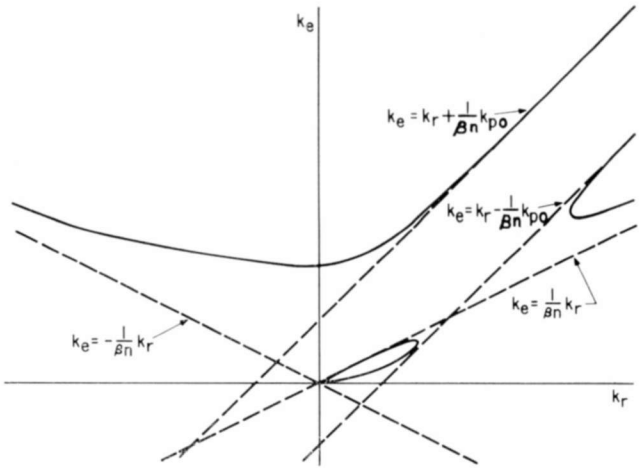
$$k_e = \pm \frac{1}{\beta n} \left[k_{p0}^2 + \left(\frac{p_m}{b}\right)^2 \right]^{1/2}. \tag{19}$$

This is the ordinate intercept for the real locus discussed in the preceding paragraph.

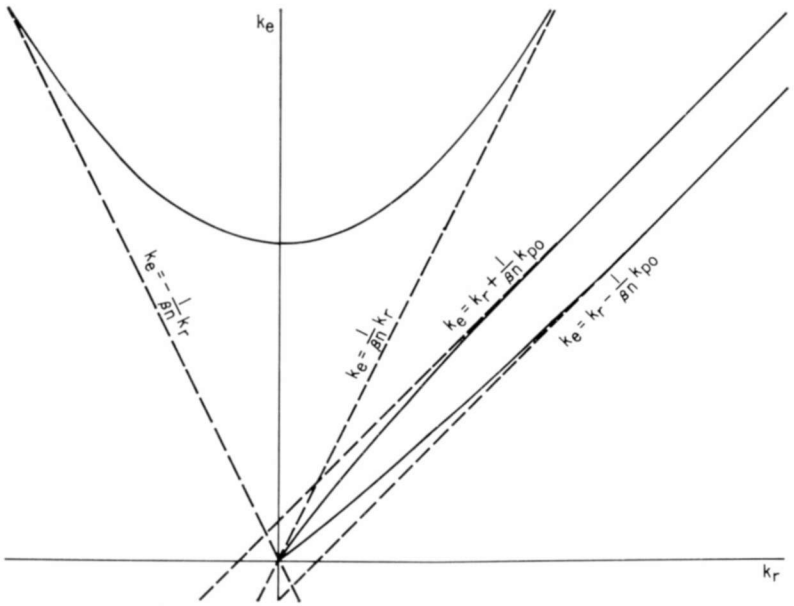
In the immediate region of the origin, both k_e and k_r are small. We thus ignore all k_e and k_r terms in Equation (18) of greater than second order and get

$$k_r^2 + \frac{2\left(\frac{p_m}{b}\right)^2 k_e k_r - \left[k_{p0}^2 + \left(\frac{p_m}{b}\right)^2 \right] k_e^2}{\left(\frac{k_{p0}}{\beta n}\right)^2 - \left(\frac{p_m}{b}\right)^2} = 0. \tag{20}$$

It is now evident that for real solutions to exist in this region, other



(a)



(b)

Fig. 3—Qualitative real mode solution (a) for $\beta n > 1$ and (b) for $\beta n < 1$.

than at the origin,

$$(k_{p0}b)^2 \geqq [(\beta n)^2 - 1]p_m^2. \tag{21}$$

Under these conditions, and in light of the double root at the origin, it appears that a loop of real solutions exists in the restricted region near the origin. It is further evident that the loops become smaller for successively higher modes, until when p_m is large enough, the locus is a degenerate double root point at the origin. For usually obtainable physical cases, $k_{p0}b$ is small, so that there is no loop present even for the lowest-order mode except when βn is close to unity. A loop is included in Figure 3a to complete the qualitative picture.

When $\beta n < 1$, a different configuration occurs. Since the channel does not penetrate into the fast-wave region except near the origin, most of the fast-wave region is far enough from the channel so that Equation (17) holds to a good approximation. Thus, the real solutions in this region are slightly perturbed, hyperbolic waveguide modes.

In the vicinity of the origin, Equation (19) is valid and the inequality of Equation (21) now holds under all conditions. Furthermore, there is always a double real root at the origin. Therefore, there are now always two real loci diverging from the origin. As one moves far out into the channel, $[k_r^2 - k_e^2(\beta n)^2] \cong k_r^2 \to \infty$ and, hence, $1 - K^2 \to 0$. Thus at high frequency, the pair of mode solutions become asymptotic to the fast and slow Hahn–Ramo lines.

The configuration is qualitatively shown for one order of mode in Figure 3b. It should be noted that for higher-order modes, and hence larger p_m, the fast and slow space-charge modes lie closer to the synchronous line, with the same asymptotic behavior.

Physical Interpretation

The configuration shown within the channel in Figure 3b is characteristic of the reduced fast and slow space-charge waves propagated by a finite electron beam in a drift tube.[16] Aside from the usually small perturbation caused by their mutual presence, these modes are independent of the waveguide modes appearing in the fast-wave region. Hence, no coupling of modes occurs between the beam and the waveguide, and no gain of propagating waves is to be expected when $\beta n < 1$.

When $\beta n > 1$, the channel lies in the fast-wave region. Here, mode synchronism and coupling between the space-charge and waveguide modes can and does occur, as is evidenced in Figure 3a. As βn becomes greater than unity, the portions of the fast and slow space-charge

[16] R. W. Gould and A. W. Trivelpiece, "A New Mode of Wave Propagation on Electron Beams," *Proc. Symposium on Electronic Waveguides*, Vol. VII, 215-228, Brooklyn Polytechnic Institute, April 1948.

modes confined to the slow-wave region in the channel couple to form the lower loop, and as βn becomes larger, they degenerate to the double root point at the origin. In the fast-wave region, the fast and slow beam modes couple to the wave guide mode to form, respectively, the left- and right-hand real loci shown in that area. The right-hand locus, in particular, indicates the presence of convective instability[17] and, hence, the possibility of achieving traveling-wave gain in the system.

It should be observed that the lower portion of the convective instability configuration, which according to Sturrock's criterion[17] delineates the lower limit of the gain region, is confined in this case to the slow-wave region inside the channel, or indeed, to the origin itself. Consequently, the region for gain should extend from near the origin to the area near which the channel intersects the wave-guide mode. This extension should be greater for each successively higher mode. While the implications of such nonsymmetric separation of the coupled configuration from the apparent point of mode synchronism are not completely understood at present,* it is clear that the constraint of the lower portion by the line $k_e = k_r/(\beta n)$ is a direct consequence of the propagation properties of the uniformly continuous dielectric medium that fills the waveguide and acts as a nonperiodic slowing circuit. As a result of the lack of periodicity, no space harmonics, and hence no backward waves, are present for coupling to the beam. Furthermore, the dielectric provides for synchronism at any beam velocity for which $\beta n > 1$. Thus, the condition $\beta n = 1$ represents threshold for the interaction, as it does in the more familiar case of Cerenkov interaction with an unbounded dielectric.[18]

Still another connection with familiar properties of Cerenkov radiation can be seen. As a result of coupling between the slow space-charge and waveguide modes, it is reasonable to expect that the strongest interaction would occur near the point of mode synchronism. Such will indeed prove to be the case. A wave generated at such a dispersion point will, therefore, approximately satisfy the equations

$$k_e^2(\beta n)^2 - k_r^2 = \left(\frac{p_m}{b}\right)^2, \tag{22}$$

$$k_e = k_r - \frac{1}{\beta n} k_{p0}. \tag{23}$$

[17] P. A. Sturrock, "Kinematics of Growing Waves," *Phys. Rev.*, Vol. 112, p. 1488, Dec. 1, 1958.

* Some speculation on the mode coupling involved in this problem appears in Reference (12).

[18] J. V. Jelly, *Cherenkov Radiation and Its Applications*, Chapter 1, Pergamon Press, 1958.

Equation (22) is, in effect, a Pythagorean relationship between the longitudinal wave number, k_r, the radial wave number, p_m/b, and the wave number $k_e\beta n = \omega n/c$ for a free wave in the dielectric medium. This relationship is illustrated in Figure 4.

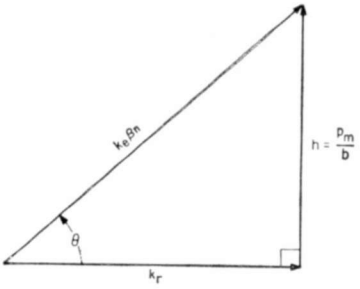


Fig. 4—Relationship among the propagation vectors.

We can then write

$$\cos\theta = \frac{k_r}{k_e\beta n}. \tag{24}$$

Substitution of Equation (23) for k_r then gives

$$\cos\theta = \frac{1 + \dfrac{1}{\beta n}\dfrac{k_{p0}}{k_e}}{\beta n}. \tag{25}$$

Since, for physically realizable conditions, $(1/\beta n)\,(k_{p0}/k_e) \ll 1$, we can write, to a good approximation,

$$\cos\theta = \frac{1}{\beta n}. \tag{26}$$

Equation (26) is the well known angular dependence of Cerenkov radiation* that emanates from a cylindrical source as a conical wave in phase synchronism with the generating electron beam. Such waves are the natural symmetric propagating modes of the cylindrical waveguide.† Slowing by the dielectric medium permits their coupling to the electron beam.

* See Reference 18, Chapter 1.
† See Reference 15, p. 329ff.

Complex Solutions

As is evident from Equation (18), the dispersion relation is an algebraic equation of fourth degree in k. Thus, for any choice of k_e there should be four values of k that satisfy the equation for each mode. At high frequencies, all four roots are real. However, there must be two pairs of complex conjugate root loci between the upper real loci and those near the origin. The real part of one must emanate from the lowest point of the real locus to the left of the channel, while the other must proceed from the lowest point of the right-hand real locus. Both imaginary parts rise from zero along the loci. The presence of the real solutions near or at the origin will cause the imaginary parts of at least one to peak and then fall to zero as it reaches the lower branch point.

To determine the behavior of the complex solutions at the origin, let $k = k_r + ik_i$ and set $k_e = k_r = 0$. Equation (18) then becomes

$$k_i^4 + \left[\left(\frac{k_{p0}}{\beta n}\right)^2 - \left(\frac{p_m}{b}\right)^2\right] k_i^2 = 0. \tag{27}$$

For this equation,

$$k_i = \pm 0$$

$$k_i = \pm\left[\left(\frac{p_m}{b}\right)^2 - \left(\frac{k_{p0}}{\beta n}\right)^2\right]^{1/2}$$

are the roots. We can now identify the loci and the appropriate sign of the imaginary parts. In the absence of the beam ($k_{p0} \to 0$), Equation (17) holds and represents unperturbed waveguide modes. Beginning at the cutoff frequency of each mode and proceeding down the frequency axis of such a dispersion plane are a locus of pure imaginary solutions representing evanescent waves. These arise at cutoff with $k_i = 0$, and peak at the origin at a value of $k_i = \pm(p_m/b)$. The negative sign corresponds to waves propagating in the $+z$ direction. We can thus identify the solution of Equation (27)

$$k_i = \pm\left[\left(\frac{p_m}{b}\right)^2 - \left(\frac{k_{p0}}{\beta n}\right)^2\right]^{1/2}$$

as the value at the origin of this evanescent locus in the presence of the electron beam. In this case the perturbation has displaced the locus into the left half dispersion plane and the solution is now complex.

The negative imaginary part is still identified with the $+z$ propagating wave. It should be noted that for the usual physical case, $k_{p0}b \ll p_m$ and the perturbation is not great, especially when $\beta n > 1$. However, it is possible, particularly when $\beta n < 1$, for the imaginary part of the evanescent locus to peak, and then, when $(k_{p0}b)^2 \geqq (p_m\beta n)^2$, to fall to zero at or near the origin.

The other roots, $k_i = \pm 0$, are then the values of the imaginary part of the lower terminal of the complex locus of the convective instability, as expected. The positive sign is thus the correct choice for the locus, giving the gain of the wave as it progresses in the $+z$ direction.

THE COMPUTER EXPERIMENT

For quantitative solutions to Equation (16), use was made of a digital computer* and automatic root-tracing procedures.[19-21] Runs were made for several sequences of parametric variation.

General Considerations

A convenient presentation may be made by normalizing the variables of Equation (16) to the waveguide radius. Multiplying through by b^2 gives

$$[(k_e b)^2(\beta n)^2 - (kb)^2]\left[1 - \frac{(k_{p0}b)^2}{(\beta n)^2(k_e b - kb)^2}\right] = p_m{}^2. \tag{28}$$

By using the relationship for total average beam current

$$I = \pi b^2 \rho_0 u_0 \tag{29}$$

and referring to Equation (9), we find

$$k_{p0}b = \left[\frac{eI(1-\beta^2)^{1/2}}{m_0\epsilon_0\pi c^3\beta}\right]^{1/2}$$

---

* In this case, the RCA 601 at the David Sarnoff Research Center, Princeton, N. J.

[19] R. W. Klopfenstein, "Zeros of Nonlinear Functions," *Jour. Assoc. Computing Machinery*, Vol. 8, p. 366, July 1961.

[20] R. L. Crane and R. W. Klopfenstein, "A Predictor-Corrector Algorithm with an Increased Range of Absolute Stability," *Jour. Assoc. Computing Machinery*, Vol. 12, p. 227, April 1965.

[21] T. C. Hilinski and R. W. Klopfenstein, Private Communication.

$$= \left[2.3467 \times 10^{-4} I \left(\frac{1}{\beta^2} - 1 \right)^{1/2} \right]^{1/2} \tag{30}$$

Here we have taken into account the relativistic mass of the electron,

$$m = m_0 (1 - \beta^2)^{-1/2}.$$

Equation (28) can now be solved for the variables $[(k_e b), (kb)]$, in terms of the parameters (βn), $(k_{p0} b)$ and p_m. Expressing $(k_{p0} b)$ in terms of the physically more meaningful quantities I and βn, however, will require further specification of n or β.

The Dispersion Plot and the Complex Loci

Examples of the lowest-order-mode solutions $[(k_e b), (k_r b)]$ are shown in Figure 5a for $\beta n > 1$ and in Figure 5b for $\beta n < 1$. The parameters chosen are:

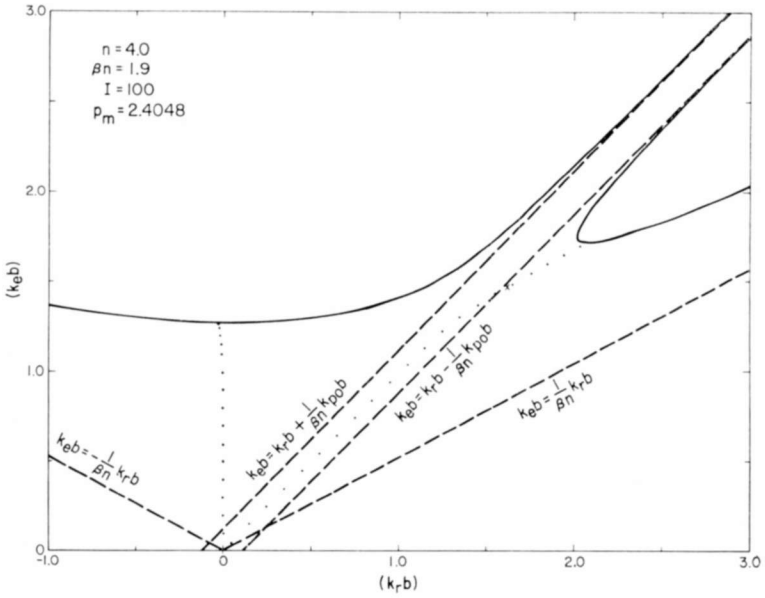
$n = 4.0$	$n = 4.0$
$\beta n = 1.9$	$\beta n = 0.5$
$I = 10^2$ amperes	$I = 10^2$ amperes
$p_m = 2.404826$	$p_m = 2.404826$

With the exception of the current, the parameters are realizable in laboratory experiments. The unusually large current was chosen to widen the channel for better viewing of details.

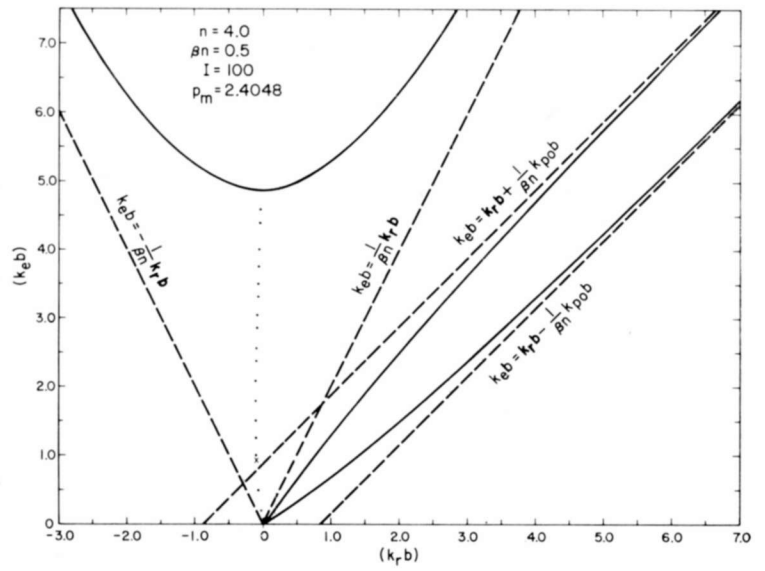
It can be seen that Figure 5a agrees closely with the configuration of Figure 3a. As expected from Equation (21), the lower part of the real locus is a double root point at the origin. The dotted lines extending from the upper bunches are the real loci of the expected complex solutions. The point marked by an x, just outside the channel on the connective instability or gain locus, denotes the point having the peak value of $(k_i b)$. The corresponding values of $k_i b$ are plotted as a function of $k_e b$ for this locus in Figure 6a. The value marked Δ in the figure represents the bandwidth for half peak gain. A plot of $k_i b$ versus $k_e b$ for the evanescent locus is shown in Figure 6b.

Figure 5b, like Figure 3b, shows the characteristic dispersion configuration of a beam in a drift tube. The great width of the channel, as compared with Figure 5a for the same current, results from the smaller implied value of β.

Physically, the width of the channel is a measure of the charge density of the beam as reflected in k_{p0}. At the lower velocity, the beam must be of greater density to produce the same current as when $\beta n = 1.9$.



(a)



(b)

Fig. 5—Computer dispersion plane solution (a) for $n = 4.0$, $\beta n = 1.9$, $I = 10^2$, and $p_m = 2.4048$, and (b) for $n = 4.0$, $\beta n = 0.5$, $I = 10^2$, and $p_m = 2.4048$.

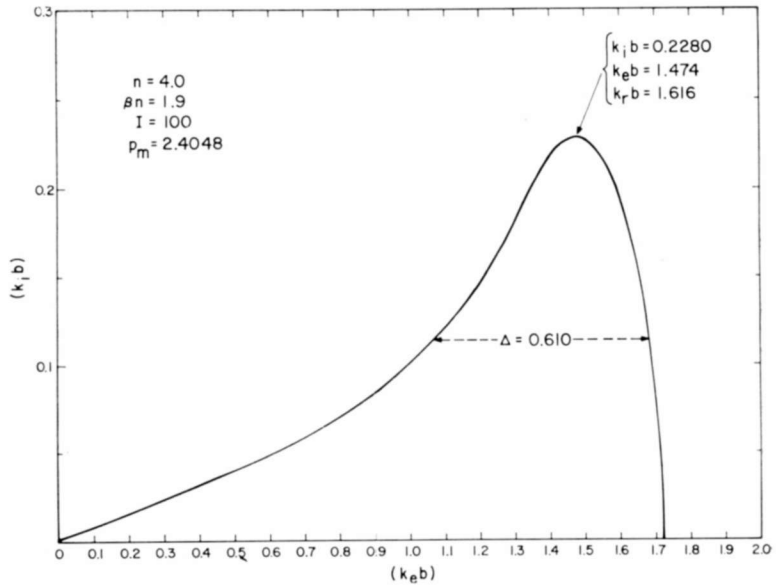


Fig. 6a—Gain versus frequency characteristic for $n = 4.0$, $\beta n = 1.9$, $I = 10^2$, and $p_m = 2.4048$.

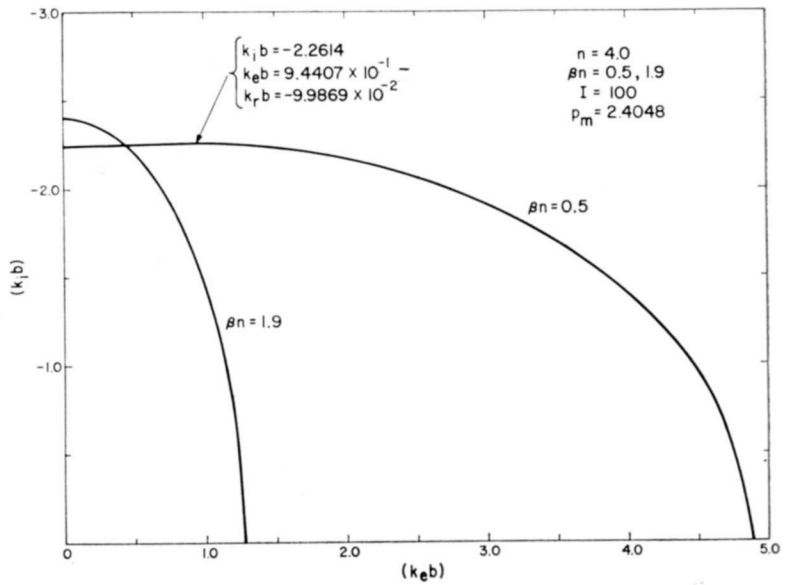


Fig. 6b—Evanescent solutions versus frequency for $\beta n = 1.9$ and $\beta n = 0.5$, with $n = 4.0$, $I = 10^2$, and $p_m = 2.4048$.

It should be noted that even though no mode coupling is present, the presence of the beam has produced a slight asymmetry of the real locus of the waveguide mode and a marked displacement of the real part of the evanscent locus from the $k_e b$ axis. A plot of $k_i b$ versus $k_e b$ for this locus is shown in Figure 6b. There is a slight peak in this plot, in accordance with the discussion of Equation (27). The position of the peak value in Figure 5b is indicated by a cross.

Variation of Average Beam Current

For the case $n = 4.0$, $\beta n = 1.9$, $p_m = 2.404826$, the complex gain locus has been computed for values of I in powers of 10 from 10^{-2} to 10^2 amperes. Plots of $k_i b$ versus $k_e b$ for these loci are shown in Figure 7a. The values of $k_i b$, $k_e b$, $k_r b$, and $k_i b / k_r b$ for the peak values of $k_i b$

Table I—Peak Gain and Associated Characteristics as a Function of Average Beam Current for $n = 4.0$, $\beta n = 1.9$, and $p_m = 2.4048$.

I	$k_i b$	$k_e b$	$k_r b$	Δ	$\frac{k_i b}{k_r b}$	$\frac{\Delta}{k_e b}$
*10^{-3}	5.319×10^{-3}	1.488	1.489	0.0163	3.572×10^{-3}	0.0110
10^{-2}	1.146×10^{-2}	1.488	1.499	0.0351	7.671×10^{-3}	0.0236
10^{-1}	2.455×10^{-2}	1.488	1.505	0.0744	1.631×10^{-2}	0.0500
1	5.244×10^{-2}	1.488	1.518	0.1544	3.455×10^{-2}	0.1038
10	1.106×10^{-1}	1.485	1.551	0.3148	7.131×10^{-2}	0.2120
10^2	2.280×10^{-1}	1.474	1.616	0.6100	1.411×10^{-1}	0.4138

* extrapolated values

are shown in Table I, along with the bandwith Δ and $\Delta/(k_e b)$ at peak. Since $k_r = 2\pi/\lambda_g$, where λ_g is the guide wavelength, $k_i b/(k_r b)$ is the gain per guide wavelength to a factor of 2π.

The peak values of $k_i b$ are plotted as a function of I in Figure 7b. It is seen that the plotted points fall closely along the curve

$$G = G_0 I^{1/3} \tag{31}$$

where G is the peak gain per unit length, and G_0 is a proportionality constant for the particular parameters selected.

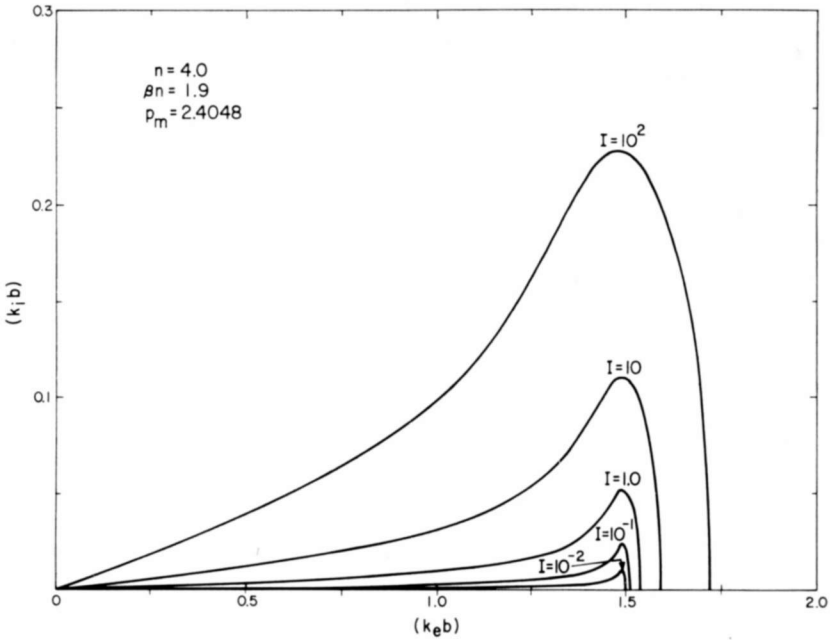


Fig. 7a—Gain characteristics as a function of average beam current, for $n = 4.0$, $\beta n = 1.9$, and $p_m = 2.4048$.

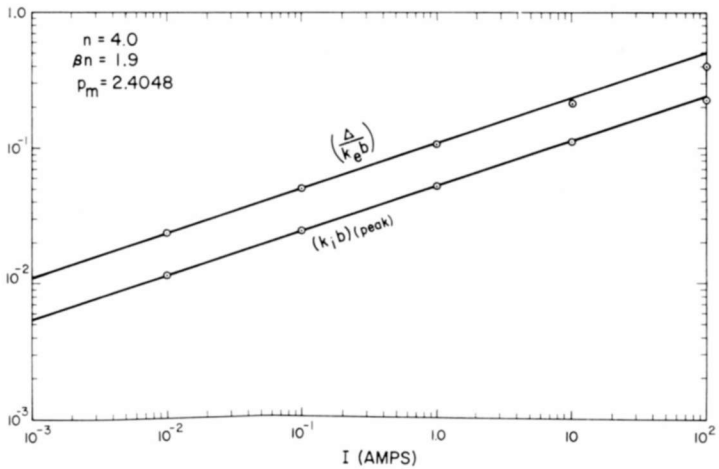


Fig. 7b—Peak gain and bandwidth versus average beam current, for $n = 4.0$, $\beta n = 1.9$, $p_m = 2.4048$.

Equation (31) is the relationship predicted by the Pierce small-signal theory.[22] Actually, the calculated points are seen to diverge from Equation (31) with increasing current. This effect results from the increasing channel width, which is ignored in the simple Pierce theory. Even so, the percent difference between Equation (31) and the calculated points is only 8.3% at $I = 10^2$, 3.6% at $I = 10$, 1.4% at $I = 1.0$ and less than 1% for lower values of I. The curve shown in Figure 7b was drawn on the assumption that the calculated value at $I = 10^{-2}$ ampere differs only negligibly from Equation (31). Thus, the value of G_0 for the chosen parameters is

$$G_0 = 5.319 \times 10^{-2} \text{ neper ampere}^{-1/3}.$$

Using this value, the value for $I = 10^{-3}$ ampere was calculated.

Note that the values of $k_e b$ in Table I are practically constant, and may be considered invariant with I. The slight variation present is again a function of the channel width, which is small for the usually encountered beam currents. As will be seen shortly, the value of $k_e b$ at peak gain is a strong function only of βn, and is thus governed primarily by the point of intersection between the beam synchronous line and the waveguide mode. Of somewhat greater variation are the values for $k_r b$. These are within 10% of $k_e b$, and understandably converge toward the values of $k_e b$ with decreasing current. Judging from the angle of intersection of the beam synchronous line with the waveguide mode, the variation of $k_e b$ with current would be expected to increase for values of βn closer to threshold, while that of $k_r b$ would be expected to decrease. Finally, note that Δ, and therefore $\Delta/(k_e b)$, also varies approximately as $I^{1/3}$, as plotted in Figure 7b.

Variation of βn

In Equation (28), β and n always appear explicitly as the factor βn. Consequently, a solution can be considered to hold for any physically meaningful combination of β and n consistent with the chosen values of βn, $k_{p0} b$, and p_m. However, Equation (30) shows that $k_{p0} b$ is also implicitly a function of I and β. Therefore, holding $k_{p0} b$ constant and varying βn conveniently implies either a variation of n with I and β constant, or a variation of β with n constant and I varying, so as to maintain the chosen value of $k_{p0} b$.

Solutions were calculated over the complex gain locus of the lowest

[22] J. R. Pierce, *Traveling-Wave Tubes*, Chapter II, D. Van Nostrand Co., New York City, 1950.

order mode of Equation (28) for values of βn ranging from 1.05 to 10.0, with $k_{p0}b = 2.0851 \times 10^{-2}$ and $p_m = 2.404826$. The value of $k_{p0}b$ resulted from the choice of $I = 1$ ampere, $\beta n = 1.9$, and $n = 4.0$, which were used previously. For each value of βn, the associated values of $k_e b$ and $k_r b$ were found with the peak value of $k_i b$. They appear in Table II, together with values of Δ, $k_i b/(k_e b)$, and $\Delta/(k_e b)$.

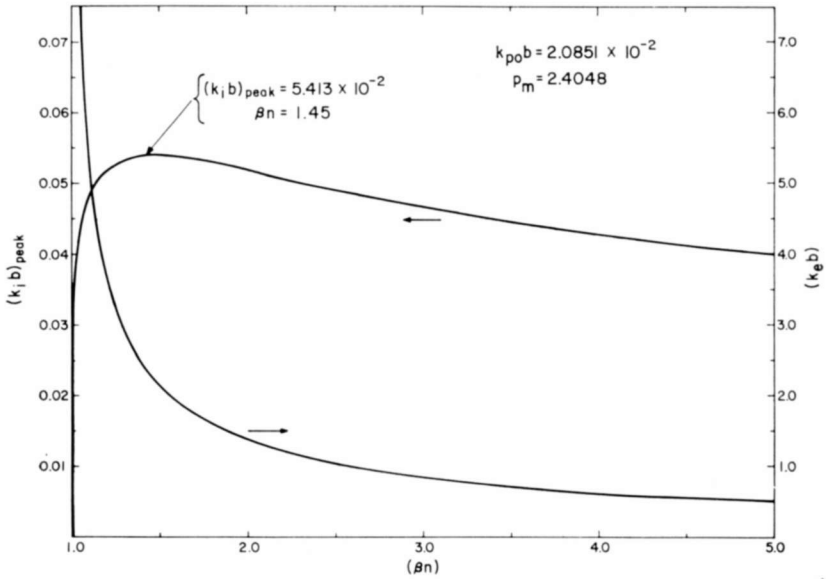


Fig. 8a—Peak gain and associated frequency variable versus βn for $k_{p0}b = 2.0851 \times 10^{-2}$; $p_m = 2.4048$.

The peak values of $k_i b$ versus βn are plotted with the accompanying values of $k_e b$ in Figure 8a. Since the interaction ceases at $\beta n = 1$, a value of $(k_i b)_{peak} = 0$ was assumed at that point. It is evident from the plot that the peak gain maximizes at $\beta n = 1.45$ and then falls off slowly with increasing βn. Thus for a value of $\beta = 0.475$, a value of $n = 3.05$ will yield the maximum gain at constant current.

When varying β while holding n constant, it is more meaningful from the laboratory standpoint to hold I constant instead of $k_{p0}b$. Such a curve may be generated directly from the plot in Figure 8a, with the help of Equation (30) for $k_{p0}b$ and the current dependence given by Equation (31). For a chosen value of $k_{p0}b$, the current varies with β as

$$I = \frac{(k_{p0}b)^2}{2.3467 \times 10^{-4}} \left(\frac{1}{\beta^2} - 1 \right)^{-1/2}. \tag{32}$$

Table II—Peak Gain and Associated Characteristics as a Function of βn for $k_{p0}b = 2.0851 \times 10^{-2}$ and $p_m = 2.4048$

βn	k_ib	k_eb	k_rb	Δ	$\frac{k_ib}{k_rb}$	$\frac{\Delta}{k_eb}$
1.05	4.401×10^{-2}	7.473	7.499	2.835	5.889×10^{-3}	0.3794
1.10	4.846×10^{-2}	5.291	5.325	1.599	9.100×10^{-3}	0.3021
1.15	5.072×10^{-2}	4.272	4.307	1.113	1.178×10^{-2}	0.2605
1.20	5.206×10^{-2}	3.597	3.625	0.8528	1.436×10^{-2}	0.2371
1.25	5.276×10^{-2}	3.250	3.291	0.6970	1.603×10^{-2}	0.2145
1.30	5.344×10^{-2}	2.913	2.950	0.5839	1.812×10^{-2}	0.2005
1.35	5.396×10^{-2}	2.656	2.690	0.4785	2.006×10^{-2}	0.1945
1.40	5.406×10^{-2}	2.461	2.496	0.4153	2.166×10^{-2}	0.1687
1.45	5.413×10^{-2}	2.285	2.316	0.3630	2.337×10^{-2}	0.1588
1.50	5.410×10^{-2}	2.151	2.183	0.3216	2.478×10^{-2}	0.1495
1.55	5.397×10^{-2}	2.033	2.067	0.2878	2.611×10^{-2}	0.1416
1.60	5.385×10^{-2}	1.924	1.956	0.2595	2.753×10^{-2}	0.1349
1.65	5.366×10^{-2}	1.832	1.864	0.2349	2.879×10^{-2}	0.1282
1.70	5.333×10^{-2}	1.753	1.787	0.2155	2.984×10^{-2}	0.1229
1.75	5.319×10^{-2}	1.672	1.702	0.1960	3.125×10^{-2}	0.1172
1.80	5.277×10^{-2}	1.611	1.646	0.1822	3.206×10^{-2}	0.1131
1.85	5.270×10^{-2}	1.545	1.576	0.1664	3.344×10^{-2}	0.1077
1.90	5.244×10^{-2}	1.488	1.518	0.1544	3.455×10^{-2}	0.1039
1.95	5.214×10^{-2}	1.435	1.464	0.1438	3.561×10^{-2}	0.1002
2.00	5.188×10^{-2}	1.388	1.419	0.1334	3.656×10^{-2}	0.0961
2.10	5.132×10^{-2}	1.302	1.332	0.1171	3.853×10^{-2}	0.0899
2.20	5.077×10^{-2}	1.227	1.256	0.1085	4.042×10^{-2}	0.0884
2.25	5.035×10^{-2}	1.190	1.216	0.0980	4.141×10^{-2}	0.0823
2.30	5.022×10^{-2}	1.161	1.190	0.0916	4.220×10^{-2}	0.0789
2.40	4.963×10^{-2}	1.102	1.129	0.0832	4.396×10^{-2}	0.0755
2.50	4.908×10^{-2}	1.048	1.074	0.0738	4.570×10^{-2}	0.0705
2.60	4.860×10^{-2}	1.002	1.031	0.0668	4.714×10^{-2}	0.0667
2.70	4.810×10^{-2}	0.9588	0.9873	0.0607	4.872×10^{-2}	0.0633
2.80	4.762×10^{-2}	0.9189	0.9458	0.0554	5.035×10^{-2}	0.0603
2.90	4.715×10^{-2}	0.8830	0.9099	0.0508	5.182×10^{-2}	0.0575
3.00	4.666×10^{-2}	0.8493	0.8748	0.0468	5.334×10^{-2}	0.0551
4.00	4.276×10^{-2}	0.6206	0.6496	0.0232	6.634×10^{-2}	0.0374
5.00	3.981×10^{-2}	0.4907	0.5132	0.0137	7.757×10^{-2}	0.0280
7.50	3.478×10^{-2}	0.3235	0.3435	0.0056	1.013×10^{-1}	0.0174
10.00	3.154×10^{-2}	0.2416	0.2591	0.0027	1.217×10^{-1}	0.0113

With the choice of $k_{p0}b = 2.0851 \times 10^{-2}$, used in Figure 8, the current becomes

$$I = 1.8526\left(\frac{1}{\beta^2} - 1\right)^{-1/2}. \tag{33}$$

The peak gain for a value of β at any other current, I_1, can now be found from Equation (31) in the form

$$(k_i b)_{I1} = R_G (k_i b)_I, \tag{34}$$

where

$$R_G = \left(\frac{I_1}{I}\right)^{1/3}. \tag{35}$$

If we conveniently choose $I_1 = 1$,

$$R_G = \left(\frac{1}{I}\right)^{1/3}. \tag{36}$$

Since the accuracy of Equation (31) decreases with increasing current, care must be taken not to consider values of β that will result in overly large values of current. For accuracy within 5%, however, the procedure should be good for values of β in excess of 0.9. Inasmuch as the entire formulation neglects far relativistic effects, this accuracy will suffice for the theory.

Therefore, one may proceed to generate the new curve by first choosing a desired value of n. This will determine the values of β along the βn absissa of Figure 8a. At each value of β, Equation (33) will then give the beam current implied by the curve for $k_{p0}b = 2.0851 \times 10^{-2}$, and the appropriate peak gain for unit current can finally be found from Equations (34) and (36). Plots of the variation in β at unit current for any choice of n up to 10 can thus be generated from the gain curve represented in Figure 8a or Table II. Such curves for $n = 4$ and $n = 2$, with $I = 1$, are shown in Figure 8b and Tables III and IV. They, too, exhibit a value of β for each n that gives maximum peak gain. For convenience in future calculations, Equations (33) and (36) are plotted in Figure 8c and tabulated in Table V. Since the associated values of $k_e b$ and $k_r b$ are not strong functions of I and are approximately equal, their values may be read to good approximation for each βn from the $k_e b$ curve of Figure 8a and Table II. This is seen

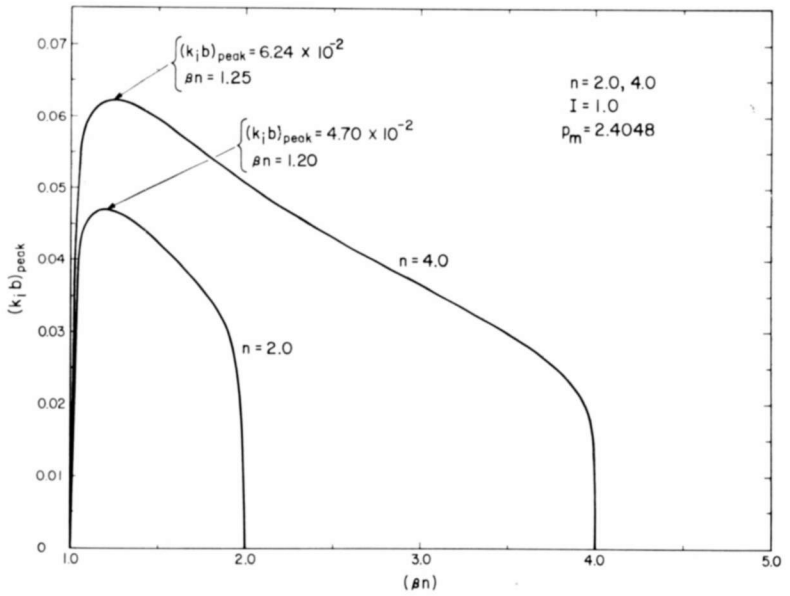


Fig. 8b—Peak gain versus βn for $n = 4$ and $n = 2$, with $I = 1.0$, $p_m = 2.4048$.

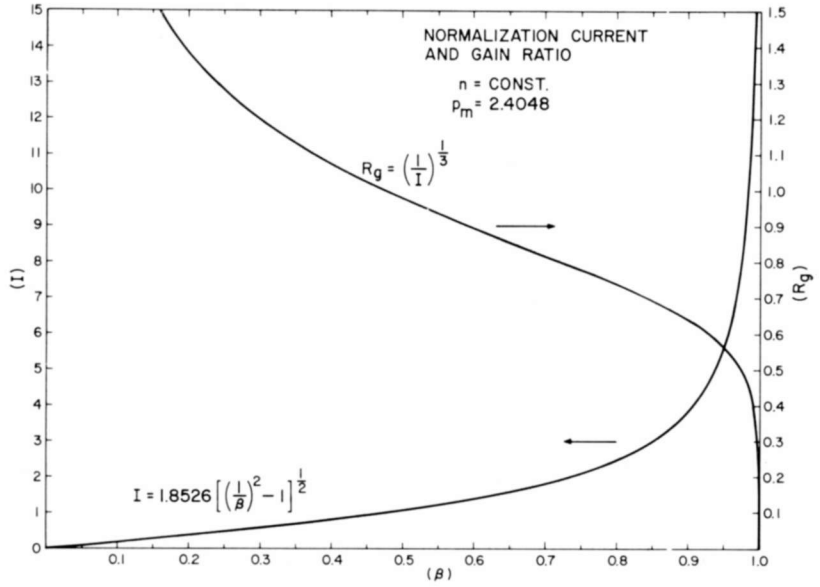


Fig. 8c—Current and gain normalizing factor versus βn for $k_{p0}b = 2.0851 \times 10^{-2}$.

by comparing the values in Table II with those in Tables III and IV, which were calculated by computer directly from Equation (28).

When considering variations of β, one must remember that by definition, $k_e = \omega/(c\beta)$. *If* $k_i b = G$, $k_r b = Q$, and $k_e b = F$ are associated

Table III—Peak Gain and Associated Characteristics as a Function of βn for $n = 4.0$, $I = 1.0$, and $p_m = 2.4048$

βn	$k_i b$	$k_e b$	$k_r b$	$\frac{k_i b}{k_r b}$
1.05	5.419×10^{-2}	7.469	7.503	7.222×10^{-3}
1.10	5.950×10^{-2}	5.288	5.328	1.117×10^{-2}
1.15	6.136×10^{-2}	4.268	4.310	1.424×10^{-2}
1.20	6.205×10^{-2}	3.594	3.627	1.711×10^{-2}
1.25	6.238×10^{-2}	3.209	3.248	1.921×10^{-2}
1.30	6.199×10^{-2}	2.910	2.952	2.073×10^{-2}
1.40	6.097×10^{-2}	2.459	2.497	2.442×10^{-2}
1.50	5.945×10^{-2}	2.149	2.184	2.722×10^{-2}
1.60	5.771×10^{-2}	1.923	1.957	2.949×10^{-2}
1.70	5.586×10^{-2}	1.752	1.788	3.124×10^{-2}
1.80	5.400×10^{-2}	1.611	1.646	3.281×10^{-2}
1.90	5.241×10^{-2}	1.486	1.516	3.457×10^{-2}
2.25	4.669×10^{-2}	1.191	1.216	3.840×10^{-2}
2.50	4.317×10^{-2}	1.049	1.074	4.020×10^{-2}
3.00	3.657×10^{-2}	0.8490	0.8699	4.204×10^{-2}
3.25	3.332×10^{-2}	0.7770	0.7945	4.194×10^{-2}
3.50	2.992×10^{-2}	0.7165	0.7326	4.084×10^{-2}
3.80	2.457×10^{-2}	0.6561	0.6711	3.661×10^{-2}
3.95	1.913×10^{-2}	0.6292	0.6399	2.990×10^{-2}
3.99	1.456×10^{-2}	0.6225	0.6305	2.309×10^{-2}

solutions for a given choice of parameters n, βn, and p_m, then either ω or b must still be chosen to complete the determination of k_i, λ_g, ω, and b. Should b be specified, k_i and λ_g will be independent of β, while ω will vary. However, choosing ω will make $b = Fc\beta/\omega$, while $k_i = G\omega/(Fc\beta)$, and $\lambda_g = 2\pi Fc\beta/(Q\omega)$. These latter functions of β must be kept in mind when computing gain and guide wavelength from data such as those given in Figure 8b and Tables III and IV.

Table IV—Peak Gain and Associated Characteristics as a Function of βn for $n = 2.0$, $I = 1.0$, and $p_m = 2.4048$

βn	$k_i b$	$k_e b$	$k_r b$	$\frac{k_i b}{k_r b}$
1.05	4.215×10^{-2}	7.480	7.505	5.616×10^{-3}
1.10	4.558×10^{-2}	5.239	5.267	8.654×10^{-3}
1.15	4.678×10^{-2}	4.238	4.268	1.096×10^{-2}
1.20	4.698×10^{-2}	3.638	3.668	1.281×10^{-2}
1.25	4.667×10^{-2}	3.218	3.249	1.436×10^{-2}
1.30	4.610×10^{-2}	2.907	2.938	1.569×10^{-2}
1.50	4.242×10^{-2}	2.147	2.170	1.955×10^{-2}
1.70	3.732×10^{-2}	1.749	1.771	2.107×10^{-2}
1.90	2.967×10^{-2}	1.487	1.504	1.973×10^{-2}
1.999	1.353×10^{-2}	1.389	1.397	9.685×10^{-3}

Table V—Current and Gain Normalizing Factor as a Function of β for $k_{p^0} b = 2.0851 \times 10^{-2}$, $p_m = 2.4048$, and n = constant.

β	I	R_G
0.00	0.000	∞
0.05	9.275×10^{-2}	2.209
0.10	1.862×10^{-1}	1.751
0.15	2.811×10^{-1}	1.527
0.20	3.782×10^{-1}	1.383
0.25	4.783×10^{-1}	1.279
0.30	5.826×10^{-1}	1.197
0.35	6.922×10^{-1}	1.130
0.40	8.085×10^{-1}	1.073
0.45	9.355×10^{-1}	1.023
0.50	1.070	9.778×10^{-1}
0.55	1.220	9.359×10^{-1}
0.60	1.389	8.962×10^{-1}
0.65	1.585	8.577×10^{-1}
0.70	1.816	8.197×10^{-1}
0.75	2.101	7.808×10^{-1}

Table V (continued)

β	I	R_G
0.80	2.470	7.398×10^{-1}
0.85	2.989	6.942×10^{-1}
0.90	3.825	6.394×10^{-1}
0.95	5.636	5.619×10^{-1}
0.99	1.300×10	4.253×10^{-1}
1.00	∞	0.000

Variation of p_m

The interaction for higher modes can be found by solving Equation (28) with values of p_m corresponding to the higher zeros of $J_0(X)$. This was done for $p_m = 5.5201$ and 8.6537, with $n = 4$, $\beta n = 1.9$, and $I = 10^2$. These values correspond to the second- and third-order modes ($m = 2,3$) and may be directly compared with the lowest mode presented above. The gain loci k_ib versus k_eb are plotted in Figure 9 for all three modes. The values of k_ib, k_eb, k_rb, $k_ib/(k_rb)$, Δ, and $\Delta/(k_eb)$ evaluated at peak gain are shown in Table VI, along with values of k_ib, k_eb and k_rb for each higher mode at the peak values for the lower modes.

Table VI—Peak Gain and Associated Characteristics for the Three Lowest Modes ($n = 4.0$, $\beta n = 1.9$, and $I = 10^2$)

p_m	k_ib	k_eb	k_rb	Δ	$\frac{k_ib}{k_rb}$	$\frac{\Delta}{k_eb}$
2.4048	2.280×10^{-1}	1.474	1.616	0.610	1.411×10^{-1}	0.4138
5.5201	3.117×10^{-1}	3.405	3.593	0.860	8.675×10^{-2}	0.2526
8.6537	3.666×10^{-1}	5.346	5.563	1.040	6.590×10^{-2}	0.1945

p_m	(k_ib) at peak 1	(k_eb) at peak 1	(k_rb) at peak 1	(k_ib) at peak 2	(k_eb) at peak 2	(k_rb) at peak 2
2.4048						
5.5201	5.237×10^{-2}	1.474	1.474			
8.6537	3.145×10^{-2}	1.474	1.474	9.054×10^{-2}	3.405	3.407

Conclusions

The picture revealed by the foregoing discussion is close to that presented by conventional traveling-wave-tube theory. The interaction of paramount interest is that between the slow space-charge wave of the beam and the circuit mode, which in this case is a mode of the dielectric-loaded waveguide. Also present are evanescent modes, which are somewhat displaced versions of the evanescent waves of the unperturbed waveguide below cutoff.

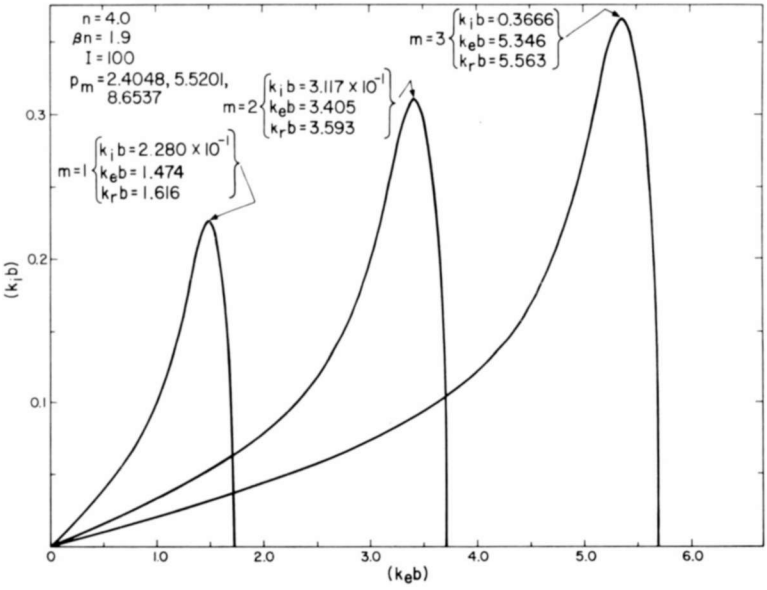


Fig. 9—Gain versus frequency characteristics for the lowest three modes, $p_m = 2.4048$, $p_m = 5.5201$, and $p_m = 8.6537$, with $n = 4.0$, $\beta n = 1.9$, $I = 10^2$.

Unlike the usual traveling-wave-tube theory, however, the complex roots denoting gain extend from the region of mode synchronism for each mode to the immediate region of the dispersion plane origin. Over most of this frequency range, the gain is small, peaking near the interaction region close to the slow Hahn–Ramo line. This extended gain region is effected by an extreme asymmetry about the apparent point of mode synchronism in the convective instability configuration in the dispersion plot. This asymmetry results from the propagation characteristics of the continuous and isotropic volume dielectric, which differ from those of more conventional itterated slowing circuits. The continuous property also precludes space harmonics and, hence, backward waves in the circuit.

While the relationship is not exact, the peak gain is found to vary as the cube root of the average beam current. Such is the case for the analogous traveling-wave-tube theory. Likewise, the bandwidth for half-peak gain varies with the same dependence, but with somewhat lesser accuracy.

It is further found that the peak gain maximizes with variation of the parameter βn, and that simple transformations to cases with constant current can be made from a given curve with $k_{p0}b$ held constant. For the lowest mode, the values of the frequency variable $k_e b$ and wave number $k_r b$ at peak gain are found to vary strongly only with changes in βn.

Two characteristics typical of Cerenkov radiation are observed in the system. The condition $\beta n = 1$ is found to be the threshold for the interaction. When $\beta n < 1$, the dispersion characteristics revert to those of a beam in a drift tube. In application, this necessitates a higher operating beam voltage than used in most traveling-wave tubes. The typical Cerenkov angle of radiation was also found to result, to close approximation, from conditions of synchronism between the beam and waveguide modes. The resultant conical wavefronts are actually the indigenous modes of the waveguide, and it is the volume propagation properties of the dielectric that place this type of mode in a dispersion region that permits coupling to the beam. Such is not the case with the usual itterated slowing structures.

Aside from possible further implications of the asymmetry in the convective instability configuration, it thus appears that the Cerenkov traveling-wave interaction does not fundamentally differ from the more familiar forms of traveling-wave interactions. The differences that do appear all relate to the specific properties of the slowing circuit, and it is in the convenience of these particular features that the potential value of such a device may lie.

Acknowledgment

The author is indebted to the capable and diligent work of T. C. Hilinski of RCA Applied Research and to the effective guidance and personal interest of R. W. Klopfenstein of RCA Laboratories, Princeton. Their efforts made possible the foregoing numerical results.

DATA PROCESSING FOR THE MASS SPECTROGRAPHIC ANALYSIS OF SOLIDS

By

J. Rogers Woolston

RCA Laboratories
Princeton, N. J.

Summary—*Because solids mass spectrographs produce their output information on a photographic plate, an arduous series of calculations based on densitometric measurements of the plate is necessary to obtain the analytical results. A program has been written for the RCA 601 computer which accepts the densitometric data, performs extensive error checks on these data, applies various correction factors, calculates the analytical results, and prints a complete report addressed directly to the customer who submitted the sample for analysis. Use of this program has reduced analytical time for such calculations by an order of magnitude, while permitting more data to be evaluated for each analysis.*

When digital computers are used to perform these calculations, it is unnecessary to use "linearization" techniques as is done for manual calculation. Such techniques usually neglect the limited and variable saturation transmission of the photoplate's emulsion, and introduce error if measurements outside the "linear" portion of the calibration curve (usually taken as 80% to 20% transmission) are used. An empirical relationship published by C. W. Hull accurately describes the entire transmission range, and has been used as the basis of this data-processing system.

Introduction

PRESENT-DAY solids mass spectrographs customarily employ ion-sensitive photoplates as the ion detector. Such photoplates exhibit a nonlinear response of darkening as a function of ion exposure. Broadly speaking, there are three methods for extracting analytical results from photographic plates exposed in a spark source mass spectrograph: (1) by visual inspection, whereby an experienced observer determines the exposure level at which each impurity becomes detectable (this produces semiquantitative results, and is widely used for this purpose); (2) by densitometry of the plate with manual (slide rule or desk calculator) calculation; and (3) by densitometry and electronic calculation, either with analog or digital computers or both. Methods (2) and (3) are used to obtain quantitative results, and both require a calibration curve of the photographic emulsion. In some cases, this curve can be derived from the data as a part of the calculation; however, in the case of analog computation a predetermined calibration curve is required.

The calibration curve can be obtained either by "functionless" methods or by "functional" methods. The first category would include the direct plotting of a large number of data points covering the entire optical transmission range of the emulsion in small increments, and the Churchill two-line method[1,2] which requires many pairs of measurements of two lines of known intensity ratio (such as isotopes whose abundance ratio falls within the requisite range). Neither of these two procedures is practical for all routine analytical work. Direct plotting is too time-consuming and demands more precision and accuracy in determining very short exposures than present instruments provide. The Churchill two-line method is applicable only for elements possessing an appropriate isotopic distribution; however, Kennicott[3,4] has described a useful computer program incorporating this method.

The functional methods can also be split into two subcategories—those that involve a linearization process, and those that incorporate an empirical function. Linearization methods such as the Seidel function[5-7] or the modified probability deviate function[8] are useful chiefly for manual graphical calculation techniques. All such linearization relationships fail to take into account the finite, variable, and mass-dependent saturation transmission of the emuulsion, and therefore some error is introduced when strong, low-transmission spectral lines are used for the calculation.* As a result, such methods are usually restricted to the "linear" portion of the curve, which ranges from about 80% to 20% transmission (or 0.1 to 0.7 density units). This range

---

[1] J. R. Churchill, "Techniques of Quantitative Spectrographic Analysis," *Ind. Eng. Chem.*, Anal. Ed., Vol. 16, p. 653, 1944.

[2] J. F. Duke, *Ultrapurification of Semiconductor Materials*, M. S. Brooks and J. K. Kennedy, eds., p. 309, Macmillan, New York, 1962.

[3] P. R. Kennicott, "A Computer Program for the Quantitative Interpretation of Mass Spectrograph Plates," Twelfth Ann. Conf. on Mass Spectrometry and Allied Topics, Montreal, June 7-12, 1964.

[4] P. R. Kennicott, "A Computer Program for the Quantitative Interpretation of Mass Spectrograph Plates," General Electric Report No. 64-RI-3766G, Research Information Section, Schenectady, New York.

[5] E. A. Baker, "Law of the Blackening of the Photographic Plate at Low Densities," *Proc. Roy. Soc. Edin.*, Vol. 45, p. 166, 1925.

[6] H. Kaiser, "The Various Methods of Evaluating Spectra in Quantitative Spectrochemical Analyses," *Spectrochim. Acta*, Vol. 2, p. 1, 1941.

[7] R. A. Sampson, "Estimation of the Continuous Spectrum of Stars," *Monthly Notices, Roy. Astron. Soc.*, Vol. 83, p. 174, 1923; "Spectrophotometric Measurements of the Effective Temperatures of Stars," Vol. 85, p. 212, 1925.

[8] J. M. McCrea, "Photographic Response and the Cumulative Law of Error," *Spectrochim. Acta*, Vol. 21, p. 1014, 1965.

* As pointed out by McCrea,[8] this error is avoided if the densitometer is adjusted to read 0% transmission when measuring a saturated area. This is often inconvenient, however, and in many cases no such saturated area exists on the photoplate formed by ions of the appropriate mass.

restriction is undesirable from an analytical standpoint, since it is important to obtain information on as many sample components as possible on a given exposure, as well as to obtain information on a given component from as many exposures as possible.

The empirical function given by Hull[9] does take into account the saturation transmission of the emulsion, and represents accurately the entire transmission range of Ilford Q2 emulsions as a function of exposure, and therefore permits all analytical data to be evaluated.

Mathematical Representation of Calibration Curve

The equation given by Hull[9] may be expressed as

$$K_x a_{x,i} E_x = \left(\frac{100 - T}{T - T_\infty} \right)^{1/R_x}_{x,i} \tag{1}$$

where, for a given sample component, x,

K_x is directly proportional to its concentration in the total ion beam striking the plate, and to the photoplate sensitivity,

$a_{x,i}$ is the abundance of isotope i of the sample component,

E_x is the exposure, indicated by the beam-monitor integrator, in nanocoulombs,

T is the percent transmission of spectral line i,

T_∞ is the percent transmission of spectral line i for an infinite exposure (saturation),

R_x is proportional to the maximum slope of the photographic response curve.

The fit of this equation to a set of directly plotted data points is illustrated in Figure 1. The data were obtained by a sequence of carefully controlled exposures of a beam of Cs thermal ions. Two of the data points, as indicated, together with a measurement of T_∞ from a very long exposure not shown, permitted the constants K and R of the Hull equation to be determined. These values were then used to plot the solid curve, which fits the remaining data very well.

---

[9] C. W. Hull, "Photographic Quantitative Analysis with a Solids Spark Mass Spectrograph," Mass Spectrometry Conf., ASTM Committee E14, New Orleans, La., June 3-8, 1962, Paper No. 72.

Because K is proportional to concentration and to plate sensitivity, the concentration of any sample component with respect to any other (e.g., the matrix, or an internal standard) is approximately given by the ratio of the two K-values, since the plate-sensitivity factors cancel. This still neglects any necessary correction factors, such as mass-dependent response of the emulsion or relative elemental sensitivity.

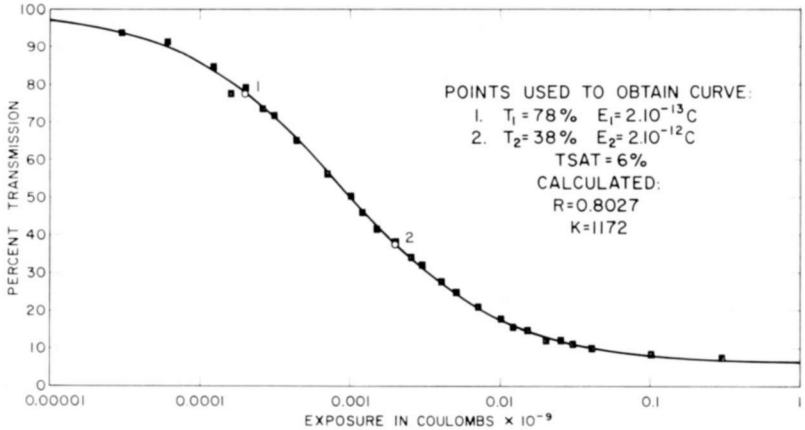


Fig. 1—Comparison of calculated calibration curve with measured data. Solid curve is calculated using $R = 0.8027$ and $K = 1172$ determined from points 1 and 2.

Thus, the basic equation to obtain concentrations of sample components, in parts per million atomic, is

$$C_x = \frac{10^6 M_x^{0.6} S_r K_x}{F_r M_r^{0.6} S_x K_r} \tag{2}$$

$$= \frac{10^6 M_x^{0.6} S_r a_{r,j} E_r \left(\dfrac{100 - T}{T - T_\infty} \right)_{x,i}^{1/R_x}}{F_r M_r^{0.6} S_x a_{x,i} E_x \left(\dfrac{100 - T}{T - T_\infty} \right)_{r,j}^{1/R_r}} \tag{3}$$

where:

subscript x refers to a given sample component (impurity) with isotope i,

subscript r refers to the matrix or other internal standard with isotope j,

C is the concentration, in ppm atomic, of the given component, x,

F is the reciprocal of the known fractional concentration of the matrix or internal standard,

M is the mass in atomic-mass units (according to Honig,[10] $M^{0.2}$ represents closely the mass response of Ilford Q2 emulsions),

S is the relative sensitivity coefficient (unless both S_x and S_r are known, they are taken to be unity).

In practice, there are two more factors that must be taken into account—the correction for background fog on the photoplate, and the correction for variable spectral line width and shape. The measured T-values used in Equation (3) are assumed to be due only to ions of the particular species considered, and it is further assumed that the spectral lines of both species are identical in width and shape. Errors will be introduced if either of these assumptions is not true. The measured transmission of a spectral line is usually the result of two additive exposures, one due to the ions that formed the line, the other, an "effective background exposure," due to a variety of possible mechanisms (heat, light, scattered ions, chemical fog, etc.). If it is assumed that both exposures obey the same physical laws, then a derivation of Equation (1) allows a "corrected line transmission" to be calculated from the measured line transmission and a measurement of the background transmission adjacent to the line;

$$T_L = \frac{100 + T_x\left[\left(\dfrac{100 - T_{LB}}{T_{LB} - T_x}\right)^{1/R} - \left(\dfrac{100 - T_B}{T_B - T_x}\right)^{1/R}\right]^R}{1 + \left[\left(\dfrac{100 - T_{LB}}{T_{LB} - T_x}\right)^{1/R} - \left(\dfrac{100 - T_B}{T_B - T_x}\right)^{1/R}\right]^R} \tag{4}$$

which, for $R = 1$, $T_x = 0$, reduces to

$$T_L = \frac{T_B T_{LB}}{10^{-2} T_B T_{LB} + T_B - T_{LB}}, \tag{5}$$

where T_L is the corrected line transmission, T_{LB} the measured line transmission, and T_B the measured background transmission, all in percent.

The correction for line shape and width is more difficult to handle. Figure 2 serves to illustrate the problem. On the left-hand side of the

[10] R. E. Honig, "Analysis of Solids by Mass Spectrometry," International Conf. on Mass Spectrometry, Paris, Sept. 1964.

figure, three typical transmission profiles are reproduced as they are obtained from a densitometer. Clear, unexposed emulsion has been set to be 100% transmission, opaqueness to be 0%. On the right-hand side of Figure 2, each of the three peaks has been replotted, point-by-point, via the calibration curve, to obtain the "true intensity" profile of the line. The area above background of these intensity profiles is now directly proportional to the number of ions forming the line. These are the profiles one would obtain directly from the mass spectrograph

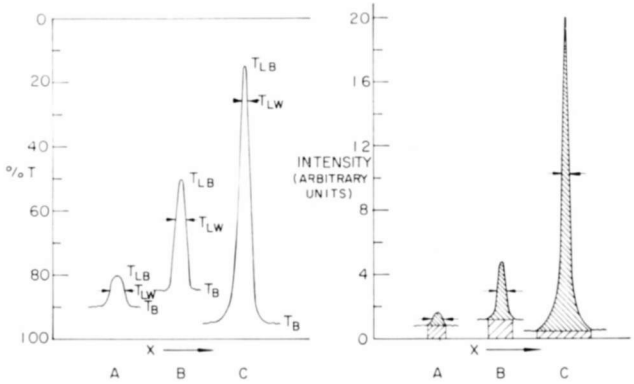


Fig. 2—Comparison of percent transmission scan with intensity profile.

with *electrical recording*. Since it is not possible to obtain intensity profiles other than by arduous, manual replotting, a simplifying assumption is often used; namely, that peak shape is substantially constant, but that peak width varies. Areas may then be obtained by the product of maximum intensity and a width measurement at some convenient geometrical position, such as half maximum intensity. The relationship between half maximum intensity and the corresponding height on the transmission profile is, of course, nonlinear, but may be calculated by another relationship derived from Equation (1).

$$T_{LW} = \frac{100 + T_\infty \left[\frac{1}{2}\left(\frac{100 - T_{LB}}{T_{LB} - T_\infty} \right)^{1/R} + \frac{1}{2}\left(\frac{100 - T_B}{T_B - T_\infty} \right)^{1/R} \right]^R}{1 + \left[\frac{1}{2}\left(\frac{100 - T_{LB}}{T_{LB} - T_\infty} \right)^{1/R} + \frac{1}{2}\left(\frac{100 - T_B}{T_B - T_\infty} \right)^{1/R} \right]^R} \tag{6}$$

which, for $R = 1$, $T_\infty = 0$, reduces to

$$T_{LW} = \frac{2T_{LB}T_B}{T_{LB} + T_B} \tag{7}$$

where T_{LW} is the percent transmission at which the width of the line must be measured on the transmission profile.

To incorporate such width measurements into the calculation of concentrations, Equation (2) would be modified as follows;

$$C_i = \frac{10^6 M_x^{0.6} S_r W_{x,i} K_x}{F_r M_r^{0.6} S_x W_{r,j} K_r}, \tag{8}$$

where $W_{x,i}$ is the line width of isotope i of sample component x and $W_{r,j}$ is the line width of the matrix or internal standard, both in arbitrary units.

A few words are in order regarding the parameters R and T_x. When two or more data points are available for the same component, the R-value may be determined for that component. These data points may be two isotopes of different abundance measured on the same exposure, or they may be the same isotope measured on different exposures. In principle, in an analysis, the R-value may be determined for every component. If this is done, in a typical analysis, different values will be found for every component. Indeed, for the same component, using many different exposures the R-values will be found to differ appreciably. This can, of course, be due to any form of inhomogeneity in the sample. Restricting determinations to isotopic pairs on the same exposures, lesser differences between R-values for the same and for different components are observed. Because R is fairly sensitive to slight error in either of the data points used for its determination, the evidence available is not yet sufficient to resolve the question of whether or not there is some fundamental relationship between R (or gamma) and the mass or chemical properties of ions. Therefore, adopting some average R-value when analyzing a given sample may be preferable to using individual R-values for each component. Evidence has shown, however, that different photoplates do exhibit significantly different average R-values; that R is sensitive to plate processing techniques; that multiply charged (higher energy) species do produce higher R-values than singly charged species; that molecular ions result in lower R-values; and that the value 1.0 is a good overall average value.

Direct measurement of T_x is usually possible only for the major constituents of a sample (on the longer exposures). Experimental data have shown that T_x is mass dependent. While the exact functional

relationship is not known, T_∞ is appreciably less for low-mass components than it is for the high-mass components, because of the greater penetration of the low-mass ions into the emulsion. The influence of T_∞ on the accuracy of calculated concentrations depends on how closely the measured line transmission approaches T_∞, and on the magnitude of T_∞ itself. For $T_\infty = 2\%$ or less, negligible error is introduced for all calculations with lines of 15% transmission or higher (even if T_∞ is taken as 0%). Experience has shown, however, that occasional batches of Ilford Q2 plates produce T_∞ values of 10% transmission or higher, and that T_∞ is very sensitive to plate-processing techniques.

Program for the RCA 601 Computer

Straight manual calculation of the above equations is too time consuming to be practical. The use of graphical linearization methods is helpful, but even then it is a lengthy operation if there are many sample components. Typically, a manual analysis of a 30-component sample will involve a total of about 5 hours work, of which less than one half hour is spent doing the densitometry. With digital electronic computation, on the other hand, the total time is reduced to that required to do the densitometry. Furthermore, with electronic computation, it is usual to collect and process many data for each component at a small cost of a few extra minutes spent in densitometry, whereas with manual means only one or two data points are employed for each component. With many points, results are improved statistically, and evidence of inhomogeneity, surface contamination, etc., is more apparent.

A program, based on Equation (1), has been written in assembly-system language for the RCA 601 computer. This language was chosen, in preference to FORTRAN, in order to obtain maximum possible efficiency from the 601, to permit use of all 64 characters of the 601 character set, and because the program would be available on a library tape and therefore would not have to be card-loaded for each usage.

Several decisions were made at the outset: (a) a card-based system is used; (b) each card contains all data necessary for a specific calculation and can contain data for up to three discrete data points (spectral lines) for a given component; (c) individual R-values are calculated and used for each component when more than one data point is provided, but low and high limits on R are established at the beginning of each run, and if the calculated value falls outside these limits the nearest limit-value is used instead (if the high and low limit values are made to coincide, the limit-value is used for all calculations); (d) all data are entered on the cards in fixed-point form with as great a

flexibility and simplicity as possible; (e) the program performs all possible error checks on the input data and the calculations; and (f) two outputs are provided, one containing details of the calculation of each component (for the analyst), and one containing a printed report text and the analytical results (for the customer).

The procedure for conducting a mass spectrographic analysis using this computer program is as follows:

1. Prepare and run sample, and process photoplate
2. Perform the densitometry on the photoplate
   a. adjust 100% transmission to be clearest unexposed emulsion, and 0% transmission to be opaqueness
   b. manually (visually) identify spectral lines
   c. scan line and adjacent background
   d. manually record the following data on data sheet (see Figure 3):
      element symbol
      relative sensitivity factor
      comments
      isotope for each data point
      line transmission for each data point
      background transmission for each data point
      exposure for each data point
      saturation transmission
3. Keypunch data
4. Submit to computer
5. Examine output, correct errors, and rerun if necessary

Figure 4 is a simplified flow chart of the computer program itself, and is useful in understanding the program logic and sequence of events. The first block (upper left) establishes the R-value limits and an assumed R-value that is used when only one data point is given. R_{ASSUMED} is also used in the background correction equation, in the half-intensity width equation, and for calculating detection limits. In addition to these, the exponent used to obtain the mass-effect correction is established at this point. While this exponent is usually set to be 0.6, any value may be used, including zero (which results in unity mass-effect factor for all elements). After these values are established, the first, or reference element, data card is processed through the following sequence of blocks shown in Figure 4. In the "EDIT DATA" block all numerical values are converted to usable floating-point form, while checking their legitimacy (no non-digit characters other than one decimal point) and that all necessary values are present (blank

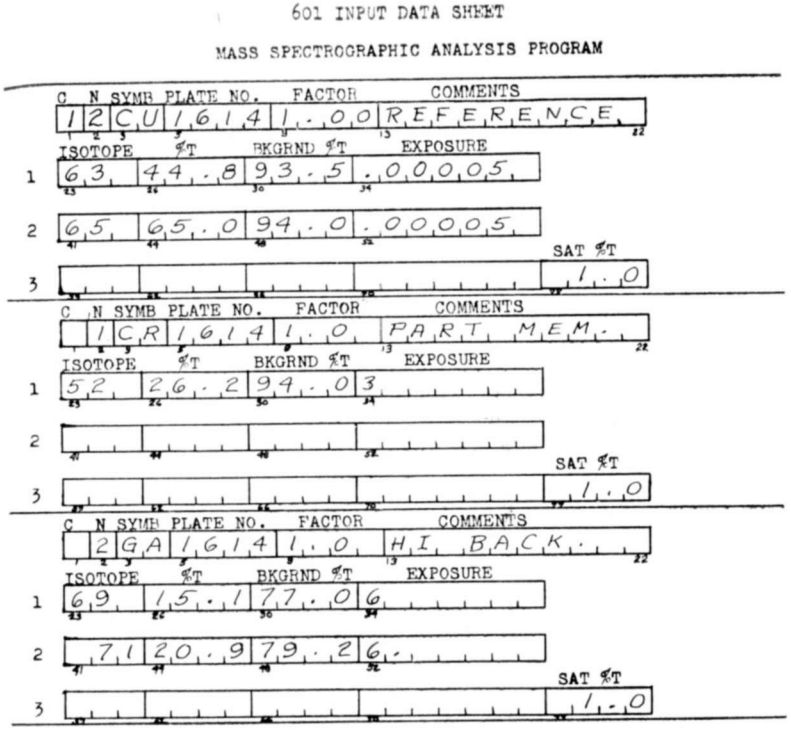


Fig. 3—Sample data sheet.

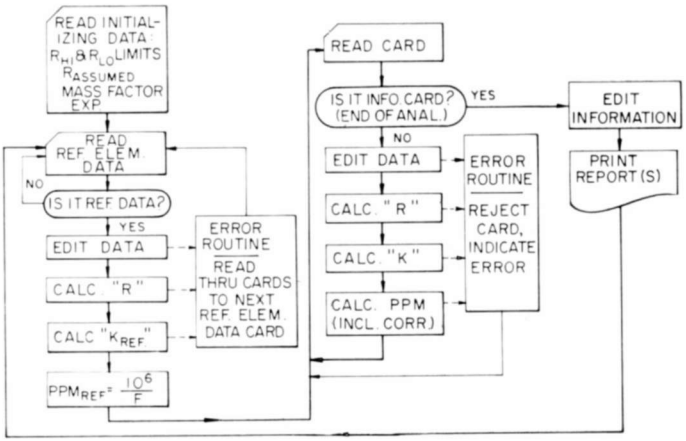


Fig. 4—Simplified flow chart of RCA 601 computer program for mass spectrographic analysis.

background transmission is treated as 100%). Also, certain logical tests are made on the data, i.e., the background transmission must be a larger number than the line transmission, saturation transmission must be less than all line transmissions, the element symbol must be valid, the isotopes must correspond to the element symbol. Another function of this step is to obtain the isotopic abundances of the given isotopes from tables internal to the program. Still a part of this step (EDIT DATA) is the calculation of the appropriate mass correction factor, correction of line transmissions for background, and calculation of half-intensity width transmissions.

The next step is the calculation of the R-value. If only one data point is given, the step is bypassed and the R_{ASSUMED} value is passed on to the following step. If two data points are given, an R-value is calculated from the relationship

$$R = \frac{\log\left[\left(\dfrac{100 - T_m}{T_m - T_x}\right)\left(\dfrac{T_n - T_x}{100 - T_n}\right)\right]}{\log\left(\dfrac{(aE)_m}{(aE)_n}\right)}, \tag{9}$$

where the subscripts m and n refer to the two data points used.

If three data points are given, three R-values are calculated from all three possible pairs of points, and then their average is taken. The resultant value is tested against the high and low limits and, if it falls outside the acceptable range, the nearest limit value is used instead for the following calculation of K.

In the "CALC. K_{REF}" block a K-value is calculated for each data point using the following version of Equation 1:

$$K_x = \frac{1}{a_{x,i}E}\left(\frac{100 - T}{T - T_x}\right)_{x,i}^{1/R_x} \tag{10}$$

The R-value used is, of course, that determined in the previous step. If two or three data points are given, the separate K-values are averaged geometrically. The resultant value is multiplied by the appropriate mass-effect factor and is stored for use in calculating the concentrations of all the remaining components in the analysis. Finally, in the PPM_{REF} block, the concentration of the reference component is calcu-

lated by dividing 10^6 by F, the reciprocal of the known fractional concentration of the reference component of the sample.

The data cards for the remaining sample components are then processed serially in exactly the same fashion up to the calculation of the ppm concentration. Here, the calculated average K-value is divided by K_{REF}, and relative sensitivity and line-width corrections may be applied. In the current form of the program, no provision has been made for including line-width measurements on the data card. Therefore, when line-width corrections are desired, the following procedure is used: a first processing of the data provides the necessary half-width transmissions; the line widths are then measured on the original densitometer charts at these transmissions, and the original exposure values are then either increased or decreased by a factor corresponding to the ratio of the measured width for each line to that for the reference component. The final concentration value is rounded off to two significant figures and stored.

The error-detection routines are an extremely important feature. Without such routines one of two things can occur—either an erroneous calculation will be made, or the entire run will be terminated by the computer monitoring system. In the latter case, as many runs could be required as there are bad data cards to detect all the errors in a group of cards, and experience has shown that statistically 3 to 4 bad cards per 100 may be expected in routine analytical work. With error-detection routines, if an error is detected on the reference-data card, all cards for that analysis are ignored, and processing continues with the first, or reference, card of the next analysis. When an error is detected on a data card other than the reference, however, only that card is rejected, and processing continues with the next data card. For every error detected, a message is printed in the output indicating the exact nature of the error, and the contents of the erroneous data card are printed out. At the present time, there are 67 discrete errors detectable by the 601 program. The most frequently encountered errors are incomplete data (e.g., exposure missing), incorrect isotope (does not correspond to element symbol), incorrect element symbol, number of data points on card incorrectly indicated, data out of order (e.g., reversed exposures), background transmission lower than line transmission, F-factor omitted (the program assumes unity in this case and prints a message indicating this action), reference data missing or not placed first in the card deck (the program ignores all data cards placed ahead of the reference data card). Most errors are detected during the editing phase. Some, however, which might result in an attempt to divide by zero or to take the logarithm of a negative number, are

detected during the calculations. Error-detection routines cannot catch numerical errors of measurement or transcription unless some logical condition is violated as a result.

A special card with a code consisting of a "2" in column 1 is used as the final card of each analysis in a batch. This is referred to as the "information" card, since it contains such items as the sample identification, customer's name, report format desired, number of copies of the report to be printed, and various bookkeeping items (problem number, etc.). When this card is encountered, the program optionally prints as many copies as desired (up to 9) of any one of a number (maximum of 64) of standard report texts, including the analytical results within the body of the text, then proceeds to the next analysis in the batch.

Figures 5 and 6 illustrate the computer output. The first part of the output for an analysis, as shown in Figure 5, is printed progressively as each data card is read and processed. In addition to displaying the input data that were punched on the card, the calculated concentration in parts per million atomic for each sample component is shown, together with the corrected line transmission, half-intensity (or linewidth) transmission, the mass factor (MF), and the R- and K-values. The T_∞ value is identified here as "TSAT". Several data card possibilities are illustrated in this sample analysis, including one card containing an error (tin has no mass 109 isotope). The error has been corrected on the following card in Figure 5. The R-values shown for copper, gallium, and tin are quite typical, and are acceptable. That shown for silver is not. The accompanying message indicates that the upper-limit value of 1.25, instead of the calculated R-value, was used to calculate the K-values. This unusually high value, 1.8797, was undoubtedly due to an inhomogeneous distribution of silver in the sample: silver appears to have two rather different concentrations in the two exposures used for the data. This difference can be seen quantitatively in the two K-values, $K1$ and $K2$, which are directly proportional to the concentrations. For each element, a K-value is calculated for each data point given. Where there are two or three data points, the individual K-values are averaged geometrically and multiplied by the mass factor, and then shown as "KAVG". For a two-point card, $K1$ will always equal $K2$ unless the calculated R-value falls outside the limits, as was the case for silver. An important aspect of the use of a program such as this in routine analytical work is the opportunity for the analyst to rapidly survey this part of the output to detect abnormalities in the results that have analytical significance.

```
        M A S S   S P E C T R O G R A P H Y
        TRACE IMPURITY ANALYSIS

  PLATE NO. 1814
MASS FACTOR CALC. AS (AT.WT. / 30)EXP .6
REFERENCE ELEMENT FACTOR =1.00

    COPPER          1.00000000  06  PPM ATOMIC       REFERENCE                                          MF= 1.568     FACTOR=1.00
      063   T1=44.8%  T1B= 93.5%   E1=.000050UNITS     T1COR=46.2%   T1LW=60.4%            R=  1.1227    K1=33799.1903
      065   T2=65.0%  T2B= 94.0%   E2=.000050UNITS     T2COR=67.7%   T2LW=76.8%                        K2=33799.1908
          TSAT=01.0%                                                                                  KAVG=53022.7045


    CHROMIUM        30.0000000   PPM ATOMIC       PART MEM.                                          MF= 1.390     FACTOR=1.00
      052     I=26.2%   TB= 94.0%    E=3.00000UNITS      TCOR=26.6%    TLW=40.6%   R(ASSUMED)=  1.0000    K=    1.13978641
          TSAT=01.0%                                                                                  KAVG=    1.58537747


    GALLIUM         47.0000000   PPM ATOMIC       HI BACK.                                           MF= 1.658     FACTOR=1.00
      069   T1=15.1%  T1B= 77.0%   E1=6.00000UNITS     T1COR=15.7%   T1LW=24.7%            R=  1.0256    K1=   1.51059317
      071   T2=20.9%  T2B= 79.2%   E2=6.00000UNITS     T2COR=22.0%   T2LW=32.7%                        K2=   1.51059316
          TSAT=01.0%                                                                                  KAVG=    2.50546991


B0  GALLIUM                                                    DETECTION LIMIT =     .3000=    MF= 1.658     FACTOR=1.00
      069(ASSUMED)      TB= 77.0%    E=6.00000(ASSUMED)    TSAT=01.0%(SPECIFIED)  R(ASSUMED)=  1.0000    K=     .00971475


    SILVER          68.0000000   PPM ATOMIC                                                          MF= 2.155     FACTOR=1.00
      107   T1=59.1%  T1B= 90.4%   E1=1.00000UNITS     T1COR=63.0%   T1LW=71.4%            R=  1.8797    K1=   1.27654257
      107   T2=18.0%  T2B= 89.5%   E2=3.00000UNITS     T2COR=18.3%   T2LW=29.5%                        K2=   2.22034559
          TSAT=01.0%                                                                                  KAVG=    3.62823654
* RAVG EXCEEDS ESTABLISHED UPPER LIMIT OF 1.25 .  THIS VALUE HAS BEEN USED TO CALCULATE THE K-VALUES.


EXIT  40  I2 INCORRECT
THE DATA CARD REPRODUCED BELOW IS INCORRECTLY PREPARED OR PUNCHED.
C   N   SY  P_.#  -F--  -COMMENTS-  I1-  -T1-  T1B-  --E1---  I2-  -T2-  T2B-  --E2---  I3-  -T3-  T3B-  --E3---  TSAT
    3   SN  1814  1.00              (118  70.7  90.6  3.00000] (109  83.0  91.0  3.00000] (120  65.0  91.4  3.00000] 01.0


    TIN             21.0000000   PPM ATOMIC                                                          MF= 2.282     FACTOR=1.00
      118   T1=70.7%  T1B= 90.6%   E1=3.00000UNITS     T1COR=76.2%   T1LW=79.4%           R1=  1.0469    K1=    .48072034
      119   T2=83.0%  T2B= 91.0%   E2=3.00000UNITS     T2COR=90.4%   T2LW=86.8%           R2=  1.0706    K2=    .50006940
      120   T3=65.0%  T3B= 91.4%   E3=3.00000UNITS     T3COR=69.1%   T3LW=75.9%           R3=  1.1484    K3=    .48904593
          TSAT=01.0%                                                                    RAVG=  1.0886 KAVG=    1.11806366
```

Fig. 5—First part of computer output (intended for the analyst).

TO: RCA REVIEW
FROM: J. R. WOOLSTON
PROBLEM NO. 1000 - MASS SPECTROGRAPHIC ANALYSIS OF COPPER CALIBRATION STANDARD
PLATE NO. 1614

THIS SAMPLE WAS ANALYZED WITH THE MS7 MASS SPECTROGRAPH FOR GENERAL IMPURITIES EXCLUDING CARBON, NITROGEN, AND OXYGEN. (THE ION SOURCE WAS NOT BAKED, THUS, THE HIGH GAS BACKGROUND PRECLUDES A DETERMINATION OF THESE ELEMENTS.)

THE VALUES REPORTED IN THE FOLLOWING TABLE (IN PARTS PER MILLION ATOMIC) HAVE BEEN COMPUTER CALCULATED AND IN MOST CASES REPRESENT AN AVERAGE OF THREE DETERMINATIONS. THE WORDS "NOT DETECTED" SIGNIFY THAT THAT ELEMENT WAS SPECIFICALLY SOUGHT AND WAS FOUND TO BE ABSENT AT A CONCENTRATION GREATER THAN ITS DETECTION LIMIT. THE CALCULATED DETECTION LIMITS DEPEND ON THE LONGEST EXPOSURE MADE DURING THE ANALYSIS, THE ABUNDANCE OF THE PRINCIPAL ISOTOPE, AND THE SENSITIVITY OF THE PHOTOPLATE. THE TWO DIGITS FOLLOWING A VALUE WHOSE DECIMAL POINT IS DISPLACED TO THE LEFT ARE THE POWER OF TEN BY WHICH THAT VALUE MUST BE MULTIPLIED (FOR EXAMPLE: 1.0000 06 = 1,000,000). COMMENTS CONCERNING INDIVIDUAL VALUES ARE INDICATED BY CODES WHICH ARE DEFINED IN THE LEGEND BELOW THE TABLE OF RESULTS.

RESULTS:

ELEMENT	AMOUNT	COMMENTS	DETECTION LIMIT
CHROMIUM	30.0000	PART MEM.	.1100
NICKEL	INTERFERENCE	I	.9500*
COPPER	1.0000 06	REFERENCE	.1500
GALLIUM	47.0000	HI BACK.	.3000*
SILVER	68.0000		.2700
TIN	21.0000		.4500
ANTIMONY	22.0000		.2600
LEAD	37.0000		.4000
BISMUTH	18.0000		.2100
PLUTONIUM	NOT DETECTED		.2300

NOTE: THE SYMBOL "*" APPEARING AFTER A DETECTION LIMIT INDICATES THAT THAT VALUE HAS BEEN SPECIALLY CALCULATED AS A RESULT OF HIGH BACKGROUND, INTERFERENCE AT THE MAJOR ISOTOPE, ETC.

IT SHOULD BE NOTED THAT LACKING INFORMATION TO THE CONTRARY, THESE RESULTS ARE BASED ON THE ASSUMPTION OF EQUAL RELATIVE SENSITIVITY FOR ALL ELEMENTS. THE ABSOLUTE ACCURACY OF ANY VALUE IS ESTIMATED AS +OR- A FACTOR OF THREE. RELATIVE ACCURACY OR PRECISION IS ESTIMATED TO BE +OR- 50% OR BETTER.

Fig. 6—Second part of computer output (report for the customer).

Figure 6 shows the second part of the computer output, from 0 to 9 copies of which are printed when the type 2, or information card at the end of each analysis is read and processed. One copy of this part may be sent directly to the customer if the analyst is satisfied with the results, thus obviating the need for a typist as well as saving further time. The analytical results are inserted into the body of the report text in order of increasing atomic number, regardless of the order of the data cards.

In addition to the ppm atomic values, detection limits for each element may be optionally calculated, as shown. General, or standard, detection limits are based, automatically, on the most abundant isotope for each element, on the longest analytical exposure used, and on the plate sensitivity (as determined via the reference component K-value and the F-factor). Detection limits assume a 100% background and a threshold of detectability of 98%. In cases where any of these conditions are not appropriate, such as excessive background or interference of the major isotope, a procedure is included for calculating special detection limits. For example, the necessary "zero-points" data card for gallium is included in Figure 5. On such a card only those data that would differ from a standard detection limit need be included, the remainder are assumed. As shown for gallium in Figure 5, the background at mass 69 (due to the intense nearby copper lines) was only 77.0% transmission. Zero-point cards are also used for elements that were specifically sought and were not found in the analysis. This is illustrated in Figure 6 for nickel and plutonium. In such a case the words "Not Detected" are entered in the ppm atomic column by the computer, unless, as for nickel, a special symbol has been punched in the comments field in which case any of a number of special messages may be printed.

A few special features of the program deserve mention. Optionally, for each component, a tabulation of the calibration curve may be printed for plotting purposes. The tabulation provides T in 1% increments, the associated E in nanocoulombs, and $\log_{10} E$, based on the calculated R and K of Equation (1). There are five numbered but nameless "dummy elements" included in the program's internal tables that are useful in reporting concentrations of molecules or compounds which are occasionally of analytical interest. A special routine provides for the calculation of the ratios of multiply charged to singly charged species. Up to 50 separate analyses may be batched together as one computer run, thus effecting a considerable economy of computer time. After the final analysis of a batch has been processed, a special "statistics" report is automatically printed containing all the information

from each analysis type 2 card, together with a count of the number of data cards processed, the number of errors detected, the computer execution time for each analysis, and the total execution time for the batch. This report is useful chiefly for bookkeeping purposes, but can be used for automatic billing for the analysis service.

A few specifications of the program for the RCA 601 may be of interest to those contemplating similar programs. The total storage used is about 9000 48-bit words. Of this about 2500 (5000 half-words) are used for the instructions, about 2000 in working arrays, and the remainder in internal tables, constants, and temporary storage. Execution times are 60 milliseconds for an average data card, about 1.8 seconds for a 30-component analysis. Thus a typical batch of 15 30-component analyses requires about 30 seconds. Print-out time for such a batch (about 100 pages), is about 5 minutes. There is no limit on the number of data cards that may be included in one analysis; however, for a given element, only the results for the last card appear in the report listing. Every card processed, of course, appears in the first part of the output (for the analyst).

Discussion

The program has been of great value to the mass spectrographic analysis service group at RCA Laboratories. An average of over 400 analyses per year are handled by a staff of two persons, one to operate the MS7 mass spectrograph, the other to perform the densitometry, prepare the data for the computer, and evaluate the results. Over 90% of these analyses are processed via the computer program. While the computer cannot improve the quantitative accuracy of the results, from a practical point of view its use permits one to make all known or necessary corrections to the data for every analysis processed, which would not be attempted manually for a large volume of work.

Of the several data corrections described, the two most difficult to handle are line width and relative sensitivity: the first because of the added labor involved, the second because the needed information is usually not available. Future work, therefore, should be directed at these two points. With analog techniques or digital sampling techniques, it should be possible to obtain intensity areas or integrals directly, as shown on the right-hand half of Figure 2. Obviously, additional hardware would be involved.

As the state-of-the-art of solids mass spectrography advances, it can be assumed that tables of reliable relative sensitivities will become

available. As they do, such tables can easily be incorporated into data processing programs as are isotopic abundance tables at present.

When large analytical work loads are involved, any reduction in routine manual operations will be profitable. As an example, the present, manual recording and keypunching of data with the 601 program could be largely automated by coupling a keypunch to the densitometer directly.

Finally, while the work described is for a specific application, the mass spectrographic analysis of solids, the principles involved are applicable in general to a much wider range of instrumental analytical techniques.

ACKNOWLEDGMENTS

I wish to express my appreciation for the many helpful suggestions and assistance rendered by A. H. Simon and T. M. Stiller during the writing of the RCA 601 program.

THE BRIDGE CELL— A NEW SUPERCONDUCTIVE MEMORY CELL FOR RANDOM-ACCESS WORD-ORGANIZED MEMORIES*

By

R. W. Ahrons

RCA Electronic Components and Devices
Somerville, N. J.

Summary—*The bridge cell is a thin-film memory cell based on the phenomenon of superconductivity. It differs from other superconductive memory cells in that each bit is stored in the form of a persistent current whose plane is perpendicular, instead of parallel, to the plane of the thin film. This method of storage leads to inherently high packing densities. An equivalent-circuit model is used to explain the operation of the bridge cell. The model comprises two inductances in parallel with a mutual coupling equal to the lower value of the two inductances. A resistance is in series with the inductance having the lower value. The analysis of this model explains the operation of the cell.*

The cell can be made by evaporation through masks or by a combination of evaporation and preferential etching. Experiments on evaporated samples have demonstrated that sense voltages higher than one millivolt can be obtained. The cell can be used as an analog memory having an output proportional to the digit-line driving current. The new bridge cell most naturally fits a word-organized memory array. A single cryotron-tree type of decode matrix can be used with the bridge cell, but greater operating speeds can be obtained with a two-level decode matrix.

Introduction

IN SUPERCONDUCTIVE thin-film memories, such as the continuous-sheet memory or the conventional cryotron memory, the loops formed by the persistent (stored) currents are oriented parallel to the substrate. The bridge cell is a new type of thin-film cell for *random-access* memories in which the persistent-current loop formed in the superconductive material is perpendicular to the plane of the substrate. In thin-film memories, bit densities are expressed in terms of the number per unit of surface area, because the thickness of the substrate dominates the third dimension. Thus, the bridge cell, with perpendicular loop storage, is inherently capable of higher packing densities than an equivalent conventional cryotron cell.

* The research on the bridge cell was supported by the Department of the Navy, Office of Naval Research, under Contract No. NONR-3879-(00).

This paper presents the theory of the operation of the bridge cell, describes the experimental verification of this operation, and discusses the potential of the cell for operation in different types of memory arrays.

Operation of the Bridge Cell

For simplicity, the explanation of the operation of the bridge cell is limited to first-order theory; the explanation, however, is sufficiently broad so that an adequate understanding of cell operation can be obtained.

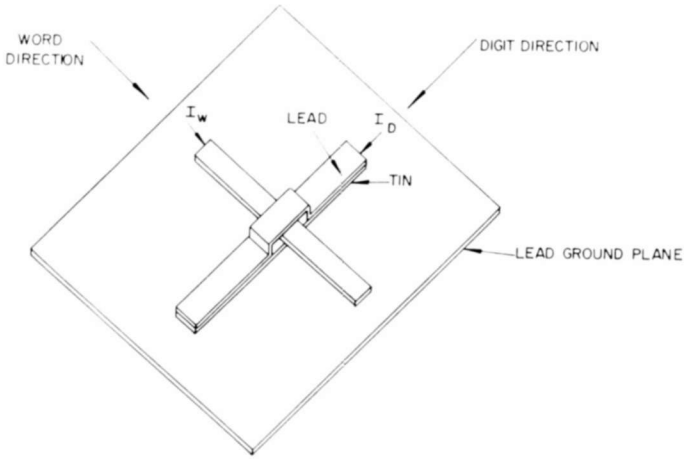


Fig. 1—The basic bridge cell.

The basic bridge cell is shown in Figure 1. The line in the digit direction consists of a lead strip superimposed on a tin strip. These strips comprising the digit line are separated in the vicinity of the word line (lead) so that the lead portion bridges over and the tin portion runs under the word line. Figure 2 shows a cross section of the bridge area and the relation of the word and digit lines to the ground plane. There can be no magnetic field above the lead portion of the digit line because of the influence of the ground plane. Assuming no initial stored flux, there can be no flux, and hence no field, in the bridge area because of the theorem which states that fluxoid is conserved in an area totally enclosed by a superconductor. As a result, all the field must be contained between the ground plane and the tin strip, and the current in the digit line, I_D, flows on the bottom surface of the tin strip, i.e., adjacent to the ground plane. The word line is

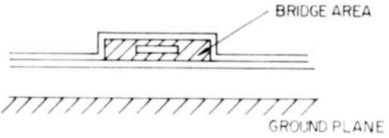


Fig. 2—Cross section of the bridge cell.

orthogonal to the bridge and acts as the control for the cell, as in a crossed-film cryotron. The field generated by a current in the word line, I_w, induces a superconductive-to-normal transition in the tin strip beneath the word line. When the conductive properties of tin become normal, all the current I_D is switched into the lead (upper) strip and remains there, even after the current is removed from the word line. When the current I_D is removed from the digit line, current continues to circulate around the bridge area. A circulating current can be used to represent a "1" and the absence of current can be used to represent a "0"; or a positive circulating current can represent a "1" and a negative circulating current can represent a "0".

Figure 3 shows equivalent circuits of the bridge cell and of a conventional cryotron cell. In these equivalent-circuit models, L_1 represents the inductance of the tin portion of the digit line, and L_2 the inductance of the lead portion. The control function of the word line is represented by the switch. When the switch is closed, the resistance in series with L_1 is shorted out to represent the superconductive state of the tin. The transition to the normal (conductive) state is made when the switch is opened, thus putting the resistance in series with L_1. The basic difference between these equivalent circuits is the mutual coupling, M, in the bridge cell. The influence of M is shown by the

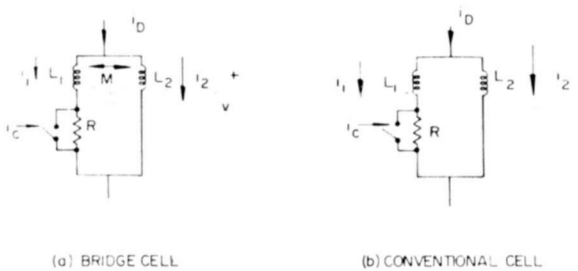


Fig. 3—Equivalent circuit of the bridge cell and of a conventional cryotron cell.

equations below, which define the voltage, v, across each parallel branch of the bridge-cell equivalent circuit when the switch is closed:

$$L_1 \frac{di_1}{dt} + M \frac{di_2}{dt} = v, \tag{1}$$

$$M \frac{di_1}{dt} + L_2 \frac{di_2}{dt} = v. \tag{2}$$

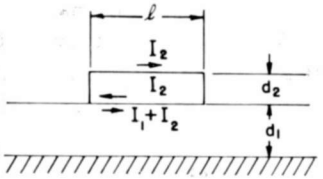


Fig. 4—Side view of the bridge cell.

Basic energy relationships can be used to obtain the following relationship for L_1, L_2, and M:

$$\frac{1}{2} L_1 I_1^2 + \frac{1}{2} L_2 I_2^2 + M I_1 I_2 = \frac{1}{2} \mu_0 H_1^2 l d_1 w_1 + \frac{1}{2} \mu_0 H_2^2 l d_2 w_2, \tag{3}$$

where I_1, I_2, d_1, d_2, and l are as defined in Figure 4, μ_0 is the permittivity of free space, w_1 is the width of the lower (tin) strip, and w_2 is the width of the upper (lead) strip. The magnetic fields are given by

$$H_1 = \frac{I_1 + I_2}{w_1}, \quad H_2 = \frac{I_2}{w_2}, \tag{4}$$

where $w_1 \geqq w_2$. Substitution of Equations (4) into Equation (3) yields

$$\frac{1}{2} L_1 I_1^2 + \frac{1}{2} L_2 I_2^2 + M I_1 I_2 = \frac{1}{2} \mu_0 \frac{I_1^2 l d_1}{w_1} + \frac{1}{2} \mu_0 I_2^2 l \left(\frac{d_1}{w_1} + \frac{d_2}{w_2} \right) + \mu_0 \frac{I_1 I_2 l d_1}{w_1}. \tag{5}$$

From Equation (5),

$$L_1 = \frac{\mu_0 l d_1}{w_1} \tag{6}$$

$$L_2 = \mu_0 l \left(\frac{d_1}{w_1} + \frac{d_2}{w_2} \right) \tag{7}$$

$$M = \frac{\mu_0 l d_1}{w_1}. \tag{8}$$

Note that $M = L_1$. For simplicity, the effects of penetration depths in tin and lead have not been considered.

The current i_2 in the L_2 branch of the equivalent circuit represents the stored current, I_s. If v is eliminated in Equations (1) and (2), the following equation, which can be solved for i_2, is obtained:

$$(L_1 - M) \frac{di_1}{dt} = (L_2 - M) \frac{di_2}{dt}. \tag{9}$$

From Equations (6) and (8), it is apparent that

$$(L_2 - M) \frac{di_2}{dt} = 0,$$

and from Equations (7) and (8), that

$$L_2 - M \neq 0. \tag{10}$$

Thus,

$$\frac{di_2}{dt} = 0,$$

and $i_2 = I_s = \text{constant}$.

If $i_2(t)$ is zero, the value of stored current, I_s, is also zero. Furthermore, when $I_s = 0$, any current applied to the digit lines immediately flows in the L_1 branch, which represents the lower part of the tin line. Thus, the equivalent-circuit model is in complete agreement with the physical model.

In the conventional cell, any current of amplitude I_D applied to the digit line splits in the proportion

$$I_1 = \frac{L_1}{L_1 + L_2} I_D, \tag{11}$$

$$I_2 = \frac{L_2}{L_1 + L_2} I_D. \tag{12}$$

Sequence of Cell Operation

The sequence of operation for both the bridge cell and the conventional cell is shown by the switching waveforms in Figure 5. At time

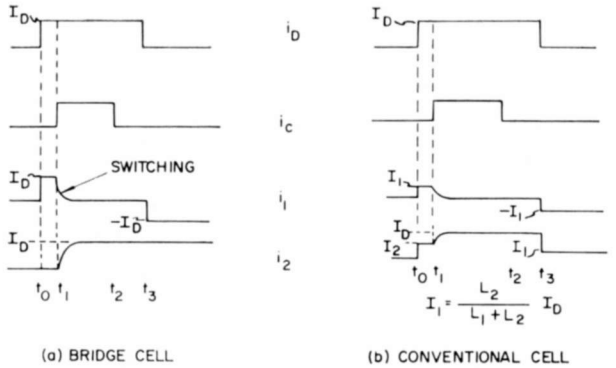


Fig. 5—Switching waveforms for the bridge cell and for a conventional cell.

t_0, a step current of amplitude I_D is applied to the digit line (i_D). This current flows in the bottom of the bridge (i_1). When a current is applied to the word line at time t_1, the current I_D in the lower part of the bridge is switched to the upper part of the bridge (i_2). At t_2, the word current is removed. At t_3, the digit current is removed, and a stored current continues to circulate in the bridge area. In the bridge cell, unlike the conventional cell, the amplitude of the stored current is equal to the amplitude of the current in the digit line, I_D.

Sensing and Speed of Operation

In the bridge cell, the sensing of the stored information is accomplished as a result of driving the lower part of the bridge normal by the application of a current to the word line. The voltage across the bridge is used as the sense signal. Figure 6 shows an idealized sense voltage. The equations for the voltage across the bridge cell (with the

switch open, as shown in the equivalent model of Figure 3) can be used to derive the equation for the sense voltages. With the switch open, the voltage across each parallel branch is

$$L_1 \frac{di_1}{dt} + M \frac{di_2}{dt} + Ri_1 = v, \tag{13}$$

$$M \frac{di_1}{dt} + L_2 \frac{di_2}{dt} = v. \tag{14}$$

Fig. 6—Idealized sense-voltage waveform.

During reading, $i_D = 0$. Thus,

$$\frac{di_1}{dt} = \frac{-di_2}{dt}. \tag{15}$$

Equations (13), (14), and (15) are combined to obtain

$$(L_1 + L_2 - 2M) \frac{di_1}{dt} + Ri_1 = 0,$$

or

$$i_1 = -I_s e^{-t/\tau}, \tag{16}$$

where

$$\tau = \frac{L_1 + L_2 - 2M}{R} = \frac{L_2 - L_1}{R}.$$

From Equations (14) and (15), the voltage is

$$v = -(L_2 - M) \frac{di_1}{dt} = I_s R e^{-t/\tau}. \tag{17}$$

The peak amplitude of the sense voltage, V_s, can therefore be expressed

as $V_s = I_s R$. The characteristic time constant for the decay of the sense voltage is τ. This same τ applies to the switching shown in Figure 5.

For the conventional cell, the sense voltage is given by

$$V_s = I_s R \frac{L_2}{L_1 + L_2}, \tag{18}$$

and the time constant is given by

$$\tau = \frac{L_1 + L_2}{R}. \tag{19}$$

The bridge cell operates faster and gives a greater output than its conventional counterpart because: (1) it stores more current for a given I_D, and (2) there is no voltage division between the inductances.

Another feature unique to the bridge cell is the increased switching gain effect. This increased gain results from the fact that the field produced when current is routed to the upper (lead) branch of the bridge aids the field produced by the word current when it switches the lower branch from superconducting to normal.

Experimental Results

A two-cell experimental sample, shown in Figure 7, was fabricated and tested to provide experimental verification of the operation of the bridge cell. The width of the word line is 10 mils, and the width of the digit line for both the upper and the lower portions of the bridge is 20 mils. The operating waveforms for this sample are shown in Figure 8. The waveforms to the left of the figure show coincident word and digit currents and the resulting sense voltage (on an expanded time scale); the waveforms to the right show the noncoincident condition. The sense trace corresponds to the time of the leading edge of the first word-current pulse. In the experiments, the word and digit current pulses were repeated at a rate of the order of 100 cycles per second. Thus, the current was stored in the cells for about 100 milliseconds before the reading.

Digit-Current Pulse Amplitude

Figure 9 shows the storage of a negative digit-current pulse. The total digit current is stored; thus, analog information can be represented by the amplitude of the digit pulse in both the negative and

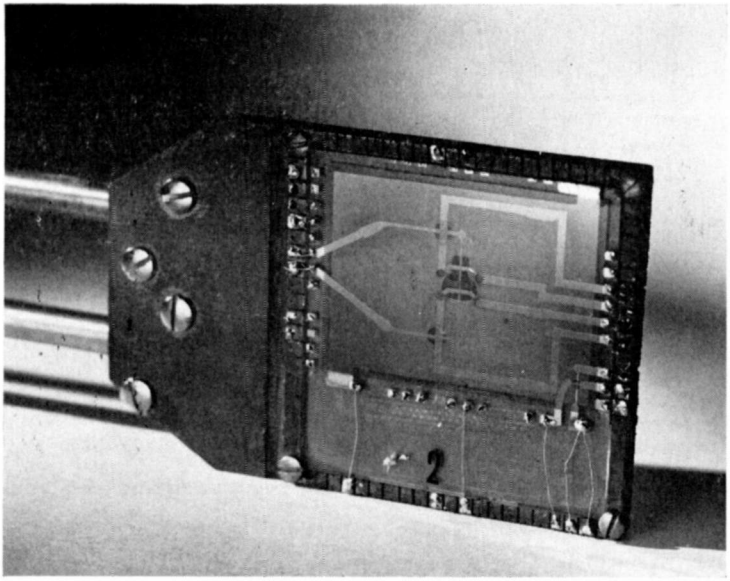


Fig. 7—Photographs of a two-cell experimental sample.

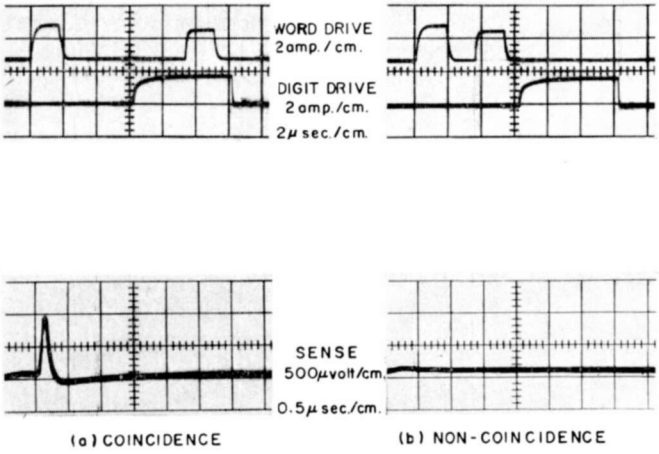


Fig. 8—Waveforms of word current, digit current, and sense voltage for both the coincident and the noncoincident conditions.

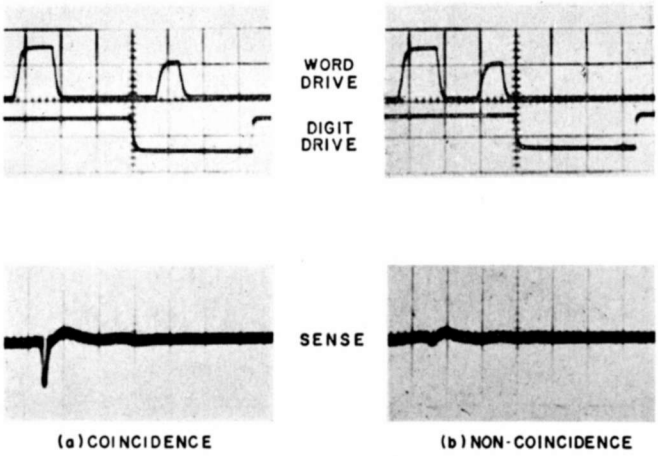


Fig. 9—Waveforms showing storage of a negative digit-current pulse.

positive directions. The series of photographs in Figure 10 shows the variation of the sense pulse with respect to the amplitude of the digit-current pulse, and Figure 11 shows peak amplitude of the sense signal as a function of the amplitude of the digit-current pulse. These data indicate that it may be possible to use the bridge cell for a memory

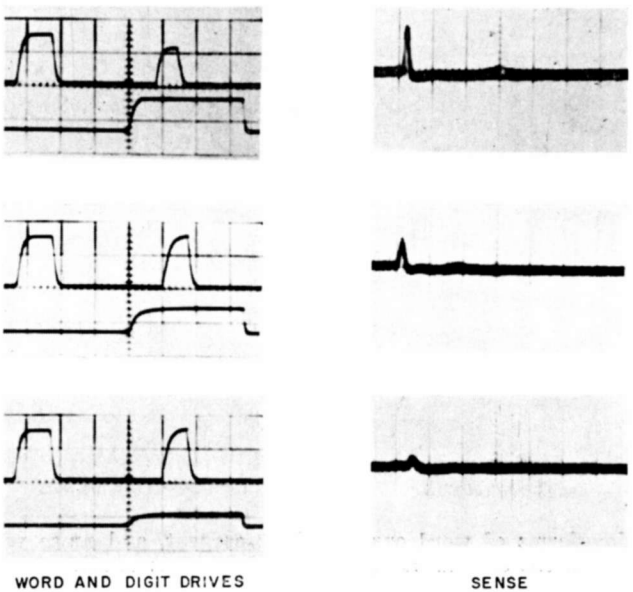


Fig. 10—Waveforms showing effect of amplitude of digit drive current.

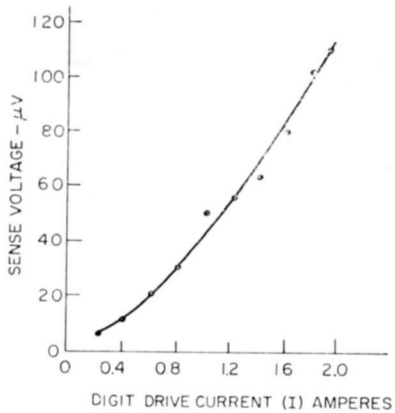


Fig. 11—Amplitude of the sense voltage as a function of the amplitude of the digit drive pulse.

having quantized levels of storage. The use of quantized levels of storage and redundancy may permit a more economical method of building an ultra-large-capacity memory. There is no doubt that the cell can be used as a neural element.

Superconductive-to-Normal Transition

The read–write cycle permits the superconducting-to-normal phase curve to be plotted easily. If the word-current write pulse is decreased until the sense pulse is at the threshold of detection, the combination of control (word) current and gate (lower branch of bridge) current at which superconductive-to-normal transition occurs can be recorded. A plot of this normal-to-superconducting transition is shown in Figure 12.

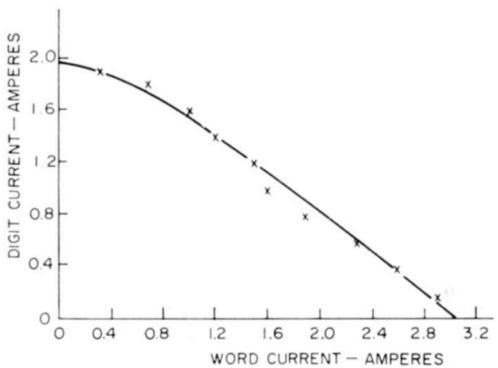


Fig. 12—Curve showing superconducting-to-normal transition for the bridge cell.

Amplitude of the Sense Voltage

The expression for the amplitude and shape of the sense signal given by Equation (17) assumes that the rise time of the word-current read pulse is much shorter than the time constant, τ, of the cell. For the experimental sample this assumption was not true; Equation (17), therefore, does not represent the observed sense signals. The amplitude of the sense signal, however, is a function of the slope of the leading edge of the word-current read pulse, as shown by the curve in Figure 13.

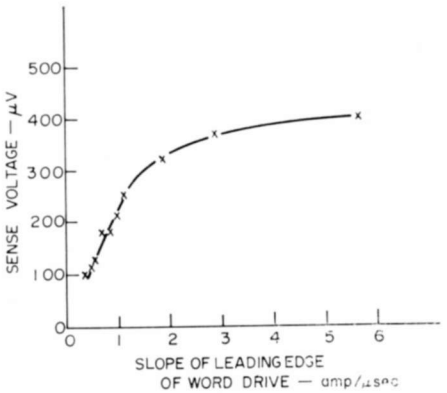


Fig. 13—Amplitude of the sense voltage as a function of the slope of the leading edge of the word-drive read pulse.

When the cell is in an array, the ultimate speed of operation is governed by the speed of the decoding matrix, and not by the speed of the cell, the latter being in the nanosecond region. Thus, one is most interested in the sense output when the word-current pulses are rise-time limited. This condition is expressed by

$$\frac{d\phi}{dt} = v(t)$$

$$\int_0^\infty v(t)\,dt = \phi_0 = (L_2 - L_1)\,I_s. \tag{20}$$

When a flux ϕ_0 is trapped, the area within a sense pulse is independent of drive rise time. If a square sense pulse of duration t_0 is assumed, the sense voltage is given by the equation

$$V_s = \frac{(L_2 - L_1) I_s}{t_0} . \tag{21}$$

Equation (21) gives a coarse approximation for sense voltage when t_0 can be approximated.

When the cell time constant dominates the speed of operation, the amplitude of the sense output, as given by Equation (17), is proportional to I_sR and is independent of the inductance. However, when the rise time of the word-current pulse dominates, the amplitude of the sense signal is proportional to I_sL (the stored flux) and is independent of the resistance for a given rise time. The stored flux can be increased in three ways: (1) by an increase in the length of the bridge, (2) by an increase in the height of the bridge, and (3) by a decrease in the width of the bridge or a decrease in just the width of the upper section. In the last case, the inductance of the bridge is inversely proportional to the width and directly proportional to the length, but the maximum stored current decreases proportionally with the width of the lower section of the bridge. Since the maximum stored current is determined by the current at which the cell self-switches, the maximum sense output does not decrease when narrower geometries are used. An increase in the length of the bridge increases the sense output voltage, but at a cost of a reduction in packing density. In the experimental cell, the length of the bridge is 200 mils. The height of the bridge was varied between 12,000 and 128,000 angstroms. The success of samples with bridge heights greater than 100,000 angstroms demonstrates that a line made of lead can traverse a 100,000-angstrom insulation rise without the lead cracking or stretching too thin to support its current. Sense signals as high as 1000 microvolts were obtained.

Speed of Operation

The potential of the bridge cell for high-speed operation, in the order of 5 to 10 megacycles, is illustrated in Figure 14. Figure 14(b) shows the coincidence of two read pulses of the order of 100 to 200 nanoseconds. Figure 14(a) shows the coincidence of two wider pulses. Note that the sense voltage has not changed; therefore the same current was stored in both cases. Because the read pulses for Figure 14(a) were more triangular in shape, the cell must switch in less than 100 nanoseconds.

Bridge-Cell Memory Arrays

The bridge cell most closely fits the word-organized array. Because the stored current loop is perpendicular to the substrate, the array is

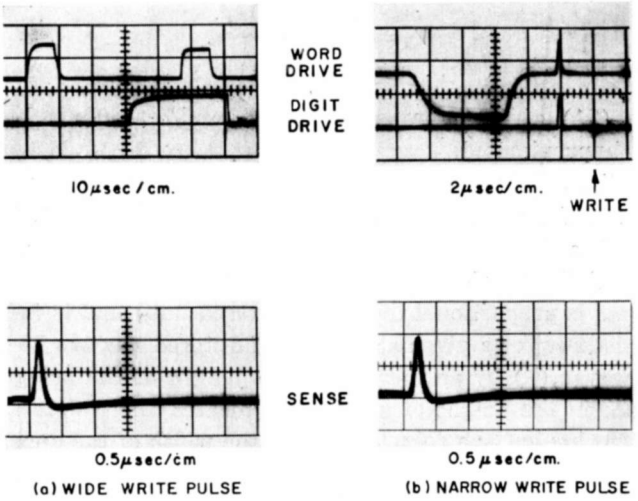


Fig. 14—Waveforms showing potential of bridge cell for high-speed operation.

formed by a set of lines, as shown in Figure 15. One of the main advantages of a word-organized array, as compared to a bit-organized array, is the elimination of the necessity for close cell-to-cell tolerance. In most bit-organized arrays, this tolerance is ideally only 1 to 1.4 times the minimum drive current.

The tolerances on the digit-drive current for the word-organized bridge-cell array are as follows:

For the digit-drive current, I_D:

Self-switch of cell $> I_D >$ minimum necessary for sensing

For the word-drive current, I_W:

Self-switch of drive lines $> I_W >$ minimum necessary for worst cell

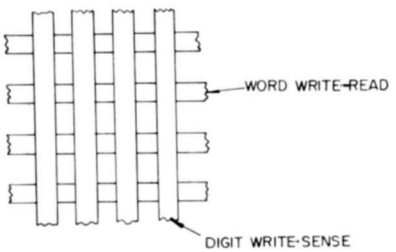


Fig. 15—Schematic of a word-organized memory array.

Because the separation of minimum and maximum values can easily be an order of magnitude, tolerances can be an order of magnitude greater than those exhibited by the coincident-current cell in a bit-organized array. However, there are three disadvantages: (1) the word-organized memory requires more elements in the switching matrix (however, the number of elements in cryogenic application does not appear to be a problem), (2) there is a greater number of interconnections between memory planes (this is a property of word-organized memories), and (3) the address lines are longer than for the bit-organized array and the propagation delay for these long lines can limit the speed in ultra-large-capacity memories.

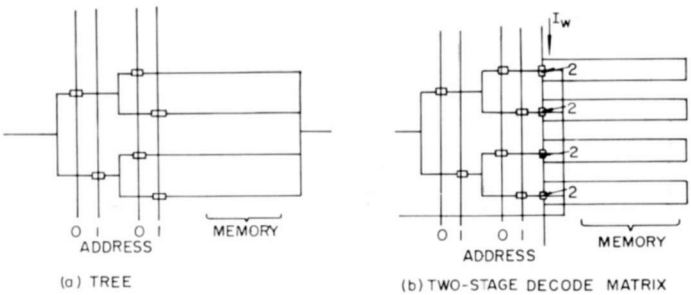


Fig. 16—Schematics of two decode-matrix circuits for a bridge-cell memory array.

Decode Matrix

Although the simple cryotron tree shown in Figure 16(a) can be used as the decode matrix in bridge-cell arrays, the two-stage decode matrix of Figure 16(b)[1] provides a higher operating speed. In the two-stage circuit, the tree excites one of the cryotrons (labeled 2) and thereby causes a shift in the current, I_W, over the memory area. The tree of Figure 16(a) has a time constant[1]

$$\tau = (\log_2 W)\,\frac{dL_C}{R} \quad \text{(word organized)} \tag{22}$$

$$\tau = (\log_2 W^{1/2})\,\frac{W^{1/2}L_C}{R} \quad \text{bit organized} \tag{23}$$

[1] R. W. Ahrons, "Superconductive Memories," Ph.D. Thesis, Polytechnic Institute of Brooklyn, June 1963.

where W is the number of word locations, d is the number of bits per word, L_C is the inductance per cell, and R is the resistance of a cryotron gate. For the circuit of Figure 16(b), the effective time constant can be approximated by

$$\tau = (\log_2 W)\frac{L_t}{R_1} + k\frac{dL_C}{R_2}, \tag{24}$$

where R_1 is the resistance of the cryotron gates in the tree, R_2 is the resistance of cryotron gate 2, L_t is the effective inductance of the control

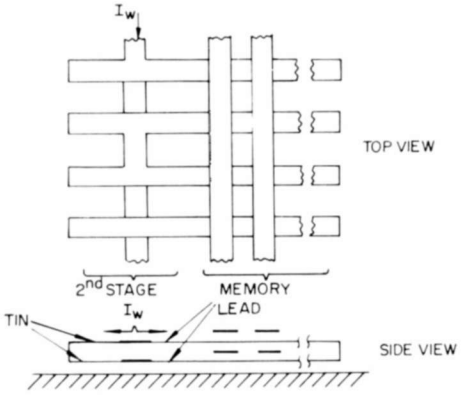


Fig. 17—Schematic of the second stage of the decode matrix and of the memory.

of cryotron 2 and the internal portion of the tree, and k should be a constant in the order of unity magnitude. Some superconducting memories cannot use the two-stage decode matrix of Figure 16(b) because this matrix requires a return line for each memory line. For a small array, the ratio dL_C/R_2 is the dominant term of Equation (24); this term is not a function of the number of words. Because d is usually two or more orders of magnitude less than $W^{1/2}$ for large arrays, the ratio dL_C/R_2 is considerably less than the time constant of Equation (23). For large arrays, $L_t \ll W^{1/2}L_C$. When the two-stage decode matrix is used, the operation of the word-organized array should be much faster than that of the bit-organized continuous-sheet memory.

The memory and second stage of the two-stage decode matrix can be implemented as shown in Figure 17. The current in the tin is switched by a control that is not shown. The memory return line is beneath the memory. Thus the memory still requires a minimum of surface area on the substrate.

Word–Bit-Organized Array

In a word-organized array, all the bits per word are located on one plane. In a bit-organized array, only one bit of the word is located on each plane. The compromise word-bit-organized array accommodates as many bits of the word on the array plane as convenient for interconnection. Figure 18 shows an example of one word line of such a memory array. In the example, the word line is divided into four words each having four bits. Any current switched into the word line remains on the bottom until one control is excited. The word drive current is then switched to the upper line only in the region of the word and

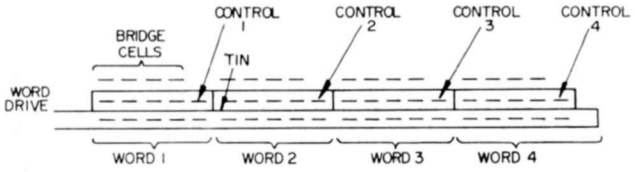


Fig. 18—Side view of one word line of a word-bit-organized array.

thereby excites only those bridge cells associated with the word. The remaining bits of that particular word are on other substrates. The controls of Figure 18 use a decode matrix composed of cryotrons similar to one of the matrices of Figure 16. Thus, in the word-bit-organized array, one direction on the plane is a word direction and the other direction is a mixture of words and bits.

Conclusion

The use of the bridge cell in the word-organized array allows very wide tolerances from cell to cell; however, interconnection of the array becomes more complex. The speed of the decoding matrix and the lengths of sense and address lines may limit the operating speeds in ultra-large-capacity arrays. In the word-bit-organized array, the interconnection problem is lessened, but at some sacrifice in operating speed, which is limited by the decode matrix and packing densities.

Acknowledgment

The author acknowledges stimulating discussions on the operation of the bridge cell with many members of the cryogenic group on the staff at RCA Laboratories. He also thanks Jack Hewitt, Mario Lavalva, and Wayne Robbins for their part in the making and testing of the memory cells.

DESIGN AND PERFORMANCE OF A 20-KILOWATT LATCHING NONRECIPROCAL X-BAND FERRITE PHASE SHIFTER

By

W. W. Siekanowicz,* W. A. Schilling,* T. E. Walsh,*
I. Bardash,** and I. Gordon***

Summary—*This paper describes the design, construction, and electrical performance of a 20-kilowatt, 8.5 to 9.5 gigacycle, nonreciprocal, latching phase shifter consisting of three bits that provide differential phase shifts of 45, 90, and 180 degrees. The small-signal design is based on computer results that show the differential phase shift as a function of ferrite geometry, remanence, dielectric constant, and frequency. Expressions are also given for calculation of r-f fields in terms of incident power.*

Of particular importance in the construction of the phase shifter is the use of a soldered bond between the ferrite and the waveguide surfaces that provides excellent mechanical, thermal, and electrical performance. The experimental results show 20 kilowatts of peak power, an average insertion loss of about 0.7 db, a difference of less than 11% between measured and computed results, and a high degree of constructional reproducibility indicated by less than ±2.5-degree deviations of the differential phase shift from the average.

Introduction

During the last few years, there has been extensive development of latched ferrite devices[1] and, in particular, nonreciprocal phase shifters of the type illustrated in Figure 1(a). These devices, which are being developed for electronically steerable phased-array radars, have the following features: rapid change of phase shift (typically several microseconds), low switching energy (ranging from tens to hundreds of microjoules), zero holding power, and small volume and weight.

The heart of the nonreciprocal phase shifter is the ferrite toroid shown in Figure 1(b). The phase shifter is operated by passage of a direct current through the magnetizing wire that produces a circumferential magnetizing field H, indicated by the arrows in Figure 1(b).

* RCA Electronic Components and Devices, Princeton, N. J.
** RCA Missile and Surface Radar Division, Moorestown, N. J.
*** RCA Laboratories, Princeton, N. J.

[1] M. A. Treuhaft and L. M. Silber, "Use of Microwave Ferrite Toroids to Eliminate External Magnets and Reduce Switching Power," *Proc. IRE*, Vol. 46, p. 1538, Aug. 1958.

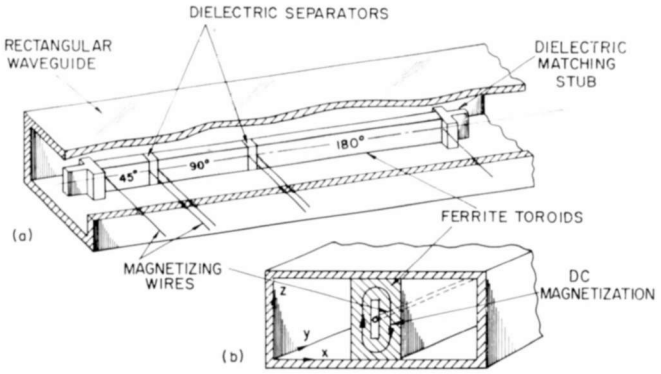


Fig. 1—Three-bit nonreciprocal latched ferrite phase shifter: (a) schematic representation and (b) cross-sectional view of the phase shifter.

Because the magnetic path is closed entirely within the magnetic material, demagnetizing effects are absent, and the toroid operates at the remanent magnetization of $4\pi M_r$ when the direct current is turned off, as illustrated in Figure 2.

The phase shift through the device is determined by the properties and the geometry of the ferrite, and orientation of d-c magnetization (Figure 1(b)) with respect to the direction of r-f propagation. Strong r-f interaction occurs between the spin dipoles in the ferrite and the r-f wave for clockwise d-c magnetization and r-f propagation in the forward ($+y$, Figure 1(b)) direction.[2] The interaction is weaker for

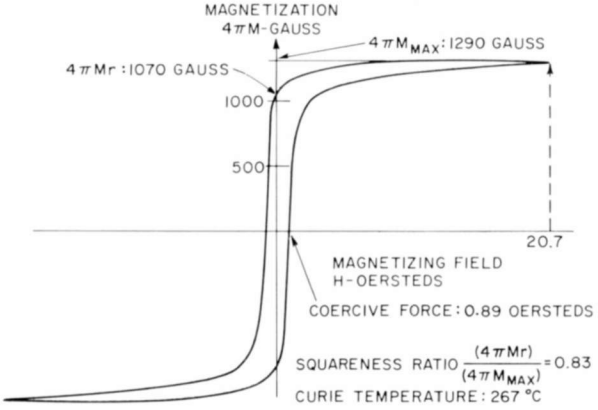


Fig. 2—Hysteresis loop and magnetic data for the toroid material.

[2] A. G. Fox, S. E. Miller, M. T. Weiss, "Behavior and Application of Ferrites in the Microwave Region," *Bell System Tech. Jour.*, Vol. 34, p. 5, Jan. 1955.

counterclockwise magnetization. The difference between these interactions results in different phase shifts. This phenomenon is directly related to the opposite signs of the off-diagonal components of the ferrite permeability tensor. For the three-bit phase shifter shown in Figure 1(a), the differences between the phase shifts for clockwise and counterclockwise directions of d-c magnetization are 45, 90, and 180 degrees.

This paper describes the theoretical and practical designs, construction, and electrical performance of a 20-kilowatt, X-band, three-bit phase shifter of the type illustrated in Figure 1. Computer results used for determination of optimum toroid dimensions are presented,

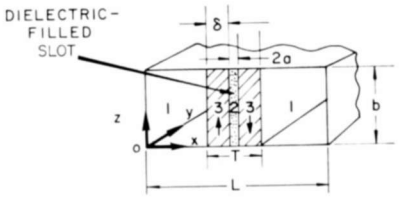


Fig. 3—Simplified model used for mathematical analysis.

as well as plots of r-f fields for the ferrite material used in the actual device. The construction and performance of the phase shifter are then described. Mathematical discussions relevant to the paper are included in the Appendix.

Calculation of Differential Phase Shift

The small-signal propagation constants β for the fundamental TE mode in the simplified configuration shown in Figure 3 were calculated for positive and negative directions of propagation from the characteristic equation, Equation (1) in the Appendix, published previously by Lax and Button[3] and by Levey and Silber.[4] The numerical results presented below were obtained from Equation (1) by use of the RCA 601 digital computer; r-f losses are neglected, and a zero applied d-c magnetic field, H, is assumed in the components of the susceptibility tensor.*

[3] B. Lax and K. J. Button, *Microwave Ferrites and Ferrimagnetics*, p. 434, McGraw-Hill Book Co., Inc., New York, 1962.

[4] L. Levey and L. M. Silber, "A Fast-Switching X-Band Circulator Utilizing Ferrite Toroids," *1960 IRE WESCON Convention Record*, Vol. 4, Part 1, pp. 11-20.

* See Reference (3), pp. 149-151.

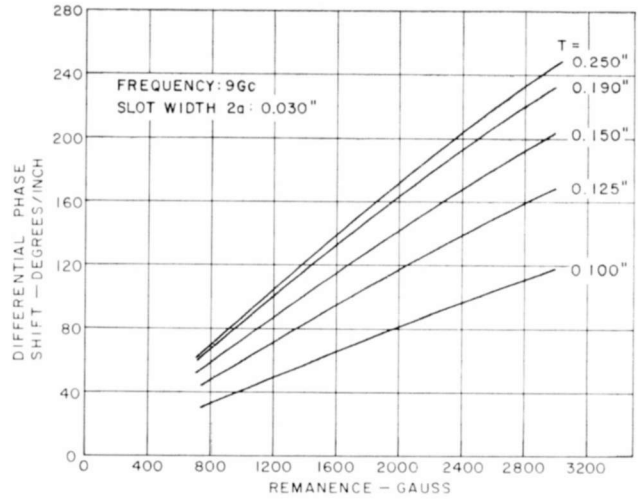


Fig. 4—Differential phase shift as a function of remenance for several values of toroid width (K_m=15, L=0.803 inch).

Figures 4 through 7 show the computed differential phase shifts for (a) a reduced waveguide width of 0.803 inch, which was prescribed by a limitation on the horizontal dimensions of the final device, (b) a relative ferrite dielectric constant K_m of 15, which is very nearly equal to the value used in the actual phase shifter, and (c) an air-filled

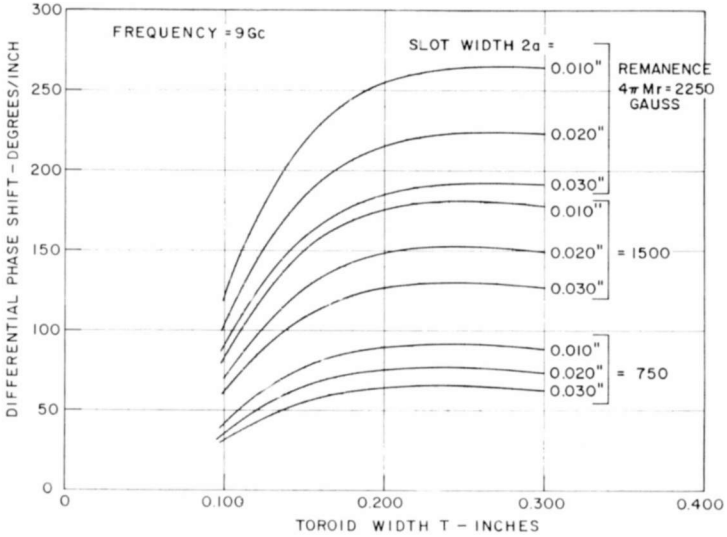


Fig. 5—Differential phase shift as a function of toroid width for several values of remanence and slot width (K_m=15, L=0.803 inch).

center slot (region 2 in Figure 3) and spaces (region 1) between the ferrite toroid and the narrow walls of the waveguide.

Figure 4 shows the differential phase shift as a function of remanence $4\pi M_r$ for several values of the toroid width T. The curves show that this phase shift is approximately proportional to the remanence.

The effect of toroid width T on the differential phase shift is shown

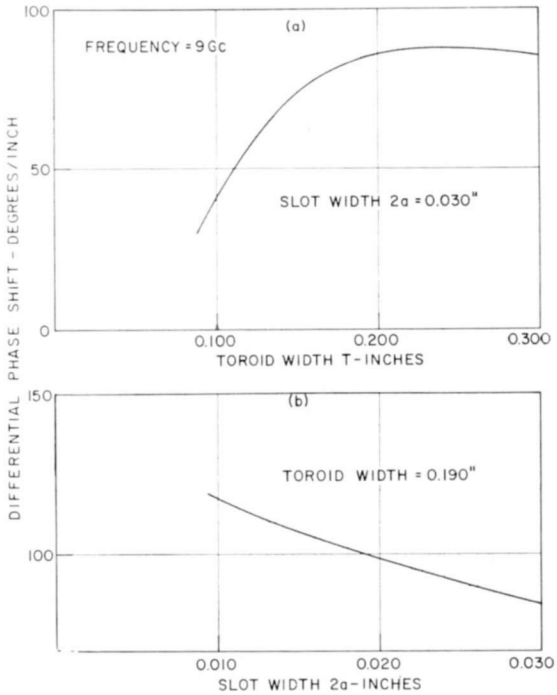


Fig. 6—Differential phase shift (a) as a function of toroid width, and (b) as a function of slot width at a remanence of 1000 gauss (K_m=15, L=0.803 inch).

more clearly in Figure 5. The phase shift has a broad maximum in the vicinity of a toroid width of 0.250 inch, and this value is very nearly independent of remanence and slot width.

Figure 6(a) shows the differential phase shift as a function of toroid width, T, for a slot width of 0.030 inch and a remanence of 1000 gauss. This remanence is approximately equal to the value used in the actual phase shifter. A final dimension of 0.190 inch was selected for the actual device as the best compromise between large differential phase shift, small cross-sectional area (and therefore volume), and low r-f losses. This dimension corresponds to a parameter D of 0.383

($D = 2\pi\delta/\lambda_0$, where λ_0 is the free-space wavelength), which is close to the minimum loss condition computed by Schlömann[5] for a similar toroid geometry.

Figure 6(b) indicates that the differential phase shift increases with decreasing slot width. In this application, a slot width of 0.030 inch was used as a compromise between minimum slot width (and therefore maximum differential phase shift) and achievement of accurate slot dimensions by ultrasonic drilling. Figure 7 shows that the differential phase shift does not vary significantly with frequency in the range from 8.5 to 9.5 gigacycles.

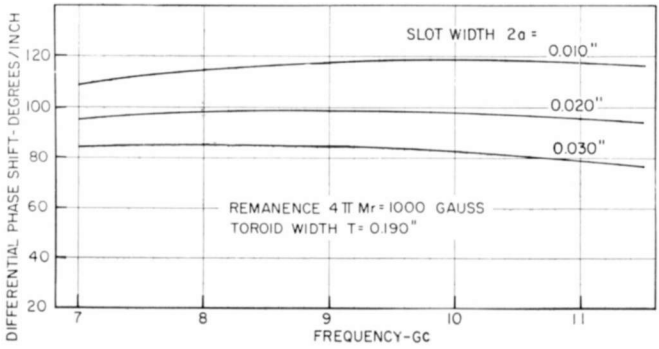


Fig. 7—Differential phase shift as a function of frequency for several values of slot width (K_m=15, L=0.803 inch).

Table I lists the parameters of the ferrite material and the toroid and waveguide dimensions used in the actual phase shifter. Table II gives propagation characteristics and r-f power distributions. The calculation of these parameters is described in the Appendix.

Figures 8 through 10 show curves of r-f field strengths in one-half of the waveguide cross section at one watt of incident r-f power for the conditions listed in Tables I and II. The calculation of these fields is described in the Appendix.* The r-f fields E_z and H_x are even functions and H_y is an odd function, with respect to the vertical center line of the waveguide cross section. For the relatively thick ferrite toroid considered in the present design, the field patterns for the

[5] E. Schlömann, "Theoretical Analysis of Twin Slab Phase Shifters in Rectangular Waveguide," presented at the 1965 G-MTT Symposium, Clearwater, Florida, May 7, 1965.

* Equations (7) through (12) in the Appendix are written in MKS units and must be multiplied by $4\pi \times 10^{-3}$ to obtain magnetic field values in oersteds as shown in Figures 9 and 10.

Table I—Characteristics of Ferrite and Physical Dimensions of Toroid and Waveguide

Frequency	9.0 gigacycles
Remanence	1000 gauss
Relative dielectric constant of ferrite	15
Diagonal component χ_{xx} of permeability tensor*	0
Off-diagonal component χ_{xy} of permeability tensor*	j 0.311
Parameter† θ for clockwise-magnetization (Fig. 3)	—j 3.21
Parameter† ρ	1.11
Waveguide width L	0.803 inch
Thickness δ (see Figure 3)	0.080 inch
Slot width 2a (see Figure 3)	0.030 inch

* See Reference (3), pp. 149-151.
† See Reference (3), p. 358.

Table II—Propagation Constants and R-F Power Distributions

Parameter	Clockwise Magnetization	Counter-clockwise Magnetization
Transverse (x) propagation constant k_a in air (radians per meter)	j 531	j 469
Transverse (x) propagation constant k_m in ferrite (radians per meter)	405	475
Axial (y) propagation constant β (radians per meter)	563	505
Per cent of r-f power flowing through the ferrite toroid	67.3	65.5
Per cent of r-f power flowing between the toroid and the waveguide wall	13.8	25.1
Per cent of r-f power flowing through the center slot	18.9	9.4
Differential phase shift = 84.3 degrees per inch		

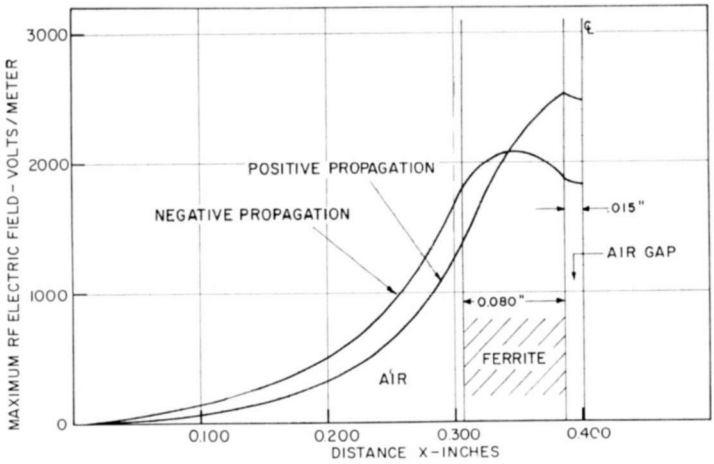


Fig. 8—R-F electric field patterns at one watt of r-f power.

positive and negative directions of propagation are significantly different.

Construction of Phase Shifter

The toroid was made of holmium-doped, gadolinium-substituted YIG.† The hysteresis loop and the major characteristics of the material are shown in Figure 2. The loop was traced for a maximum magnetiz-

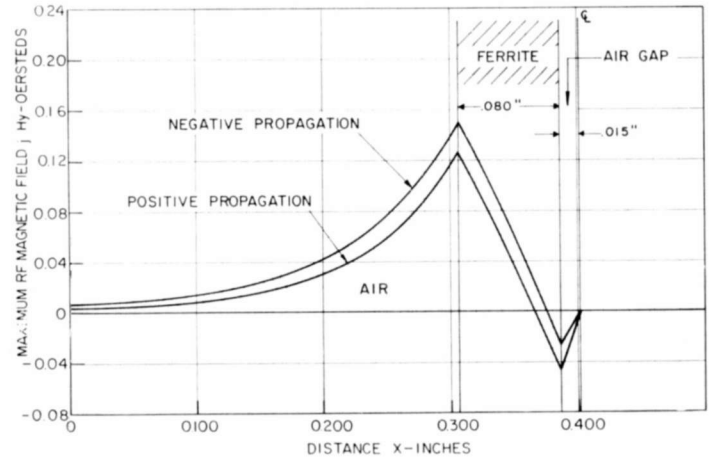


Fig. 9—Axial r-f magnetic field patterns at one watt of r-f power.

† The material was purchased from Trans-Tech, Inc., Gaithersburg Maryland.

ing field H of 20.7 oersteds, which is approximately equal to the value used in the actual device. For a remanence $4\pi M_r$ of 1070 gauss and an operating frequency of 9 gc, the ratio ω_r/ω is 0.333 (where $\omega_r = 1.76 \times 10^7 \times 4\pi M_r$ and $\omega = 2\pi \times$ frequency, both in radians per second). This ratio satisfies one of the criteria for high peak r-f power.[6]

The lengths of bits that would provide differential phase shifts of 45, 90, and 180 degrees were obtained by use of the computer results presented previously and small corrections derived from preliminary measurements.

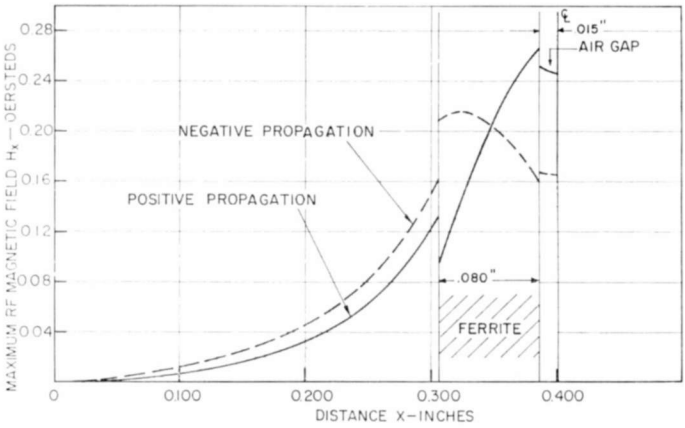


Fig. 10—Transverse r-f magnetic field patterns at one watt of r-f power.

In the development of the phase shifter, particular attention was paid to the method of attaching the toroid to the metal surfaces of the waveguide. Dielectric bonding agents, conducting epoxies, and mechanical clamping were eliminated for reasons of limited long-range stability, excessive r-f loss, low thermal conductivity, low mechanical strength, loss of differential phase shift as a result of magnetostriction, and lack of reproducibility.

The provision of excellent heat transfer between the toroid and the external surroundings was of particular importance in this application because the specifications called for an average power capability of 200 watts. To satisfy this requirement and others, a technique was developed for applying a low-loss, strongly adhering, silver coating to the surface of the toroid. For this particular application, the toroids were soldered to 0.020-inch-thick, 0.190-inch-wide copper-plated molyb-

[6] R. Seckelmann, "Microwave Digital Phase Shifters," *Proceedings 1965 Imtermag Conference*, p. 3.2-6, Fig. 7.

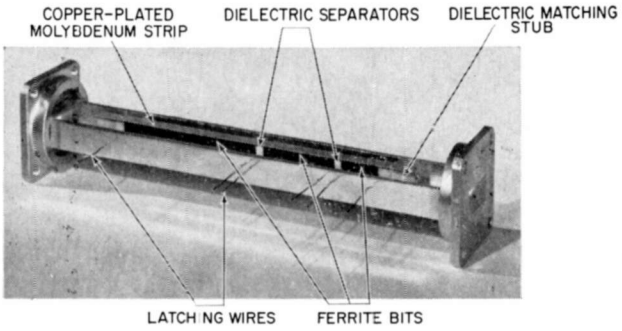


Fig. 11—Photograph of the phase shifter with the upper waveguide wall partially removed for exposure of the ferrite subassembly.

denum strips by use of a 60/40 tin–lead solder. Mechanical tests indicated that the ultimate tensile strength of the soldered bond is in excess of 1100 pounds per square inch. Soldering of the ferrite toroids, the high-alumina matching sections, and the metalized dielectric separators between the ferrite toroids to the molybdenum strips was performed simultaneously with suitable jigging. This operation results in the rigid subassembly shown in Figure 11. Nonmagnetic dielectric separators are used to eliminate magnetic coupling between the bits and to provide continuity of dielectric constant.

The molybdenum strips of the subassembly are then soldered to the grooves in the horizontal walls of the waveguide, as shown in Figure 12, to form a mechanically rigid assembly. The molybdenum strips, which have a coefficient of thermal expansion approximately equal to that of the ferrite material, are used to minimize axial stresses on the toroids that might otherwise cause changes of phase shift during temperature cycling as a result of magnetostriction.

The completely soldered phase-shifter assembly approaches the

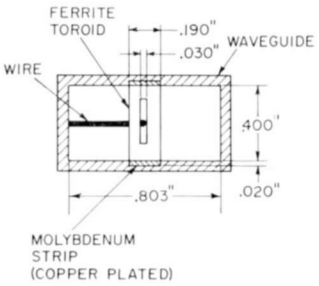


Fig. 12—Cross sectional view of the phase shifter.

practical limit of average-power-handling capacity and, as will be shown later, provides low insertion loss.

Electrical Performance

Figure 13 shows a comparison between measured and computed values of the differential phase shift per inch across the frequency range from 8.5 to 9.5 gigacycles. The measured curve represents average values obtained from measurements on ten phase shifters. The computed curve represents the portion of the lowest curve shown in Figure 7 extending from 8.5 to 9.5 gigacycles.

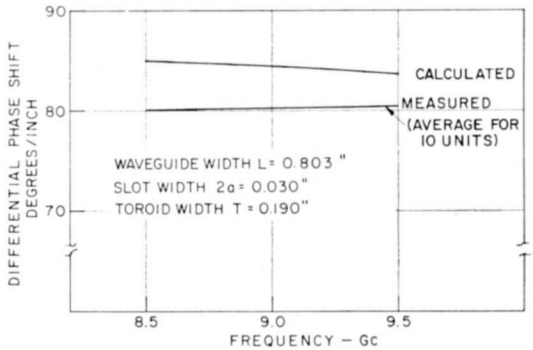


Fig. 13—Measured and calculated differential phase shifts as functions of frequency.

When a correction is made for the difference between the actual remanence of 1070 gauss shown in Figure 2 and the value of 1000 gauss used for the computed curve, the maximum difference between the measured and the calculated figures is less than 11 percent of the theoretical values. This difference is to be expected because the d-c magnetization in the corner regions of the toroid is curved (Figure 1(b)), and, as a result, those regions are only partially effective in providing differential phase shift.

Figure 14 shows the differential phase shift as a function of frequency for each of the three bits in the phase shifter. The curves show the average values and the deviations from the average obtained from measurements on ten phase shifters. The maximum deviation was 2.5 degrees. In addition, the average phase shift was within 2 degrees of the nominal value.

The standing-wave ratio was maintained below 1.25 across the frequency range from 8.5 to 9.5 gigacycles. It was also found that

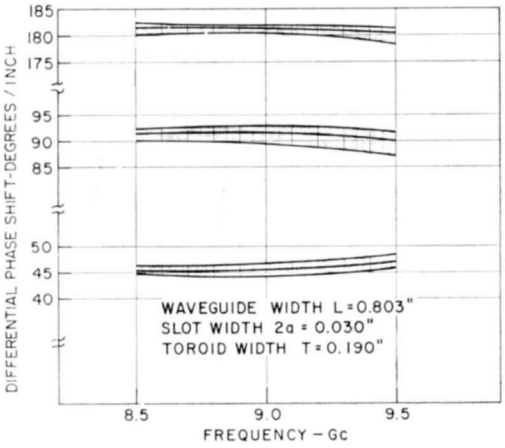


Fig. 14—Differential phase shift as a function of frequency. Heavy line indicates average of ten units; shaded portions indicate the spread.

the soldered joints provided a lower VSWR than other methods of attachment. The average, minimum, and maximum insertion losses of the complete phase shifter were 0.7, 0.4, and 1.2 db, respectively. These results show that the technique used for soldering the ferrite assembly to the waveguide walls is capable of providing an excellent r-f match and low insertion loss.

Figure 15 shows the insertion loss as a function of incident power. The results show that the insertion loss remains essentially constant at 0.9 db up to a power level of about 20 kilowatts.

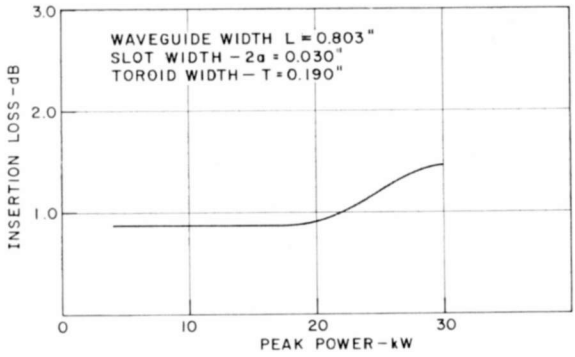


Fig. 15—Insertion loss as a function of peak r-f power input (duty cycle = 0.001, pulse length = 2.5 μsec.)

Conclusions

A 20-kilowatt, nonreciprocal, three-bit (45, 90, and 180 degree differential phases), latching phase shifter has been developed for operation at X-band. The important feature of the design is the use of a completely soldered assembly. Measurements have shown that such an assembly provides an average insertion loss of 0.7 db, and improves the r-f match. The soldered joint also approaches the ultimate design for average-power-handling capacity.

It was also demonstrated that the design provides a high degree of phase reproducibility. Measurements performed on ten units have shown that deviations of the differential phase shift from the average value were less than ±2.5 degrees across the frequency band from 8.5 to 9.5 gigacycles.

Theoretical results were also presented for the differential phase shift, field patterns, and calculation of r-f field intensities in terms of r-f power. The difference between measured and theoretical figures was less than 11 percent of the theoretical values.

Appendix

The propagation constants for the twin-slab nonreciprocal phase shifter (Figure 3) can be calculated from the following characteristic equation, which applies for a dielectric-filled center slot:[3,4]

$$\tan k_a\left(\frac{L}{2}-a-\delta\right)= \frac{\frac{k_a}{\rho}\left[\frac{k_d \sin k_d a}{\rho}-\left(k_m \cot k_m\delta + j\frac{\beta}{\theta}\right)\cos k_d a\right]}{\left(\frac{\beta^2}{\theta^2}-k_m{}^2\right)\cos k_d a-\frac{k_d}{\rho}\left(k_m \cot k_m\delta - j\frac{\beta}{\theta}\right)\sin k_d a}. \quad (1)$$

In this equation, the expressions for the parameters ρ (the reciprocal of the relative ferrite permeability) and θ have been given by Lax and Button.[3] The symbols k_d, k_m, and k_a denote, respectively, the transverse (x) propagation constants in the dielectric-filled center slot, the ferrite, and the spaces between the toroid and the narrow walls of the waveguide. The expressions for these constants have been published previously[4] and include the relative dielectric constants K_m and K_d for the ferrite and the material filling the center slot, respectively. The dimensions L, a and δ are indicated in Figure 3.

For the latched toroid, the applied d-c magnetic field H is zero, and the χ_{xx} component of the susceptibility tensor is therefore zero.[8] This condition also leads to significant simplification of expressions for χ_{xy}, θ, and ρ, as shown below:

$$\chi_{xy} = \frac{j|\gamma|4\pi M_r}{\omega} = \frac{1}{\theta} = \frac{1}{j\theta_i}, \tag{2}$$

$$\rho = \frac{1}{1+\chi_{xy}^2}, \tag{3}$$

where $|\gamma|$, the gyromagnetic ratio, is equal to 1.76×10^7 radians per second-oersted, and θ_i is a real number.

In the MKS system, the expressions for the r-f electric (E) and magnetic (H) field intensities in the phase shifter are as follows:

$$E_{z1} = [A \sin k_a x]\, e^{j(\omega t - \beta y)}, \tag{4}$$

$$E_{z2} = \left[B \cos k_d \left(\frac{L}{2} - x \right) \right] e^{j(\omega t - \beta y)}, \tag{5}$$

$$\mathrm{E}_{z3} = [C\, e^{jk_m x} + De^{-jk_m x}]\, e^{j(\omega t - \beta y)}, \tag{6}$$

$$H_{x1} = \left[\frac{A\beta}{\omega\mu_0} \sin k_a x \right] e^{j(\omega t - \beta y)}, \tag{7}$$

$$H_{x2} = \left[\frac{B\beta}{\omega\mu_0} \cos k_d \left(\frac{L}{2} - x \right) \right] e^{j(\omega t - \beta y)}, \tag{8}$$

$$H_{x3} = \frac{\rho}{\omega\mu_0} \left[C \left(\beta - \frac{k_m}{\theta} \right) e^{jk_m x} + D \left(\beta + \frac{k_m}{\theta} \right) e^{-jk_m x} \right] e^{j(\omega t - \beta y)}, \tag{9}$$

$$H_{y1} = \left[-j \frac{Ak_a}{\omega\mu_0} \cos k_a x \right] e^{j(\omega t - \beta y)}, \tag{10}$$

$$H_{y2} = \left[-j \frac{B}{\omega\mu_0} k_d \sin k_d \left(\frac{L}{2} - x \right) \right] e^{j(\omega t - \beta y)}, \tag{11}$$

$$H_{y3} = \frac{\rho}{\omega\mu_0} \left[C \left(\frac{\beta}{\theta} + k_m \right) e^{jk_m x} + D \left(\frac{\beta}{\theta} - k_m \right) e^{-jk_m x} \right] e^{j(\omega t - \beta y)}. \tag{12}$$

In these expressions, the permeability μ_0 is 1.257×10^{-6} henry per meter, x denotes distance from the narrow wall of the waveguide, and the subscripts 1, 2, and 3 refer to the regions indicated in Figure 3. The constants A, B, C, and D denote maximum values and are determined by the incident power.

When the r-f electric field E_{z2} is 1 volt per meter ($B = 1$, Equation (5)) at the center of the waveguide cross section, the r-f powers P_1, P_2, and P_3 flowing through regions 1, 2, and 3, respectively, are given by the following equations. When k_a and k_d are imaginary, as in most practical cases, the expressions for P_1 and P_2 are

$$P_1 = \frac{|A|^2|\beta|b\left(\frac{L}{2} - a - \delta\right)}{2\omega\mu_0}\left[\frac{\sinh 2|k_a|\left(\frac{L}{2} - a - \delta\right)}{2|k_a|\left(\frac{L}{2} - a - \delta\right)} - 1\right] \tag{13}$$

$$P_2 = \frac{|\beta|ab}{2\omega\mu_0}\left[1 + \frac{\sinh 2|k_d|a}{2|k_d|a}\right]; \tag{14}$$

when k_m is imaginary, P_3 is given by

$$P_3 = \frac{\rho b}{2\omega\mu_0}\left[|C|^2\left(\frac{|\beta|}{|k_m|} - \frac{1}{\theta_i}\right)\left(e^{-2|k_m|[(L/2)-a-\delta]} - e^{-2|k_m|[(L/2)-a]}\right) + |D|^2\left(\frac{|\beta|}{|k_m|} + \frac{1}{\theta_i}\right)\left(e^{2|k_m|[(L/2)-a]} - e^{2|k_m|[(L/2)-a-\delta]}\right) + 4|\beta|\delta R_c(C^*D)\right]. \tag{15}$$

When k_m is real,

$$P_3 = \frac{\rho b}{\omega\mu_0}\left\{\left(|C|^2 + |D|^2\right)\beta\delta + \left[\sin 2k_m\left(\frac{L}{2} - a\right) - \sin 2k_m\left(\frac{L}{2} - a - \delta\right)\right]\left[\frac{\beta}{k_m}\operatorname{Re}(DC^*) + \frac{1}{\theta_i}\operatorname{Im}(DC^*)\right] - \left[\cos 2k_m\left(\frac{L}{2} - a\right)\right.\right.$$

$$-\cos 2k_m\left(\frac{L}{2}-a-\delta\right)\Bigg]\left[\frac{\beta}{k_m}\,\mathrm{Im}\,(DC^*)-\frac{1}{\theta_i}\,\mathrm{Re}\,(DC^*)\right]\Bigg\} \tag{16}$$

where Re and Im denote real and imaginary parts, respectively.

When B is equal to one volt per meter, the expressions for the constants A, C, and D are as follows:

$$A=\frac{\left[\rho\left(k_m{}^2-\frac{\beta^2}{\theta^2}\right)\sin k_m\,\delta\,\cos k_d\,a+k_d\,k_m\cos k_m\,\delta\,\sin k_d\,a-jk_d\,\frac{\beta}{\theta}\sin k_m\,\delta\,\sin k_d\,a\right]}{k_m\,k_a\cos k_a\left(\frac{L}{2}-a-\delta\right)} \tag{17}$$

$$C=\frac{\left[\rho\left(k_m-\frac{\beta}{\theta}\right)\cos k_d\,a-jk_d\sin k_d\,a\right]}{2k_m\,\rho}\,e^{-jk_m[(L/2)-a]} \tag{18}$$

$$D=\frac{\left[\rho\left(k_m+\frac{\beta}{\theta}\right)+jk_d\sin k_d\,a\right]}{2k_m\,\rho}\,e^{jk_m[(L/2)-a]} \tag{19}$$

The expressions for these constants, and the characteristic Equation (1), were obtained by making E_z and H_y continuous at the outside $(x=(L/2)-a-\delta)$ and inside $(x=(L/2)-a)$ surfaces of the toroid.

In the field patterns shown in Figures 8 through 10, positive signs were used for β; the sign of θ was changed for Equations (9), (12), (16), (17), (18), and (19).

The r-f field intensities at one watt of r-f power (Figures 8 through 10) were determined as follows: First the constants A, C, and D were calculated from Equations (17) through (19) and the total r-f power $(P_1+P_2+P_3)$ was calculated for an r-f electric field intensity of one volt per meter $(B=1)$ at the center of the waveguide. The desired field intensities were then obtained from Equations (4) through (12) by use of the calculated values for the constants A, $(B=1)$, C, and D, and these equations were divided by the square root of the previously determined total r-f power.

The r-f field intensities at a total r-f power of P_T watts can be obtained by multiplying the field intensities plotted in Figures 8 through 10 by the square root of P_T.

RCA Technical Papers†

Third Quarter, 1965

Any request for copies of papers listed herein should be addressed to the publication to which credited.

"The Amplification and Time Delay Characteristics of the Plasma Density Wave in Germanium," Y. Kuniya, *Japanese Jour. Appl. Phys.* (July) 1965

"Automatic Correction of Timing Errors in Magnetic Tape Recorders," W. J. Hannan, J. F. Schanne, and D. J. Woywood, *Trans. IEEE PTGMIL* (July/October) 1965

"A 'C'-Band Rutile Traveling-Wave Maser," L. C. Morris and D. J. Miller, *IEEE Jour. Quantum Electronics* (July) 1965

"Characteristics and Applications of RCA Insulated-Gate Field-Effect Transistors," D. M. Griswold, *Trans. IEEE PTGBTR* (July) 1965

"Comment on the Paper 'Potential Distribution and Negative Resistance in Thin Oxide Films'," A. G. Revesz, *Jour. Appl. Phys.* (Communications) (July) 1965

"The Communications Explosion," D. Sarnoff, *Signal* (July) 1965

"Destructive Circuit Malfunctions and Corrective Techniques in Horizontal Deflection," C. F. Wheatley, *Trans. IEEE PTGBTR* (July) 1965

"Development of the RCA Family of 90-Degree Rectangular Color Picture Tubes," A. M. Morrell, *Trans. IEEE PTGBTR* (July) 1965

"The Efficiency of a Multicontact Hall-Effect Gyrator at Finite Magnetic Fields," A. Newton and Coauthor, *Proc. IEEE* (Correspondence) (July) 1965

Author's Comments on "The Impedance of an Electric Dipole in a Magneto-Ionic Medium," H. Staras, *Trans. IEEE PTGAP* (Communications) (July) 1965

"Indium Antimonide Thin-Film Transistor," V. L. Frantz, *Proc. IEEE* (Correspondence) (July) 1965

"Infrared Dielectric Constant and Ultraviolet Optical Properties of Solids with Diamond, Zinc Blende, Wurtzite, and Rocksalt Structure," M. Cardona, *Jour. Appl. Phys.* (July) 1965

"Intrinsic Hysteresis in Type II Superconductors," M. Cardona, J. Gittleman, and B. Rosenblum, *Phys. Letters* (July) 1965

"A New Negative Resistance Effect in Indium Antimonide," L. F. Eastman, *Proc. IEEE* (Correspondence) (July) 1965

"The Performance of Color-Television Picture-Tube Phosphor Screens," A. E. Hardy, *Trans. IEEE PTGBTR* (July) 1965

"On the Polymorphic Modifications of Phthalocyanines," J. M. Assour, *Jour. Phys. Chem.* (July) 1965

"Rare Earth Sesquiselenides and Sesquitellurides with the Sc_2S_3 Structure," J. P. Dismukes and J. G. White, *Inorganic Chem.* (July) 1965

"Techniques for Measuring the Vertical Tracking Angle of Stereophonic Phonograph Pickups," J. G. Woodward, *Jour. Aud. Eng. Soc.* (July) 1965

† Report all corrections to *RCA Review*, RCA Laboratories, Princeton, New Jersey, 08540.

"Time-Resolved Spectral Output of Pulsed GaAs Lasers," T. Gonda, H. Junker, and M. F. Lamorte, *IEEE Jour. Quantum Electronics* (July) 1965

"Wave Propagation Through Quasi-Optical Irregularities," D. A. DeWolf, *Jour. Opt. Soc. Amer.* (July) 1965

"Electron Spin Resonance of Vanadyl Phthalocyanine," J. M. Assour, J. Goldmacher, and S. E. Harrison, *Jour. Chem. Phys.* (1 July) 1965

"Beam Deflection Tube Oscillates at Two Frequencies," W. M. Austin, *Electronic Design* (July 5) 1965

"Possible Two-Stream Instabilities of Drifted Electron-Hole Plasmas in Longitudinal Magnetic Fields," B. Vural and M. C. Steele, *Phys. Rev.* (5 July) 1965

"Cyclotron Harmonic Resonances Observed by Satellites," I. P. Shkarofsky and T. W. Johnston, *Phys. Rev. Letters* (12 July) 1965

"Tunneling in III-V Compound p-n Junctions," R. M. Williams and Coauthor, *Phys. Rev. Letters* (26 July) 1965

"Absorption of Transverse Microwave Phonons by Al-Pb Superconducting Diodes," B. Abeles and Y. Goldstein, *Phys. Letters* (August) 1965

"Additional Comments on 'Excess Current in Gallium-Arsenide Tunnel Diodes'," R. D. Gold, *Proc. IEEE* (Correspondence) (August) 1965

"Conduction Mechanism in Single Crystal β-Indium Sulfide In_2S_3," W. Rehwald and G. Harbeke, *Jour. Phys. and Chem. Solids* (August) 1965

"Crystal Growth and Electro-Optic Effect of Bismuth Germanate, $Bi_4(GeO_4)_3$," R. Nitsche, *Jour. Appl. Phys.* (August) 1965

"The Effect of the Aperture Illumination on the Circular Aperture Antenna Pattern Characteristics," A. F. Sciambi, *Microwave Jour.* (August) 1965

"Electron Emission from Solidifying Uranium Surfaces," P. R. Sahm, *Jour. Appl. Phys.* (Communications) (August) 1965

"Hybrid-Coupled VHF Transistor Power Amplifier," R. M. Kurzrok, S. J. Mehlman, and A. Newton, *Solid State Design* (August) 1965

"Infrared Fiber Optics," N. Aron and L. Arlan, *Electronic Design News* (August) 1965

"Long Interval Time-To-Pulse-Height Converter," R. Wasserman, *Rev. Sci. Instr.* (Notes) (August) 1965

"Optimum Focusing of Space Camera Systems," L. S. Herczeg, *Jour. S.M.P.T.E.* (August) 1965

"Recombination and Meteor Trail Decay," B. B. Robinson, *Jour. Geophysical Research* (August) 1965

"Simple Force Magnetometer," R. L. Harvey, *Rev. Sci. Instr.* (August) 1965

"Universal Coaxial Transmission Line," B. K. Kellom, *Broadcast News* (August) 1965

"Observations of the Triplet State in Phthalocyanines," W. F. Kosonocky, S. E. Harrison, and R. Stander, *Jour. Chem. Phys.* (1 August) 1965

"Collective Behavior of Polar Impurities in Ionic Crystals," W. Zernik, *Phys. Rev.* (2 August) 1965

"Polarized Absorption and Emission in an Octacoordinate Chelate of Eu^{3+}," J. Blanc and D. L. Ross, *Jour. Chem. Phys.* (15 August) 1965

"The Overlay Transistor, Part I: New Geometry Boosts Power," D. R. Carley, P. L. McGeough, and J. F. O'Brien, *Electronics* (August 23) 1965

"The Overlay Transistor, Part II: Putting the Overlay to Work at High Frequencies," D. J. Donahue and B. A. Jacoby, *Electronics* (August 23) 1965

"Seebeck Coefficient in N-Type Germanium-Silicon Alloys: 'Competition' Region," A. Amith, *Phys. Rev.* (30 August) 1965

"Applications of Differential Equations in General Problem Solving," R. W. Klopfenstein, *Assoc. for Computer Mach.* (September). 1965
"An Avalanching Silicon Diode Microwave Amplifier," L. S. Napoli and R. J. Ikola, *Proc. IEEE* (Correspondence) (September).. 1965
"The Capabilities and Prospects of Television Camera Tubes in Applications for Astronomy," A. D. Cope, E. Luedicke, and L. E. Flory, *Jour. S.M.P.T.E.* (September) 1965
"A Cassegrainian Feed for Wide-Band Satellite Communications," P. Foldes, S. Komlos, N. K. Chitre, R. Schwerdtfeger, and T. Szirtes, *RCA Review* (September) 1965
"Controlled Sublimation Technique and Its Utilization for the Crystal Growth of Hexamine," G. E. Gottlieb, *Jour. Electrochem. Soc.* (September) 1965
"Cross-Polarized Electromagnetic Backscatter from Turbulent Plasmas," R. S. Ruffine and D. A. deWolf, *Jour. Geophysical Research* (September) 1965
"Deformation of and Stress in Epitaxial Silicon Films on Single-Crystal Sapphire," D. J. Dumin, *Jour. Appl. Phys.* (September) 1965
"Domain Wall Velocities and the Surface Layer in $BaTiO_3$," D. R. Callaby, *Jour. Appl. Phys.* (September) 1965
"Etching of Dislocations on the Low-Index Faces of GaAs," M. S. Abrahams and C. J. Buiocchi, *Jour. Appl. Phys.* (September) 1965
"High-Efficiency, High-Order, Idler-Less Frequency Multipliers Using Hyperabrupt Varactors," E. Markard and S. Yuan, *RCA Review* (September) 1965
"A Programable Integrating Television Camera for Astronomical Applications," L. A. Boyer, L. E. Flory, J. M. Morgan, and W. S. Pike, *Jour. S.M.P.T.E.* (September) 1965
"Project Ranger Television System," B. P. Miller, *RCA Review* (September) 1965
"Push-Pull Saturated Core Tunnel Diode Inverters," R. Feryszka and P. Gardner, *RCA Review* (September) 1965
"A Solid-to-Solid Diffusion Technique," J. Scott and J. Olmstead, *RCA Review* (September) 1965
"Special Tube Types," L. W. Aurick, *S-9 Magazine* (September)... 1965
"Spurious Signals in Satellite Command Systems," J. C. Blair, *Trans. IEEE PTGEMC* (September) 1965
"Stabilized Operation of Xenon Flashlamps Using Ultraviolet Illumination," R. J. Pressley and H. Y. S. Tang, *Rev. Sci. Instr.* (Notes) (September) 1965
"Thermally Stimulated Conductivity of Neutron-Irradiated Silicon," J. J. Wysocki, *Jour. Appl. Phys.* (Communications) (September) 1965
"Traveling-Wave Characteristics of Cerenkov Interaction," J. J. Stekert, *Proc. IEEE* (Correspondence) (September) 1965
"Coherent Light Radiation by a Lorentz Field in GaAs," K. K. N. Chang and H. J. Prager, *Phys. Rev. Letters* (6 September).. 1965
"Stark Effect of 4 States and Linear Crystal Field in $BaClF:Sm^{2+}$," Z. J. Kiss and H. A. Weakliem, *Phys. Rev. Letters* (6 September) 1965
"Insulating Ferromagnetic Spinels," P. K. Baltzer, H. W. Lehmann, and M. Robbins, *Phys. Rev. Letters* (13 September) 1965
"Attitude Control for the TIROS Weather Satellite," W. P. Manger, Int. Fed. of Automatic Control, *Proc. Symposium on Automatic Control in the Peaceful Uses of Outer Space*, Stavanger, Norway (June 22) 1965
"Design and Performance of a 20-KW Latching Nonreciprocal X-Band Ferrite Phase Shifter," W. A. Schilling, W. W. Siekanowicz, E. Walsh, I. Bardash, and I. Gordon, *Nat. Electronics Conf. Record*, Chicago, Ill. (Oct. 23) 1965

"Designing Equipment for AutomaticTestability, F. Liguori *Conf. Record, Automatic Support Systems Symposium*, St. Louis, Mo. (June 7-9) 1965

"The Development of Camera Tubes for Recording Astronometric Images," A. Cope and E. Luedicke, *Proc. Imperial College 3rd Int. Symposium on Photoelectronic Imaging Devices*, London, England (Sept. 20) 1965

"Dielectric-Tape Camera System for Meteorological Applications," L. Freedman, *Proc. XVI Int. Astronautical Cong.*, Athens, Greece (Sept. 12) 1965

"Digital Computer Control Systems from the Control-Oriented Point of View," A. S. Robinson, *Conf. Record, Int. Fed. of Inf. Proc. Socs.*, New York City (May 24-29) 1965

"Electron Cyclotron Plasma Thrustor for Orbit Corrective Missions," H. Hendel, *Conf. Proc., Advisory Gr. for Aeronautical Research and Development*, NATO, Pisa, Italy (Sept. 9) 1965

"Emulation on RCA Spectra 70 Systems," W. R. Lonergan, *IEEE WESCON*, Part 4, *Computers & Data Processing* 1965

"Extending Integrated Circuits," M. Kidd and L. C. Drew, *Conf. Record, East Coast Conf. on Aerospace and Navigation Electronics*, Baltimore, Md. (Oct. 27) 1965

" 'Hall' Drift Velocity at High Electric Fields in Indium Antimonide," M. Glicksman and W. A. Hicinbothem, Jr., *Proc. 7th Int. Conf. of Phys. and Semicon.*, Paris (July 12-24, 1964) 1965

"High Speed Optical Character Readers," A. J. Torre, *IEEE WESCON*, Part 4, *Computers & Data Processing* 1965

"High-Speed Solid-State Automatic Signal Attenuator and Its Application," W. Farrell, *Proc. 11th Ann. IES Tech. Mtg.*, Sherman House, Chicago, Ill. (Apr. 21-23) 1965

"Instabilities and Growing Waves in Solid State Plasmas," M. Glicksman, *Proc. 7th Int. Conf. of Phys. and Semicon.*, Paris (July 12-24, 1964) 1965

"Integrated Circuits for Use in the RCA Spectra 70 Series Computer," R. D. Lohman and S. E. Basara, *IEEE WESCON*, Part 4, *Computers & Data Processing* 1965

"Ladder Network Synthesis by Similarity Transformation," T. G. Marshall, Jr., *Proc. of Allerton Conf. on Circuit and System Theory* (Oct. 20) 1965

"Lightweight Converter Using Step-Width Reglator," J. J. Klein, R. C. Kee, and E. F. Dion, *Conf. Record, East Coast Conf. on Aerospace and Navigation Electronics*, Baltimore, Md. (Oct. 27) 1965

"Memories in Present and Future Generations of Computers," J. A. Rajchman, *IEEE WESCON*, Part 4, *Computers & Data Processing* 1965

"Message Channel Monitoring," C. J. Moore, *Proc. Nat. Comm. Symposium*, Utica, N. Y. (Oct. 11) 1965

"Micoelectronic Airborne Data Link," E. J. Mozzi, *Proc. IEEE Conf. on Military Electronics*, Washington, D. C. (Sept. 22) .. 1965

"A Microelectronic Digital Frequency Synthesizer for a UHF Communications Transceiver," E. D. Menkes and J. W. Harmon, *Proc. Nat. Comm. Symposium*, Utica, N. Y. (Oct. 11) 1965

"Microwave Power by Parametric Frequency Multiplication in Transistors," M. Caulton, H. Sobol, and R. L. Ernst, *Nat. Electronic Conf. Record*, Chicago (Oct. 25-27) 1965

"MOS Micropower Complementary Transistor Logic," M. M. Mitchell and R. W. Ahrons, *IEEE WESCON*, Part 2, *Integrated Circuits* 1965

"Multiple Access Considerations for Communications Satellites," F. Assadourian and D. Jacoby, *Proc. 1965 IEEE Conf. on Military Electronics*, Washington, D. C. (Sept. 22) 1965

"Performance Analysis of 3- and 4-Coil FM Tuners Using N-P-N Silicon Transistors," R. V. Fournier, C. H. Lee, and R. T. Peterson, *Nat. Electronics Conf. Record*, Chicago, Ill. (Oct. 25-27) 1965

"The Performance of the TIROS IX Wheel Satellite," A. Schnapf, *Proc. XVI Int. Astronautical Cong.*, Athens, Greece (Sept. 13) 1965

"Programmable Integrating Television System for Use with the Stratoscope," L. E. Flory, W. Pike, J. Morgan, and L. Boyer, *Proc. Imperial College 3rd Int. Symposium on Photoelectronic Imaging Devices*, London, England (Sept. 23) 1965

"Progress in Solid-State Microwave Power Sources," F. Sterzer, *IEEE Symposium on Microwave Theory and Techniques*, Clearwater, Fla. (May 5-7) 1965

"Radar Cross-Section at Long Waves," H. Benecke, *AMRAC Proc., XII, Part 1* 1965

"Random-Access Mass Memory System," J. F. Gates, *Conf. Record IFIP Cong.*, New York (May 24-29) 1965

"RCA Spectra 70, Basic Design and Philosophy of Operation," A. D. Beard, *IEEE WESCON*, Part 4, *Computers & Data Processing* 1965

"Reporting Financial Data to Management in a Research Environment," W. R. Breeze, *Amer. Management Assoc.* 1965

"Saturn V Diphase Data Modem Development," L. E. Langley and J. K. Oliver, *Proc. 1965 IEEE Conf. on Military Electronics*, Washington, D. C. (Sept. 22) 1965

"Scratchpad-Oriented Design of the RCA Spectra 70," A. T. Ling, *Conf. Record, Fall Joint Computer Conf.*, Las Vegas, Nev. (Nov. 30-Dec. 2) 1965

"Shock Capabilities of Electro-Dynamic Shakers," J. McClanahan and J. Fagan, *Proc. 11th Ann. IES Tech. Mtg., Sherman* House, Chicago, Ill. (Apr. 21-23) 1965

"Short Cuts to EMI Filter Design," A. W. DiMarzio, *IEEE Nat. Symposium on Electromagnetic Compatibility*, New York City (June 28-30) 1965

"Simple Threshold Learning," J. Sklansky, *Conf. Record, Int. Fed. of Inf. Proc. Socs.*, New York City (May 24-29) 1965

"Simplifying the Use of Automatic Test Equipment with a Compiler," R. C. Miller, *Proc. IEEE Automatic Support Systems Symposium*, St. Louis, Mo. (June 7-9) 1965

"System Adaptive Techniques," M. Masonson, *Proc. 1965 IEEE Conf. on Military Electronics*, Washington, D. C. (Sept. 22) 1965

"Temperature Coefficient of Vacuum-Deposited Thin-Film Resistors as Patterned with Abrasive-Jet and Electron-Beam Milling," E. Casterline and H. Schwarz, *IEEE Microelectronics Symposium, Conf. Record*, St. Louis, Mo. (May 24-26) 1965

"Three-Axis Attitude Control System Utilizing a Single Reaction Wheel and Magnetic Torquing," H. Perkel, *Proc. XVI Int. Astronautical Cong.*, Athens, Greece (Sept. 14) 1965

"Transistor and Nuvistors in a Two-Meter Transceiver," R. M. Mendelson, *RCA Ham Tips* (Spring & Summer) 1965

"Transportation Environments Revisited," A. S. Baron and A. Schilling, *Proc. 11th Ann. IES Tech. Mtg.*, Sherman House, Chicago, Ill. (Apr. 21-23) 1965

"Unilateralization and Widebanding of Parametric-Admittance Frequency Converters," J. Klapper, *Nat. Electronic Conf., Conf. Record*, Chicago (Oct. 25-27) 1965

"Wideband Direct Recording with High Frequency Bias," W. B. Hendershot, *11th Conf. on Magnetism & Magnetic Materials, Conf. Record*, San Francisco (Nov. 16) 1965

AUTHORS

R. W. AHRONS received the B.S. and M.S. degrees from the Massachusetts Institute of Technology in 1954. He received the Ph.D. degree from Polytechnic Institute of Brooklyn in 1963, majoring in Computers and Communications. Following graduation from MIT, he joined the research staff of RCA Laboratories in Princeton, N. J., where he worked in the fields of solid-state applications to television receivers and video tape systems. In 1961, he transferred to the computer research laboratory where he was engaged in the research application of superconductivity to memory and logic. He assumed his present position as an Engineering Leader in Electronic Components and Devices, Somerville, N. J., in 1964. In this capacity, he is responsible for the circuit application of new solid-state devices. Dr. Ahrons is a nonresident instructor for the Massachusetts Institute of Technology, where he teaches mathematics and solid-state circuitry to MIT's cooperative course students.

Dr. Ahrons is a senior member of the Institute of Electrical and Electronics Engineers, Eta Kappa Nu, Tau Beta Pi and Sigma Xi.

I. BARDASH graduated from Cornell University in 1958 with a B.E.E. degree and from the University of Pennsylvania in 1963 with an M.S.E.E. degree. He is presently enrolled in the Graduate School of Arts and Sciences at the University of Pennsylvania where he is working towards a Ph.D. Following his graduation from Cornell he was employed by American Bosch Arma Corp., Garden City, New York, where he was a member of the applied research group. In November of 1959, he joined RCA as a microwave engineer with the Missile and Surface Radar Division. Since that time he has been concerned with the research and development of ferrite microwave devices. For the past three years he has been studying problems associated with the design and fabrication of latching digital ferrite phase shifters for use in electronically steerable phased arrays.

Mr. Bardash is a member of the Institute of Electrical and Electronic Engineers, and is a registered professional engineer in the state of New Jersey.

IRWIN GORDON received the B.S. degree in Ceramics from Rutgers University in 1948, the M.S. degree in 1951, and the Ph.D. degree in 1952 from the same institution. From 1948 to 1952 he held a research assistantship in the School of Ceramics at Rutgers University working on single crystal synthesis and studies of various silicates and rare earth aluminates. In 1952 Dr. Gordon joined the technical staff of RCA Laboratories where he has worked on the development of magnetic materials for various applications. This has included materials for permanent magnets, multi-function composite materials, recording-tape pigments, small size high-frequency antennas, and non-spinel type magnetic materials for use in the VHF and UHF regions. Most recently he has investigated ferrites for use in laminated computer memories, as well as ferrites for microwave use.

He is a member of the American Ceramic Society, the N. J. Ceramic Association, Sigma Xi, Keramos, and The American Society for Testing Materials.

PETER MARK received the Ph.D. degree in Solid State Physics from New York University in 1958. In that year he joined the Research Division of the Polaroid Corporation and was active in the development of several electro-optical devices. While at Polaroid, Dr. Mark began a research program on the transport properties of organic crystals. From 1961 to 1962, Dr. Mark was a visiting lecturer at the Munich Institute of Technology. In Munich, he continued the organic crystals research program which lead to the development of the technique of transient injection currents for the investigation of charge transport in insulators. Since joining the RCA Laboratories at the David Sarnoff Research Center, Princeton, N. J., he has been engaged in work on the photoelectronic properties of insulators with emphasis on injection currents, free carrier diffusion and surface phenomena. Dr. Mark is a member of the American Physical Society and Sigma Xi.

WILLIAM A. SCHILLING received the B.E.E. degree from Manhattan College in New York City in 1956 and the M.E.E. degree in microwave engineering from New York University in 1960. In 1956, he joined the Western Union Telegraph Company, where he worked on S and C band microwave relay systems. During 1962, he joined Waveline, Inc., where he designed all types of waveguide and coaxial components. In 1964, he joined the microwave applied research group, RCA Electronic Components and Devices at Princeton, N. J., to do advanced development work on ferrite devices. He has been engaged in the development of high-power ferrite phase shifters and new, improved ferrite materials for the microwave region.

Mr. Schilling is a member of the Institute of Electrical and Electronics Engineers and Eta Kappa Nu.

W. W. SIEKANOWICZ received the B.S. degree from the Imperial College of Science and Technology, London University, England; the M.S. degree from Columbia University, New York, N. Y., and the D.E.E. degree from the Polytechnic Institute of Brooklyn, N. Y., in 1948, 1950 and 1960, respectively. He joined the microwave tube engineering group of RCA Electronic Components and Devices, Harrison, N. J., in July 1950, and transferred to the microwave applied research group, Princeton, N. J. in 1956. He has specialized in the fields of microwave tubes and ferrite devices. Dr. Siekanowicz is a member of Sigma Xi.

J. J. STEKERT was awarded the B.S. degree in physics by Union College in 1957. He received the M.S. in physics in 1959 and the Ph.D. in 1965 at the University of Pennsylvania. In previous years, he participated in research projects in the fields of biophysics and infrared detection, and during military service in the years 1954 and 1955, he was engaged primarily in radio communications. Since joining RCA in 1961 he has worked in the plasma physics group of the applied research department. In addition to a continuing interest in Cerenkov microwave generation, he has done varied work in electromagnetic interactions with plasmas and plasma dynamics.

Dr. Stekert is a member of the American Institute of Physics, the Institute of Electrical and Electronics Engineers, the American Association of Physics Teachers, and the Franklin Institute.

THOMAS E. WALSH received the B.S. degree in Electrical Engineering from Fairleigh Dickinson University in 1958 and the Ph.D. degreee in Physics in 1963 from the Johns Hopkins University where he held a teaching Assistantship. From 1961 to 1963 he was a research assistant at the Johns Hopkins Laboratory of Astrophysics and Physical Meteorology where he was engaged in research on the atmospheres of the Earth and the Jovian planets. In 1963, he joined the microwave applied research group of Electronic Components and Devices, Princeton, N. J., where he has been engaged in the development of tunnel-diode amplifiers, microwave light modulators and detectors, solid-state power sources, and ferrite devices. Dr. Walsh is a member of Sigma Xi.

J. ROGERS WOOLSTON received the B.S.E. degree from Princeton University in 1955, in a combined program of electrical engineering and physics. He spent the summers of 1952 and 1953 at RCA Laboratories, and the summer of 1954 at Bell Laboratories, working primarily on semiconductor devices. Since 1955, he has been with RCA Laboratories, working initially on materials research in both thermoelectrics and intermetallic compound semiconductors. Since 1958, he has specialized on the analysis of solids by mass spectrometry, including the design, development, and improvement of equipment and techniques for thermal vaporization analysis and for r-f spark source analysis. Mr. Woolston has written several programs for the RCA 601 computer as aids to both analysis and research with the MS7 mass spectrograph. He is a member of the Institute of Electrical and Electronics Engineers and the American Society for Testing Materials.

L. W. ZELBY graduatd from the Moore School of Electrical Engineering, University of Pennsylvania, with a BSEE in 1956 and obtained an MS in EE in 1957 at the California Institute of Technology. Subsequently, he returned to the University of Pennsylvania where he received a Ph.D. in Electrical Engineering in February 1961. Dr. Zelby joined RCA in 1960 and participated in the early studies of Cerenkov interaction in the microwave region. In 1961 he returned to the University of Pennsylvania and joined the faculty of the Moore School where he is now an Associate Professor. He continued his work at RCA in the capacity of consultant, primarily on the Cerenkov study in which he has actively participated to the present time. His major fields of endeavor have been electromagnetic propagation, electrodynamics of moving media, surface waves, and plasma physics.

Dr. Zelby is a registered professional engineer in the state of Pennsylvania, and a member of the American Physical Society, the Institute of Electrical and Electronics Engineers, and the Franklin Institute.