RCA REVIEW

a technical journal

Published quarterly by

RCA LABORATORIES

in cooperation with all subsidiaries and divisions of

RADIO CORPORATION OF AMERICA

VOLUME XXVIII | MARCH 1967 | NUMBER 1

CONTENTS

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

RCA REVIEW

BOARD OF EDITORS

DETERMINANTS OF ELECTRONICALLY STEERABLE ANTENNA ARRAYS

By

W. T. Patton

RCA Missile and Surface Radar Division
Moorestown, N. J.

Summary—This paper examines factors that determine the design of an electronically steerable antenna array. The considerations that direct the use of an array rather than a reflector antenna are enumerated. Arrays are classified according to the type of steering system used and critical factors are given for each type. The behavior of array components such as antenna elements, phase shifters, time delayers, and beam formers are described and the determining relationships between array design factors and array system performance are developed.

Introduction

THE use of arrays of antennas is as old as radio science. Arrays of two or more elements for coverage control are common in broadcast practice, and curtain arrays and the multi-unit steerable array (MUSA)[1] are examples from communications practice. The first radar antennas were arrays, but the parabolic reflector antenna attained a dominant role in the development of radar systems because of its simplicity and ease of construction. In the past ten years, interest in the array antenna has renewed as multiple-target-handling requirements increased and techniques for rapid phase shifting were developed. This attention is manifest in the rate at which the volume of literature on array radar and array components has grown since 1958. In this paper we examine some of the factors that favor an array antenna over a reflector antenna and that differentiate among the various types of array design.

The Choice Between Arrays and Reflector Antennas

In general, a reflector system is preferred if it can be designed to satisfy the requirements of a given antenna problem. The principal factors influencing the antenna efficiency of a reflector system are

[1] H. T. Friis and C. B. Feldman, " Multiple Unit Steerable Antenna," *Proc. IRE*, Vol. 25, p. 841, July 1937.

spillover loss (energy radiated by the primary feed not intersecting the main or subsidiary reflectors), aperture blockage (shadowing of the main reflector by the feed, subreflectors, and support structures), illumination taper (distribution of primary energy over the aperture defined by the main reflector), and reflector errors. The dissipation loss in the feed system is generally a minor contributor to loss of efficiency in a reflector antenna system, but it is the principal contributor to loss of efficiency in the array system. The use of reflector antennas with small apertures (less than five wavelengths in diameter) tends to be restricted by large aperture blockage and by the decreasing ability of the reflector to collimate the feed energy. The usual geometric-optics approximations used in reflector design are not valid for small apertures. For very large apertures, the reflector antenna is limited by the ability to hold close mechanical tolerances in very large moving structures. Thus array antennas are often used where either relatively small apertures or very large apertures are required. This discussion, however, is devoted largely to the choice between arrays and reflector antennas in the size range where both are feasible.

Some of the antenna system requirements that determine the choice of an electronically scanned array over a moving reflector antenna are beam agility, multiple-target-handling capability, very high radiated power, and illumination control.

Beam Agility

The large mechanical inertia of a moving reflector antenna system restricts it to applications utilizing a continuous beam motion, such as a search mode, where the motion is essentially uniform throughout the search zone, or a track mode, where the system follows the single target along its trajectory. An electronically scanned array system, on the other hand, can place each radar pulse or a group of pulses at any angle within the coverage zone of the system, independent of the direction of any previous pulse. This capability allows a single array to perform both a search function and a track function on multiple simultaneous targets that might require several separate reflector systems.

Multiple-Target-Handling Capability

A moving-reflector tracking system can track several objects that are within its beamwidth. However, because of mechanical inertia, it will usually track one such object as a primary target while secondary targets may enter and leave the antenna beam. The antenna beamwidth required for a given problem is usually dictated by angular

resolution and gain requirements. When the angular distribution of targets in a multiple-target problem is only a little greater than the antenna beamwidth, a multiple-beam reflector system can be used. But when the angular distribution is relatively large, an electronically scanned array must be used.

Very High Radiated Power

The power that can be radiated by a reflector-antenna system is limited by the amount of power that can be generated in a single source or handled in a single transmission line. An array has no such limitation; it may use as many sources as it has radiation elements. The power radiated from an array can be proportional to its area, whereas the power radiated from a reflector system is limited by its single feed.

Illumination Control

The illumination taper used with reflector antenna systems is usually a compromise between spillover efficiency and aperture efficiency. Uniform illumination, which would yield the highest aperture efficiency, cannot be obtained because of spherical divergence from the primary feed (space taper) and because of the beam shape of the primary feed. In a typical design, the reflector intercepts a large part of the primary energy (high spillover efficiency) but the primary energy at the edge of the reflector is low compared with that at its center (low aperture efficiency). In array design the illumination taper is completely under the control of the designer and is usually determined by consideration of beam shape, aperture efficiency, and side-lobe level requirements.

Types of Scanning Systems

One of the factors that makes electronically scanned arrays so versatile and at the same time so fertile a field for study and discussion is the variety of systems and types of components that can be applied. They can be classified by mission, function, power level, types of key components, location of active components, beam-forming system, or scanning system. Many types of scanning systems for arrays have been proposed but they are essentially combinations of the three principal techniques discussed in this section. These are (1) time-delay scanning, (2) phase scanning, and (3) frequency scanning.

In discussing the characteristics of these scanning techniques and the basis for choice among them, it is convenient to recall that the pattern of a rectangular planar array can often be separated into the product of patterns of two linear arrays. The pattern of a linear

array of N regularly spaced isotropic elements can always be represented[2] as an N^{th} order polynomial in the variable

$$x = e^{i\psi},$$

where

$$\psi = \frac{2\pi d}{\lambda} \sin\theta - \delta$$

d = interelement spacing,

λ = free-space wavelength at frequency of operation,

θ = observation angle measured from the array normal,

δ = steering phase delay inserted in the network feeding the element.

The pattern of a uniformly illuminated line array is fundamental to array theory, since all other patterns can be expressed as linear superpositions of such patterns with phase tapers differing by a multiple of 2π.

A few remarks about the uniform array pattern at this point will aid in later discussions. The normalized pattern of a uniform array of N elements can be written in closed form as

$$P(\psi) = \frac{\sin\dfrac{N\psi}{2}}{N\sin\dfrac{\psi}{2}}.$$

First we note that this function is periodic and has a maximum value whenever $\psi = 2\pi n$, $n = 0, \pm 1, \cdots$. The number n is the spectral order of the array pattern, with zero associated with the main beam of the array and all larger orders (positive or negative) termed grating lobes. There are also $N-2$ minor lobes (side lobes between spectral orders, with the largest of these taking on the value $(2/3)\pi (\approx 13.5$ db) for large N. If we plot this function in terms of a variable $u = \sin\theta$ with $\delta = 0$, we find that it has zeros whenever $u = n\lambda/(Nd)$ except where n/N is an integer (see Figure 1). These exceptions are

[2] S. A. Schelkunoff, "A Mathematical Theory of Linear Arrays," *Bell System Tech. Jour.*, Vol. 22, No. 1, p. 80, 1943.

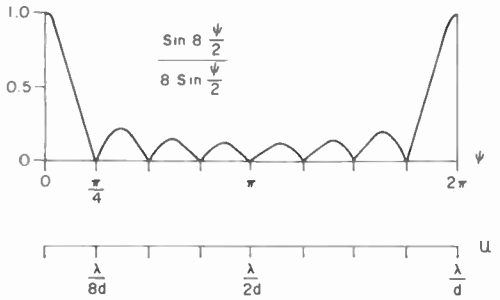

Fig. 1—Normalized pattern of a uniform array of 8 elements.

the grating lobes mentioned above. With this brief discussion as background, we now describe the various steering schemes.

Time-Delay Scanning

The direction of the main beam is that for which $\psi = 0$. The direction is changed by varying δ so that

$$\delta = (2\pi d/\lambda) \sin \theta_0 = \frac{2\pi f d}{c} \sin \theta_0$$

where θ_0 is the angle to which the array is steered. Varying δ has the effect of translating the pattern in the u coordinate by an amount $u' = \sin \theta_0$. If the phase increment δ is made to vary with frequency in the manner indicated, the beam maximum will stay in the desired direction over a broad range of frequencies. This type of steering is accomplished by switching in sections of delay line of length $l = d \sin \theta$ or time-delay increments of $t = (d \sin \theta)/c$, as shown in Figure 2.

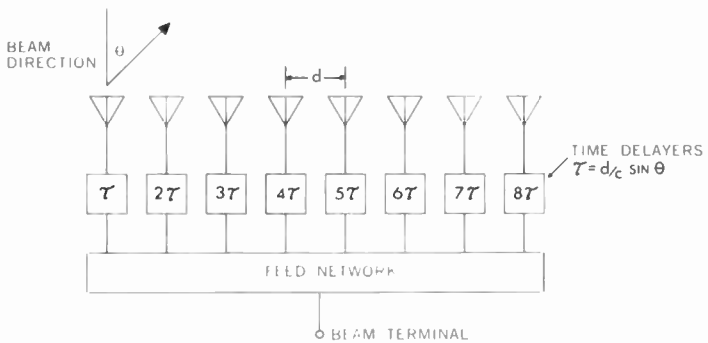

Fig. 2—Time-delay scanning.

This time delay is the difference between the time of arrival of a wave front at one element and its arrival at the next element. With time-delay steering, the energy contributions from all elements arrive at the sum point in the combiner network at the same time.

Time-delay scanning offers the greatest instantaneous signal bandwidth of the various scanning techniques to be discussed. In essence, this steering technique in itself does not limit the signal bandwidth. Any attempt to discuss the limitations of this technique here and to justify consideration of other more-narrow-band techniques would only serve to "date" this paper, since the discussion would depend largely on the present state of technology. The major considerations, however, are the r-f dissipation loss in the delay lines and switches, the power-handling limitations imposed by the delay lines and switches, and the cost of these components.

Phase Scanning

For phase scanning, the phase increment δ is computed (modulo 2π) for the center frequency according to

$$\delta = \frac{2\pi d}{\lambda_0} \sin \theta_0 - 2\pi n$$

such that

$$0 \leq \delta < 2\pi.$$

The steering is accomplished by "phasers" (phase shifters) in the network that feed each radiating element; these phasers are adjustable over the range from 0 to 2π radians, independent of frequency (see Figure 3). This scanning method is indistinguishable from time-delay scanning for single-frequency continuous-wave applications. For short-pulse and frequency-modulated applications, however, we find that the beam direction varies with frequency. The variable ψ in the direction θ_0 takes on the values

$$\psi = \frac{2\pi d}{c} (f - f_0) \sin \theta_0$$

at the frequency f. The response of the antenna system in the direction θ_0 is down by $2/\pi$ (about 4 db) when

$$f - f_0 = \pm \frac{c}{2Nd \sin \theta}.$$

The bandwidth between 4-db points then is

$$f_2 - f_1 = 1/\tau$$

where

$$\tau = \frac{Nd \sin \theta}{c}$$

is the time between the arrival of the wave front at the first and last

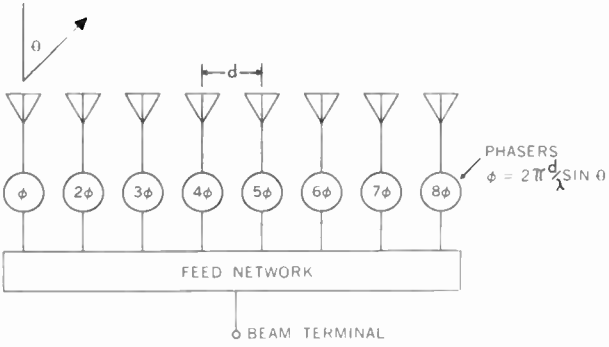

Fig. 3—Phase scanning.

elements of the array. This is sometimes called the array filling time.

The bandwidth of a phase-scanned array decreases with increasing size and with the angle of scan from the array normal. As an example, for a 10% bandwidth when an array is scanned to 45 degrees, the array must be less than about 14 wavelengths on a side.

Frequency Scanning

In a frequency-scanned array, the phase increment δ is obtained by a length of delay line, l, between elements, as shown in Figure 4. The length of this delay line is adjusted so that its electrical length is an integral number, r, of 2π radians at the center frequency, f_0, and so that the array is scanned to angle θ_m at the frequency f_m, or

$$\frac{2\pi l}{\lambda_0} = 2\pi r$$

$$\frac{l}{\lambda_m} - \frac{l}{\lambda_0} = \frac{d}{\lambda_m} \sin \theta_m.$$

From these relations we derive the ratio of delay-line length to element spacing in terms of the maximum scan angle and frequency as

$$\frac{l}{d} = \frac{f_m \sin \theta_m}{f_m - f_0}.$$

This ratio is sometimes called the "wrap-up ratio" and is roughly given by the ratio of the sine of the maximum scan angle to the per cent deviation in frequency required to scan.

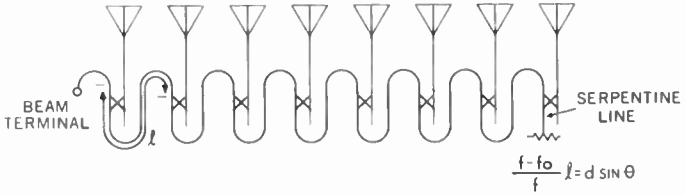

Fig. 4—Frequency scanning.

A frequency-scanned array establishes a correspondence between a direction in space and a frequency. In this case, as with phase scanning, we define the bandwidth as the interval between those frequencies that scan the $2/\pi$ points on the main beam in a given direction. Again

$$f_2 - f_1 = 1/\tau.$$

Here

$$\tau = \frac{N(l - d \sin \theta)}{c}.$$

As an example, consider a frequency-scanned array covering plus to minus 45 degrees in a 10% total frequency band. The wrap-up ratio is

$$\frac{l}{d} = 14$$

and the array filling time is

$$\tau = \frac{Nd}{c}(14 - \sin \theta).$$

In this example the bandwidth of a frequency-scanned array varies only slightly with the steering direction, but is about one-twentieth of the bandwidth of a phase-scanned array of the same size and maximum scan angle.

A frequency-scanned array tends to be a light-weight, low-cost system and is preferred for applications where its narrow bandwidth can be tolerated. A phase-scanned array has much greater instantaneous bandwidth (one setting of phase shifters) than the frequency-scanned system, plus extended tunable bandwidth (phase shifter setting adjusted for different center frequencies); however, it must carry the added cost and weight of the phasers and power-divider networks. The largest bandwidth is provided by time-delay scanning, with the largest cost and weight contributed by the time-delay units.

The optimum steering system for large bandwidths may be a hybrid array of phased subarrays with relatively few time-delay units controlling the subarrays. With this technique, the size of a subarray can be made small enough to give adequate antenna bandwidth independent of the size of the array. This technique also allows a choice of the number of elements to be driven by a single power amplifier, so that the optimum amplifier size can be used.

Selection of Phasers

An important decision in the design of transmitter arrays is whether to use one amplifier for each radiating element or one amplifier feeding many elements. This decision is usually made on the basis of the power to be radiated by each element, the availability of phase shifters at that power level, and the relative costs of providing many small amplifiers or a few large amplifiers. The range of choice between suitable phase-shifting techniques is largest when one amplifier per element is used. However, this approach is usually more costly than one in which high-level r-f phasers are used and one amplifier drives many elements.

The phasers used in electronic scanning arrays may be divided into two classes, continuous and digital. The digital device is composed of sections, each of which has two discrete phase-shift states determined by its physical parameters and selectable by a control signal. The differential phase shift (difference between phases at the two states) of these devices is relatively independent of small variations in the control signal. By contrast, the various phase settings of a continuous phaser are obtained by varying the control signal, since the insertion phase is directly dependent on the magnitude of the control signal. Phase accuracy in the digital phaser is dependent primarily on its

construction; phase accuracy in the continuous phaser depends both on construction and on the accuracy of the control signal. Digital phasers are most appropriate for electronic scanning arrays, since they are usually controlled by a special-purpose digital computer (sometimes called a beam-steering controller). Continuous phasers can, of course, be made to act as digital phasers by quantizing the control signal.

High-Power Phasers

High-power r-f phasers can be further classified as diode phasers or ferrite phasers depending upon the active element in the device. Diode phasers generally fall within one of two categories employing

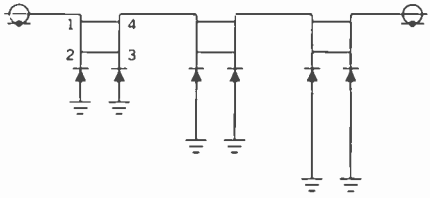

Fig. 5—Reflection diode phaser.

either reflection in a 90-degree hybrid circuit or propagation through a periodically loaded transmission circuit. Ferrite phasers can be either reciprocal (phase shift independent of direction through the device) or nonreciprocal (phase shift different for the two directions of transmission) and can be of either the reflection or transmission types. Diode phasers are generally reciprocal in character.

The operation of reflection-diode phasers (Figure 5) depends on the fact that when equal impedances are placed at ports 2 and 3 of a 90-degree 3 db hybrid coupler, the reflected energy appears at port 4 with a phase dependent on the phase of the reflection coefficient at ports 2 and 3. In practice, the reflections at ports 2 and 3 are obtained by a length of shorted transmission line, and a diode switch is used to move the apparent short position on this line by a fixed amount. Thus, each hybrid junction and diode-pair circuit acts as one bit of a digital phaser. Alternatively, a pair of varactor diodes can be employed to vary the phase of the reflected wave as a function of applied bias voltage to obtain a continuous phaser.

In diode phasers of the transmission type (Figure 6), the diode switch is part of a lumped periodic loading circuit in the transmission line. In digital operation, the diode is switched from its forward-bias

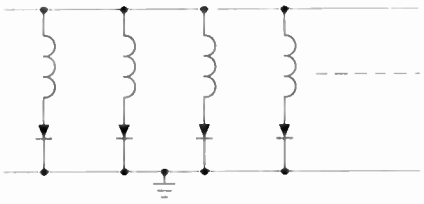

Fig. 6—Transmission diode phaser.

to its reverse bias condition to vary the reactive loading and, therefore, the phase delay through the section. Here again varactor diodes can be employed for continuous phasing. The transmission-diode phaser is better suited to high-power applications than the reflection type since the load can be more easily distributed between many diodes in a single device.

Ferrite phasers can be designed so that the variable permeability of the material is effectively a scalar quantity, or they can be designed to utilize the tensor nature of ferrite permeability. In the first type, shown in Figure 7, the ferrite is generally a rod located in the center of a waveguide and magnetized longitudinally. The phase shift is dependent on the magnitude of the applied field. Phasers of this type are generally continuously variable and reciprocal. A principal disadvantage of this type of ferrite phaser is the high inductance of the magnetizing coil and the requirement for maintaining rather large currents during operation of the device.

This disadvantage of the variable permeability phaser is avoided in the closed-loop twin-slab phaser (Figure 8), which operates on the remanent magnetization of a ferrite material having a square hysteresis loop. It is essential for this type of operation that a low-reluctance closed magnetic path be available for the remanent flux.

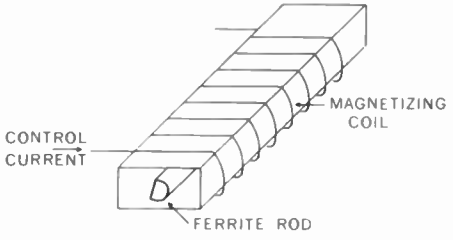

Fig. 7—Variable-permeability ferrite phaser.

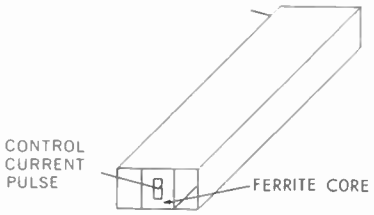

Fig. 8—Nonreciprocal latching ferrite phaser.

Such devices utilize the tensor character of ferrite permeability and are nonreciprocal. The differential phase shift is obtained by switching the remanent magnetization from one direction to the other. These phasers are called "latching" because they remain in one state with no further expenditure of energy until a switching pulse reverses the magnetization to the second state where it again remains until subsequent switching pulses are applied. An interesting feature of this phaser is that the round-trip phase shift is the same for either state. The disadvantage of a nonreciprocal phaser is evident in its use in a duplexed array, where it must be reversed after each transmission to receive the return signal. If the switching pulse does not drive the ferrite to saturation, the differential phase shift depends on the strength of the switching pulse. In this way it can be operated as a continuous variable phaser.

Nonreciprocal ferrite phasers are also built using a switching three-port ferrite circulator as the active element. This device, shown in Figure 9, is a reflection-type phaser with the differential phase adjusted by the phase of the reflection at the second port of the circulator. The transmitted energy then either takes the 1-3 reference path through the circulator or the 1-2-3 delayed path, depending on the sense of circulation. Such devices are built in a latching configuration and operate on the remanent magnetization in a closed magnetic loop.

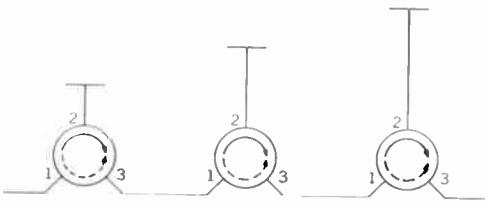

Fig. 9—Switching circulator phaser.

Low-Level Phasers

When phasing is to be accomplished at low levels (one amplifier per element) all the r-f phasing techniques available at high power can be used, and, in addition, heterodyne phasing techniques can be employed. It may also be desirable to do the phasing and beam forming at an intermediate frequency.

The heterodyne phasing technique operates by mixing a control frequency f_c with the operating frequency f_0 and sending the difference frequency $(f_0 - f_c)$ to a second mixer, where it is combined with the control frequency delayed by an amount τ (see Figure 10). The phase of the sum frequency f_0 at the second mixer is $2\pi f_c \tau$, proportional to

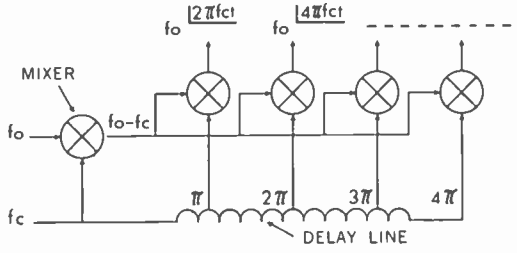

Fig. 10—Heterodyne phaser.

the control frequency and the time delay. In an array, the control frequency is delayed by an amount $n\tau$. The output phase at the operating frequency is then continuously variable in proportion to the control frequency.

Phasing and beam forming at an intermediate frequency open up many additional possibilities, such as conventional lumped-constant phasers, phase-locked variable-frequency oscillators, and others. This approach also opens up the field of so called adaptive or self-phasing arrays. A discussion of this area, however, is beyond the scope of this paper.

The Required Number of Phaser Bits

The important consideration in the design of an electronic scanning array is the number of bits of control and the number of bits of phase required for each phase shifter. The requirement for the number of control bits can easily be derived from the fact that a phase taper of 2π radians across the array will displace the beam of a uniformly illuminated aperture by $2/\pi$ beam width. Thus if the minimum desired

displacement is q fractional beam widths, then the number of control bits, p, must satisfy $2^p > N/q$, where N is the number of elements along a diameter of the array. For example, eight control bits are required to steer a 50-element line array in increments of one-fifth beam width. The smallest control bit is smaller than the practical error in the setting of a phaser. Therefore, the number of bits, b, provided in the phaser will usually be smaller than the number of control bits required. The difference between the number of control bits and the number of phaser bits is the number of round-off bits, r, which indicates the maximum number of elements, 2^r, that will receive the same phaser setting.

Table I

Number of bits	Maximum loss (db)
2	0.915
3	0.212
4	0.056
5	0.014

The use of a number of phaser bits smaller than the number of control bits in effect breaks the array into at least $N/2^r$ subarrays, each containing at most 2^r elements on one side, with all elements at the same phase. The imperfect phasing resulting from a finite size of phase quantization can produce both a loss in gain and a pointing error in beam direction. Both of these conditions are maximum at the first beam position off the broadside direction.

The loss in directivity due to a phase-error distributed uniformly over the interval $-\delta$, δ (where δ is half the smallest phaser bit) is given by $(\sin\delta/\delta)^2$ where $\delta = \pi/2^b$. The loss function can be approximated by the first two terms of its series expansion

$$\frac{\sin\delta}{\delta} \approx 1 - \frac{\delta^2}{6}$$

and the series expansion of the log function gives for the loss

$$L = 20 \log \frac{\sin\delta}{\delta} = 1.446\delta^2.$$

The results of this calculation are given in Table I. From this table it can be seen that four bits of phase are usually adequate from a gain consideration.

At the first beam position off broadside, the array pattern is formed as a product of a subarray pattern pointed broadside and the pattern of an array of isotropic sources, whose spacing is the length of a subarray, pointed in the desired direction. This pattern is written as

$$P(\sin\theta) = \frac{\sin\left[\left(2^r \frac{\pi d}{\lambda}\right)\sin\theta\right] \sin\left\{\frac{N}{2^r}\left[2^r \frac{\pi d}{\lambda}\sin\theta - \frac{\pi}{2^b}\right]\right\}}{2^r \sin\left[\frac{\pi d}{\lambda}\sin\theta\right] \quad \frac{N}{2^r}\sin\left[2^r \frac{\pi d}{\lambda}\sin\theta - \frac{\pi}{2^b}\right]}$$

The function is maximum at

$$\sin\theta_{\max} = \frac{\lambda}{2^p d} \frac{\frac{N^2}{2^{2r}}}{1 + \frac{N^2}{2^{2r}}}$$

but the desired direction is given by

$$\sin\theta = \frac{\lambda}{2^p d}.$$

The difference between these is the approximate error in radians;

$$\theta_\epsilon = \frac{\lambda}{Nd} \frac{2^r}{2^b N}.$$

Expressed as a fraction of the bandwidth of the array, the error is

$$\frac{Nd}{\lambda}\theta_\epsilon \approx \frac{2^r}{2^b N}.$$

This can be interpreted as the product of the smallest fractional part of 2π radians available in the phaser times the largest fractional part of the array co-phased at the first beam position. For highly accurate arrays, the required pointing accuracy may well determine the number of bits required. It should be remembered, however, that whenever the required phase increment per element equals the smallest phaser incre-

ment, there is no loss or pointing error due to phase quantization. In the previous example of a 50-element line array steered in increments of one-fifth beam width, a maximum pointing error of 1/200 beam width is obtained with three round-off bits. Five bits must then be provided for the phaser, resulting in a maximum loss due to phase quantization of 0.014 db.

Selection of Time Delayers

Since phase is the product of frequency and time delay, the techniques used to produce phasers can, in principal, be used to produce time delayers. The critical difference between phase and time delay, as used in arrays, is that phase is computed on the basis of modulo 2π, i.e., only the fractional part of the total number of cycles of phase is used. In fact, phasers employ time delay in lengths less than one wavelength, while time delayers require transmission paths nearly as long as the longest dimension of the array. Short delay lengths allow dispersive media to be employed in phasers and provide low loss. Large delay lengths, however, require nondispersive media and lead to higher losses and additional weight.

It is in wide instantaneous bandwidth arrays that time-delay steering is required. Thus, the time delayer must be designed to function over large bandwidths. It is essential that the time delayer have a linear phase–frequency characteristic over its bandwidth, since this determines the range accuracy and resolution of the radar. The many switches separated by relatively long delay lengths require an exceptionally good match at each switch junction. For example, just two connectors of 1.2 VSWR can produce a phase error of about 0.5 degree, and two sets of these connectors can produce a maximum phase error of about 2 degrees. Thus, care must be taken in the design of a time delayer, not only to reduce junction mismatch but also to space junctions so that their phase errors do not accumulate.

The insertion loss through a time delayer can vary several decibels from its minimum-delay to its maximum-delay state. This variation in the output-power level can exceed the tolerance of the power amplifier driven by the time delayer, or can lead to an undesirable amplitude taper across the array. It is therefore necessary to compensate the insertion loss to reduce its variation to a fraction of a decibel.

Because of the requirement for a linear phase–frequency relationship in the device, r-f time delayers generally operate by selecting alternate TEM transmission lines of different lengths, with either a diode switch (Figure 11) or a switchable circulator (Figure 12).

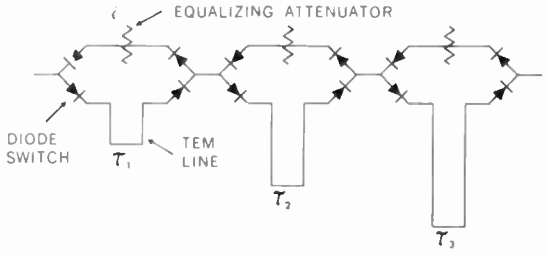

Fig. 11—Diode time delayer.

Reflection-type devices can also be used, but these tend to be more sensitive to junction-match conditions and are more difficult to equalize with respect to insertion loss. Two-path time delayers can be equalized by making the loss in the short path equal to that of the long path. Some of the errors induced by junction match can be reduced by the use of isolators in the two paths. The use of isolators is particularly important when circulator switches are used (otherwise, the two paths act as a resonant ring with respect to leakage through the input circulator).

The choice between diode-switched and ferrite-switched time delayers depends largely on the frequency and the mission of the array. Diode delayers tend to have the shortest switching time, whereas circulator delayers tend to have less loss and higher power-carrying capacity. Diode switches are particularly attractive at the lower microwave frequencies where their loss is least.

Selection of an Array Feed Network

The characteristics of an array are determined almost as much by the way it is fed as by the way it is steered. There are many ways to feed an array, and we discuss here some of the more important tech-

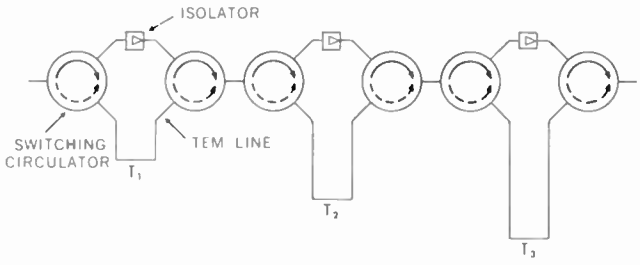

Fig. 12—Switching circulator time delayer.

niques. The term "feed network" will be used to mean the method of distributing, with equal phase, the r-f excitation from a given source to the many antenna elements, or, in the receiving sense, the method of collecting at a given terminal the energy from the many antenna elements. Feed networks can be classified as either single or multiple beam-forming networks, and as either transmission-line or optical feeds. The transmission-line feeds can be further classified as parallel or serial feeds, depending on whether or not essentially equal transmission path lengths exist between the input and output terminals.

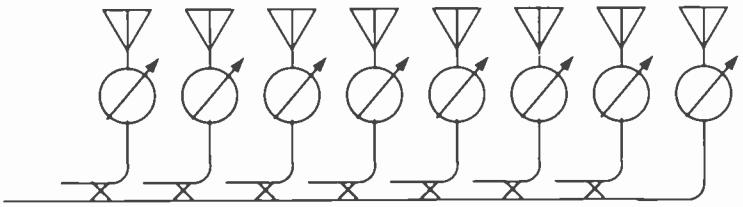

Fig. 13—Reactive corporate-feed network.

Single-Beam Transmission-Line Parallel Feed

The single-beam transmission-line parallel feed (Figure 13) generally delivers equal-amplitude inphase excitation to the phasers from a single source. This is accomplished by a number of levels of N-way power division, and is sometimes called a corporate feed because of its resemblance to a corporate organization chart. Low-level power division (close to the antenna elements) is generally binary because of the advantage of symmetrical 2:1 power dividers and the availability of hybrid junctions for use where a matched feed is required. Higher levels may require three-way, five-way, or higher power dividers to accommodate the number of elements in the array in case it is not a power of two.

Matched power dividers (those appearing matched when looking into an output port with the input port matched) are most easily realized for 2:1 power division. The realization here is the familiar hybrid tee or side-wall and top-wall hybrid couplers, which provide a matched fourth port to terminate unequal reflections at the output ports. This is contrasted with reactive N-way power dividers that have a reflection coefficient of $1 - (1/N)$ looking into an output port. When N is large, a wave entering an output port is almost totally reflected if

all other output ports are not simultaneously driven at the same phase. When reactive power dividers are used, the energy reflected from an antenna element will be re-reflected at the divider and form a second antenna beam, having passed a total of three times through the phasers at the elements. In the case of reciprocal phasers, this multiple reflection will produce a large spurious beam steered three times as far from broadside as the desired beam; this spurious beam may also be well above the desired side-lobe level. In the case of nonreciprocal phasers, the round-trip phase shift is generally independent of phaser

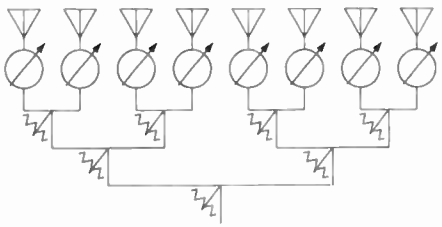

Fig. 14—Matched corporate-feed network.

setting, and the energy reflected from the elements passes back through the power divider without reflection. Thus the choice between a matched feed and a reactive feed depends largely on the choice between reciprocal and nonreciprocal phasers.

Parallel feeds are generally designed for greater bandwidth than is the case for serial feeds. For the most part, power-divider junctions of one design are used throughout the parallel feed, while serial feeds require a special junction design for each element in a line of the array. However, parallel feeds are often larger, heavier, and more expensive than serial feeds.

Single-Beam Transmission-Line Serial Feeds

An example of the serial-feed configuration is a frequency-scanned array. The excitation provided to each element in the array is adjusted by the coupling between the line feeding that element and the serial-feed line. The coupling values vary from element to element, because the power level in the feed line decreases as more elements are fed. Thus, individual coupler designs are necessary for each element in the line. Another phenomenon typical of serial feeds is the tendency for the small reflections from the various branch-line junctions to add at

certain frequencies or at certain beam positions, giving a high VSWR and reduced gain.

When a serial feed is used with a phase-steered array, the length of the feed line is considerably shorter than that required for frequency scanning, and the bandwidth is considerably greater (see Figure 15). The bandwidth of a serial-feed configuration can be made as great as that of a parallel feed by compensating the difference in path length along the serial feed line. Compensation is achieved by additional length between the line and the element. This technique, however, leads again to a large, expensive structure and negates the principal advantage of serial feeding.

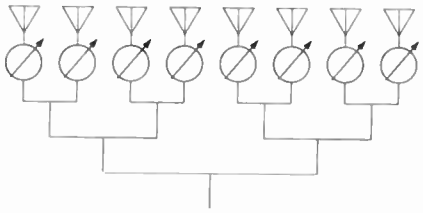

Fig. 15—Serial-feed network.

The parallel feed tends to be large and to have a large bandwidth, while the serial feed tends to be smaller and lighter with a smaller bandwidth. The optimum feed for a particular array system, depending on the bandwidth requirements, might be a combination parallel–serial feed; this could be designed to use the maximum path-length difference between elements consistent with the design bandwidth to reduce the size and weight of the feed structure.

Single-Beam Optical Feed

An optically fed array generally consists of a single feed illuminating many pick-up (primary) antennas that send the collected energy through steering phasers to radiating (secondary) antennas. The collimating function (providing equi-phased excitation at the secondary antennas) is generally accomplished by optical techniques such as by parabolic reflectors or by lens structures (Figure 16). In some configurations, the collimation is accomplished by the steering phasers, but usually at the cost of decreased array bandwidth. Optical feeding can also be accomplished from the radiating side of the array, with the same element performing both the collecting and radiating func-

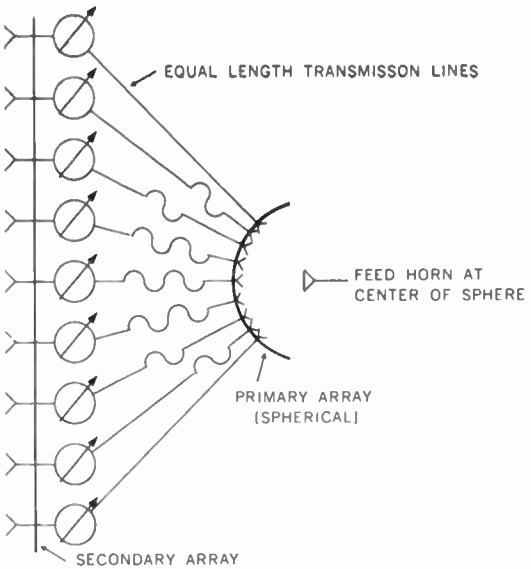

Fig. 16—Spherical–planar optical-feed system.

tions (Figure 17). This is accomplished at the cost of aperture blockage by the feed source.

Optical feeds, whether they be horns feeding elements of a subarray, large transmission lenses, or reflector arrays, have a great advantage in weight and cost. These techniques, however, have many

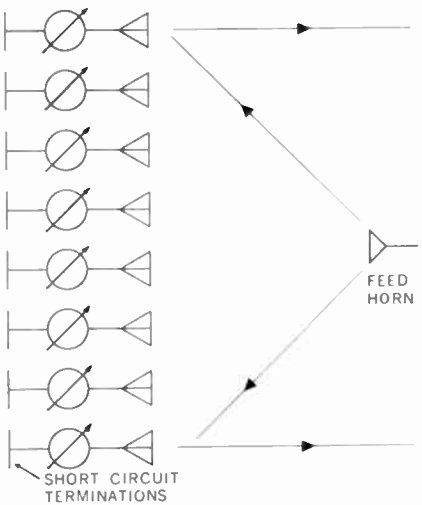

Fig. 17—Reflection array optical feed.

of the disadvantages common to optical antenna systems. On the other hand, an electronically scanned beam can be added to an optical antenna system to provide flexibility at the cost of added weight and greater loss.

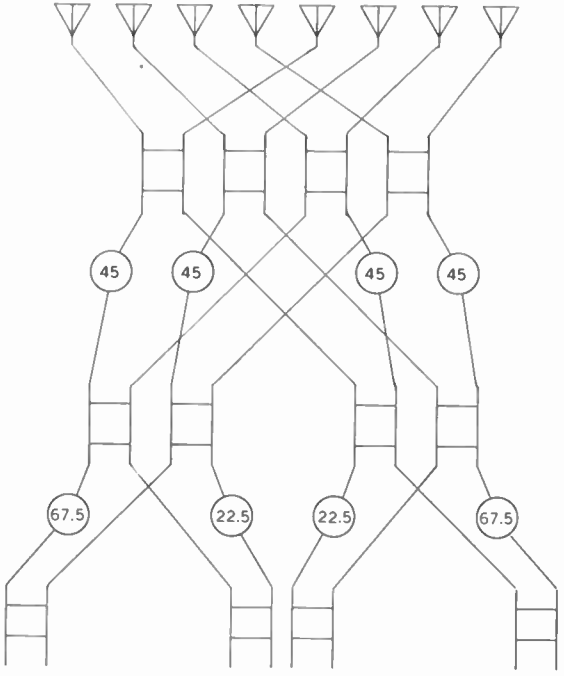

Fig. 18—Parallel multiple-beam-forming matrix.

Parallel Multiple-Beam Feeds

The parallel multiple-beam-forming network shown in Figure 18 is a network of hybrid couplers and fixed phase shifters that provides a number of isolated beam ports equal to the number of elements in the array. Each beam port is associated with an illumination phase taper across the array differing in increments of 2π radians from all others. These beams have the property of being essentially orthogonal, which implies that the total power radiated in any combination of these beams is the sum of the powers radiated in each beam. Another aspect of beam orthogonality is that energy supplied at any one beam port is not coupled to any other beam port by the beam-forming network.

Adjacent beams produced by the parallel beam-forming network cross over at about the 4-db point on the main beam ($2/\pi$ voltage ratio). Since the length of the array in wavelengths varies with frequency, the beam width varies inversely with frequency. This changing beam width, together with a constant crossover level, implies that the pointing angle of off-broadside beams must change. The amount of change is proportional to the distance of the beam from broadside in beam widths. This phenomenon gives us another point of view on the bandwidth characteristics of phase scanning discussed earlier.

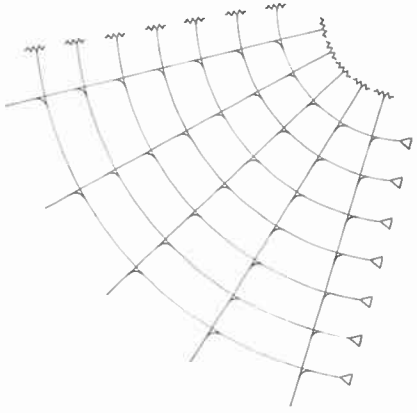

Fig. 19—Serial multiple-beam-forming network.

A multiple-beam network can be operated with separate transmitters and/or receivers on each beam port, giving simultaneous coverage of the scan sector of the radar with a single aperture. Alternatively, a single transmitter–receiver circuit can be switched to the various beam ports. This provides scanning by beam switching rather than by phasing. One disadvantage of this technique is that steering is restricted to increments of one beam width unless a complex handover network is used for interpolation between beam positions.

Serial Multiple-Beam Feeds

A serial-feed multiple-beam former is constructed by extending the element lines beyond the couplers at a given serial feed line instead of terminating them in matched loads (Figure 19). These lines are coupled to a second serial-feed line and subsequent feed lines, with the element line lengths adjusted to provide a one-wavelength taper at the center frequency between the edge elements. Each serial feed line then forms

a separate beam terminal, with the beams separated in time delay rather than in phase. The various beams cross over at the 4-db point only at the design center frequency. The cross-over level is higher at the low-frequency end of the band and lower at the high-frequency end of the band. The beams are not orthogonal except at the center frequency. This produces a loss over the frequency band equivalent to that due to frequency scanning in the parallel beam former.

The choice between serial beam formers rests on considerations much like those for the corresponding single-beam feeds. Additional considerations are the better interbeam isolation provided by the parallel beam former and the fact that the serial beam former requires N^2 couplers for an N-element N-beam network, while a parallel beam former requires $N(\log_2 N)/2$ couplers.

Multiple-Beam Optical Feeds

Multiple-beam optical feeds can be designed in either a transmission-lens configuration or a reflector configuration. Wide-angle coverage or a wide field of view with a narrow beam generally requires a spherically symmetrical optical system. This requirement is satisfied by a Luneberg transmission lens or by a spherical reflector. In these systems, multiple feed elements are placed on the optical focal surface, and collector or transfer elements are located on a concentric spherical pickup surface. The collector elements are connected through equal-length cables to radiating elements located at corresponding positions on a remote radiating surface of the same contour as the pickup surface. Both systems can be folded by the use of reflectors, with the same elements serving both the feed and collecting functions.

From purely optical considerations, it should be possible to space the beams produced by an optical beam former as closely as desired. However, inter-beam coupling and the limitations imposed by practical feed-element sizes limit these techniques to beam spacing the same as or larger than that achieved by transmission-line multiple beam formers.

Multiple-beam optical feeds have many of the same advantages and disadvantages as single-beam optical feeds. The differences in size, weight, and cost, however, are less marked in the multiple-beam case.

Radiating-Element Considerations

A primary consideration in designing an antenna element for an electronically scanned array is the area on the array face that can be occupied by an element. It is important that this area be as large as

practicable, because it determines the number of antenna elements, phasers, couplers, and other components that will be required in an array of a given size. As the antenna elements are made larger, secondary lobes appear in the pattern response of the array that can be as large or larger than the main lobe. These secondary lobes are called grating lobes or spectral orders. They result from the inherent ambiguity of multiples of 2π radians in the relative phase between elements, and they depend on the periodic spacing of array elements.

An interesting question in the design of array radiating elements is whether the array is more properly considered as a continuous aperture or as a collection of individual elements. This question is partly answered by the appearance of grating lobes. When the elements are so small that no grating lobes appear when the array is illuminated at any angle in its field of view, the gross aperture characteristics dominate and the character of the individual elements is of secondary importance. If grating lobes are visible, however, the character of the individual large elements is of primary importance in determining the performance of the array. Emphasis here is placed on periodic arrays of small elements, since these considerations apply in general to many different varieties of element types. The performance of arrays of large elements is so unique to the arrangement and element type that a general discussion is impossible. However, techniques for reducing the number of elements are discussed briefly.

Element Area

It was shown earlier that the pattern of a linear array is periodic in the variable $u = \sin \theta$ with period λ/d, where d is the element spacing. It was also shown that the operation of beam steering is a translation along the u axis of the pattern function. The observable radiation pattern of the array is determined by those values of the pattern function for which u corresponds to a real angle in space. Observable pattern response occurs for values of u between one and minus one. The interval $(-1, 1)$ on the u axis is called the visible interval or visible space. If the array is to have only one spectral order in visible space at any one time for any scan condition, the period of the pattern function must be larger than the length of the visible interval. For this condition, we must have

$$\frac{\lambda}{d} > 2.$$

The spacing between elements must be less than one-half wavelength.

Generally, however, the array will be scanned over only part of the visible interval, say to θ_m from the normal. A shorter period,

$$\frac{\lambda}{d} > 1 + \sin \theta_m,$$

will keep all higher spectral orders out of the visible interval. Thus, the general requirement for element spacing is

$$\frac{d}{\lambda} < \frac{1}{1 + \sin \theta_m}.$$

This relationship shows that if the array is not scanned from the normal ($\theta_m = 0$), a full wavelength spacing may be used.

The variable u is, in fact, the direction cosine of the observation angle with respect to the axis of the line array. If an array of these line arrays is arranged along an orthogonal axis, we can designate the direction cosine on that axis by the variable v. The composite pattern of this linear array of linear arrays is the planar-array pattern obtained as the product of the two linear-array patterns. The planar-array pattern is periodic in u and v in u-v space with period λ/d_u in the u direction and period λ/d_v in the v direction.

The visible interval on the u-v plane is a circle about the origin of unit radius (Figure 20(a)). Inside this circle, the scan coordinates are related to spherical polar coordinates about the array normal by

$$u = \sin \theta \cos \phi$$

$$v = \sin \theta \sin \phi.$$

When the beam is scanned to the point (u', v') grating lobes occur at the points $(u' + i[\lambda/d_u],\ v' + j[\lambda/d_v])$ for all integer values of the indices i and j. As with the linear array, we are assured that no grating lobes fall within visible space if the spacing of elements in either direction is

$$\frac{d}{\lambda} = \frac{1}{1 + \sin \theta_m},$$

However, this spacing is inefficient and leads to more elements than are absolutely necessary to prevent grating lobes.

If a second array of the same spacing as the first is interposed half-way between the elements of the first in both directions (Figure

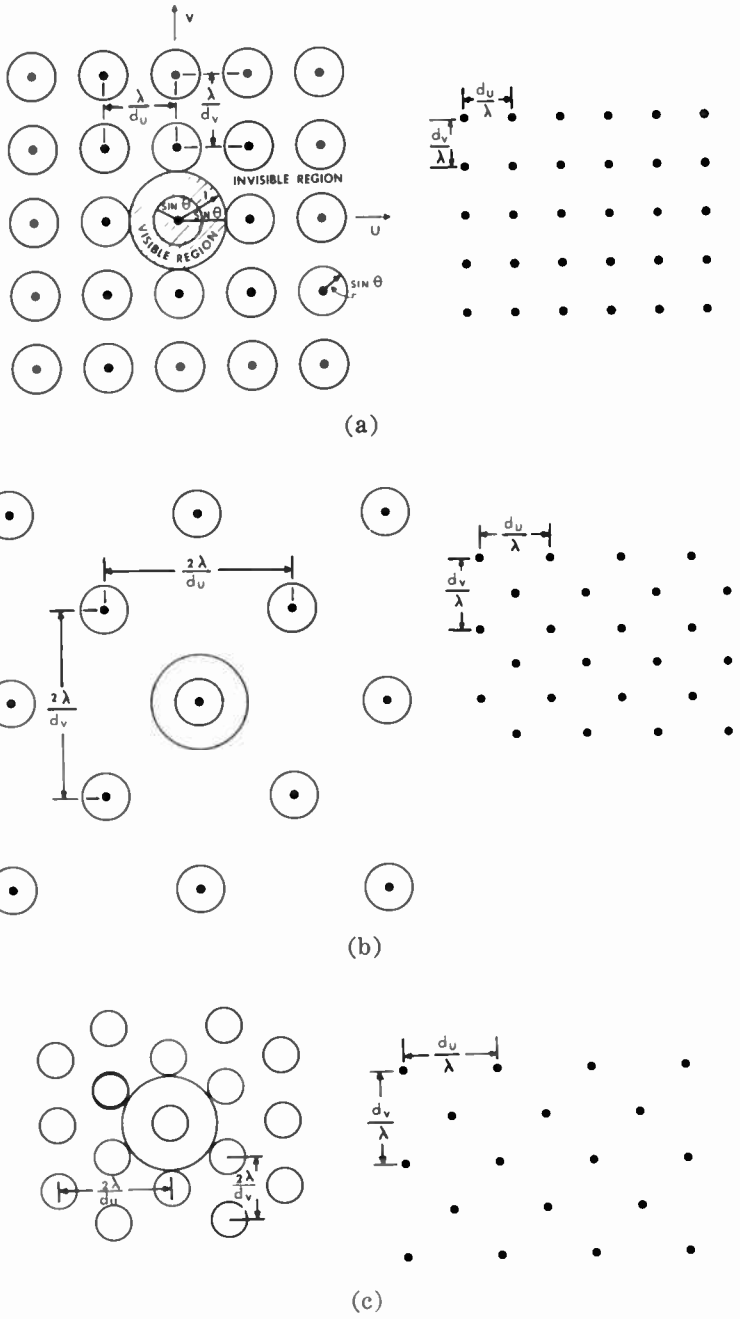

Fig. 20—The relation between element spacing and grating lobes in a planar array.

20(b)), alternate grating lobes are eliminated.[3] The grating lobes in this case occur at the points $(u' + i[\lambda/d_u],\ v' + j[\lambda/d_v])$ where the sum of the indices i and j is an even integer. Six grating lobes can now be placed within the distance $1 + \sin\theta_m$ of the main beam (Figure 20(c)). Let

$$2\frac{\lambda}{d_u} = \sqrt{\left(\frac{\lambda}{d_u}\right)^2 + \left(\frac{\lambda}{d_v}\right)^2} = 1 + \sin\theta_m.$$

Then

$$\frac{\lambda}{d_v} = \sqrt{3}\frac{\lambda}{d_u} = \frac{\sqrt{3}(1 + \sin\theta_m)}{2}.$$

This staggered element arrangement allows an element area of

$$A_e = \frac{2}{\sqrt{3}}\frac{\lambda^2}{(1 + \sin\theta_m)^2},$$

while the square element arrangement allows an element area of only

$$A_e = \frac{\lambda^2}{(1 + \sin\theta_m)^2}.$$

The design procedure for maximizing the element area for a given scan sector is equivalent to the following. First the angular coverage area is laid out on a unit sphere. The pole of the sphere is chosen to minimize the polar distance to the points of extreme scan or to the scan region of greatest importance. The outline of this area is then projected on the equatorial plane. A guard zone is established about this contour at distance $1 + (\lambda/D)$, where D is the diameter of the array. The additional guard distance λ/D is included to keep the edge of the grating lobe beam from entering visible space. Next, the element spacing is adjusted so that as many grating lobes as possible coincide with the guard contour. When the coverage is an octant of a sphere, six grating lobes can be placed on the contour. For asymmetrical

[3] E. D. Sharp, "A Triangular Arrangement of Planar-Array Elements Reduces the Number Needed," *I.R.E. Trans. on Antennas and Propagation*, Vol. AP-9, No. 2, p. 126, March 1961.

coverage contours, fewer can be placed in the minimum position. This procedure will lead to the maximum element spacing for a given array problem.

Element Gain

The two important characteristics of array-antenna elements are gain and impedance, where gain usually means the available gain per element of the array and impedance is the driving point impedance or reflectance at the element terminals when it and all other elements are active. This discussion is primarily concerned with elements in planar arrays. However, a brief discussion of linear-array-element performance will be offered for the insight it will provide.

It is well known that the directive gain of a continuous-broadside line array approaches, as a limit, twice the length of the array in wavelengths. Directive gain is defined as the ratio of effective radiated power (4π times the maximum power radiated per unit solid angle) to the total radiated power. An array of N isotropic sources of spacing d/λ (approximating a continuous line array) would therefore have a directive gain of $2Nd/\lambda$ if N is sufficiently large and d/λ is small. The directive gain per element then is $2d/\lambda$. The directive gain of a single isotropic radiator is unity. Thus, for spacings less than one-half wavelength, the array must react on the element to reduce its gain, while for spacings greater than one-half wavelength, the array increases the directive gain of the element. The directive gain at a full-wavelength spacing is twice that of the isolated element. When the spacing exceeds one wavelength, two grating lobes appear in visible space, increasing the radiated power about threefold and reducing the directive gain of the element to 2/3. The element directive gain increases to 4/3 at two wavelengths spacing and drops to 4/5 beyond. As the spacing is increased further, the maximum departure of the element directive gain in an array from its isolated gain decreases further. This behavior shows that the array dominates element performance for spacings less than a wavelength, while the element is dominant for much larger spacings (Figure 21). A much more rigorous discussion of elements in linear arrays has been given by Tai[4] and others.

In a planar array, the directive gain of an element is just 4π times the area (A_e) in square wavelengths occupied by the element when the spacing is so small that no grating lobes appear. When the array is scanned away from the array normal, the projected area in the direc-

[4] C. T. Tai, "The Nominal Directivity of Uniformly Spaced Arrays of Dipoles," *Microwave Journal*, Vol. 7, No. 9, p. 51, Sept. 1964.

tion of the beam determines its directivity. Directive gain, then, is

$$g_d = \frac{4\pi A_e}{\lambda^2} \cos\theta.$$

The gain figure most often used, however, is the available gain of the element, which considers the ability of the element to receive energy from its feed line as well as its directivity. This figure is measured

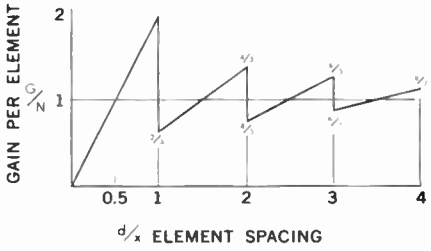

Fig. 21—Variation of element gain with element spacing in a very large linear array of isotropic radiators.

by the driving-point reflectance ρ of the element. The available gain then, is

$$g_a = \frac{4\pi A_e}{\lambda^2} \cos\theta \, [1 - |\rho(\theta,\phi)|^2].$$

The factor $1 - |\rho(\theta,\phi)|^2$ is the transmission loss between the aperture and space. It is determined by the element match and by the aperture polarization effects. The reflectance ρ of the element is shown as a function of the scan coordinates θ and ϕ because it varies in both magnitude and phase with the direction of scan.

It follows from the superposition theorem of linear systems tha the available gain of an element in an active array is the same as its available gain measured with all other elements terminated by the generator impedance of their feed lines. This provides a useful tool for evaluating the performance of an element in an array.

Aperture Impedance

When considering the impedance of an array radiating element the assumption of a large array is usually made. Attention is then largely

restricted to central elements or elements that are surrounded by a large number of other like elements. All such elements appear to be in essentially the same environment and should therefore respond alike. This assumption is valid for the majority of elements in typical arrays except where discontinuities occur in the form of supporting structure.

The impedance response of simple radiating elements in an array is determined to a first order by the impedance of a continuous-current sheet approximated by the array of elements. The simple current-sheet impedance will be called the aperture impedance of the array. The impedance of an actual array differs from this value because of superimposed "slow-wave" current sheets induced by the periodicity of the array; these usually contribute only to the element reactance. The effect of these higher-order current sheets is generally small as long as grating lobes are kept out of the visible interval. More complicated element structures can sometimes support "free" slow-wave modes of propagation that serve to magnify greatly the effects of the higher-order current sheets. This effect usually degrades the array performance but offers the possibility, with careful design, of improving the wide-angle impedance characteristics[5] and gain of the array.

A sheet of electric current flowing in the x direction on a perfectly conducting surface at $z = 0$ will produce an electromagnetic field above the plane whose components are

$$H_x = 0,$$

$$H_y = -J_x,$$

$$H_z = \frac{\beta}{\kappa} J_x,$$

$$E_x = -\frac{(\beta^2 + \kappa^2)}{\kappa k} \zeta J_x,$$

$$E_y = \frac{\alpha\beta}{k\kappa} \zeta J_x,$$

$$E_z = \frac{\alpha}{k} \zeta J_x,$$

where J_x is the surface current density, $\zeta = \sqrt{\mu/\epsilon}$ is the intrinsic

[5] F. G. Magill and H. A. Wheeler, "Wide-Angle Impedance Matching of a Planar Array by a Dielectric Sheet," *IEEE Trans. on Antennas and Propagation*, Vol. AP-14, No. 1, p. 49, Jan. 1966.

impedance of space, and $k = 2\pi/\lambda$ is the radian wave number, with components α, β, κ in the x, y, and z directions. In terms of the usual spherical-polar coordinate system,

$$\alpha = k \sin\theta \cos\phi$$

$$\beta = k \sin\theta \sin\phi$$

$$\kappa = k \cos\theta.$$

The aperture impedance is given by the reaction between the electric field and the electric current as

$$Z = \frac{-E_x}{J_x} = \zeta \frac{\beta^2 + \kappa^2}{\kappa k} = \frac{\sin^2\theta \sin^2\phi + \cos^2\theta}{\cos\theta} \zeta.$$

In this case, the aperture impedance is purely resistive and reactance variation at the aperture is due entirely to higher-order current sheets.

Another model of interest, because it simulates the behavior of slots or small-horn radiators, is that of a sheet of magnetic current, K_y, flowing in the y direction along a perfectly conducting surface at $z = 0$. Field components of this source are

$$E_x = -K_y,$$

$$E_y = 0,$$

$$E_z = \frac{\alpha}{\kappa} K_y,$$

$$H_x = \frac{\alpha\beta}{k\kappa\zeta} K_y,$$

$$H_y = \frac{\alpha^2 + K^2}{k\kappa\zeta} K_y,$$

$$H_z = \frac{\beta}{\kappa\zeta} K_y$$

and the aperture impedance is

$$Z = \frac{-K_y}{H_y} = \frac{k\kappa\zeta}{\alpha^2 + \kappa^2} = \frac{\cos\theta}{\sin^2\theta \cos^2\phi + \cos^2\theta} \zeta,$$

which again is purely resistive.

A slightly more complicated model is afforded by an electric current sheet displaced a distance d from a perfectly conducting sheet. For this case the impedance obtained for the electric sheet on an impeding surface is modified by the factor $(1 - \exp\{-j2\kappa d\})/2$, yielding

$$Z = \frac{\sin^2\theta \sin^2\phi + \cos^2\theta}{\cos\theta}\left(\sin^2 2\pi \frac{d}{\lambda}\cos\theta + j\frac{\sin 4\pi \dfrac{d}{\lambda}\cos\theta}{2}\right)$$

This model, representing an array of dipole elements above a ground plane, provides an example of a reactive component of aperture impedance arising from the fundamental current sheet itself.

For the magnetic current sheet model as a simplest example, the array gain function becomes

$$g_a = \frac{4\pi A_e}{\lambda^2}\frac{4\cos^2\theta}{(1+\cos\theta)^2}$$

in the E plane $(y = 0)$. The aperture gain has fallen to one half its normal value for a scan to 57°. This compares to a scan of 60° for projected area considerations alone. These considerations generally limit the maximum scan capability of an array to about 57°.

Array Thinning Techniques

The number of elements required to produce a beam of width θ_B, when an element spacing of $\lambda/2$ is assumed for wide-angle scanning, is approximately

$$N = \frac{10^4}{\theta_B{}^2}.$$

A typical array radar may require a one-degree transmitter beam with emphasis on gain and power and a one-third degree receiver beam with emphasis on angular resolution. The transmitter would then require about 10^4 elements and the receiver would require about 10^5 elements. The magnitude of these numbers has led to an investigation of techniques for reducing the number of active elements and phaser elements, particularly in the receiving array. Because of the power and gain requirements of a transmitting array, thinning techniques are not

ordinarily applicable to this or to duplexed arrays, where the same aperture performs both receiver and transmitter functions.

One popular technique for reducing the number of elements in an array is that of density tapering. Density tapering lowers the level of illumination near the edge of an array and reduces the side-lobe level of the array pattern. A conventional technique, known as amplitude tapering, accomplishes this by a reduction in the gain or amplitude of the edge elements. A reduction in edge illumination can also be accomplished by using fewer elements, all having the same gain or amplitude in the edge region, with a net saving in the total number of elements used.

There is no unique way of achieving a given density distribution, and, on the average, the several techniques give essentially the same result. One technique is the pseudo-random thinning process, which is based on a regular grid of elements spaced to prevent grating lobes. A random decision is made as to which elements are to be excited; the probability that a given element will be excited is equal to the normalized value of the desired illumination function at the element location. As a result, nearly all elements in the center of the array are illuminated, but only a few near the edge of the array.

Several physical and statistical constraints must be considered in the design of a density-tapered array. First, the average side-lobe level of an antenna is inversely proportional to antenna gain, and hence, for density tapering, it is inversely proportional to the number of elements activated. Second, when the side-lobe level to be achieved is lower, it is less likely that edge elements will be activated, and therefore the number of active elements will be smaller. Finally, the peak side-lobe level can be expected to be at least 3 db above the average side-lobe level but will usually be less than 10 db above the side-lobe level of the continuous illumination function being approximated by density tapering. This technique can produce very narrow beams, with side-lobe levels below —20 db using only 10^3 elements.

Conclusions

Many different classifications or subclassifications of phased arrays have been discussed in terms of types of scan, control elements, feed systems, element distribution, and other parameters. Some of the principal characteristics of the major techniques in each class have been described. However, an array design for a particular mission at a particular time will result from an optimum combination of techniques in each class, depending on the then current state of technology.

To illustrate the influence of technology on the development and

deployment of array radars, we consider briefly the order of development in the present rise of array technology. Among the earliest of the present array radars systems was the frequency-scanned array of moderate size, with a beam width of one to two degrees. This system has the advantage of relatively low cost and weight, and the disadvantage of narrow signal bandwidth and a practical size restriction. The need for more capable array systems, together with a lack of high-power high-speed phasers, led to the development of arrays using a transmitter for every element in the array, with phase control at low power levels. For time-delay phasing, this may well be the ultimate realization of array bandwidth and power capabilities.

The development of the digital latching ferrite phaser provided the high-power high-speed phaser needed for the development of arrays of intermediate capability. The development of high-power diode phasers at the lower microwave frequencies closely parallels that of the ferrite phaser at the higher frequencies. Arrays based on these high-power phasers range in capability from single-transmitter phase-scanned arrays to hybrid arrays using multiple transmitters feeding phase-scanned subarrays and low-level time-delay scanning for fine beam-positioning control. This latter technique has most of the capabilities of the fully time-delayed array, extends the technique to the higher microwave frequencies, and effects a considerable reduction in the number of active components required in the array.

Scientific and engineering technology has developed to the point that electronically scanned array radars can be built to provide a broad spectrum of capabilities. The major challenge in the development of future array technology will be the reduction in the cost of array components.

EXPERIMENTAL INVESTIGATION OF MULTIWATT ARGON LASERS

I. GOROG AND F. W. SPONG

RCA Laboratories
Princeton, New Jersey

Summary—*Because of the lack of reliable quantitative theory relating optical output power and efficiency to the operating parameters of the ionized argon laser, an experimental investigation was made to determine the quantitative dependence of the output power on discharge current, gas pressure, external axial magnetic field, and discharge-tube bore size. The experimental curves obtained are given. Since the curves are self-explanatory, discussion is confined to pointing out salient features of the data. Explanations of a more theoretical nature are offered only where it is felt they are unavailable in the literature. All the work reported centers on multiwatt systems. The maximum optical output power measured was 6.86 watts from a 4-mm bore 18-inch long tube. The best efficiency figure obtained was* 1.02×10^{-3}.

GENERAL DISCUSSION

THIS PAPER presents results of an experimental investigation of the dependence of the optical output power of multiwatt argon lasers on bore size, manifold pressure, magnetic field, and discharge current. Water-cooled quartz discharge tubes were employed. The bulk of the experiments were performed with 3, 4, and 6-mm inside diameter and 18-inch-long tubes. One series of measurements was performed on a 4-mm bore 11-inch-long tube and one on an 8-mm bore 11-inch-long tube.

All the tubes had standard L-cathodes of 1 inch diameter located either on the tube axis (in which case a small hole drilled through the cathode body provided passage for the laser light) or at a 90° angle to the tube axis. From the point of view of discharge stability, the preferred cathode position is on axis and close to (< 1 inch) the entrance to the narrow-bore active section. The disadvantage of this arrangement is that, due to the axial magnetic field, the cathode can easily be damaged by excessive ion bombardment (especially at low pressures, high magnetic fields, and high currents). Reliable cathode operation was achieved by locating the magnet either on the axis approximately 3 inches from the narrow bore entrance, or off the axis approximately 3 inches from the entrance and 3 inches from the axis. The first of these arrangements resulted in more stable operation and the second allowed lower pressures and higher fields without cathode damage.

The anodes were water-cooled copper coils, located on the tube axis. The exact positioning of the anode is less important than that of the cathode, but it is not entirely abitrary. If the anode is located too close to the exit of the narrow-bore section and asymmetrically with respect to the tube axis, or if the tube is located asymmetrically with respect to the magnet axis, the discharge can be bent towards the wall of the anode bulb, which is not water cooled, resulting in local heating and melting of the quartz. If the cylindrical symmetry is maintained, but the anode is located too close to the exit and the tube is operated at low pressures and high fields, the plasma jet emerging from the narrow-bore section will not be terminated by the anode, but will extend beyond it. It may act as a plasma torch and melt the anode bulb behind the anode. If the anode is located too far away from the exit, instabilities may develop. The best results were obtained by locating the anode approximately 1 inch from the exit, symmetrically with respect to the tube axis.

The instabilities were observed by monitoring the a-c component of the discharge current. The oscillation frequencies were observed to be in the 10-25 kHz region. Furthermore, it was observed that the presence of instabilities was associated with certain operating regions and not with specific thresholds; in many instances it was found that by decreasing or increasing either the magnetic field or the discharge current, the oscillations could be quenched.

Pressure was monitored by a Pirani-type gauge and is subject to the usual pressure-gauge uncertainties. Furthermore, the gas manifold pressure was measured, and not the pressure in the active narrow-bore section. The two may be significantly different due to pump-out effects. The range of pressure values was ~ 0.05 to 1 torr. The range of magnetic field values was 0-2500 gauss. The range of discharge current values was 0-50 amperes. The power supply employed was an unfiltered three-phase full-wave rectifier, having a maximum capacity of 50-amperes at 280 volts d-c.

The laser optics employed consisted of two 2-meter-radius external mirrors, one with high reflectivity ($R > 99\%$) and one partially transmitting ($R \simeq 95\%$ or $R \simeq 90\%$). The cavity length was approximately 150 cm.

Optical power was measured by a bolometer-type laser radiometer (Westinghouse Model RN-1). The power values given always imply total power from all lines.

It is unlikely that the efficiency can be increased significantly above 10^{-3}. Other problem areas, which can be solved by proper engi-

neering, are related to the cavity optics and can be divided into two parts:

(1) General cleanliness of both the mirrors and Brewster windows. In one set of experiments more than a factor of two decrease in output power was observed after extended operation in a normal laboratory environment.

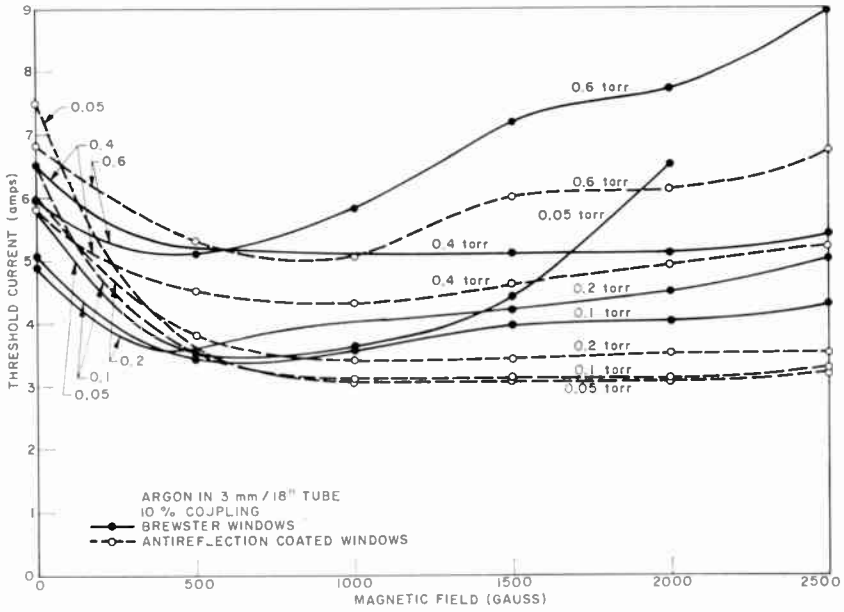

Fig. 1—Threshold current as a function of axial magnetic field with manifold pressure as a parameter.

(2) Mirror heating effects. If the laser is operated at 5 watts power level, corresponding to 100 watts standing-wave power in each direction, a marked beam narrowing occurs. The maximum power level a standard mirror can sustain is estimated at approximately 1,000 watts/cm^2.

Threshold Data

Figures 1 through 6 show the threshold currents for the lowest-threshold (4880 Å) argon-laser line as a function of axial magnetic field. The parameters are pressure and tube geometry. The lowest measured total tube drop was 69 volts (4-mm bore 11-inch-long active section, 0.2 torr manifold pressure 1,000 gauss axial magnetic field, and 8 amperes discharge current) and the highest measured

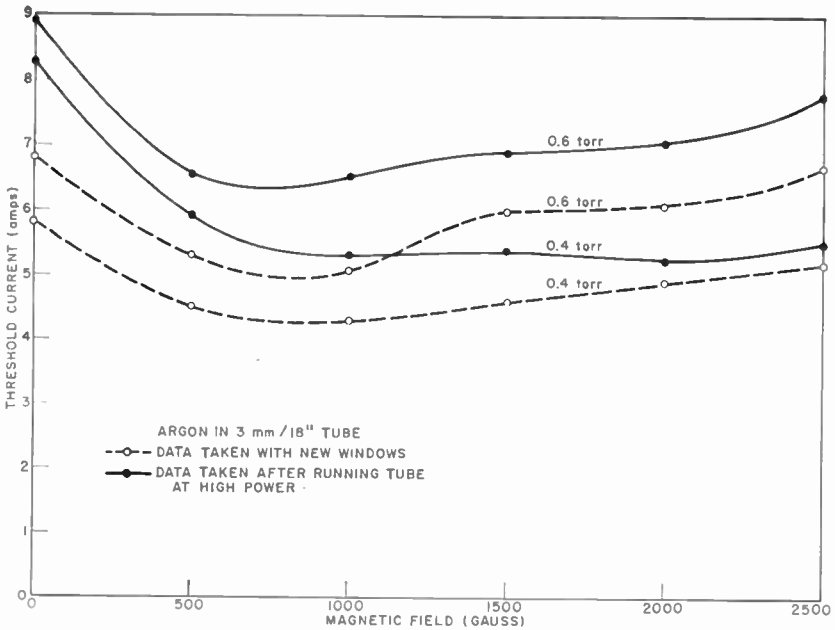

Fig. 2—Threshold current data showing degradation of antireflection coatings.

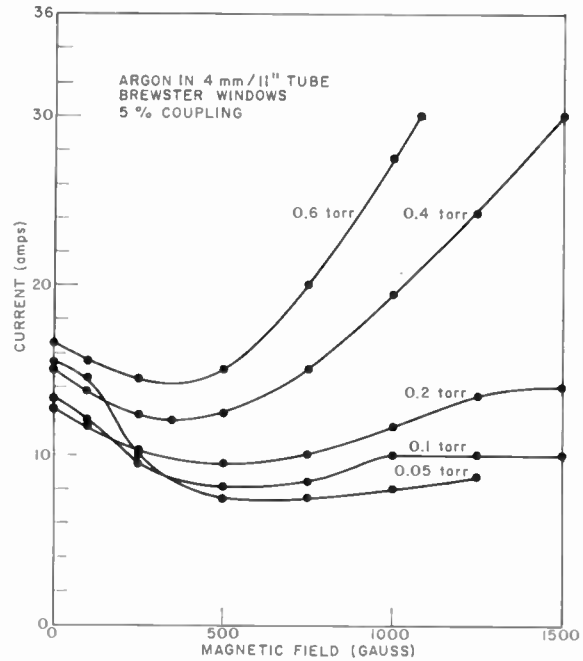

Fig. 3—Threshold current data for argon laser with 4-mm bore, 11-inches long.

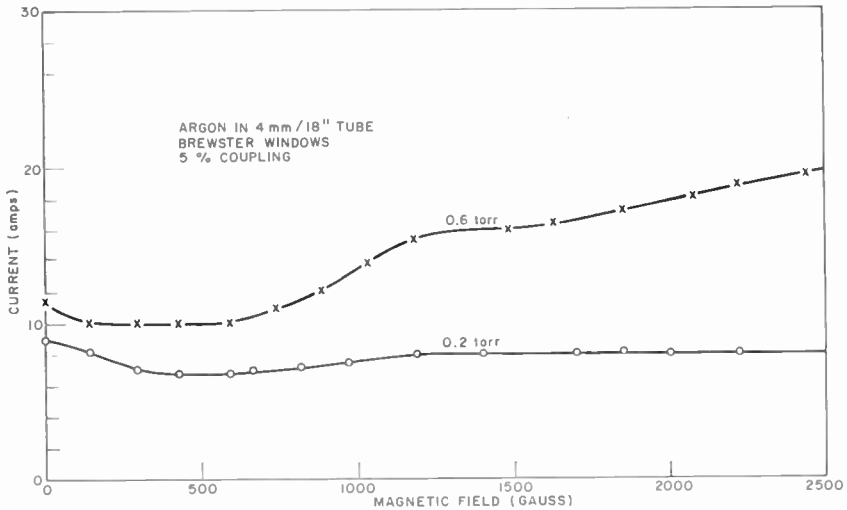

Fig. 4—Threshold current data for argon laser with 4-mm bore, 18-inches long.

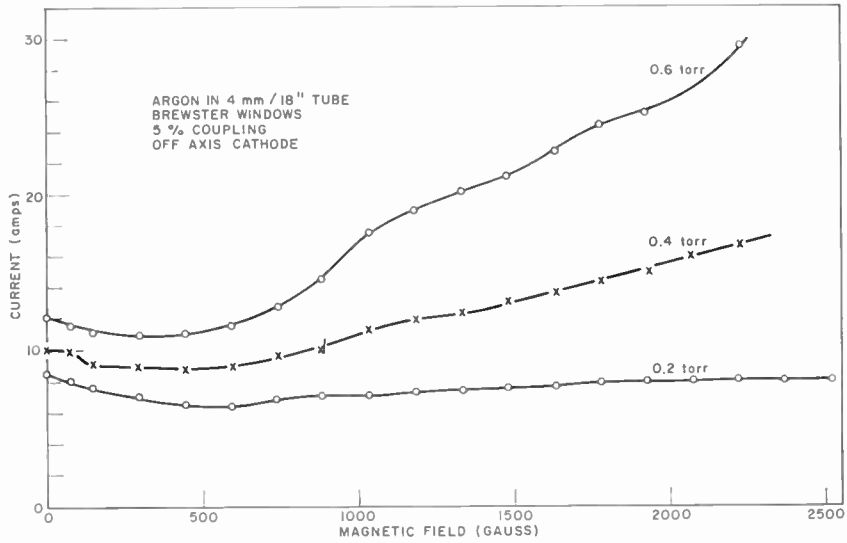

Fig. 5—Threshold data for argon laser with 4-mm bore, 18-inches long, with off-axis cathode.

threshold voltage was 233 volts (3-mm bore 18-inch-long active section, 0.6 torr manifold pressure, no axial magnetic field, and 5.95 amperes discharge current).

The question of Zeeman effects was studied by designing a tube with demountable end flanges to which either Brewster windows or antireflection coated perpendicular windows could be attached. The antireflection coated windows show significant deterioration with

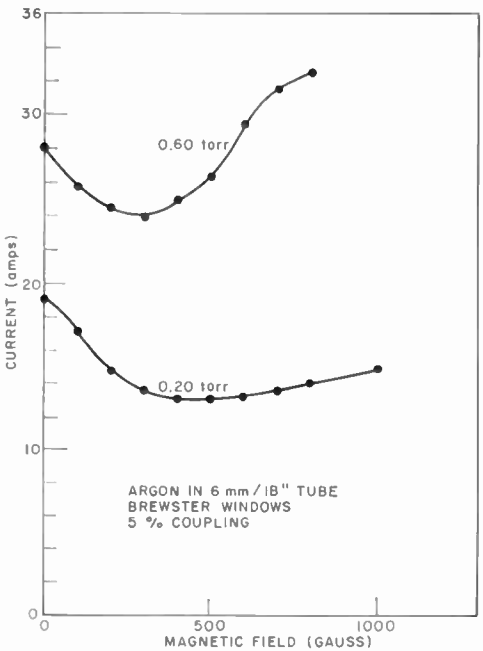

Fig. 6—Threshold data for argon laser with 6-mm bore, 18-inches long.

operating time (Figure 2). Therefore, the absolute position of the curves in Figure 1 is indicative only of the range of threshold currents. Of physical significance, however, is the rate of change of the threshold current with magnetic field at field values above optimum. In the 0.2-0.4 torr pressure region (which corresponds to the optimum for high-power operation, as shown in Figures 8-11), the threshold behavior is not very different whether antireflection coated or Brewster-angle windows are employed. This finding is in good agreement with the power measurement results of Labuda, et al.[1] At very low

[1] E. F. Labuda, E. I. Gordon, and R. C. Miller, *IEEE Jour. of Quantum Electronics*, QE-1, p. 273, 1965.

or very high pressures the improvement is noticeable, but these regions are of lesser practical importance.

In a previous study of the high-pressure, high-magnetic-field threshold behavior,[2] it was found that for $p \gtrsim 0.2$ torr and $B \gtrsim 400$ gauss at constant discharge current, I, the pressure versus magnetic field threshold curves resemble hyperbolas,

$$K(I) = (p - p_o)\ (B - B_o)\ ,$$

where the parameter K is a monotonically increasing function of I, or inversely, I is a monotonically increasing function of K. Then

$$\left.\frac{\partial I}{\partial B}\right|_{p\ =\ \text{constant}} = \frac{dI}{dK}\frac{\partial K}{\partial B} = \frac{dI}{dK}\,p$$

and so

$$\left.\frac{\partial I}{\partial B}\right|_{p_1} > \left.\frac{\partial I}{\partial B}\right|_{p_2} \quad \text{when } p_1 > p_2$$

That is, the optimum field value is more critical at high pressures than at low pressures. Compare, for example, the rate of increase of threshold current with magnetic field at 0.6 torr and 0.2 torr in Figures 1 through 6.

Pump-out effects may play an important role in the pressure versus magnetic field behavior. The occurance of pump-out toward the anode[3] at these operating parameters indicates that it is caused by ionic losses to the discharge tube wall.[4] Also, it has been found that the optimum operating pressure of segmented tubes[5] is significantly lower than that of continuous-bore tubes (~ 0.1 torr versus ~ 0.4 torr), which indicates that large pressure differences must exist between the manifold and the narrow-bore active section.

If one assumes that, at the current densities of interest, the wall temperature plays a secondary role in establishing the particle density in the active section, then one must conclude that the large difference

[2] I. Gorog and F. W. Spong, "High Pressure, High Magnetic Field Effects in Continuous Argon Lasers," *Appl. Phys. Letters*, Vol. 9, p. 61, 1 July 1966.

[3] E. I. Gordon, E. F. Labuda, and W. B. Bridges, "Continuous Visible Laser Action in Singly Ionized Argon, Krypton, and Xenon," *Appl. Phys. Letters*, Vol. 4, p. 178, 15 May 1964.

[4] G. Francis, *Handbuch der Physik*, Springer-Verlag, Berlin, Vol. 22, p. 198, 1956.

[5] K. G. Hernqvist and J. R. Fendley, private communication.

of optimum pressure between the segmented and continuous-bore tubes is due to the stronger pump-out in the latter case, and furthermore, that the optimum particle density in the active section is significantly lower than that corresponding to the optimum measured manifold pressure. Then, as the magnetic field is increased beyond its optimum value at a given pressure, its main effect on laser performance is to

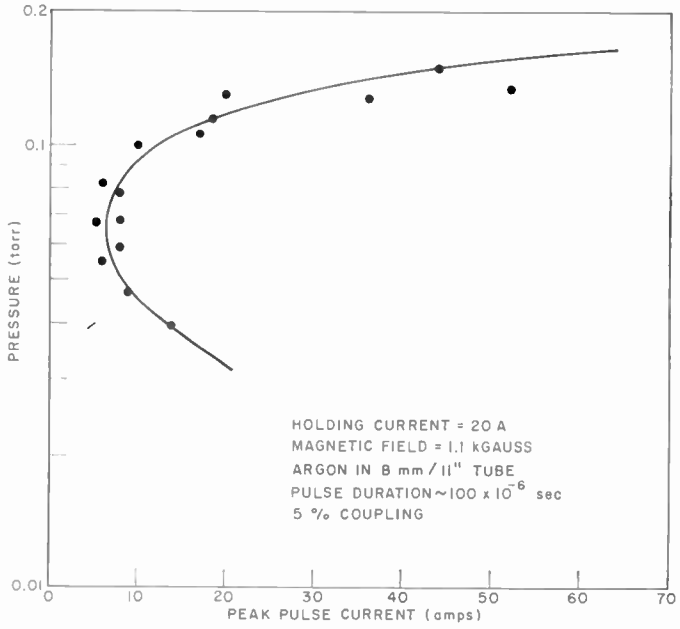

Fig. 7—Pulsed argon-laser threshold current as a function of pressure.

increase the particle density in the active section through the reduction of wall losses and consequent reduction of pump-out. Simultaneously the electron temperature decreases. This relation explains the high-field behavior. The reduction of threshold current between zero field and optimum field can be accounted for by noting that in a two-step excitation process, the inversion density, and therefore the gain, is a monotonically increasing function of the wall-loss time constant. In this regime, therefore, if the wall losses are reduced, a lower current is sufficient to maintain a constant gain.

The behavior of minimum-threshold optimum pressure as a function of magnetic field is of interest but is difficult to determine for continuous systems, because of the excessive cathode damage at low pressure and high magnetic fields. In a 4-mm tube at zero field this pressure is somewhere between 0.05 and 0.2 torr; at high fields it is

probably lower (see Figure 3). To eliminate the complication caused by pump-out, pulsed threshold measurements are useful. Figure 7 shows the results of an experiment in which the discharge has been maintained by a constant direct current and the pressure dependence of the threshold has been investigated by a superimposed short pulse. The optimum pressure is in the 0.06-torr range. Long-pulse* experiments performed without a holding current indicate that the optimum pressure is still lower.[6]

Table I—Typical Operating Data for Experimental Tubes

Optical Power Output (Watts)	Pressure (Torr)	Magnetic Field (Gauss)	Current (Amperes)	Voltage (Volts)	Tube Power Input (Watts)	Efficiency
3-mm bore/18-inch tube						
3.30	0.95	1,260	27.5	180	4,950	6.67×10^{-4}
2.19	0.72	1,260	22.5	168	3,780	5.79×10^{-4}
1.48	0.50	1,000	17.5	159	2,780	5.32×10^{-4}
0.99	0.40	1,000	15.0	155	2,320	4.25×10^{-4}
4-mm bore/11-inch tube						
1.67	0.60	500	50.0	109	5,450	3.07×10^{-4}
0.96	0.50	500	40.0	103	4,120	2.32×10^{-4}
4-mm bore/18-inch tube						
6.86	0.66	740	50.0	144	7,200	9.50×10^{-4}
5.60	0.66	740	40.0	138	5,520	1.02×10^{-3}
4.34	0.45	740	35.0	124	4,350	1.00×10^{-3}
1.17	0.26	592	20.0	114	2,280	5.15×10^{-4}
6-mm bore/18-inch tube						
1.47	0.37	667	50.0	101	5,050	2.92×10^{-4}
0.57	0.27	400	40.0	97	3,880	1.47×10^{-4}

Power Measurements

Measurements of the total optical-power output (all lines) are given in Figures 8 through 19. Table I gives typical values of the parameters. The curves are organized into three groups.

Figures 3 through 11 show power as a function of manifold pres-

* By choosing a pulse length longer than all excitation and thermalization time constants, but shorter than the pump-out time constants, the laser operation can be considered as c-w, but the discharge to be free of pump-out effects.

[6] S. A. Ahmed, private communication.

sure at constant magnetic field and discharge current. The discharge current value was chosen to be high enough to correspond to multiwatt operation, but low enough to allow a pressure scan without cathode

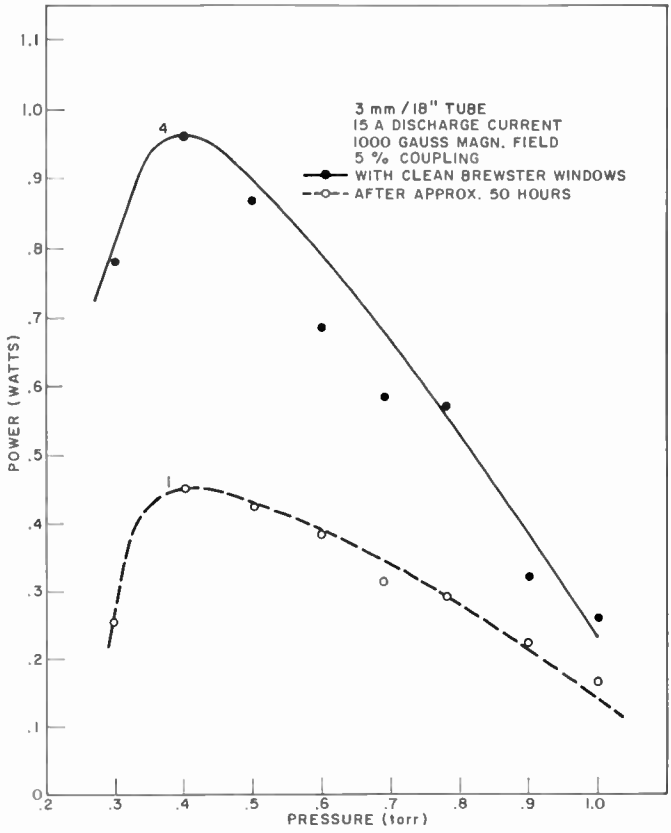

Fig. 8—Optical output power versus pressure measured after approximately 50 hours (curve 1) and then after Brewster windows have been cleaned (curve 4) (3-mm bore/18-inch-long tube).

deterioration. The value of the magnetic field was selected after preliminary measurements to correspond approximately to the optimum.

Figures 12 through 15 show power as a function of axial magnetic field at constant current and constant manifold pressure. The values of the discharge currents are the same as the ones on the corresponding pressure-scan curves. The constant manifold pressures were determined by the optima of Figures 8 through 11.

Figures 16 through 19 show power as a function of discharge

current. In Figures 16 and 18 the pressure and magnetic-field values for each data point were adjusted to correspond to the approximate optimum at the particular current value.

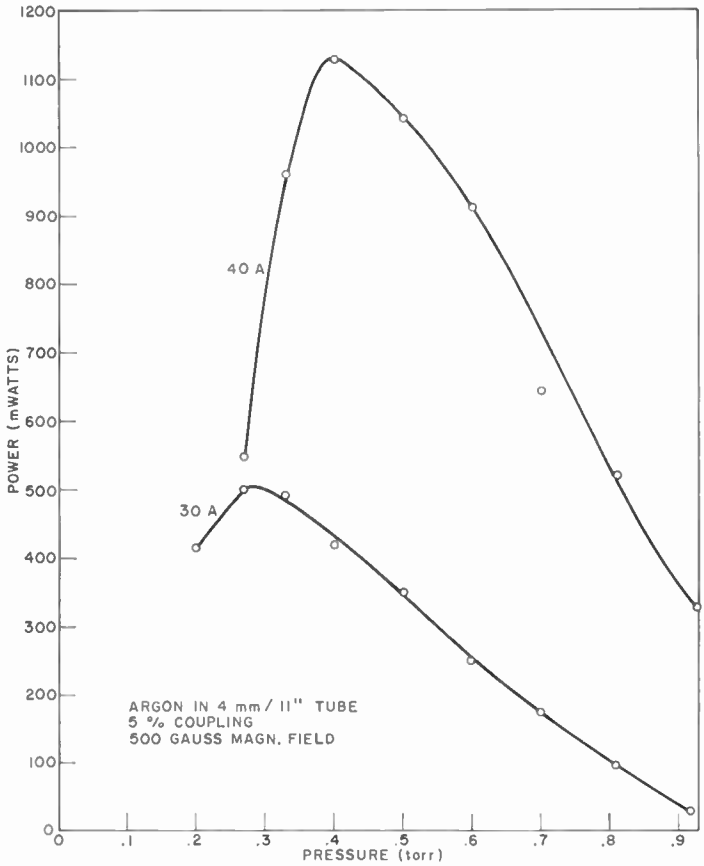

Fig. 9—Optical output power versus pressure for argon laser with 4-mm bore, 11-inches long.

Optimum pressure was found to be a very sensitive function of excitation current; in the 3-mm/18-inch tube it ranges from 0.26 torr at 10 amperes to 0.95 torr at 27.5 amperes; in the 4-mm/18-inch tube the lowest pressure used was 0.2 torr (corresponding to higher than optimum) up to 15 amperes, and the highest was 0.66 torr at 50 amperes. The optimum magnetic field at optimum pressure was found to be a slowly varying function of excitation current: for the 3-mm/18-inch tube it is in the range of 1,000 to 1,200 gauss, and for the

4-mm/11-inch tube it is 600 to 700 gauss. In none of these measurements was coupling truly optimized; it was kept a constant 5%, which is in the range of the optimum at high currents for most of the tubes used.

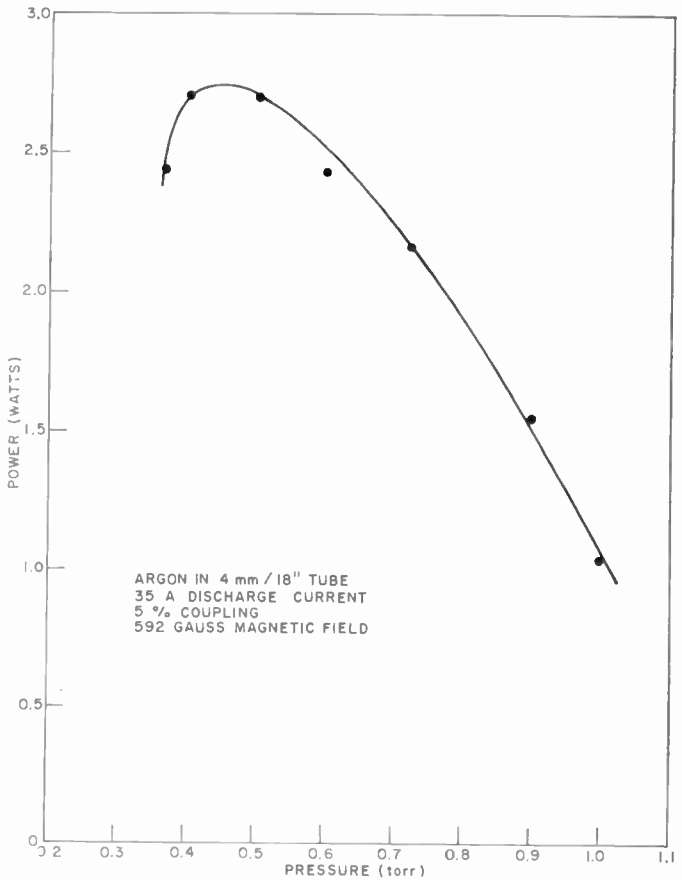

Fig. 10—Optical output power versus pressure for argon laser with 4-mm bore, 18-inches long.

When using the results of these measurements, one should remember that the output power is a strong function of the small unavoidable losses at the optical surfaces of the cavity, i.e. at the mirrors and the Brewster windows. These losses vary with the age of the mirrors, from tube to tube, and with the age of a tube. Hence, differences can easily occur when comparing different tubes, or even the results of measurements taken on the same tube at different times.

Nevertheless, the general functional dependence on the parameters should be valid, and, by exercising sufficient care, the differences can be kept within a few percent at high power (high gain) and close to optimum coupling.

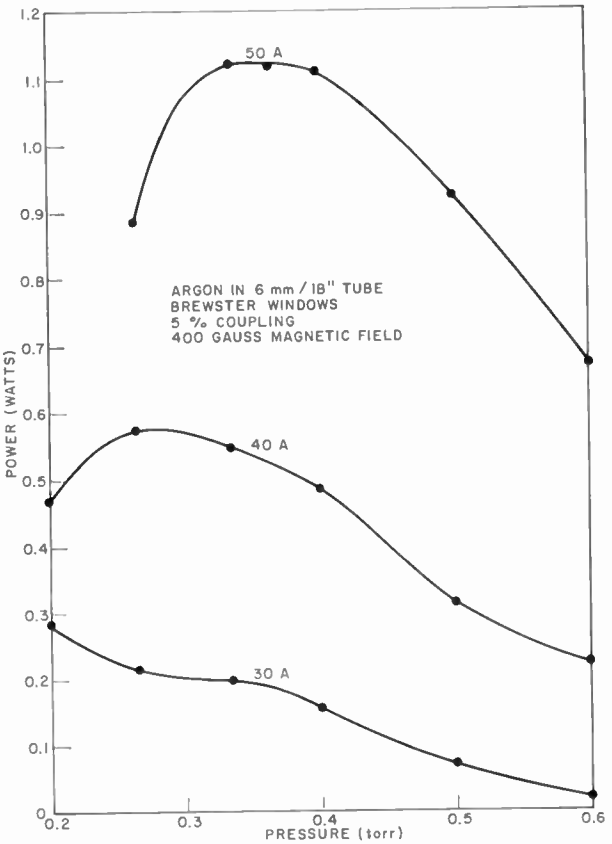

Fig. 11—Optical output power versus pressure for argon laser with 6-mm bore, 18-inches long.

A surprisingly large part of the short-term (50-hour running time) power degradation was found to be associated with the Brewster windows. Virtually complete recovery of the initial power was obtained by careful cleaning of the windows, which were assembled with demountable O-ring seals for this purpose. Such complete recovery was unexpected, because visual inspection indicated that the outer surfaces become much dirtier than the inner surfaces. Also, the outer surfaces were exposed to exactly the same environment as

the mirrors, which could not be cleaned. A reasonable explanation for the short-term degradation is that the windows become electrically charged and attract more dust than the uncharged mirrors.

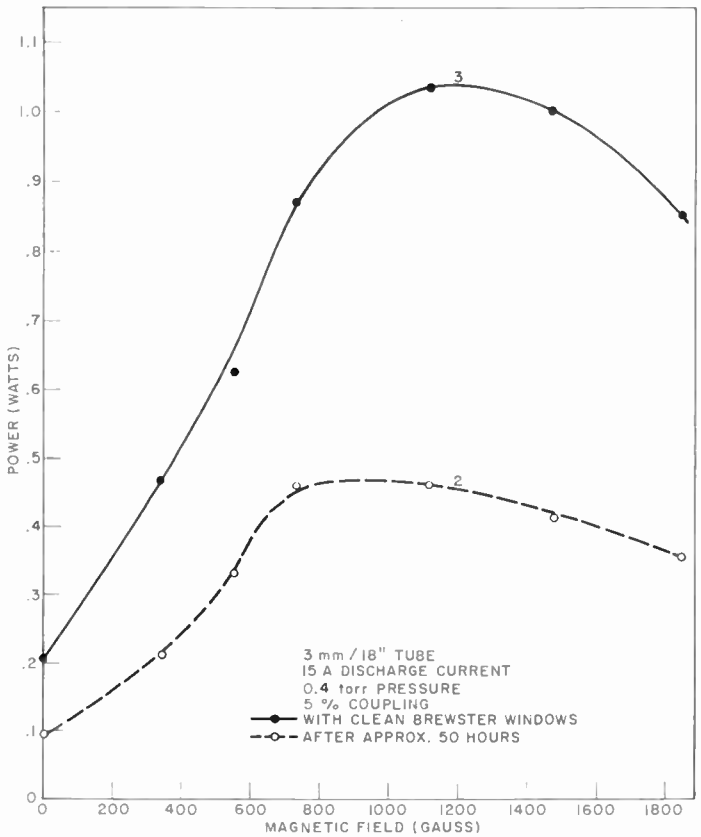

Fig. 12—Optical output power versus magnetic field measured after approximately 50 hours (curve 2) and then after Brewster windows have been cleaned (curve 3) (3-mm bore/18-inch-long tube).

The above remarks are illustrated by the curves in Figures 8 and 12. The dotted curves were taken after running the tube for approximately 50 hours, after which it was disassembled, the windows cleaned, the tube reassembled, and the measurements represented by the solid curves were performed. The numbers on the curves in Figures 8 and 12 (1, 2, 3 and 4) indicate the time sequence of these measurements.

Apart from giving quantitative information concerning the values of actual operating parameters and the available power output, Figures

8 through 15 also indicate the following important functional behavior.

(a) The optimum pressure depends on discharge current. As pointed out above in connection with the power versus dis-

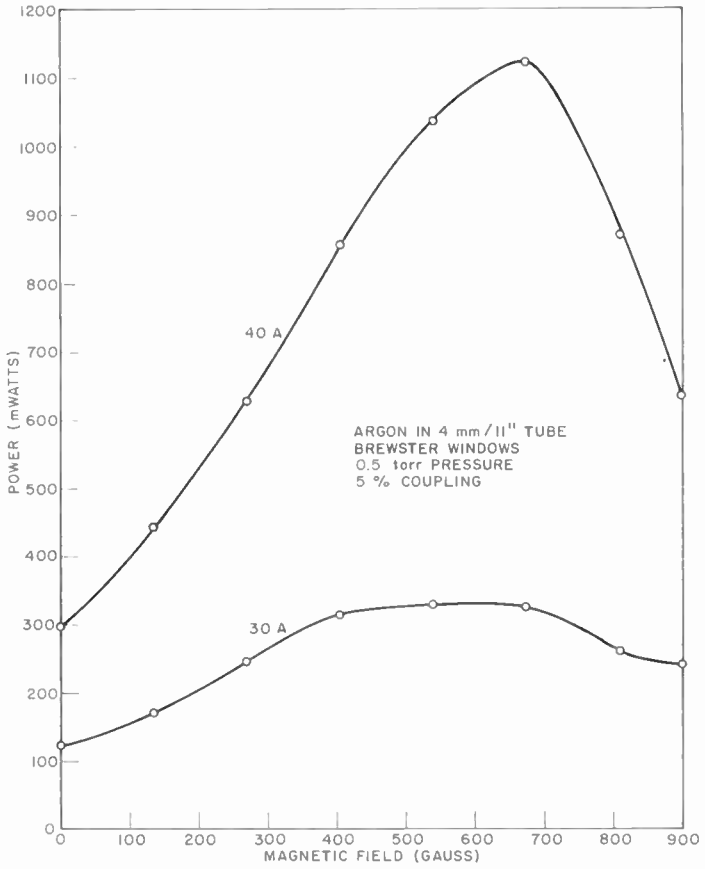

Fig. 13—Power versus magnetic field for argon laser with 4-mm bore, 11 inches long.

charge current curves, the higher the discharge current, the higher the optimum pressure required. This is clearly indicated by Figures 9 and 11, where pressure scans were taken at various current values.

(b) The optimum pressure depends on discharge-tube bore diameter. The larger the bore diameter, the smaller the optimum pressure. From Figure 10, the optimum pressure at 35 amperes for a 4-mm/18-inch tube is 0.45 torr; extrapolating

the data on Figure 11 indicates that at the same discharge current the optimum pressure for a 6-mm/18-inch tube is less than 0.3 torr.

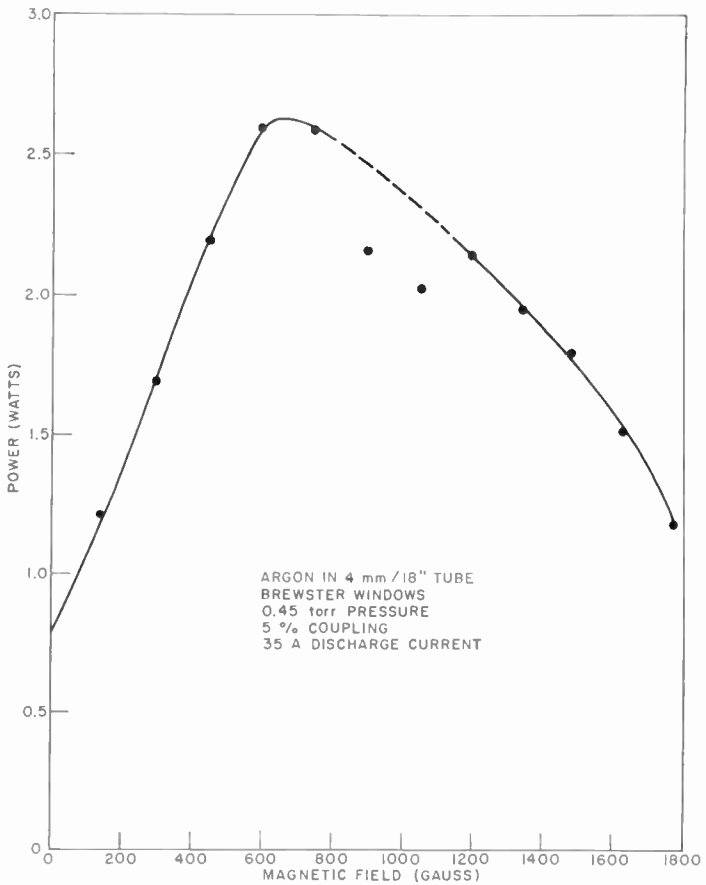

Fig. 14—Optical output power versus magnetic field for argon laser with 4-mm bore, 18 inches long.

(c) The optimum pressure depends on tube length. The longer the tube, the higher the optimum pressure. The optimum pressure at 40 amperes discharge current for a 4-mm/11-inch tube, as shown in Figure 9, is around 0.4 torr; in Table I, the corresponding value for the 4-mm/18-inch tube is 0.66 torr.

These three relationships confirm the suggestion that pump-out must play a dominant role in establishing the particle density in the

narrow-bore active section, and that the pressure value associated with that particle density and the gas temperature is significantly lower than the measured manifold pressure.

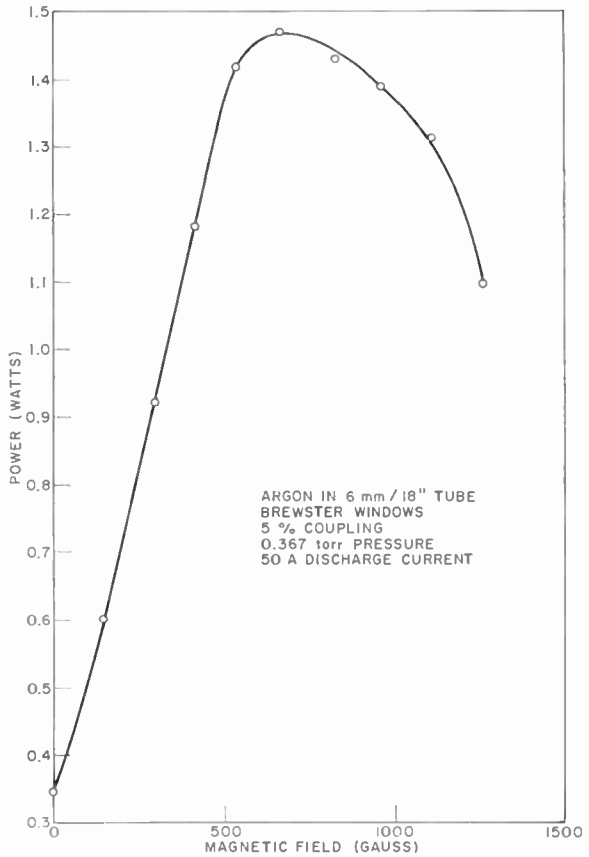

Fig. 15—Optical output power versus magnetic field for argon laser with 6-mm bore, 18 inches long.

(d) The optimum magnetic field value depends on discharge-tube bore diameter. For small-bore tubes, the smaller the bore diameter the larger the optimum magnetic field. Figure 12 shows that the optimum field value for a 3-mm/18-inch tube is approximately 1,200 gauss; Figure 14 shows that for a 4-mm/18-inch tube it is 700 gauss. Note, however, from Figure 15, that the difference in the optimum field value between the 4-mm and the 6-mm bore tubes is very small.

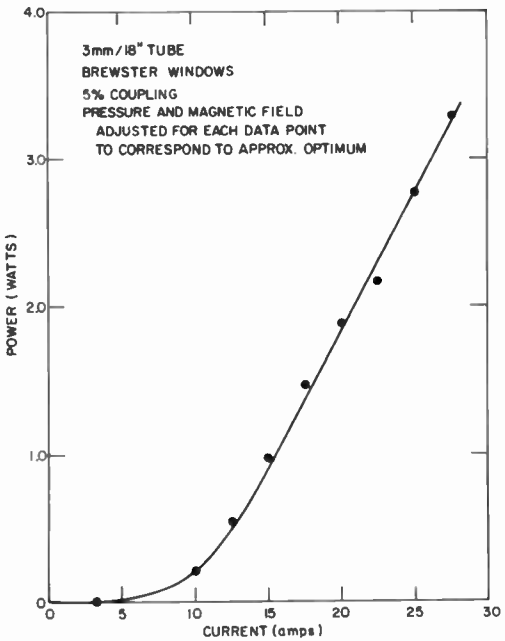

Fig. 16—Optical output power versus discharge current with pressure and magnetic field adjusted to approximately optimum for each data point (argon laser with 3-mm bore, 18 inches long).

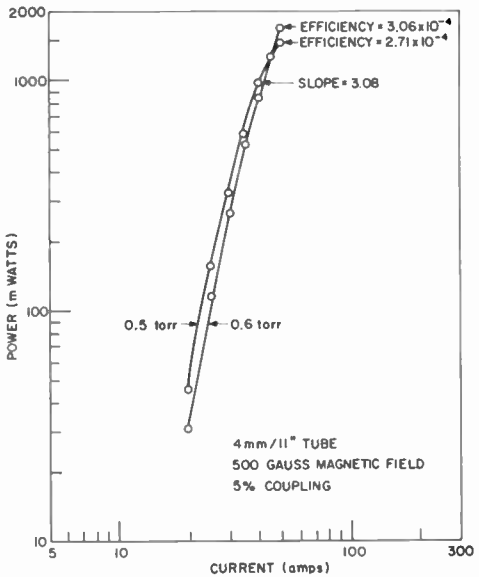

Fig. 17—Optical output power versus discharge current for argon laser with 4-mm bore, 11 inches long.

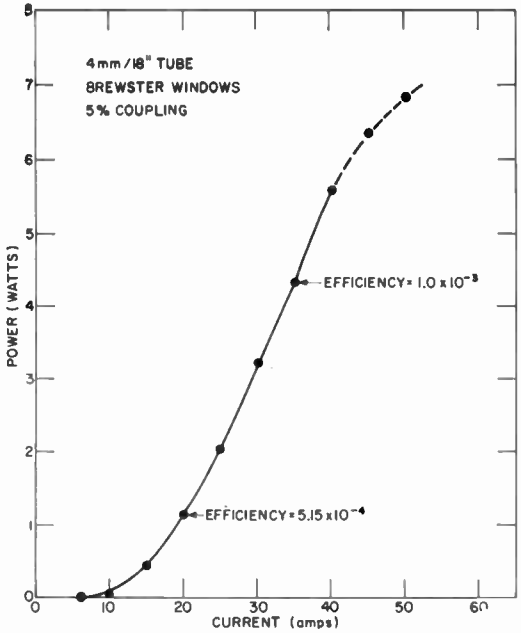

Fig. 18—Optical output power versus discharge current with pressure and magnetic field adjusted to approximately optimum at each data point (argon laser with 4-mm bore, 18 inches long).

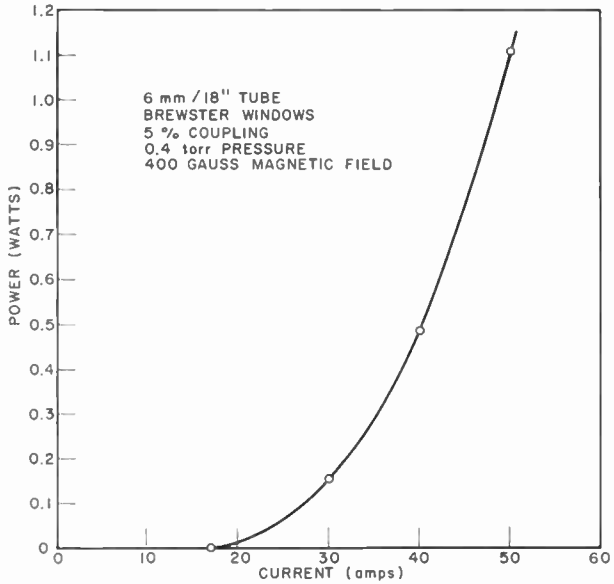

Fig. 19—Optical output power versus discharge current for argon laser with 6-mm bore, 18 inches long.

The strong dependence of the optimum field value on the bore size for small tube diameters suggests that in these cases the transverse ionic motion is collision free[7] (Tonks–Langmuir regime) and that a transition to collision-controlled motion[8] (Schottky regime) occurs around 6 mm diameter. This suggestion, however, should be considered only as speculation. The proper theoretical picture of high-current-density discharges ($\sim$ 500 amperes/cm^2) in magnetic fields is not known. From the agreement between the computed axial voltage gradient (based on an approximately known gas pressure and temperature) and the measured one, Hernqvist and Fendley[5] concluded that the Langmuir picture is the correct one, at least in the zero axial magnetic field case. Bickerton and von Engel,[9] however, point out that in discharges that show good agreement with the collision-free theory at no magnetic field, collisions can play an important role in an axial field of a few hundred gauss. Furthermore, all available positive column theories assume a low percentage ionization, and the partial pressure of the electrons is always ignored. Gudzenko et al[10] measured the electron density in an argon laser (2.6-mm bore/30-cm long) by adding a small amount of hydrogen and measuring the Stark broadening of the hydrogen lines. They found that the electron density at 400 amperes/cm^2 is 8×10^{13} cm^{-3}.* The collision-free theory predicts[5] $T_e = 5.2 \times 10^4$ °K. The electron partial pressure is then $p_e = n_e k T_e = 0.43$ torr, which is certainly not negligible in comparison to the neutral gas pressure. (Note that this comment holds even if the error in the density measurements is an order of magnitude.)

---

[7] L. Tonks and I. Langmuir, "A General Theory of the Plasma of an Arc," *Phys. Rev.*, Vol. 34, p. 876, Sept. 15, 1929.

[8] For discussion of the Schottky theory of the positive column in the absence of an axial magnetic field, see: J. D. Cobine, *Gaseous Conductors*, p. 236 Dover Publications, Inc., New York, 1958. A good account of the collision controlled positive column in an axial magnetic field is found in B. Lehnert, *Second United Nations Int. Conf. on the Peaceful Uses of Atomic Energy*, Vol. 32, p. 349, 1958.

[9] R. J. Bickerton and A. von Engel, "The Positive Column in a Longitudinal Magnetic Field," *Proc. Phys. Soc.*, Sect. B, Vol. 69, p. 468, April 1956.

[10] L. I. Gudzenko, V. N. Kolenikov, N. N. Sobolev, and L. A. Shelepin, "Spectroscopic Studies for Use in Ion Lasers," paper presented at the 1966 Int. Quantum Electronics Conf., Towne House, Phoenix, Ariz., April 1966.

* If the electron density is computed from the constants given in Reference (5) for a 2-mm-bore tube at 400 amperes/cm^2 current density, then $n_e = 10^{14}$ cm^{-3}.

MAGNETIC FIELD EFFECT ON ACOUSTROELECTRIC GAIN IN SEMICONDUCTORS

By

M. C. Steele

RCA Laboratories
Princeton, New Jersey

Summary—The effect of a magnetostatic field on piezoelectric acoustoelectric amplification has been calculated. It is found that the electron drift velocity needed to maximize the gain is significantly reduced by the magnetic field for high-mobility materials such as InSb. The results are discussed in terms of a simple magnetoresistance model. Finally, an explanation of the observed low-field microwave emission from InSb is offered in terms of the present calculation.

Introduction

THE DISCOVERY[1] of amplification of ultrasonic waves in cadmium sulfide has stimulated much further experimental and theoretical work in the field of acoustoelectric effects. In this paper, the effect of a magnetostatic field on such ultrasonic amplification is investigated theoretically. It is found that magnetic fields can greatly reduce the electron drift velocity needed to maximize the acoustic gain in semiconductors that have high electron mobilities, such as indium antimonide.

The Dispersion Equation

We consider the case of a piezoelectric solid of the III-V compound semiconductor type (e.g., InSb, GaAs). The geometrical configuration for the problem is given in Figure 1. It shows a shear wave, with wave vector $\vec{k}$, propagating in the z direction, along the <110> axis of the crystal, while the lattice displacement is along the x axis, which is a <100> axis. Under these conditions, there is a piezoelectric field in the direction of sound propagation, which is also taken parallel to the d-c drift velocity v_0 imparted to the electrons. Setting the magnetostatic field, B_0, in the x-z plane at an angle θ to the z-axis simplifies the problem without making the approach any less general.

[1] A. R. Hutson, J. H. McFee, and D. L. White, "Ultrasonic Amplification in CdS," *Phys. Rev. Letters*, Vol. 7, p. 237, 15 Sept. 1961.

The dispersion equation is then derived with the following assumptions:

(1) Only first-order perturbations of the form exp $\{i(\omega t - kz)\}$ are considered (small-signal approximation).
(2) Only slow waves of angular frequency ω are considered, so that $\nabla \times \mathbf{E}_{rf} \approx 0$ and the r-f magnetic field ≈ 0.

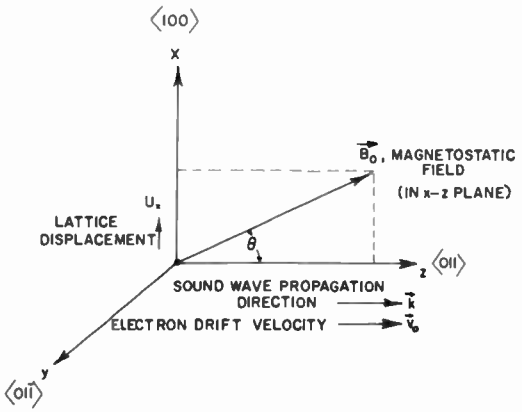

Fig. 1—Geometry of crystal orientation, sound-wave propagation, electron drift velocity and magnetostatic field used to derive the dispersion equation.

(3) The electrons have a constant drift velocity, v_o, and a constant scattering frequency, ν.
(4) The scattering frequency is taken to be much greater than the absolute value of the Doppler-shifted frequency, $\omega - kv_0$.
(5) The sound wavelength, λ, is much less than any dimension of the crystal sample.

With these assumptions the following dispersion equation is derived from Maxwell's equations, the piezoelectric relations, the sound-wave equation, and the electronic equations of motion:

$$(\omega^2 - k^2 v_s^2)\left[\frac{\omega_R}{\phi} + i\left(\omega - kv_0 - \frac{i D_n k^2}{\phi}\right)\right] - i K^2 k^2 v_s^2 \left(\omega - kv_0 - \frac{i D_n k^2}{\phi}\right) = 0 \qquad (1)$$

where

K^2 = piezoelectric electromechanical constant,
v_s = velocity of the sound shear wave,
D_n = electron diffsion coefficient (for $B_0 = 0$),
$\omega_R = \omega_p^2/\nu$ = dielectric relaxtion frequency,
ω_p = electron plasma frequency in the solid of dielectric constant ϵ,
$\phi = (1 + \omega_c^2\tau^2)/(1 + \omega_c^2\tau^2 \cos^2\theta)$,
ω_c = electron cyclotron frequency,
$\tau = \nu^{-1}$ = electron collision time.

Equation (1) reduces identically to that derived by White[2] when the dimensionless parameter $\phi = 1$. This occurs when either $B_0 = 0$ or $\theta = 0$. The other interesting property of ϕ is that it becomes independent of B_0 when $\omega_c\tau \gg 1$ (i.e., in strong magnetic fields). In that case

$$\lim_{\omega_c\tau \gg 1} \phi = 1/\cos^2\theta \ . \tag{2}$$

Amplification Coefficient

The solution of Equation (1) is obtained using the approximation

$$\frac{kv_s}{\omega} = 1 + i\alpha \ , \tag{3}$$

where α, the gain per radian, is $\ll 1$. The gain per cm is $\alpha\omega/v_s$. When $\alpha > 0$, the sound wave is amplified. Equations (1) and (3) lead to

$$\alpha \simeq \frac{\dfrac{-K^2}{2}\left(1 - \dfrac{v_0}{v_s}\right)\left(\dfrac{\omega_R}{\omega\phi}\right)}{\left(\dfrac{\omega_R}{\omega\phi}\right)^2\left(1 + \dfrac{\omega^2}{\omega_R\omega_D}\right)^2 + \left(1 - \dfrac{v_0}{v_s}\right)^2} , \tag{4}$$

where $\omega_D = v_s^2/D_n$ = diffusion frequency. We require $v_0 > v_s$ to have $\alpha > 0$. Equation (4) is identical to the equation derived by White[2] if $\phi = 1$. All the effects of the magnetic field appear in the parameter ϕ.

[2] D. L. White, "Amplification of Ultrasonic Waves in Piezoelectric Semiconductors," *Jour. Appl. Phys.*, Vol. 33, p. 2547, Aug. 1962.

Several aspects of Equation (4) are important.

(1) For fixed B_0, the maximum gain per cm, $\alpha\omega/v_s$, occurs at a frequency ω_m given by

$$\omega_m^2 = \left(\frac{\omega_R}{\phi}\right)(\omega_D\phi) = \omega_R\omega_D , \tag{5}$$

which is identical to White's[2] expression *without* a magnetic field. We shall show later in this paper that this result follows from simple magnetoresistance considerations for the special case of $\theta = \pi/2$.

(2) At the frequency of maximum gain ($\omega_m^2 = \omega_R\omega_D$), $\alpha\omega_m/v_s$ is maximized by the condition

$$\phi^2\left(1 - \frac{v_o}{v_s}\right)^2 = 4\left(\frac{\omega_R}{\omega_D}\right). \tag{6}$$

From Equations (4) and (6), the maximum gain per cm is

$$\left(\frac{\alpha\omega}{v_s}\right)_{max} = \frac{K^2}{8}\left(\frac{\omega_m}{v_s}\right). \tag{7}$$

With $B_0 = 0$, the maximum gain per cm at ω_m occurs when

$$\left(1 - \frac{v_o}{v_s}\right)^2 = 4\left(\frac{\omega_R}{\omega_D}\right), \tag{8}$$

and has exactly the same maximum value as Equation (7).

Effect on Critical Drift Velocity

The important magnetic field effect appears to be a significant lowering of the electron drift velocity needed to maximize the gain. Consider the case of n-type InSb at 77°K with parameters such as those used by Buchsbaum et al;[3] $n = 2 \times 10^{14}/\text{cm}^3$, $\mu = 2 \times 10^5$ cm^2/volt sec, $\nu = 5 \times 10^{11}$/sec, $\omega_p \simeq 1.3 \times 10^{12}$/sec, $\omega_R = 3.4 \times 10^{12}$/sec, $\omega_D \simeq 4.1 \times 10^7$/sec, $\omega_m = 1.2 \times 10^{10}$/sec, $v_s = 2.3 \times 10^5$ cm/sec. The orientation used by Buchsbaum et al was that of Figure 1. If Equation (8) were

[3] S. J. Buchsbaum, A. G. Chynoweth, and W. L. Feldmann, "Microwave Emission from Indium Antimonide," *Appl. Phys. Letters*, Vol. 6, p. 67, 15 Feb. 1965.

still valid, the drift velocity for maximum gain (for no magnetic field) would be $v_0 = 1.3 \times 10^8$ cm/sec. This electron drift velocity is sufficiently high to produce copious quantities of electron–hole pairs by band-gap impact ionization. Such avalanche breakdown would obscure the acoustoelectric effect completely. This situation would become even more pronounced if a higher initial carrier concentration were used. We suggest that this may explain the failure to observe any significant acoustoelectric effects in n-type InSb *without* applied magnetic fields.

Equation (8) however, is no longer valid for such high drift velocities, since our assumption

$$\nu \gg | \omega_m - k v_o | \tag{9}$$

is violated. To study the acoustoelectric effect in such a regime requires an extension of the problem to include the Doppler-shifted terms in the equations of electron motion. This has been done for the case of no magnetic field by Vural and Bloom[4] in a general way. However, to apply the results to the specific case of InSb would require the use of a computer.

Let us now consider the same InSb crystal in a magnetic field sufficiently strong to lower the optimum drift velocity to the point where Equation (9) becomes valid. For simplicity, we assume $\theta = \pi/2$. To make ν at least ten times greater than $| \omega_m - k v_o |$ requires that $| 1 - (v_o/v_s) | \leqq 4.2$. Equation (6), then, gives a *minimum* magnetic field (at $\theta = \pi/2$) of 6000 gauss. At fields higher than this, our calculation would be valid. In addition, at such fields, $v_o = 1.2 \times 10^6$ cm/sec, well below the value needed for avalanche breakdown. In summary, the application of a sufficiently strong magnetic field to high-mobility semiconductors allows the acoustoelectric effect to be observed in a regime where the simplified analysis is valid. Recent experiments by Ross and Bray[5] with InSb tend to confirm this conclusion.

We also suggest the possibility that the microwave emission observed by Buchsbaum et al[3] may be due to acoustoelectric amplification. In their experiments, the magnetic field would play a critical role by lowering the optimum drift velocity (needed for maximizing the acoustic gain) to values well below avalanche breakdown.

Magnetoresistance Explanation

We have noted that if $\phi = 1$, Equation (1) and (4) become identi-

[4] B. Vural and S. Bloom, "Streaming Instabilities in Solids and the Role of Collisions," *Trans. IEEE PTGED*, Vol. ED-13, p. 57, Jan. 1966.

[5] J. B. Ross and R. Bray, "Magnetic Enhancement of Acoustic Interaction in III-V Semiconductors," *Bull. Amer. Phys. Soc.*, Series II, Vol. 12, p. 119, Jan. 1967.

cal to the usual relations for the acoustoelectric effect in the absence of a magnetic field. It can be shown that our results are directly obtainable from those of White[2] if the following transformations are made in his equations:

$$\omega_R \rightarrow \omega_R' = \frac{\omega_R}{\phi}$$
$$\omega_D \rightarrow \omega_D' = \omega_D \phi \tag{10}$$

For the special case of $\theta = \pi/2$, $\phi = (1 + \omega_c^2\tau^2)$; and we can ascribe the transformations to a simple magnetoresistance effect. The relaxation frequency is proportional to the mobility μ while the diffusion frequency is proportional to $1/\mu$. If we note that the r-f mobility varies as $1/(1 + \omega_c^2\tau^2)$, all the results follow. This variation of the r-f mobility results from being able to disregard the effect of the r-f Hall field on the r-f magnetoresistance. That is assured by the assumption that λ is much less than the transverse dimensions of the crystal. When λ becomes of the order of the transverse dimension we expect the effect to be greatly diminished by the r-f Hall field.

Conclusions

The calculation has shown that a magnetostatic field can significantly reduce the electron drift velocity needed to maximize the acoustoelectric gain in high-mobility semiconductors. For InSb at 77°K, it is estimated that a reduction of at least a factor of 10 can result for magnetic fields of a few thousand gauss. We have explained the calculated results in terms of a simple magnetoresistance effect for the special case of a transverse magnetic field. Finally, we have suggested the possibility that the effect of the magnetic field on the drift velocity may explain the microwave emission observed in InSb.

Acknowledgment

The author wishes to thank B. B. Robinson, B. Vural, A. Rose, and R. H. Parmenter for many fruitful discussions and suggestions relating to this subject.

MICROWAVE POWER COUPLING TO THE HELICON MODE IN INDIUM ANTIMONIDE

By
G. A. Swartz

RCA Laboratories
Princeton, N. J.

Summary—Microwave power coupling from a waveguide to the helicon mode in a semiconductor is improved by the placement at the semiconductor surface of cones made from certain high-dielectric, low-loss materials. Aluminum oxide cones, which are positioned at the ends of a 2-mm long InSb cylinder doped with 10^{15} n-type impurities/cm³, increased K-band power transmission through the cylinder by 11 to 13 db. Cones fabricated from several other commercially available, high-dielectric materials altered the polarization of the incident microwave fields, and were therefore useless. Pure indium antimonide discs with 10^{14} n-type impurities/cm³, which are placed at each end of the 2mm long InSb cylinder, increase the power transmission through the cylinder when the applied magnetic field is less than 10 kilogauss. Above 10 kilogauss the discs decrease the power transmission. Relative phase-shift and transmission measurements through a block of pure InSb indicate a mode change at a 10-kilogaus field that can account for the transmission decrease in the pure-doped-pure system above 10 kilogauss.

Introduction

IN RECENT YEARS numerous investigations have been made, both experimental and theoretical, of mechanisms for the generation of microwave power in solid-state plasmas.[1-5] Devices that utilize such microwave generation mechanisms will become useful if methods for coupling the microwave power from the solid-state plasma are developed. In general, the coupling method will depend on whether power is coupled to the TEM or TM mode in the plasma. Although the microwave power may be generated in the TEM or TM mode, development

[1] D. Pines and R. Schrieffer, "Collective Behavior in Solid-State Plasmas," *Phys. Rev.*, Vol. 124, p. 1387, Dec. 1, 1961.

[2] J. Bok and P. Nozieres, "Instabilities of Transverse Waves in a Drifted Plasma," *Jour. Phys. Chem. Solids*, Vol. 24, p. 709, 1963.

[3] B. Vural and M. C. Steele, "Possible Two-Stream Instabilities of Drifted Electron-Hole Plasmas in Longitudinal Magnetic Fields," *Phys. Rev.*, Vol. 139, p. 300, 5 July 1965.

[4] R. D. Larrabee and W. A. Hicinbothem, Jr., "Observation of Microwave Emission from Indium Antimonide," *Proc. Symposium on Plasma Effects in Solids*, p. 181, Dunod, Paris, 1965.

[5] S. J. Buchsbaum, A. G. Chynoweth, and W. L. Feldmann, "Microwave Emission from Indium Antimonide," *Appl. Phys. Letters*, Vol. 6, p. 67, 15 Feb. 1965.

of a coupling mechanism for only one of the two modes may be sufficient because of the possibility of mode conversion in the solid. This paper describes an investigation of a method for coupling microwave power to the helicon mode in indium antimonide. The helicon wave propagates in a TEM mode along the magnetic field and is circularly polarized in the same sense as the electron rotation around the magnetic field lines.[6,7] Power coupling is accomplished by having the wave gradually alter its phase velocity from the velocity in the waveguide to that in the indium antimonide. This gradual alteration of the phase velocity is accomplished by placement in the waveguide of high-dielectric cones at the surface of the indium antimonide or by a gradual change in the free-carrier density of the indium antimonide.

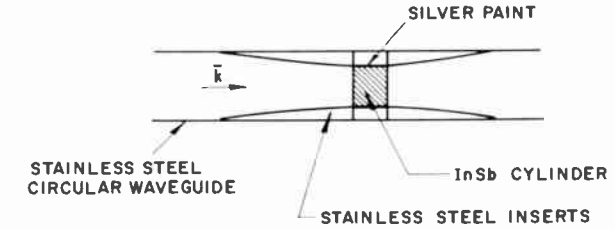

Fig. 1—Schematic drawing of InSb cylinder mounted in waveguide.

Experimental Apparatus

Power transmission and relative phase shift through a cylinder or block of indium antimonide were measured at K-band frequencies. The cylinder of InSb is mounted in circular stainless-steel waveguide by wedging it between two stainless-steel inserts, as shown in Figure 1. The inserts gradually narrow the waveguide from 0.5 to 0.375 inch, which is the diameter of the InSb cylinder. The block of InSb is soldered to a molybdenum diaphram that is held between two waveguide flanges. The arrangement is shown in Figure 2. Silver paint is applied to the periphery of the cylinder or block.

A solenoid with an inside diameter of 1.25 inches and a length of 3.5 inches is positioned coaxially with the guide, and the entire system is immersed in liquid nitrogen. A small positive pressure of helium gas is maintained in the waveguide to prevent seepage of the liquid nitrogen into the guide. The liquid-nitrogen-cooled solenoid is capable

[6] A. Libschaber and R. Veilex, "Wave Propagation in a Gyromagnetic Solid Conductor: Helicon Waves," *Phys. Rev.*, Vol. 127, p. 774, Aug. 1, 1962.
[7] J. Gremillet, "Electromagnetisme.—Propagation du mode 'helicon' dans les semiconducteurs en ondes metriques et decametriques," *Comptes Rendus*, Vol. 256, p. 2792, 25 March 1963.

of generating 6000 gauss for 10 seconds with a d-c supply or 27,000 gauss for 30 milliseconds with a 150-volt 0.1-farad capacitor bank. The waveguide is fabricated from stainless steel rather than copper to minimize the eddy currents caused by the pulsed magnetic field and to permit the current flow in the solenoid to represent the magnetic field in the sample independent of the time rate of change.

The wave incident on the semiconductor is circularly polarized by means of a thin polystyrene vane positioned between the rectangular-

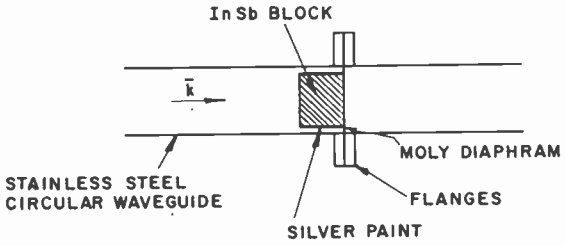

Fig. 2—Schematic drawing of InSb block mounted in waveguide.

to-circular guide transition and the sample. The vane is at a 45° angle to the electric field of the wave in the TE_{11} mode of the circular guide. Whether the wave is circularly polarized in the same or opposite sense as the helicon wave depends on the magnetic-field direction. The wave emerges from the semiconductor and is incident on a circular-to-rectangular guide transition. The detector receives only that portion of the wave that has its electric field vector parallel to the electric vector of the TE_{01} wave in the rectangular guide. Summation of the power received by the detector at two mutually perpendicular orientations of the rectangular guide yields the total power incident on the circular-to-rectangular transition.

The bridge circuit, which uses a four-probe line for simultaneously measuring the instant power transmission and phase shift,[8] is shown in Figure 3. The four-probe line is placed at a point in the circuit where a standing wave is formed by a reference signal from the klystron and the signal transmitted through the InSb sample. The four probes are positioned along the length of the waveguide, one-eighth wavelength apart. The signal from each of the probes travels equal distances to four crystal detectors. The crystal voltage from the first probe is subtracted from that of the third probe and the voltage

[8] F. J. Fitz Osborne, "A Multiple-Probe Microwave System for Plasma Studies," *Can. Jour. Phys.*, Vol. 40, p. 1620, 1962.

from the second probe is subtracted from that of the fourth probe. A plot on cartesian coordinates of the first and third probe difference-voltage as a function of the second and fourth probe difference-voltage gives a polar plot of the relative transmission through the sample plotted in the radial direction, and phase shift plotted in the azimuthal direction. A curve representing transmission versus phase shift is traced on an oscilloscope during a magnetic-field pulse. The oscilloscope is set to blank every millisecond beginning at the peak field. Since the magnetic field as a function of time is known, the transmission and phase shift as a function of magnetic field are also known.

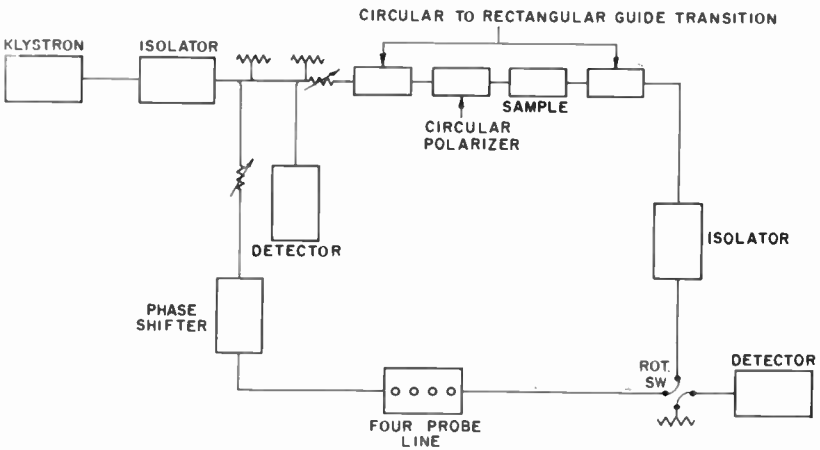

Fig. 3—Bridge circuit for measurement of phase shift through InSb sample.

Experimental Results

The power coupling to a helicon-wave mode in indium antimonide provided by cones fabricated from various high-dielectric materials was experimentally evaluated. A list of the materials and their pertinent microwave properties is given in Table I. The material that yields the most satisfactory results is a high-temperature aluminum oxide ceramic* with a dielectric constant of 9. The power transmitted at 23 GHz through two aluminum oxide cones 1.1 inches in height and placed base-to-base is 1 db below the incident power. The only other satisfactory material found was a mixture (50-50 by volume) of titanium dioxide powder (rutile) and epoxy.† The mixture had a dielectric constant of 9. The power transmitted through two base-to-

* High Temperature Alumina AV30, McDanel, Beaver Falls, Pa.
† Bipac BA-2101, Tra-Con Inc., Medford, Mass.

base epoxy-rutile cones 0.5 inch high was 3.4 db below the incident power (3 db attenuation and 0.4 db reflection). Commercially available, high-dielectric, low-loss materials such as Stycast and Eccoceram‡ altered the circular polarization of the incident wave. Since the incident wave must be circularly polarized in the same sense as the helicon wave for optimum coupling, any arbitrary alteration of the polarization by the cones cannot be tolerated.

Table I—List of Materials Used to Fabricate Cones and Their Pertinent Microwave Properties

Material	Dielectric	Loss Tangent	Comment
Aluminum oxide ceramic	9	~0.002	
Rutile-epoxy mixture	9	~0.01	
Stycast-high K500	25	<0.002	altered polarization
Stycast-high K500	15	<0.002	altered polarization
Eccoceram-high K	25	<0.001	altered polarization
Pressed rutile powder	90	~0.002	altered polarization

To measure the power coupling improvement caused by a high-dielectric cone at the surface of indium antimonide, the base of a cone was placed at each end of an InSb cylinder 2 mm long by 9.5 mm in diameter, and the measured power transmission was compared to that without the cones in position. An oscilloscope trace of the crystal voltage at the output detector as a function of magnetic field is shown in Figure 4 for transmission through the InSb cylinder without cones and in Figure 5 for transmission through the InSb cylinder with aluminum oxide cones at each end. When the transmission was measured, the isolator between sample and detector was removed. In the top curve of each figure, the circular polarization of the incident wave is in the same sense as the helicon wave; in the bottom curve of the magnetic field is reversed, so that the circular polarization is in the opposite sense to the helicon wave. One can see that the top curve indicates high transmission compared to the bottom curve. At 20 kilogauss the aluminum oxide cones improve the helicon mode transmission through the cylinder by 11 db and at 5 kilogauss the transmission is improved by 13.5 db. Table II is a tabulation of measured

‡ Emerson and Cuming, Inc., Canton, Mass.

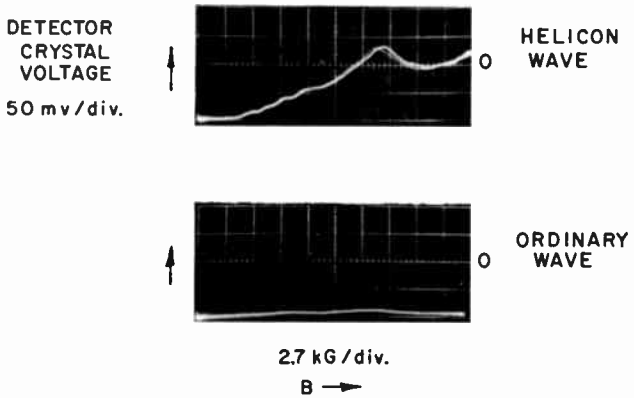

Fig. 4—Transmission through 2-mm InSb cylinder as a function of magnetic field. The incident wave is circularly polarized in the same (top curve) and opposite (bottom curve) sense as the helicon wave. Input power = 10 milliwatts.

and calculated transmissions through the InSb cylinder without cones, with Al_2O_3 cones, and with rutile-epoxy cones at 5, 10, and 20 kilogauss.

The equation for attenuation of the helicon wave and the equation for transmission through a flat interface formed by two dielectric materials are used to calculate the power transmission through the InSb-cone systems. The size resonances, which changed the transmis-

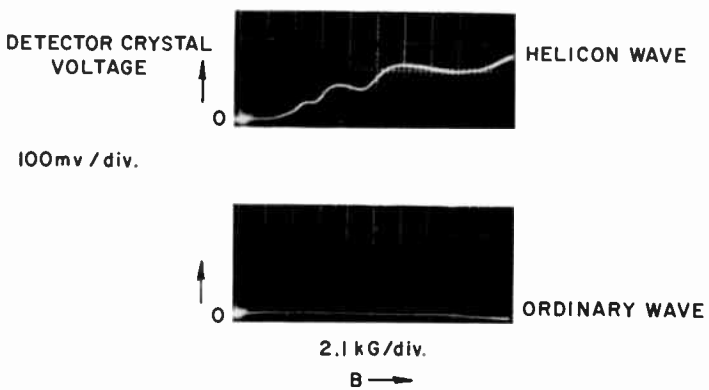

Fig. 5—Transmission through 2-mm InSb cylinder with aluminum oxide end cones as a function of magnetic field. The incident wave is circularly polarized in the same (top curve) and opposite (bottom curve) sense as the helicon wave. Input power = 2.0 milliwatts.

sion by 25%, or 1 db, are neglected. The attenuation, α, of the helicon wave is given by[6,7]

$$\alpha = \frac{\omega_p^2 \sqrt{\epsilon_L}}{2c\omega_c^2 \tau \left(1 + \frac{\omega_p^2}{\omega\, \omega_c}\right)^{1/2}}, \tag{1}$$

where ω = the signal frequency

c = the velocity of light

Table II—Measured and Calculated Transmissions through InSb Sample and Sample With Various Cone Configurations

Configuration	Transmitted Power (db below incident power)					
	$B = 5$ kG		$B = 10$ kG		$B = 20$ kG	
	cal	exp	cal	exp	cal	exp
2-mm sample only	30	26.6	19.3	21	14.1	16.8
2-mm sample with Al_2O_3	19.5	13.0	4.3	8.7	4.4	5.7
2-mm sample with a rutile-epoxy cone at each end	21.9	18.3	11.7	12	7.34	8.15
2-mm sample between 2 pure samples	25.7	23.5	14.9	21	10.4	27.5
Sample as above with a rutile-epoxy cone at each end	20.1	20.75	10.1	14	6.2	30

τ = the electron collision time

ϵ_L = the static dielectric constant of the material

ω_p = the electron plasma frequency

ω_c = the electron cyclotron frequency.

The plasma frequency is

$$\omega_p = \sqrt{\frac{ne^2}{\epsilon_L m^*}}, \tag{2}$$

where e is the electron charge, n is the electron density in the conduction band, and m^* is the electron effective mass. The electron cyclotron

frequency is

$$\omega_c = \frac{e}{m^*} B, \tag{3}$$

where B is the applied magnetic field. The transmission, T, of an electromagnetic wave from a lossless medium of dielectric ϵ_I through a flat interface to a lossless medium of dielectric ϵ_{II} is[9]

$$T = \frac{4\sqrt{\epsilon_I \epsilon_{II}}}{(\sqrt{\epsilon_I} + \sqrt{\epsilon_{II}})^2}. \tag{4}$$

The dielectric constant that is used in Equation (4) for the helicon wave in the semiconductor is[6,7]

$$\epsilon = \epsilon_L \left(1 + \frac{\omega_p^2}{\omega\, \omega_c} \right) . \tag{5}$$

For the calculation, the measured loss and dielectric constant of the cone is used. The calculated transmission is in agreement with the measured values except for the InSb-Al_2O_3 cone system at a 5 kilogauss field, where there is a 6.5 db discrepancy. With a 20-kilogauss field applied to the InSb-Al_2O_3, cone system, the transmitted power is 5.7 db below the incident power. The attenuation caused by power lost in the InSb is 1.5 db, and therefore the total attenuation caused by reflection and coupling loss is 4.2 db, or 2.1 db at each InSb surface.

A test of microwave power coupling to the helicon mode was made by a second method, which utilized a gradual change in ω_p (and, therefore, the wave phase velocity) as the wave traveled from the guide into the bulk InSb. A 1.25-mm thick disc of InSb with 10^{14} n-type impurities/cm^3 was placed at each end of the 2-mm InSb cylinder that contained 10^{15} n-type impurities/cm^3. Figure 6 shows the power transmission of the helicon wave (top curve) and ordinary wave (bottom curve) as a function of magnetic field. The measured and calculated transmission of the helicon wave through the pure-doped-pure (10^{14}/cm^3, 10^{15}/cm^3, 10^{14}/cm^3) InSb with and without rutile-epoxy end cones is given in Table II for 5-, 10-, and 20-kilogauss fields. The agreement between the measured and calculated transmission is good at 5 kilogauss, poor at 10 kilogauss, and extremely poor at 20 kilogauss. The pure InSb discs at each end of the doped InSb cylinder improve transmission by 3 db at 5 kilogauss and decrease transmission by 10 db at 20 kilogauss.

[9] G. Joos, *Theoretical Physics*, Chap. XVIII, p 352, Hafner Publishing Co., New York, 1950.

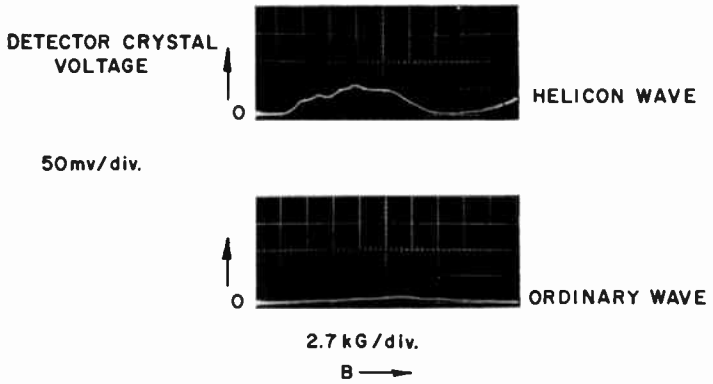

Fig. 6—Transmission through InSb cylinder with pure InSb discs at each end as a function of magnetic field. The incident wave is circularly polarized in the same (top curve) and opposite (bottom curve) sense as the helicon wave. Input power = 10 milliwatts.

An explanation for the poor agreement between experiment and theory for the pure-doped-pure InSb system at high fields is found in the transmission and phase-shift measurements through a block of pure InSb. The block is 8 mm × 7 mm in cross section and 4.8-mm long and is mounted in the K-band waveguide as shown in Figure 2. Figure 7 shows the power transmission of the helicon wave (top curve) and ordinary wave (bottom curve) as a function of magnetic field. At fields greater than 10 kilogauss, the ordinary wave is transmitted, and therefore a real dielectric constant is indicated. The dielectric

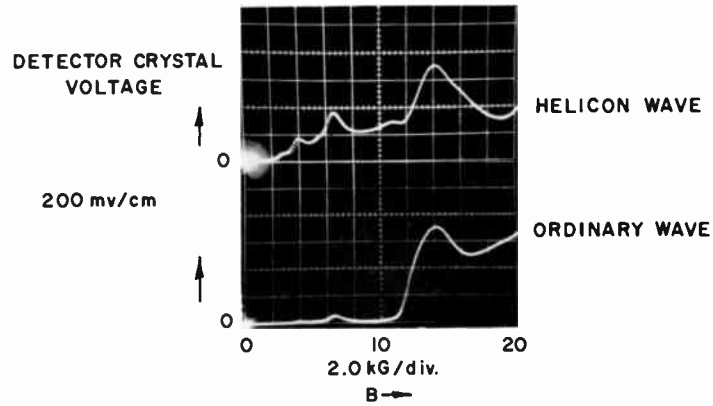

Fig. 7—Transmission through a block of pure InSb as a function of magnetic field. The incident wave is circularly polarized in the same (top curve) and opposite (bottom curve) sense as the helicon wave. Input power = 100 milliwatts.

constant for the ordinary wave is[6,7]

$$\epsilon = \epsilon_L \left(1 - \frac{\omega_p^2}{\omega\, \omega_c} \right) \tag{6}$$

and is real when $\omega_p^2/(\omega\, \omega_c) \leq 1$. With knowledge of the critical field (10 kilogauss) and the static dielectric constant (18.7), the charge carrier density is calculated to be 1.5×10^{14} electrons/cm^3.

Simultaneous phase shift and transmission measurements through the InSb block are made with the 4-probe line in the bridge circuit

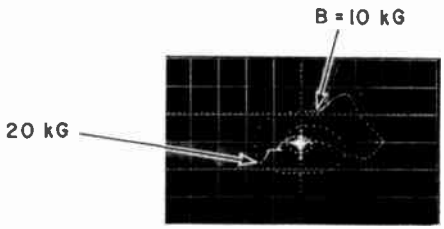

Fig. 8—Transmission in radial direction and phase shift in azimuthal direction through a block of pure InSb as magnetic field is varied from 20 kilogauss to zero. Scope trace is blanked every millisecond. Incident wave is polarized in same sense as helicon wave.

shown in Figure 3. Such measurements are shown in Figure 8 for the helicon wave; the phase shift is in the azimuthal direction and transmission in the radial direction. The magnetic field is varied from 20 kilogauss to 0 and the scope trace is blanked every millisecond, beginning at the peak of the magnetic field. At zero field, the scope trace is at the origin. Typical helicon transmission is characterized by the spiraling section of the curve. An abrupt change in the spiral nature of the curve is observed at a 10-kilogauss magnetic field. For comparision, the phase shift and transmission of the helicon wave through the 2-mm long doped InSb cylinder are shown in Figure 9. The magnetic field is varied from 27 kilogauss to 0 and the curve does not exhibit an abrupt change in the helicon spiral.

At a magnetic field greater than the critical field of 10 kilogauss, two modes of propagation are possible in the pure material. The data shown in Figure 8 indicates that the fast, or ordinary, wave mode is preferred to the slow, or helicon, wave mode even though the incident wave is circularly polarized in the same sense as the helicon wave. Mode conversion occurs, possibly at the InSb surface. Thus in the

pure-doped-pure InSb system at high fields, the incident wave, which is polarized in the same sense as the helicon wave, excites an ordinary wave in the pure material and the ordinary wave is then reflected at the interface of the doped and pure InSb.

Summary

A system for efficient coupling of microwave power from a waveguide to the helicon mode in a semiconductor has been demonstrated. Aluminum oxide cones placed at each end of a doped InSb cylinder improve the power transmission through the cylinder by 11 to 13 db. Replacement of the aluminum oxide cones by cones made from a low-loss, high-dielectric material, such as Stycast or Eccoceram, is not feasible, because these materials alter the polarization of the incident wave. Pure indium antimonide discs placed at each end of the doped InSb cylinder improve the power transmission when the magnetic field is such that only the helicon mode is allowed. At high magnetic fields the incident microwave power, which is circularly polarized in the same sense as the helicon wave, propagates through the pure InSb in the ordinary wave mode and is reflected from the doped InSb. The preference for the ordinary wave mode in pure InSb at high fields is indicated by phase-shift measurements in a block of pure InSb.

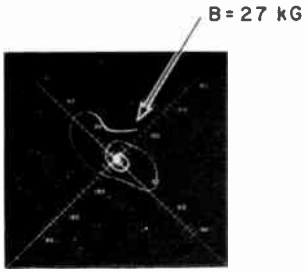

Fig. 9—Transmission in radial direction and phase shift in azimuthal direction through 2-mm cylinder of InSb as magnetic field is varied from 27 kilogauss to zero. Scope trace is blanked every millisecond. Incident wave is polarized in same sense as helicon wave.

Acknowledgement

I wish to thank M. C. Steele for suggesting the experiments reported here.

UV-SENSITIVE PHOTOEMITTERS*

By

A. H. Sommer

RCA Electronic Components and Devices
Princeton, N. J.

Summary—This paper combines a report of recent work on far-UV photoemitters with a general survey of the state of the art in this field. Sections II and III discuss UV-transmitting windows, in particular LiF, and the problems of response measurements below 2000 Å. The fabrication and properties of "solar blind" materials are discussed in Section IV, with emphasis on the alkali tellurides for the 4000 to 2000 Å range and the alkali halides and copper iodide for the range below 2000 Å.

I. Introduction

THE difference between visible and ultraviolet radiation is not of a fundamental nature but rather one that is arbitrarily created by the physiology of the human eye. Thus there is no basic difference between detectors for visible and for ultraviolet radiation, although detection does tend to become easier with increasing photon energy. In the specific case of photoemissive detectors, many considerations apply to both visible and UV radiation. There are, however, at least three special problems in the study of photoemitters for the UV that justify separate research in this field.

1. The first problem is a purely technological one. Glass, the standard material for the envelopes of photoelectric tubes used for the detection of visible and near infrared radiation, does not transmit UV radiation. The use of quartz and sapphire for wavelengths down to about 1800 Å requires special techniques that have not yet been adapted to the semi-mass-production methods applicable to phototubes made of glass. Below 1800 Å the best window material is lithium fluoride, which transmits down to 1050 Å; this material presents a number of difficulties that are discussed in this paper. Additional problems arise at wavelengths below 1050 Å where no window material is available.

* The work described in this paper was performed under NASA, Goddard Space Flight Center, Greenbelt, Md., Contract No. NAS-5-1337.

2. A second problem is that of measurement of the absolute or even of the relative sensitivity. In the wavelength range where an incandescent lamp has appreciable emission, measurements are relatively simple because tungsten lamps emit a continuous spectrum and have a very long life, so that one absolute calibration over the desired wavelength range allows accurate measurements to be made conveniently over long periods of time. Tungsten lamps have insufficient emission in the UV region, and one is forced to use mercury vapor lamps in the range where quartz transmits the radiation, and gas (e.g., hydrogen) discharge lamps below 2000 Å. These light sources tend to have strong lines superimposed on continuous spectra and to have a light output that fluctuates somewhat.

3. For most UV applications, the photoemissive material must fulfill one important requirement, namely, it must be "solar blind." In other words, it should not respond to the radiation emitted by the sun, which means it should be insensitive above approximately 3500 Å for use in the earth's atmosphere and above 2000 Å for use in space. Most of the work described in this paper is concerned with this problem. Without the "solar blind" requirement, there would be very little problem in making UV-sensitive cathodes, since, as is shown later, all photocathodes with good sensitivity to visible light have equal or higher sensitivity in the UV.

All three of these problems are a direct consequence of the high energy of the photons in the UV region. The window problem occurs because photons of higher energy can produce a larger variety of electronic transitions, thus leading to a greater probability of light absorption. As to light sources, the high photon energy of UV radiation cannot be obtained from a black body radiator because it would require temperatures above the melting or evaporation point of any practical material. Finally, the fact that all materials that are sensitive to visible light have high UV response is qualitatively obvious, because if a visible photon has sufficient energy to release an electron from a particular material, then a photon of higher energy should be at least as effective.

II. Windows for UV Radiation

For many years quartz has been the standard window material for the UV range from 4000 Å down to approximately 2000 Å. For wavelengths below 2000 Å, a number of alkali and alkaline-earth halides have been used successfully, especially CaF_2 and LiF. LiF is the material with the shortest known cutoff, 1050 Å. CaF_2, with a cutoff

at 1250 Å, has no basic advantages over LiF, although it is sometimes useful in the detection of weak signals in the range from 1250 to 2000 Å because it filters out the strong Lyman alpha line at 1216 Å. Our preliminary experiments indicated that CaF_2 has no advantage over LiF as regards either the mechanical problem of making vacuum-tight seals to glass or its chemical resistance to the ambient atmosphere. Therefore all our subsequent work was done with LiF because of its transmission at shorter wavelengths.

Sealing of LiF Windows

Most of the experiments with LiF windows were made with polished circular discs of 3/4-inch diameter and 1 mm thickness. They were sealed to glass envelopes by means of silver chloride cement, using essentially the method described by M. H. Greenblatt[1] for CaF_2. Windows sealed by this method can be made vacuum tight and they withstand temperatures up to 300° to 350°C. It should be emphasized that the silver chloride gives a successful seal only as long as it is virtually colorless. As soon as it begins to turn brownish, because of photochemical decomposition, it produces imperfect seals. Thus it is essential that the material be stored in the dark and replaced as soon as a color change is observed.

Optical Properties of LiF Windows

LiF windows have high transmission above 1600 Å (greater than 70% for one-mm thickness) and a sharp threshold at 1050 Å (determined by the energy band-gap of LiF). Between these limiting values transmission varies greatly from sample to sample. Our transmission measurements were made as follows. A simple tube was built in which the LiF window to be tested could be moved, by means of a magnet, in and out of the light path between the exit slit of a TROPEL vacuum monochromator and a CsI photocathode. Measurement of the ratio of the photocurrents obtained with and without the window in the light path provided a simple transmission measurement throughout the far UV spectrum without any need for absolute calibration of either the monochromator light source or of the CsI photocathode.

The results can be summarized as follows. Tests were performed to determine whether a window made of a cleaved LiF crystal had higher transmission than a polished window. The experiment did not indicate any significant difference. This is in agreement with the

---

[1] M. H. Greenblatt, "Sealing a Calcium Fluoride Window to Glass," *Rev. Sci. Instr.*, Vol. 29, p. 738, Aug. 1958.

finding by Patterson and Vaughan[2] that while freshly cleaved LiF windows have higher transmission, the water vapor in the atmosphere quickly reduces the transmission to a value comparable to that of polished LiF.

Two typical transmission curves obtained for polished LiF windows* are shown in Figure 1. For comparison, these curves are shown again

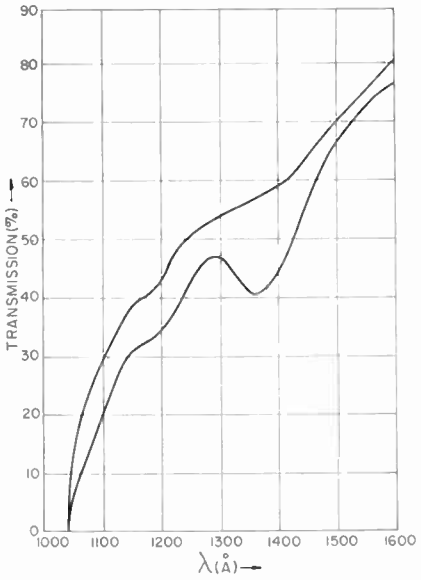

Fig. 1—Spectral transmission of two LiF windows.

as curves I and II in Figure 2 together with curves obtained with other materials and by other experimenters. Curve III refers to a window provided by L. Dunkelman of the Goddard Space Flight Center; curves IV and V refer to later improved samples from the same source. Curves VI and VII are derived from the paper by Patterson and Vaughan[2]; curve VII was measured with the same window as curve VI, but after exposure to water vapor. It is apparent that this exposure reduces transmission and also causes the structure seen in the transmission curve between 1250 and 1400 Å. This structure is also seen in our sample represented by curve II.

---

[2] D. A. Patterson and W. H. Vaughan, "Influence of Crystal Surface on the Optical Transmission of Lithium Fluoride in the Vacuum-Ultraviolet Spectrum," *Jour. Opt. Soc. Amer.*, Vol. 53, p. 851, July 1963.

* Windows were obtained from F. Cooke, North Brookfield, Mass.

The curves in Figure 2 show that the transmission of LiF below 1600 Å varies widely from sample to sample. Hence measurements made on photocathodes contained in tubes with LiF windows give ambiguous information about the sensitivity of the photocathode unless

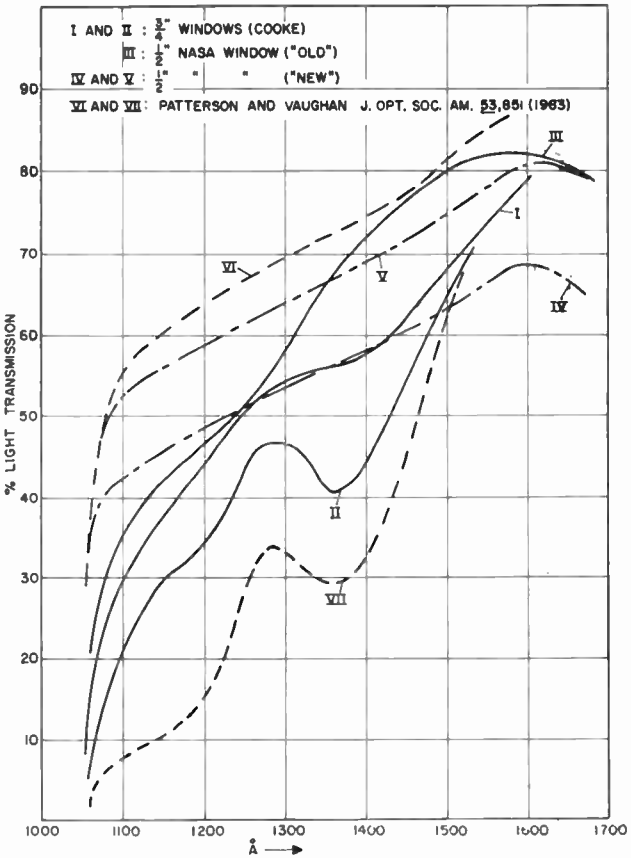

Fig. 2—Spectral transmission of LiF windows of different manufacture.

the transmission of the window is known. Even then an unavoidable uncertainty is introduced by later changes in the transmission due to the water vapor contained in the atmosphere. In space applications, an additional complication is added in that radiation damage may affect transmission.

It is worthwhile to note the great difference between the effects of water vapor and of liquid water on the characteristics of LiF windows. If the window comes in contact with liquid water, it immediately takes

on a cloudy appearance and becomes unusable. On the other hand, while water vapor has an immediate effect on the transmission of freshly cleaved LiF surfaces,[2] it seems to form a relatively inactive surface film with the result that the transmission decreases only very slowly during further exposure to the atmosphere. In a laboratory atmosphere the transmission remains almost constant over a period of months.

UV Transmission of Conducting Substrates

For many applications, in particular for image tubes and television camera tubes, it is desirable to use a semitransparent photocathode that is formed on the flat window of the tube. Since many of the photoemissive materials used in the far UV are very resistive, it is necessary to form the cathodes on substrates that combine conductivity with high UV transmission.

In the visible range, conducting tin oxide (TIC) is often used as a substrate having high light transmission. However, this material rapidly becomes absorbing at wavelengths below 4000 Å and is thus not useful in combination with LiF windows. The best alternative to TIC is a thin evaporated metal film. Our experiments indicated that tungsten films are satisfactory for this purpose. Earlier work had shown that tungsten films that pass approximately 85% of the incident light in the visible region have sufficient conductivity to act as a substrate for insulating photocathodes, but the transmission characteristics of these films in the far UV have not been reported in the literature. Therefore experiments were made in which tungsten was evaporated onto LiF windows until the transmission of visible light was reduced by a known amount. The windows with tungsten deposit were then measured for light transmission in the far UV. Figure 3 shows two curves obtained in this way. The set-up did not allow a transmission measurement of the LiF windows before tungsten evaporation, but a comparison with Figures 1 and 2 clearly indicates that the transmission of the tungsten film below 2000 Å cannot be appreciably lower than in the visible and that, therefore, such films are suitable substrates for semitransparent far-UV photocathodes.

Transmission below 1050 Å

The transmission of an insulating window material is directly related to its energy band gap, i.e., to the energy difference between the filled valence band and the empty conduction band. If the photon energy exceeds the band-gap energy, photons are absorbed; therefore transmission at wavelengths below 1050 Å, corresponding to a photon energy of approximately 12 eV, requires a window material having a

band-gap energy in excess of 12 eV. Since no material is known that has a band gap greater than that of LiF (approximately 12 eV), windowless phototubes must be used for the UV region below 1050 Å.

In this context it may be useful to discuss some practical aspects of windowless tubes. For space work there is no basic need for a

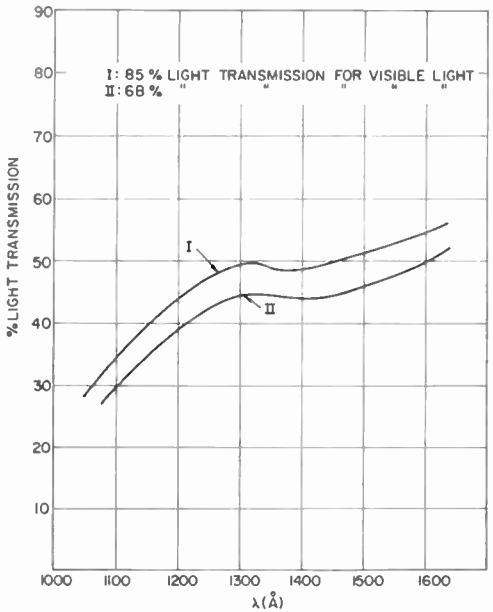

Fig. 3—Spectral transmission of thin tungsten films on LiF windows in the far UV.

window because the ambient is a high vacuum; moreover, some of the far-UV-sensitive materials, such as CsI (see Section IV), are reasonably stable in air, so that a window is not needed for chemical protection. However, in many types of tubes, the window is needed not only to maintain high vacuum and to produce the physical separation between photocathode and ambient, but also to serve as a support for the photocathode. This applies particularly to image and camera tubes where the image is focused onto a semitransparent cathode deposited on a flat window.

In the absence of a window the cathode must be used as an opaque cathode, i.e., the electrons are emitted from the side on which the radiation is incident. This causes a number of optical and electron optical design problems, particularly in a wavelength range where mirror optics have severe materials limitations. It is important to

bear in mind that for most space applications other than simple detection of photons, the design techniques for a windowless tube are quite different from those for tubes with a semitransparent photocathode.

III. Measurements of Photocathode Sensitivity in the Far UV

Absolute spectral response measurements of photocathodes in terms of photocurrent per unit of incident energy or of quantum efficiency (electrons per incident photon) are relatively simple in the visible and in the UV region down to about 2000 Å. The light sources that can be used in this range, such as incandescent lamps and gas discharge lamps, produce continuous or predominantly continuous light emission and are reasonably stable over long periods. Furthermore, the cathodes to be measured are contained in tubes with glass or quartz windows that transmit a large and constant fraction of the incident light. Finally, calibration by comparison with standard detectors or standard light sources is easy to perform.

In the UV region below 2000 Å measurements become more difficult for various reasons. In the 1000 to 2000 Å range, the commonly used light source is the hydrogen discharge. This source emits predominantly a line spectrum superimposed on a relatively weak continuum. When used in a conventional vacuum monochromator, the hydrogen discharge is produced in a part of the system that cannot be separated by a window from the section of the apparatus containing the grating and the exit slit. As a result, the whole system is exposed to air whenever the sample being measured is changed. Our work with a TROPEL vacuum monochromator showed that both the absolute and the relative amplitudes of the hydrogen emission lines vary from measurement to measurement. This effect is illustrated in Figure 4 which shows uncorrected spectral response curves taken on the same CsI cathode, but with the monochromator let down to air and re-evacuated between the two measurements. The large change in relative amplitude at the 1216 and 1608 Å lines is particularly striking.

The most likely explanation for these changes is that the relative excitation of the hydrogen lines is very sensitive to contamination in the gas itself and to changes in the nature of the surface film adsorbed on the walls of the discharge vessel. These surface-film changes are likely to affect the recombination of the hydrogen molecules. Absolute sensitivity measurements on photocathodes would, therefore, require a complete calibration of the light source for each cathode to be investigated.

Such a calibration is much more difficult than for visible light

sources, for two main reasons. First, a reliable standard detector is required. Metzger and Cook[3] have pointed out the difficulties in thermocouple measurements in the vacuum ultraviolet. The two calibration methods used most frequently in recent years are based on the use of ionization chambers[4] and the conversion of the UV radiation into visible light by means of sodium salicylate followed by a conventional photomultiplier.[5] However, uncertainties are associated with both of

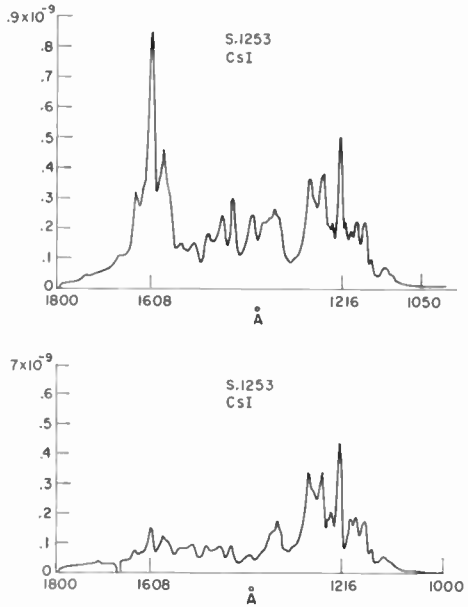

Fig. 4—Uncorrected spectral response curve of a CsI photocathode.

these methods. For instance, with ionization chambers, one must assume 100% conversion efficiency, and with the sodium salicylate, one must assume a conversion factor that is constant both with wavelength and in time.

The second difficulty in calibrating the vacuum UV light source is that, in order to avoid the uncertainty introduced by unknown or

---

[3] P. H. Metzger and G. R. Cook, "Flux Distribution of the Hopfield Helium Continuum from the Photoionization of Ar, Kr, and Xe," *Jour. Opt. Soc. Amer.*, Vol. 55, p. 516, May 1965.

[4] P. H. Metzger, "On the Quantum Efficiencies of Twenty Alkali Halides in the 12-21 eV Region," *Jour. Phys. Chem. Solids*, Vol. 26, p. 1879, 1965.

[5] P. D. Johnson, "Vacuum Ultraviolet Monochromator," *Jour. Opt. Soc. Amer.*, Vol. 42, p. 278, April 1952.

changing transmission of LiF windows in the far UV, the standard detector must be located within the vacuum chamber of the monochromator. This requires quite complex manipulations if a number of photocathode materials are to be measured within a relatively short time. Since we were mainly concerned with the relative sensitivities of different cathode materials in the far UV, the results discussed in Section IV are therefore based on relative and semiquantitative measurements obtained as follows.

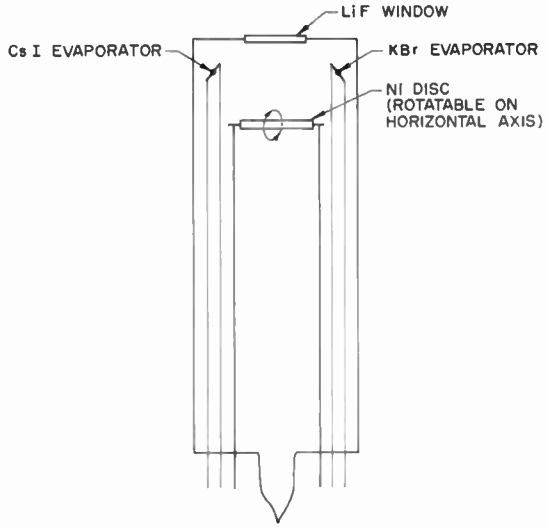

Fig. 5—Experimental tube for comparing spectral response of far-UV photocathodes.

For relative measurements, the simple tube shown diagrammatically in Figure 5 was used. This tube has a LiF window that can be brought into close contact with the LiF window attached to the exit slit of the vacuum monochromator. A one-inch-diameter nickel disc is mounted in such a way that either side can be made to face the window by means of a magnet. Cathode materials such as alkali halides can be evaporated on the disc, as indicated in the figure. By measuring the photocurrent as a function of wavelength for both sides of the disc, the relative sensitivity of two cathodes can be determined quite accurately, because any changes in the emission of the light source or in the transmission of the windows are the same for both cathodes.

For measurements below the cutoff of LiF, the front end of the tube shown in Figure 5 and the window on the monochromator exit

slit are removed and the tube is attached directly to the monochromator. Such measurements are, of course, only meaningful for cathode materials whose photoemissive characteristics are not affected by exposure to air.

Semiquantitative measurements were based on the assumption, supported by experience, that the sensitivity of CsI photocathodes to far UV is reproducible from sample to sample. Therefore, in most tests CsI was evaporated on one side of the disc and the unknown material on the other. For the absolute sensitivity of CsI, the values of Taft and Philipp[6] were accepted. They show a constant quantum efficiency of approximately 20% for photon energies in excess of 8 eV. Quite recently, other authors have published values indicating quantum efficiencies as high as 50%[4] and even 90%[7] at 11 eV. These discrepancies confirm the statement that absolute calibration in the far UV represents a major problem.

IV. Photocathode Materials for the UV

General Considerations

It was pointed out in Section I that the main practical interest in UV-sensitive photocathodes is limited to "solar blind" materials. The materials available for this spectral region fall into two main categories —metals and semiconductors (or insulators). The experimental work discussed here is concerned predominantly with the second group, but a few general statements about metallic photoemitters may be useful.

Metallic Photoemitters

In the visible range, metallic photoemitters have no practical importance for a number of reasons. First, the light absorption is low because of the high reflection of visible light. Second, the photoelectrons produced within the metal lose energy by electron scattering at such a rate that the escape depth, and thereby the quantum efficiency, is small. Third, the work function of most metals is above 3 eV, which means that the long-wavelength threshold lies below 4000 Å.

With decreasing wavelength, i.e., with increasing photon energy, the reflectivity of metals diminishes and the energy losses become less serious, because the photoelectrons have higher initial energy. More-

[6] E. A. Taft and H. R. Philipp, "Photoelectric Emission from the Valence Band of Some Alkali Halides," *Jour. Phys. Chem. Solids*, Vol. 3, p. 1, Nos. 1/2, 1957.

[7] W. F. Krolikowski and W. E. Spicer, to be published.

over, the work function of all metals is low enough to allow photoemission at wavelengths below approximately 2400 Å (5 eV). Thus, in principle, metallic photoemitters become feasible in the far UV. However, to ensure "solar blindness," metals with work functions greater than, say, 4 eV must be used, but the photoemission of metals increases so slowly with decreasing wavelength that such metals still have low quantum efficiency (below 1%) down to the LiF transmission limit. Below 1000 Å (above 12 eV) the sensitivity of metals increases sharply (see, for instance, the curves in Reference (8)), and, therefore, metallic photocathodes are useful in the region where windowless tubes must be used. There is some doubt, however, as to whether the high sensitivity of metals below 1000 Å is a property of the pure metal or is due to a thin surface layer of metal oxide or other contaminant. Most published results have been obtained under vacuum conditions where such surface films are likely to be present. In fact, a recent publication[9] indicates that metals such as Al and In that have been evaporated in ultra-high vacuum do not show the increased efficiency below 1000 Å until they have been exposed to air. Our own experiments were mainly concerned with tubes having LiF windows and therefore metallic cathodes were not investigated in detail, except for a few comparative measurements reported later.

Semiconducting and Insulating Photoemitters

It is not within the scope of this paper to discuss photoemission from semiconductors in any detail (for a recent review see Reference (10)). However, a brief summary of some of the basic considerations relevant to the experimental work reported later may be useful.

In terms of an energy-band model (see Figure 6), semiconductors and insulators are both characterized by a filled valence band separated from an (at 0°K) empty conduction band by the forbidden energy gap E_G. Semiconductors differ from insulators only by the value of E_G. If E_G is small enough at room temperature for some electrons to have enough thermal energy to be raised into the conduction band, the material is a semiconductor; if E_G is larger, the material is an insulator. Hence the difference is one of degree, with the borderline lying in the range between 2 eV and 2.5 eV. In the subsequent discussion, the two terms will therefore be used interchangeably.

[8] Chapter on "Photoelectric Emission from Solids," in *Handbuch der Physik*, Vol. 21, Pt. I (In English), Springer-Verlag, 1956.

[9] N. A. Axelrod, "Volume Photoelectric Effect in Metals," *Jour. Opt. Soc. Amer.*, Vol. 56, p. 203, Feb. 1966.

[10] Chapter on "Photoelectric Emission," by A. H. Sommer and W. E. Spicer, in *Photoelectronic Materials and Devices*, edited by S. Larach, Van Nostrand, Princeton, N. J., 1965.

Using again the simple band diagram of Figure 6, photoconductivity can occur in a material if the incident photon has an energy equal to or greater than E_G, i.e., if it has enough energy to raise an electron from the valence band to the conduction band. (Here we assume the idealized case of an intrinsic material; the effect of impurity levels is discussed later.) For photoemission, the electron requires the additional energy E_A (the electron affinity) to escape into the vacuum. Therefore the minimum photon energy for photoemission is $E_G + E_A$.

In reality, all materials contain some levels in the forbidden gap due either to impurities and/or to non-stoichiometric composition of the material. While these levels greatly affect the characteristics of photoconductors, their effect in photoemitters is relatively minor. This can readily be seen from the following consideration. Assume, for

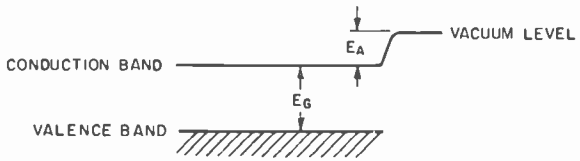

Fig. 6—Simplified energy-band model of semiconducting photoemitter.

example, that the material represented in Figure 6 is n-type doped and therefore has occupied donor levels in the forbidden gap close to the conduction band. Then, photons of relatively low energy, equal to the distance from the donor level to the vacuum level, are able to release electrons. This effect is bound to be very small, because the escape depth of electrons, determined by the energy losses of the electron as it moves toward the vacuum interface, is only of the order of a few hundred angstroms. The intrinsic material absorbs most of the incident light within this distance from the surface. However, the concentration of the donors is a very small fraction of that of the intrinsic material, with the result that only a very small fraction of the incident light is absorbed within the escape region. The photoemission from donor levels therefore has an effect only in the region of the long wavelength threshold. However, as shown later, it is important in the case of "solar blind" cathodes.

The preceding paragraphs were concerned mainly with factors determining the characteristics of photoemitters near the long-wavelength threshold. A second important property of cathode materials is the quantum efficiency in the wavelength range of peak sensitivity. In general, the efficiency tends to increase with photon energy and to

reach a substantially constant peak value above a certain photon energy, usually several electron volts above the threshold energy. However, the absolute magnitude of this peak value varies within wide limits for different cathode materials.

A qualitative prediction of the peak quantum efficiency of a material can be made if the values for band gap, E_G, and electron affinity, E_A,

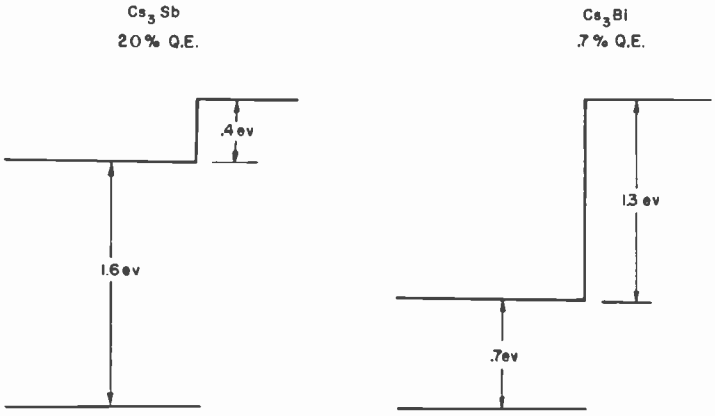

Fig. 7—Simplified energy-band models of Cs_3Sb and Cs_3Bi.

are known. If one assumes, for example, a photon energy equal to or only slightly above the value $E_G + E_A$, then two possibilities exist for a photoelectron released from the valence band. If E_A is larger than E_G, an electron raised to the conduction band can either use its remaining energy to escape into the vacuum or it can raise a second electron from the valence band into the conduction band, with the result that neither of the two electrons has sufficient energy to escape into the vacuum. On the other hand, if E_A is smaller than E_G, an electron raised to the conduction band can escape into the vacuum but does not have enough energy left for the "wasteful" release of a second electron from the valence band into the conduction band (for details see Reference (10)). In the first case, $E_G < E_A$, the probability of the electron producing a second electron–hole pair appears to be much larger than the probability of emission into vacuum. Therefore, this type of material tends to have much lower peak quantum efficiency. The situation is very convincingly illustrated by two examples of semiconducting photoemitters whose band models are shown in Figure 7. Cs_3Sb with $E_G > E_A$ has quantum efficiency as high as 20%, whereas the chemically similar compound Cs_3Bi with $E_G < E_A$ has a peak effi-

ciency approximately 30 times smaller. It will be shown later that this ratio of E_G to E_A is of considerable practical importance in the choice of UV-sensitive photocathodes.

Photocathode Materials

Because of the different techniques required for window materials as well as for measurements, it is useful to consider materials separately for the three ranges: 4000 to 2000 Å (quartz window), 2000 to 1050 Å (LiF window), and below 1050 Å (no window). The following discussion of materials for these three ranges is, for the most part, limited to solar-blind photocathodes.

Materials for the 4000 Å to 2000 Å Range

In contrast to the visible region of the spectrum, where very few photocathodes are available, photoemissive response has been observed in a large number of materials in the region below 4000 Å. Figure 8 represents a summary of some typical curves found in the literature.[11] It must be emphasized that these curves cannot be considered reproducible for different samples of the same material, since they depend

---

[11] The references from which the curves have been derived are:
CsI: E. A. Taft and H. R. Philipp, "Photoelectric Emission from the Valence Band of Some Alkali Halides," *Jour. Phys. Chem. Solids*, Vol. 3, p. 1, Nos. 1/2, 1957.
InSb (also GaAs and Bi_2Te_3): D. Haneman, "Photoelectric Emission and Work Functions of InSb, GaAs, Bi_2Te_3 and Germanium," *Jour. Phys. Chem. Solids*, Vol. 11, p. 205, Nos. 3/4, 1959.
Ni(1), HgTe: O. M. Sorokin, "Photoelectron Emission from Mercury Telluride," *Soviet Phys. Solid State*, Vol. 2, p. 2091, April 1961.
Ni(2): A. J. Blodgett, Jr. and W. E. Spicer, *Phys. Rev.*, Vol. 146, p. 390, 1966.
Ge, Te, Pt: L. Apker, E. Taft, and J. Dickey, "Photoelectric Emission and Contact Potentials of Semiconductors," *Phys. Rev.*, Vol. 74, p. 1462, Nov. 15, 1948.
Bi (also As and Sb): L. Apker, E. Taft, and J. Dickey, "Some Semimetallic Characteristics of the Photoelectric Emission from As, Sb, and Bi," *Phys. Rev.*, Vol. 76, p. 270, July 15, 1949.
PbS(1): R. A. Knapp, "Photoelectric Properties of Lead Sulfide in the Near and Vacuum Ultraviolet," *Phys. Rev.*, Vol. 132, p. 1891, 1 Dec. 1963.
PbS(2): R. M. Oman and M. J. Priolo, *Jour. Appl. Phys.*, Vol. 37, p. 524, 1966.
BaTe: R. J. Zollweg, "Optical Absorption and Photoemission of Barium and Strontium Oxides, Sulfides, Selenides, and Tellurides," *Phys. Rev.*, Vol. 111, p. 113, July 1, 1958.
BaO: H. R. Philipp, "Photoelectric Emission from Barium Oxide," *Phys. Rev.*, Vol. 107, p. 687, Aug. 1, 1957.
Na_3Sb: W. E. Spicer, "Photoemissive, Photoconductive, and Optical Absorption Studies of Alkali-Antimony Compounds," *Phys. Rev.*, Vol. 112, p. 114, Oct. 1, 1958.
K_2Te, Cs_2Te, Second Interim Report, ONR Contract Nonr-2263(00), 1958.

greatly on the surface conditions. In Figure 8, two examples are shown to illustrate the great discrepancies found in published response curves. By adopting ultra-high-vacuum techniques, the older curves for Ni and PbS (curves Ni(1) and PbS(1)) have been transformed in recent work to curves Ni(2) and PbS(2). The differences in both quantum yield and threshold wavelength are astonishing.

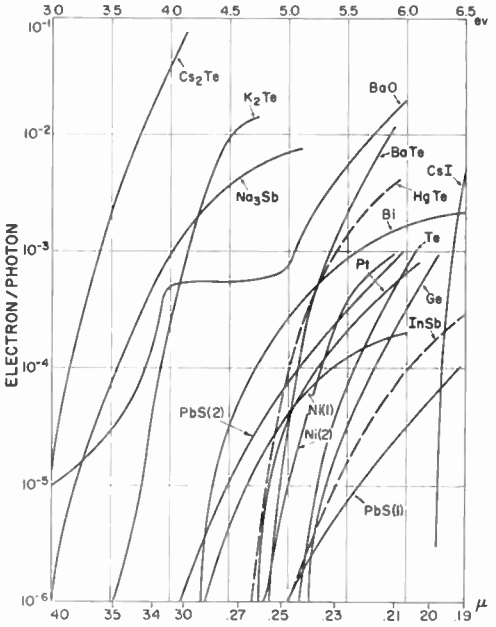

Fig. 8—Spectral response curves of photocathodes for the 2000 to 4000 Å region (see Ref. (11)).

(a) Cesium Telluride

Among the solar blind cathodes for the 2000 to 4000 Å region, the alkali tellurides have been the most successful for practical use. Cesium and rubidium telluride were studied by Taft and Apker[12] who established the most important characteristics of these cathode materials. Two typical response curves of Cs_2Te (taken from Reference (12)) are shown in Figure 9. They represent the two extremes that are obtained by the conventional method of exposing an evaporated layer of tellurium to Cs vapor. Curve I, despite its lower peak efficiency, is

[12] E. Taft and L. Apker, "Photoemission from Cesium and Rubidium Tellurides," *Jour. Opt. Soc. Amer.*, Vol. 43, p. 81, Feb. 1953.

obviously much more desirable for solar-blind applications than curve II, which has a shoulder extending well into the visible region of the spectrum. Unfortunately, a reproducible technique for obtaining cathodes with response like that shown in curve I has not been developed. It is generally assumed that the shoulder in curve II is caused by electron emission from donor levels that are attributed to an excess of Cs over the stoichiometric amount required for the compound Cs_3Te.

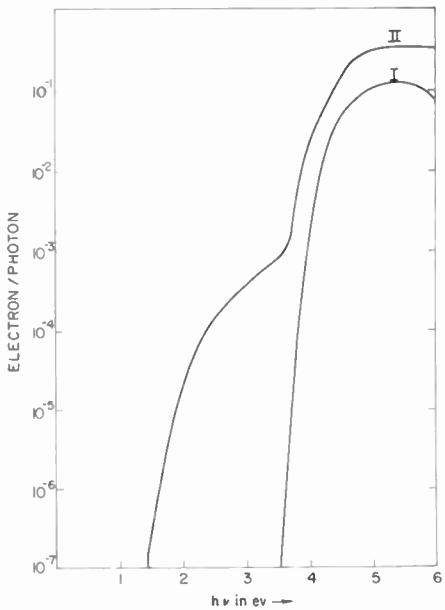

Fig. 9—Spectral response curves of Cs_3Te [from Taft and Apker, *Jour. Opt. Soc.*, Vol. 43, p. 81 (1953)].

Our work on Cs_3Te was concerned mainly with two problems—how to process semitransparent cathodes for use in photomultipliers, image tubes, etc., and how to reduce the long-wavelength response to a minimum. Regarding the first problem, the initial requirement is for a conducting transparent substrate, because of the very high resistance of Cs_3Te films. As discussed in Section II, evaporated tungsten films passing approximately 85% of the incident light have proved sufficiently conducting. In practical applications, another difficulty arises when the cathode is to be used in a tube that requires a very good vacuum during operation. It is common practice to subject such tubes to a degassing bake at 300 to 400°C. However, if the tube contains elementary tellurium as a source for evaporation of the tellurium film

that forms the base of the alkali telluride cathode, then the degassing temperature is limited to a much lower value because of the high vapor pressure of Te. This limitation can be overcome by evaporating the tellurium from a compound that decomposes at a temperature above 400°C and that contains as the second component an element of very low vapor pressure. It was found[13] that indium telluride fulfills these conditions. This compound decomposes in vacuum above 500°C, but the vapor pressure of indium at this temperature is still below 10^{-8} torr, so that only tellurium is evaporated.

Using an indium telluride bead as a source of tellurium, Cs_2Te cathodes were processed in experimental tubes with quartz or alkali fluoride windows. First, a layer of Te was evaporated onto a conducting substrate. For opaque cathodes, i.e., for radiation incident from the vacuum interface, the substrate was a metal disc (e.g., the first dynode of a photomultiplier). In this case the thickness of the Te film is not critical. For semitransparent cathodes, a thin tungsten film was used as a substrate and the thickness of the Te film was carefully controlled to obtain the best compromise between loss of photons by insufficient light absorption and loss of photoelectrons by electron release too far from the vacuum interface. The thickness was monitored by measuring the light transmission during the evaporation. It was found that Te films that transmit as much as 95% of incident visible light produce the best results.

After deposition of the Te film, the tube is heated to about 140°C and the Te is exposed to Cs vapor. During this part of the process, photoemission is monitored simultaneously under visible (tungsten lamp) and UV (mercury quartz lamp) radiation. Before the release of the Cs, the elementary Te has no sensitivity to visible light and only very low response to UV. During the addition of Cs, response to visible light is observed; subsequently the response to UV increases at a faster rate than that to visible light. The ratio of UV response to visible response reaches a maximum beyond which the response to visible light increases preferentially.

For a response as nearly "solar blind" as possible, the exposure to Cs vapor was always stopped as close to the peak ratio as possible. Even then, the long-wavelength response was unduly high. Since, as mentioned before, the long-wavelength response is attributed to excess Cs in the cathode, it appeared logical to eliminate this excess by evaporating additional amounts of Te until the exact stoichiometric two-to-

[13] A. H. Sommer, "Spectrally Selective (Solar Blind) UV Photomultipliers with Fluoride Windows," *Rev. Sci. Instr.*, Vol. 32, p. 356, March 1961.

one ratio is restored. This process has an effect in the desired direction, but it seems to be very difficult to achieve curves similar to curve I in Figure 9, either by compensating for excess Cs with additional Te or by stopping the initial addition of Cs before the stoichiometric excess is established.

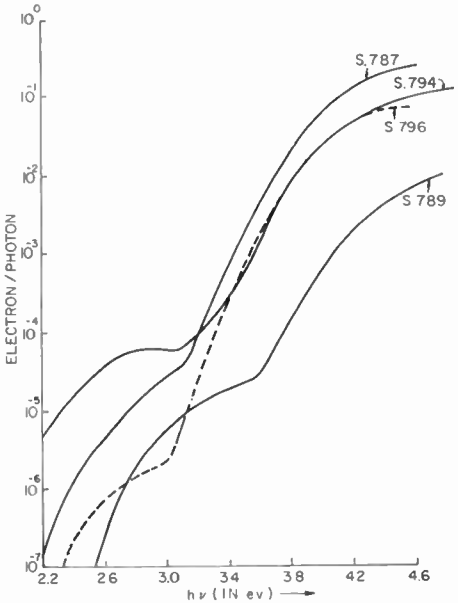

Fig. 10—Spectral response curves of typical Cs_2Te cathodes.

Typical response curves obtained with opaque Cs_2Te cathodes are shown in Figure 10. Figure 11 shows curves taken with semitransparent cathodes (including one Rb_2Te cathode). Two main conclusions can be drawn from the figures. First, the peak quantum efficiency is higher for opaque than for semitransparent cathodes (this is not surprising considering the light loss in the tungsten substrate and the limitations of semitransparent cathodes with regard to optimum thickness). Second, a great spread exists from sample to sample, not so much in peak quantum efficiency as in position and magnitude of the long-wavelength hump. In the best cathodes, this hump occurs only at a wavelength where the efficiency is three orders of magnitude below the peak value, but this is still detrimental in many practical applications because of the rapid rise in light emission from the sun in the same spectrum range.

hand, in multiplier or image tubes, which are operated at high voltages, the absence of Cs may also bring some advantages, such as reduced dark currents. Thus the K_2Te cathode may find use in some applications where low dark current is more important than high peak quantum efficiency.

(c) Other Cathodes for the 2000 Å to 4000 Å Range

Among the materials studied, sodium antimonide (Na_3Sb) was the

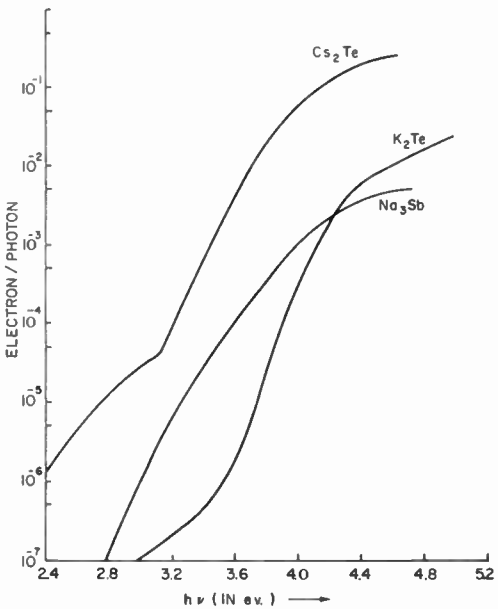

Fig. 13—Spectral response curves of Cs_2Te, K_2Te, and Na_3Sb.

only other more or less solar-blind cathode with appreciable response above 2000 Å (see Figure 8). For comparison, Figure 13 shows the response curves of the best Cs_2Te, K_2Te, and Na_3Sb cathodes that were made in this study. It is apparent that the Na_3Sb cathode has the least desirable properties in that it combines low peak quantum efficiency with a very slowly decaying response towards longer wavelengths. The low efficiency of Na_3Sb can be explained by the low ratio of band-gap energy (1 eV) to electron affinity (2.2 eV) of this material.

Materials for the 1050 Å to 2000 Å Range

(a) General Considerations

On the basis of earlier publications, experimental work on cathodes for the spectral region below 2000 Å was concentrated on the alkali

halides, previously investigated by Taft and Philipp,[6] and on copper iodide.[14-16] (Earlier summaries of far-UV detectors can be found in the papers by Dunkelman and his co-workers.[17])

Before discussing the specific properties of the materials investigated, we will review some facts that are relevant to all the photoemissive materials whose sensitivity is limited to the range below 2000 Å (photon energy above 6 eV).

(1) As stated earlier, materials with a photoemission threshold in the region of 6 eV should have a value of $E_G + E_A$ in excess of 6 eV. Because of the additional requirement for high quantum efficiency that $E_G > E_A$, materials such as germanium or cadmium sulfide that have a small band gap and large electron affinity are not likely to have high quantum efficiency, even at wavelengths far below 2000 Å. This rule has been proved experimentally with many materials of this type, but it cannot be considered to have universal validity.*

From the foregoing it is clear that suitable materials should have band gaps greater than 3 eV. Thus, the materials are insulators rather than semiconductors, and they are transparent for visible light because photons with energy less than 3 eV are not absorbed.

(2) Photoemitters with threshold wavelengths near or below 2000 Å have the interesting chemical property that they are not oxidized and thereby destroyed on exposure to oxygen or dry air. This is in sharp contrast to all photoemitters used for visible light, such as the alkali antimonides, that are only stable in high vacuum. Even the semi-solar-blind materials such as the alkali tellurides must be maintained in vacuum.

---

[14] J. Mattler, "Photoélectricité.—Une Cellule Photoélectrique Très Sensible Pour L'ultraviolet Extrême," *Comptes Rendus*, Vol. 217, p. 143, 2 Aug. 1943, and "Photoélectricité.—Rendement Quantique et Seuil Photoélectrique de L'iodure Cuivreux," p. 447, Nov. 1943.

[15] Yu. A. Shuba and I. V. Smirnova, "Photoemission of Electrons by Copper and Silver Iodides," *Soviet Phys. Solid State*, Vol. 2, p. 1201, Dec. 1960.

[16] A. M. Tyutikov and Yu. A. Shuba, "Measurement of Small Photocurrents in the Study of Photoelectric Emission," *Optics and Spectroscopy*, Vol. 9, p. 332, Nov. 1960.

[17] L. Dunkelman, W. B. Fowler, and J. Hennes, "Spectrally Selective Photodetectors for the Middle and Vacuum Ultraviolet," *Appl. Opt.*, Vol. 1, p. 695, Nov. 1962.

* I am indebted to W. E. Spicer for pointing out that the applicability of the rule depends on the type of band structure and on the scattering length for pair production as a function of energy. Thus, for instance, Spicer quotes the observation (N. B. Kindig and W. E. Spicer, *Phys. Rev.*, Vol. 138, A 561, 1965) that the quantum yield of cadmium sulfide, which has a high E_A:E_G ratio of greater than 2:1, exceeds 0.1 electron per photon at a photon energy of 11 eV.

In a qualitative way, the different behavior is not difficult to understand. The $E_G + E_A$ value of a material is a measure of the energy required to overcome the binding forces of an electron. If this energy is low, a physical consequence is that a photon of low energy can cause electron emission. The corresponding chemical consequence is that the "affinity" to the electronegative oxygen atom is enhanced. At present, not a single material is known that does not obey the empirical rule that all materials with $E_G + E_A$ greater than approximately 6 eV are insensitive to oxygen, and all materials with $E_G + E_A$ less than 6 eV are immediately oxidized. It must be emphasized that this rule does not apply to non-oxidizing chemical reactions, such as reaction with water vapor.

(3) All the investigated materials have been produced in the form of thin films deposited by evaporation of the material itself, as distinct from the cathodes with longer wavelength response that require a more complex activation process involving controlled reaction of a base layer with an alkali metal (alkali antimonides, alkali tellurides, etc.). Thus, the only variable parameter in this simple fabrication method is the thickness of the film. As with other cathodes, this parameter is very critical for semitransparent cathodes. In the case of these large-band-gap materials, it is also quite critical for opaque cathodes because the insulating films must not be so thick as to impede the supply of electrons. Generally, a thickness of several hundred angstroms gives satisfactory light absorption but does not introduce undesirable high-resistance effects.

In the experiments discussed below, the thickness of the evaporated cathode film was controlled by one of three methods. If the required thickness is known, it can be obtained by evaporating to completion a weighed bead of the cathode material from the appropriate distance. If the required thickness is not known, the sensitivity must be monitored during evaporation. This is accomplished by one of two methods. In the first, a small hydrogen discharge lamp with LiF window* is used (of course, care is taken that the LiF window of the lamp is in close contact with the LiF window of the phototube to avoid absorption of the radiation by air). As an alternative, the cathode material is evaporated while the cathode is exposed to a narrow part of the spectrum, usually the Lyman alpha line at 1216 Å, emitted by the vacuum monochromator. This second method is not only more specific, because any desired part of the spectrum can be chosen, but it also has the advan-

* Manufactured by G. B. L. Associates of Alexandria, Va.

tage that the two LiF windows can be eliminated by attaching the tube directly to the exit slit of the monochromator.

(b) Copper Iodide

The work on CuI was originally started because of the publication by Shuba,[15] which indicated a high quantum efficiency in the 10% range below 1650 Å and a threshold wavelength longer than that of the alkali halides. This combination would represent a desirable cathode material to bridge the gap between the alkali halides and tellurides.

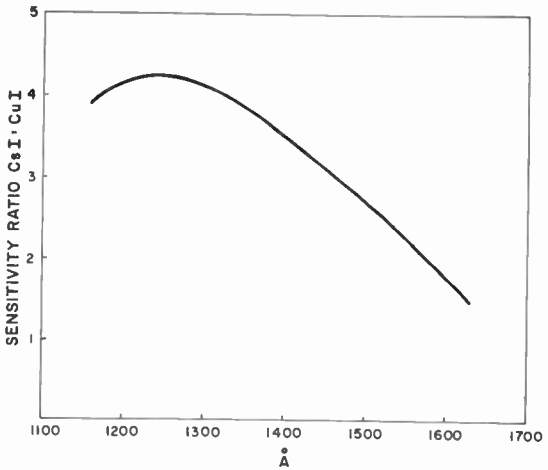

Fig. 14—Relative spectral response of CsI and CuI.

As was mentioned in Section III, we did not make absolute quantum efficiency measurements below 2000 Å, but relied on relative measurements using CsI cathodes as a standard and assuming a constant quantum efficiency of 20% for CsI at wavelengths below 1200 Å. If the higher values mentioned in Section III should be confirmed, the values for other materials quoted below would have to be corrected accordingly.

The following results were obtained with evaporated CuI cathodes:

(1) The threshold wavelength is similar to that of the alkali halides, i.e., close to 2000 Å.

(2) The quantum efficiency in the Lyman alpha range is at best a factor four below that of CsI (see Figure 14), corresponding to an upper limit of about 5%. Since a later paper by Shuba[16] refers to values in the neighborhood of 1% rather than the earlier 10%, our value between 1% and 10% does not seem unreasonable. Moreover,

light absorption data indicate that CuI has a band gap of 3 eV and a relatively large electron affinity of 3 eV, which qualitatively agrees with the lower efficiency in comparison with the alkali halides, which have a much larger E_G to E_A ratio (see below).

(3) Evaporated CuI layers invariably had higher conductivity than expected from such a large-band-gap material. The most likely explanation is that it may be impossible to maintain exact stoichiometry in these films, with the result that impurity levels are created by the excess of one or the other component. In phototubes with semitransparent cathodes, this conductivity proved sufficient to dispense with a conducting substrate.

(4) CuI cathodes appear to be relatively insensitive to short-term exposure to air, but there are indications that they are attacked by water vapor. However, systematic experiments concerned with chemical stability have not been made.

(c) Cesium Iodide

Within the uncertainty imposed by the varying light output of the TROPEL monochromator, discussed in Section III, it was found that evaporated CsI cathodes are quite reproducible in their performance and probably similar to those on which Taft and Philipp[6] have made absolute measurements (see Figure 15, taken from Reference (6)). A number of experiments with CsI cathodes are described below.

(1) Using the experimental tube shown in Figure 5, the far-UV photoemission of CsI was compared with that of a tungsten electrode. The result (see Figure 16) clearly shows the superiority of CsI throughout the investigated region of the spectrum. At the Lyman alpha line the sensitivity of CsI is more than a factor of ten higher and even below 1000 Å the factor is still over two.

(2) A similar experiment was performed with two stainless-steel electrodes. The results (see Figure 17) are similar to those obtained with tungsten. It should be mentioned that, even without any elaborate cleaning of the surface, the sensitivities of the two electrodes were almost identical, as can be seen from the measured points in Figure 17.

(3) Two experiments were made to investigate the effect of film thickness on the photoelectric characteristics. In the first, weighed beads of CsI were evaporated onto a tungsten electrode while the photoemission was monitored with the Lyman alpha line of the monochromator. During the evaporation the photoresponse increased until it became constant at a thickness at which most of the incident light is absorbed. From the weight of the bead and the geometry, it was

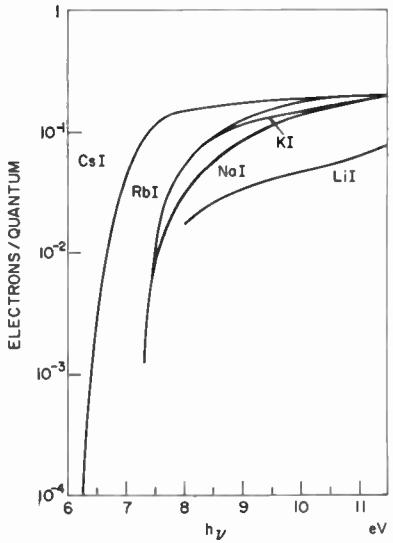

Fig. 15—Spectral response curves of alkali halides [from Taft and Philipp, *Jour. Phys. Chem. Solids*, Vol. 3, p. 1 (1957)].

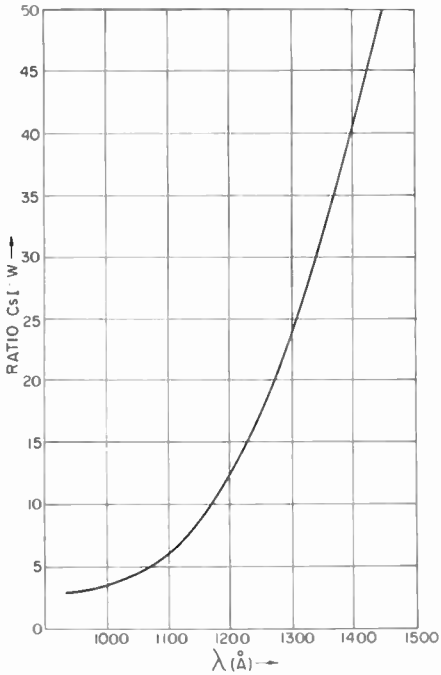

Fig. 16—Relative spectral response of CsI and W below 1500 Å.

estimated that the film was approximately 550 Å thick, corresponding to about 120 molecular layers of CsI. No high-resistance effects were observed at this thickness.

The second experiment was designed to examine the effect of thickness on the shape of the spectral response curve of a semitransparent CsI cathode. CsI was evaporated onto the LiF window of a tube, while

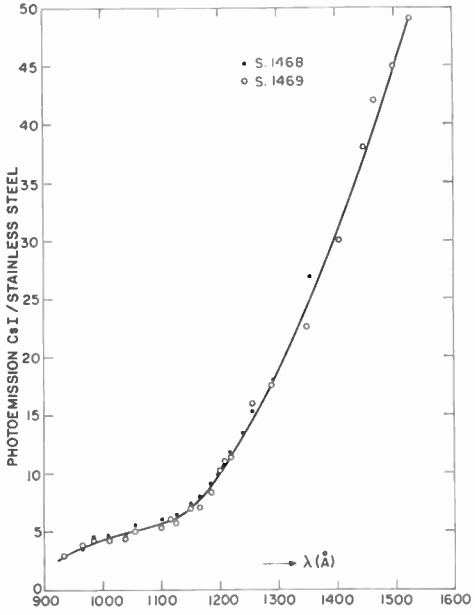

Fig. 17—Relative spectral response of CsI and stainless steel below 1600 Å.

the response to the Lyman alpha line was measured. As soon as an increase occurred, the evaporation was stopped and the sensitivity was measured at various wavelengths between 1125 Å and 1608 Å. This process was repeated until the sensitivity had gone through a maximum at all wavelengths. The effect of a total of eight evaporation steps is shown in Figure 18.

Two conclusions can be drawn from the curves. First, the peak response is reached at the same thickness of the CsI film for the whole wavelength range covered by the measurements, probably because the absorption is substantially constant in this range. Second, beyond the peak, the response drops faster at longer than at shorter wavelengths. This effect is to be expected, because the increased thickness reduces

the probability of escape for the slower photoelectrons more than for the faster electrons.

(4) The electron affinity of semiconductors can often be reduced and, thereby, quantum efficiency and threshold wavelength increased, by a surface treatment with cesium. It therefore seemed of interest to examine the effect of Cs on CsI, which has a band gap of 6 eV but an electron affinity of only a small fraction of one eV.[18] On exposure

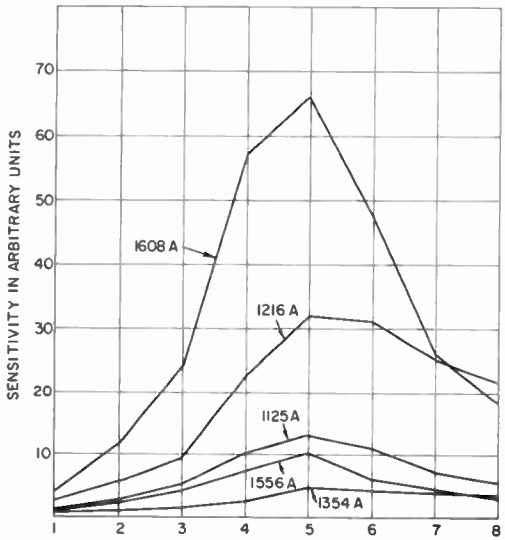

Fig. 18—Photoelectric sensitivity of a CsI film as a function of thickness (steps 1 to 8) at 5 wavelengths.

to Cs at room temperature, the sensitivity of a CsI cathode to the hydrogen discharge lamp did not change, but a small amount of sensitivity was obtained in the visible region when the cathode was illuminated with a 100-watt tungsten lamp. On removing excess Cs at 140°C the response to the UV lamp still remained constant, but the response to the tungsten lamp dropped sharply, although it never disappeared completely.

This result was expected. The Cs obviously cannot reduce the very low original electron affinity by an appreciable amount, but apparently enough Cs is retained to produce a very small amount of long-wavelength response that may have its origin in an adsorbed surface film of Cs or in a donor impurity produced by the Cs.

[18] H. R. Philipp and E. A. Taft, "Photoelectric Emission from the Valence Band of Cesium Iodide," *Jour. Phys. Chem. Solids*, Vol. 1, No. 3, p. 159, 1956.

(5) Many experiments were made to study the effect of air exposure on the photoemission of CsI cathodes. The results are somewhat ambiguous. In most cases, exposure to air had very slight effect, even when extended over weeks and months. In other cases, however, the sensitivity was reduced; in the worst cases, a previously clean deposit took on a milky appearance. The most likely explanation is a reaction with the water vapor in the ambient, which is bound to vary in concentration even in the air-conditioned laboratory atmosphere. As pointed out earlier, the binding forces of an electron in CsI are so large that the material cannot be affected by an oxidation reaction. However, CsI is so soluble in water, that a dissolving process rather than a chemical reaction is quite feasible. It is also possible that a small amount of CsI is hydrolized in a reaction similar to that suggested[2] for LiF.

From earlier work, particularly that of Taft and Philipp,[6] it appears that CsI is probably the best far-UV photoemitter, both for peak quantum efficiency and for relatively long threshold wavelength. However, some exploratory work was done on three other alkali halides, CsBr, RbI, and KBr, mainly because a cutoff at a shorter wavelength than that of CsI is desirable in some spectroscopic applications where the second-order diffraction of the grating limits the useful wavelength range to one octave. In such a case, a shorter threshold wavelength of photosensitivity permits measurements at correspondingly shorter wavelengths at the other end of the range.

CsBr—A comparison of CsBr with CsI is shown in Figure 19. In agreement with published curves,[6] the sensitivities are almost identical between 1100 and 1550 Å, but that of CsBr drops off sharply at longer wavelengths and is ten times lower than that of CsI at about 1700 Å.

The effect of exposure to air was studied in another experiment and the results are shown in Figure 20. The initial ratio of sensitivities is similar to that shown in Figure 19, except that the CsBr has a cutoff at a somewhat shorter wavelength. After exposure to air, however, the difference between the two materials is greatly reduced, indicating that the exposure tends to increase the threshold of CsBr. It is interesting to note that, after the initial effect, there is very little change during the following three weeks.

RbI—The behavior of RbI relative to CsI is shown in Figure 21. Again, before exposure to air, the sensitivity is similar to that of CsI with a sharp drop of the sensitivity above 1600 Å. On initial exposure to air, the long wavelength response increases and the difference be-

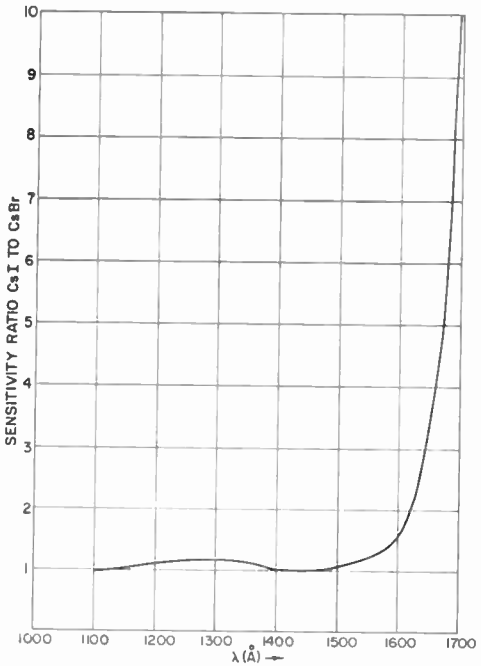

Fig. 19—Relative spectral response of CsI and CsBr.

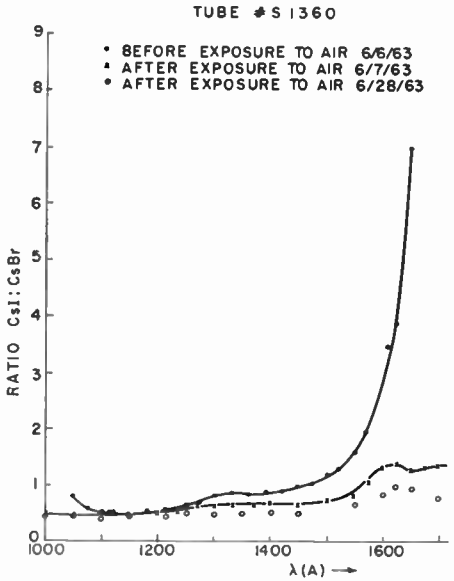

Fig. 20—Effect of air exposure on the spectral response of CsBr.

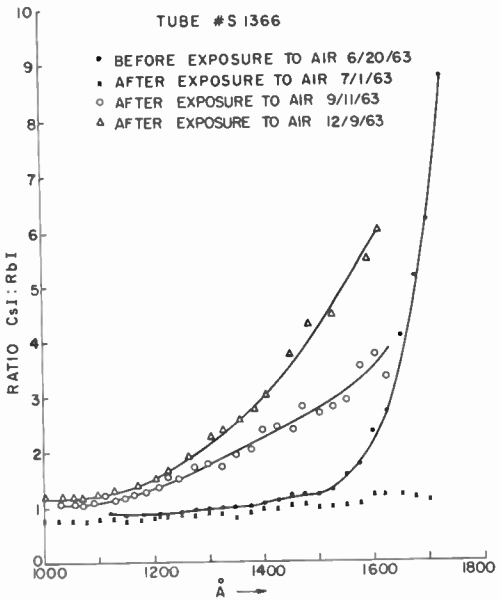

Fig. 21—Effect of air exposure on the spectral response of RbI.

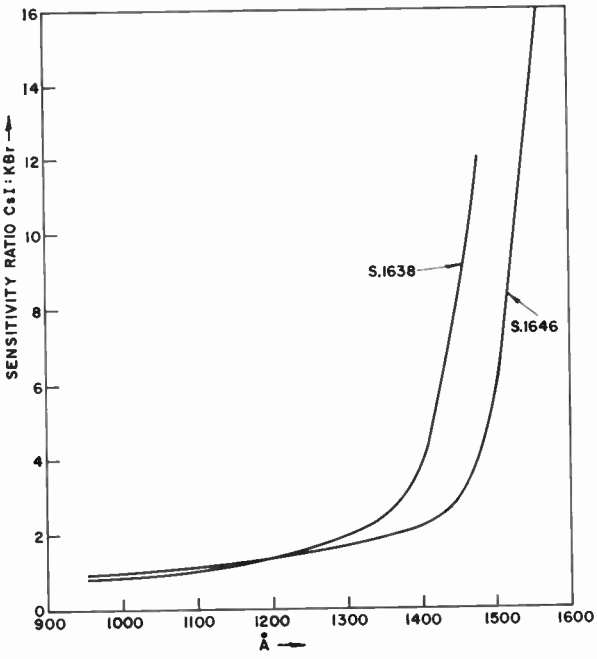

Fig. 22—Relative spectral response of CsI and KBr.

tween RbI and CsI is very slight. On longer exposure, however, the sensitivity tends to drop again and the loss of sensitivity extends to shorter wavelengths.

KBr—Measurements on two cathodes are shown in Figure 22. As with CsBr and RbI, the quantum efficiency is very similar to that of

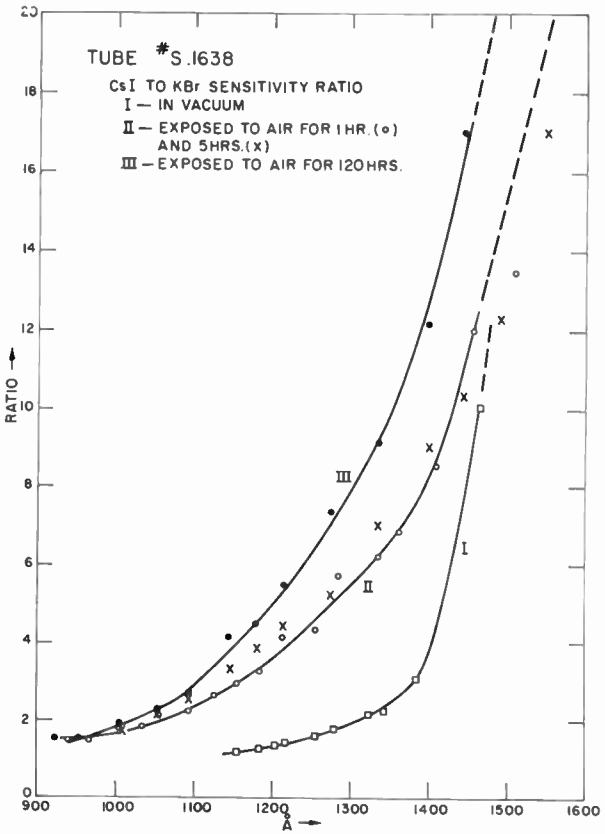

Fig. 23—Effect of air exposure on the spectral response of KBr.

CsI at wavelengths up to approximately 1300 Å. Above this value the two KBr cathodes differ slightly in that the cutoff of one is at about 1450 Å and that of the other is about 50 Å higher. On exposure to air (see Figure 23) the sensitivity of KBr seems to decrease throughout the measured spectral range within a relatively short period.

The experiments with CsBr, RbI, and KBr confirmed the findings of Taft and Philipp[6] that these materials have quantum efficiencies in

the same range as that of CsI at short wavelengths, but have shorter threshold wavelengths. The experiments with exposure to air were too few to allow definite conclusions, but the general tendency seems to be for these three alkali halides to be less stable than CsI. It also appears that the exposure to air can cause either a gain or a loss in sensitivity in the threshold region.

Materials for the Range below 1050 Å

For some applications a photoemitter that is blind to the Lyman alpha line at 1216 Å would be useful, because this line represents an unwanted background in some astronomical and related measurements in space. For a threshold at such a short wavelength, an emitter with $E_G + E_A$ greater than 10 eV is required; obviously, such materials can only be used in windowless tubes, because LiF does not transmit radiation below 1050 Å.

The two most likely materials for this range are MgO and LiF. The photoemission from MgO has been studied by Stevenson and Hensley[19] who found quantum efficiencies in excess of 10% at 1000 Å but also some sensitivity above 1200 Å. It is quite difficult to make thin films of MgO of controlled thickness; therefore, LiF seems to be a more promising material for the range below 1200 Å because, like the other alkali halides, it can be evaporated as a compound.

One experiment with LiF was made as follows. A tube of the type shown in Figure 5 was attached directly to the exit slit of the vacuum monochromator to eliminate the need for a window. Then, LiF was slowly evaporated onto a tungsten disc while the photosensitivities of both the coated and uncoated sides of the disc were monitored. Since tungsten itself is more sensitive at this wavelength than LiF, the sensitivity of the coated side decreased with continuing evaporation of LiF. The evaporation was stopped when the reading no longer changed with additional LiF, indicating that the film was thick enough to absorb all the incident radiation.

The results of the stepwise evaporation of LiF on the spectral response (in arbitrary units) are shown in Figure 24. It is apparent that with the addition of LiF the unwanted response above 1200 Å drops sharply relative to that of tungsten, while the efficiency below 1000 Å stays approximately the same. Unfortunately, it was found that on subsequent exposure to air, the response to longer wavelengths was greatly increased (see Figure 24), and the advantage of LiF as

[19] J. R. Stevenson and E. B. Hensley, "Thermionic and Photoelectric Emission from Magnesium Oxide," *Jour. Appl. Phys.*, Vol. 32, p. 166, Feb. 1961.

a Lyman alpha blind cathode over a metallic photoemitter, such as tungsten, becomes marginal. For the practical use, therefore, an LiF cathode film would have to be protected from exposure to air between deposition and use. The effect of air exposure on photoemission is probably basically the same as that on the light transmission of LiF windows, i.e., it is caused by superficial reaction with water vapor, as suggested by Patterson and Vaughan.[2]

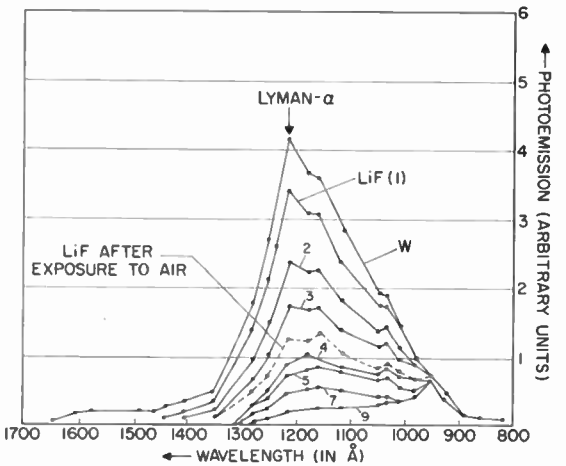

Fig. 24—Spectral response curves of LiF.

V. Conclusions Concerning Far-UV-Sensitive Photoemitters

The most important characteristics of presently available photocathodes for the far UV can be summarized as follows.

(1) All cathodes with high sensitivity to visible light, such as the S-11 and S-20, also have high efficiency in the UV.

(2) For solar-blind applications in the UV range down to 2000 Å, the best materials are cesium and rubidium telluride, but these are usually not completely insensitive to visible light. Potassium telluride is better in this respect, but at the expense of lower quantum efficiency at shorter wavelengths. All the alkali tellurides are drastically affected when exposed to air.

(3) For threshold wavelengths at or below 2000 Å, several alkali halides and copper iodide are useful. These materials are all more or less stable on exposure to air. The choice of the most suitable material depends on the requirements of the application.

The following are the most important considerations.

a. CsI has the longest threshold wavelength (about 2000 Å) of the alkali halides, and is probably the most stable on exposure to air.

b. CsBr, RbI, and KBr have peak quantum efficiency similar to that of CsI but shorter threshold wavelength. This is useful in applications where second-order diffractions interfere with measurements. CsBr, RbI, and KBr tend to be less stable in air than CsI.

c. CuI is inferior to the above alkali halides in peak quantum efficiency and in stability on exposure to air. On the other hand, it has sufficient conductivity to operate as a semitransparent cathode without the need for a conducting substrate. There is also some reason to believe that CuI is less liable to radiation damage than the alkali halides, but little information is available on this point.

(4) For Lyman alpha blind cathodes, LiF appears to be a promising material, but it does not retain its sharp threshold if exposed to air. Metals are also efficient cathodes below 1000 Å, but they all have residual sensitivity up to or above 2000 Å, because the work function of all metals is below 6 eV.

ACKNOWLEDGMENTS

H. E. McCandless of RCA Laboratories developed the lithium fluoride seals described in this report and has made invaluable contributions to the design of all our experimental tubes. The great interest shown in this work by L. Dunkelman of Goddard Space Flight Center is gratefully acknowledged.

SOURCES IN PLASMAS*

By

M. P. Bachynski

RCA Victor Company, Ltd.
Montreal, Canada

Summary—Currently, a strong interest exists in the behavior of radio-frequency sources embedded in an ionized medium. This interest arises from the fact that antennas are used aboard satellites and rockets for space exploration and therefore must perform in the natural plasma of the earth's environment. In addition, on-board radio-frequency systems are affected by induced regions of ionization such as those that surround re-entry vehicles or result from rocket exhausts, either on ascent or when used for descent. A further interest is in the application of r-f sources to measurements of the properties of laboratory plasmas.

This paper is concerned with a review of developments in this field since 1963. Considered in turn are the formulation of the boundary conditions for antennas in ionized media, the characteristics of various antenna configurations in unbounded plasmas, the nonlinear effects that result from the presence of the source in the plasma, and, in particular, the effect on the surrounding medium. Relevant experimental verification of the analytic predictions is presented wherever possible. Finally, areas of research and unsolved problems requiring urgent attention are singled out.

I—Introduction

MANY sources of electromagnetic energy exist. These may be antennas radiating either harmonic or aperiodic signals, individual charged particles or beams of charged particles, naturally generated electromagnetic energy, explosions, etc. When in a plasma, the electromagnetic fields of these sources have a marked influence on the free charges of the plasma. In addition, nonuniform regions or sheaths will develop between the material boundaries possessed by many sources and the plasma. Furthermore because of the thermal properties of the plasma, acoustic disturbances may also be generated. Thus a marked interaction can develop between a source of electromagnetic energy located in a plasma and the plasma.

A strong interest currently exists in the behavior of radio-frequency sources embedded in a plasma because of its direct relevance to space exploration, re-entry physics, and plasma diagnostics. This paper reviews work performed since 1963 with prime emphasis on radio-frequency antennas in unbounded plasmas and the types of environments in which they must operate for most practical applications.

The properties of an antenna in an unbounded plasma is determined not only by the configuration of the antenna and its current distri-

* Invited review paper presented in Session II on Electromagnetic Properties of Ionized Media at the XVth General Assembly of the International Scientific Radio Union, Munich, Germany, September 5-15, 1966.

bution but also by the nonuniform region, or sheath, that surrounds that antenna. For a source moving in a plasma, the sheath can arise and be modified not only by the usual difference in mobility between electrons and ions but also by the ion current intercepted, by the migration of the charges due to the r-f fields on the antenna, and, for long antennas, by induced potentials resulting from the antenna cutting across magnetic field lines. As will be shown, these effects make the accurate formulation of the boundary conditions and the analysis of the antenna properties a formidable task.

To obtain tractable solutions, most authors have assumed simplified boundary conditions and a variety of antenna configurations have been analysed on this basis. These include short dipoles, magnetic dipoles, cylindrical dipoles, line sources, spherical antennas, linear antennas, and biconical antennas. The interest has been primarily in the radiation and impedance characteristics of the antennas. As will be outlined, many of the analyses extend to anisotropic plasmas and thermal effects, thus showing the coupling of energy between electromagnetic and thermal or acoustic modes.

The near fields of antennas embedded in ionized media subject the surrounding plasma to very strong, nonlinear electromagnetic fields. This can result in plasma heating, breakdown, resonance, and stimulated phenomena as well as noncollisional dissipation of electromagnetic energy. This interaction gives rise to new observable effects as well as modifying the antenna characteristics. Progress that is beginning to be made in accounting for such behavior is indicated. Although most of the work on sources in ionized media has been analytic, some experiments both in the laboratory and in space have been conducted. Such results as are relevant to the theoretical developments are considered.

Finally, areas of research and unsolved problems requiring future attention are singled out.

It may be worthwhile to point out that most papers on antennas in plasmas appear in relatively few journals (in addition to proceedings of special conferences*), principally the *IEEE Transactions on Antennas and Propagation* and *Radio Science*, with a lesser number in *Canadian Journal of Physics* and *Proceedings of Institute of Electrical Engineers* (London). Numerous other journals publish an occasional article in this area, but these comprise less than 10-20% of the total.

---

* Noteworthy conference proceedings are: "Electromagnetic Theory and Antennas,"—Proc. of Symp. held at Copenhagen, June 1962, E. C. Jordan, editor, Pergamon Press, London, 1963, and "Electromagnetic Aspects of Hypersonic Flight,"—2nd Symp. on the Plasma Sheath, W. Rotman, H. K. Moore, R. Papa, editors, Spartan Books, Inc., Baltimore, 1964.

II—Boundary Conditions

The properties of a source such as an antenna in a plasma are determined not only by the configuration of the source (antenna) and its current distribution, but also by the nonuniform region, or sheath, that surrounds the antenna and marks the transition region from the antenna surface to the true plasma. The different physical processes that can give rise to such a sheath or transition region are stated and progress on understanding these effects quantitatively is indicated.

2.1 Normal Sheath

Any electrically floating body immersed in a plasma will come to equilibrium with its environment such that there is no net current to or from the body (i.e., the electrical charge on the body is constant). Since the mobility of the electrons in a thermalized plasma is much greater than that of the ions, the body will become negatively charged relative to the plasma, thereby erecting a potential barrier to reduce its electron collection to the same amount as the ion collection. There thus exists a difference between the "floating potential" of the body and the plasma potential. The potential barrier established results in a space-charge field in a region near the boundary between the plasma and the body. In this nonuniform region, charge neutrality is not achieved and, hence, a sheath region exists. The nature of the sheath will depend on the body material (conductor or insulator) and on the body geometry.

The rigorous formulation of the boundary conditions for a source in a plasma must include the properties of the sheath region. This is, in general, a formidable task, and workers, particularly those devoted to electromagnetic theory, ignore the plasma physics of the boundary conditions and resort to simplified boundary conditions. The electromagnetic boundary conditions on a metal or dielectric are well-known and are not considered here. Only the influence of the plasma sheath on the boundary conditions will be considered. The approaches taken to date can be classified as follows:

(1) The sheath region is neglected entirely (see Section 3), and the source is assumed to be in contact with a cold plasma with a sharp or "jump" boundary at the surface of the source at which neutral plasma exists. This approximation is appropriate for some idealized problems, but one is never certain of its quantitative influence.

(2) A dielectric layer is included between the source and the plasma, corresponding to a simplified model of an insulated antenna.[1] The dielectric layer is found to have an effect, particularly for radio

frequencies below the plasma frequency, where it is found that the radiation resistance remains finite and, in fact, decreases rapidly with size of insulation.

(3) The "rigidity" boundary condition for a warm plasma in which the assumption, borrowed from acoustics, is made that the normal component of the electron velocity in the plasma is zero at the interface.[2] Otherwise the source and plasma regions are separated by a sharp interface into two distinct homogeneous regions. Since it is known that charge particles do in fact cross the transition region and are absorbed by the source, this boundary condition is too restrictive physically and is inconsistent with the plasma physics of the problem.

(4) The "absorptive" boundary condition[3] assumes that nearly all electrons striking the surface are absorbed so that the electron density in the plasma is uniform nearly up to the source boundary. Physically, this occurs when the source is biased to a potential equal to the plasma potential. This boundary condition assumes a constant drain of electrons from the plasma and an infinite supply of electrons in the plasma so that its properties do not change.

(5) In the hybrid "admittance" boundary condition suggested by Cohen,[4] the normal component of the electron velocity, $\mathbf{v}$, (taking $\mathbf{z}$ as the unit vector normal to the boundary in question) is written

$$\mathbf{z}\cdot\mathbf{v} = Y_a \mathbf{z}\cdot\mathbf{E} + Y_b n_1 .$$

Y_a and Y_b are the admittances characterizing the transition, $\mathbf{E}$ is the electric field, and n_1 is the perturbation in electron density. The first term corresponds to a net transport of electrons due to a space-charge field in the sheath; the second term relates electron thermal properties. In addition, to ensure an electrically floating condition (no net charge transport), the condition

$$< (n_0 + n_1)\, \mathbf{v}\cdot\mathbf{z} > = 0$$

is necessary. Here, n_0 is the unperturbed value of electron density and the brackets signify a time average. Although these boundary conditions are a closer representation than the previous condition of what occurs at the boundary of a plasma with another medium, their direct application to antenna problems is not obvious.[5]

From the above approaches, it is clear that the application of accurate sheath conditions to "source" problems in plasmas has barely been attempted. In a way this is surprising, since a vast literature exists on sheaths that occur next to material bodies immersed in

plasmas. Solutions for potential profiles, ion currents, energy distributions, and floating potentials of boundaries for plane, cylindrical, and spherical geometries have been developed. Added complications and details have been considered, such as ion motion through the sheath, effect of mobility and fringing electric fields, and drift current. Much of the plasma physics literature on sheaths includes the kinetic theory (distribution function) approach, while the electromagnetic development has been based chiefly on the hydrodynamic or average particle equations.

2.2 R-F Sheath

A radiating source in a plasma, rather than consisting of an electrode at a fixed potential resulting in a sheath, as described above, generally consists of a large potential difference between two points (such as electrodes) that is varied at the radio frequency of the source. The near fields of an r-f source are large and hence have a profound influence on the resulting sheaths.

At radio frequencies that are low compared to the plasma frequencies, the results will be much the same as in the normal sheath, with a varying potential on the source resulting in a slow charging-up process. The source will initially collect a large electron current on the positive voltage part of the cycle, charging the source negatively until it is able to collect electron current only on the most positive part of the cycle. Ions will flow during most of the cycle.

At radio frequencies close to the plasma frequency or other characteristic frequencies of the plasma, resonance effects will occur. These are discussed in Section 4.3.

At radio frequencies that are high compared to the plasma frequency, the r-f field changes sign many times as the particles progress in the sheath region. The nonuniformity of the electric field (∇E) gives rise to a time averaged force on the electrons that results in a redistribution of the charges about the source under the influence of these r-f fields.

The theory of such time-averaged effects on charged particles in a-c fields has been summarized by Johnston.[6] More recent extensive work has been done at Saclay, France, in relation to plasma confinement schemes.[7] This gradient-field effect has been considered in connection with antenna impedance measurements,[8] rocket experiments,[9] and with heating of the ionosphere.[10] The r-f properties of sheaths have also been investigated. In particular, Butler and Kino[11] working with a laboratory plasma find that application of an r-f field results in

a thicker plasma sheath. In summary, although investigations of r-f sheath size and particle distribution have been made, almost no consideration has yet been paid to the influence of r-f sheaths on source radiation, source impedance, and nonlinear coupling phenomena.

2.3 "Ram" Sheath

A body moving with velocity V in a plasma such that its velocity is greater than the thermal velocity of the ions (V_t) and less than the electron thermal velocity (v_t) will collect ion or "ram" current by virtue of sweeping out (and collecting) the ions in its path. (This is the

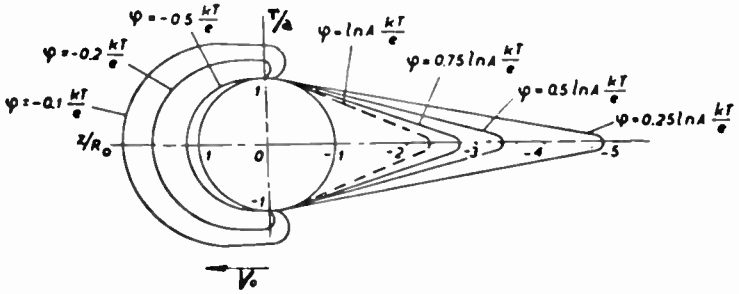

Fig. 1—Typical potential distribution in the sheath surrounding a perfectly reflecting spherical body.[15]

case for a satellite in the ionosphere.[12-14]) To a first approximation the ram current is given by

$$J_+ = neVA,$$

where n is the charge-particle number density in the plasma, e is the electron charge, and A is the cross-sectional area of the body normal to its velocity vector.

The ram ion collection modifies the normal sheath discussed earlier; it also leaves a rarefied region behind the body and a condensation region ahead of the body. This results in a complicated distribution of charged-particle concentration that can give rise to periodic variations in electron density distribution in space,[15] induced perturbations in electric and magnetic fields, longitudinal plasma (acoustic) waves, redistribution of charged particle energies, and instabilities. A typical distribution of the potential about a spherical body[15] is shown in Figure 1. The nature of the effects depends on whether the particles are specularly or diffusely reflected at the body or are absorbed.

Recent laboratory techniques using neutralized ion beams[16,17] have been applied to the simulation of the interaction between a spacecraft, such as a satellite in the ionosphere, and its local environment. Such investigations give an indication of the plasma sheath formed on a satellite whose velocity is greater than the ion velocity and the in-

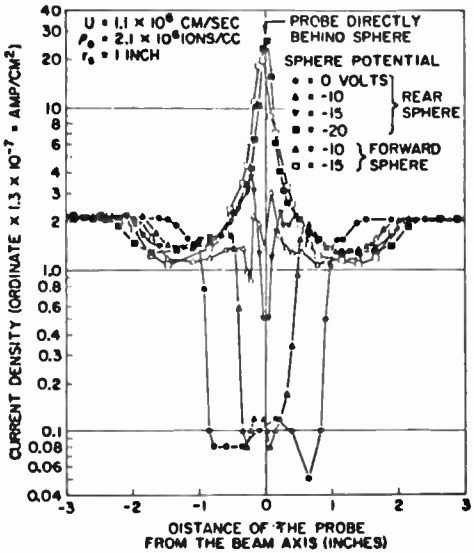

Fig. 2—Ion current density in the wake of a sphere for several potentials of the sphere.[16] (Rear sphere = 4 inches downstream, forward sphere = 7 inches downstream of center of sphere.)

fluence of the sheath on local measurements and upon the electrical drag experienced by the satellite. Hall et al[16] have made detailed measurements of the sheath potentials and ion-current densities on and in the vicinity of a spherical model illuminated by a neutralized beam of cesium ions. They find complicated sheath distributions and ion flow patterns, as shown in Figure 2, that depend upon the satellite potential. Shadow regions of ion current for a given satellite potential can become peaks of ion current for more negative vehicle potentials. The general features are in accord with the theory of Davis and Harris.[18] (The flow around discs and spheres has also been investigated by Cox and Clayden[19] using a low-density plasma tunnel.) Knechtel and Pitts[17] have made direct measurements of electrical drag on a model satellite in the laboratory using a beam of mercury plasma. They conclude that the electrical drag can contribute significantly to the satellite ion drag and that the theory of Jastrow and Pearse,[20]

which considers both the ion impact and the influence of the plasma sheath, agrees well with the measurements.

The flow of plasma past a conducting body immersed in a magnetic field has been investigated in a laboratory experiment by Kawashima.[21] Measurements have been made of the shock thickness in front of the body and of the downstream wake of a magnetic field oriented both along the flow and perpendicular to the plasma flow. In both cases, the thickness of the shock boundary is found to increase with increasing magnetic field. The downstream wake gradually disappears as the density of the plasma stream is decreased. For the parallel orientation of magnetic field, a disturbance appears upstream of the body if the plasma velocity is below the Alfven velocity. No such disturbance is apparent when the plasma is "super-Alfvenic."

The determination of the particle flux and electric and magnetic field distributions in the vicinity of such a body is essential for the interpretation of the results of various satellite measurements. A source aboard such a vehicle will find itself in this complicated sheath that can significantly alter its free-space characteristics. No attempts, either analytic or experimental, to ascertain the behavior of a radio-frequency source in such an environment have been reported.

2.4 "$V \times B$" Sheath

Long antennas on a satellite (as is the case for topside ionospheric sounding satellites) orbiting at a velocity V will cut magnetic field B lines of the earth, thereby developing a potential gradient ($\mathbf{V} \times \mathbf{B} \cdot \mathbf{L}$) along the length L of the system. This potential gradient, along with the fact that the net charge collection is zero, causes a redistribution of the particle-collection regions. The result is that the tip of one antenna goes to a positive potential, enabling it to collect all the electrons required for equilibrium; this tip exhibits the smallest sheath radius. The potential gradient is such that the remainder of the system is below the floating potential, collects only ions, and exhibits large volumes of electron depletion sheaths, as shown[16] in Figure 3.

The $V \times B$ sheath has been studied analytically[22] and by laboratory-simulation experiments[23-25] to determine the effect of the sheath on measurements of the local satellite environment made by instruments in the vicinity of the satellite body. Techniques[24] have been tested[25] whereby the deleterious effects of the sheath on local measurements can be eliminated, or at least minimized. These techniques still leave the antenna contained in the $V \times B$ sheath and no attention has been paid to the influence of the sheath on such sources, although some skewing of the beam radiation pattern due to the asymmetry of the

$V \times B$ sheath and an effect on the antenna impedance would be expected.

2.5 *Magnetic Barrier Sheath*

A sheath region is set up at the interface between a moving plasma and a magnetic field with which the plasma interacts if the magnetic field pressure exceeds the kinetic pressure of the plasma. This results because the ions (in the streaming plasma), which have larger mass and, hence, greater kinetic energy, penetrate further into the magnetic field than the electrons. This charge separation gives rise to an electric field E_s that accelerates the electrons in the $\mathbf{E}_s \times \mathbf{B}$ direction

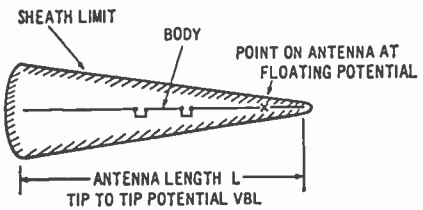

Fig. 3—General shape of the sheath around a body in a plasma resulting from an induced $V \times B$ potential.[23]

and causes them to execute broad turns. The different depths of penetration of the electrons and ions result in a transition region, or sheath. The sheath is of importance in the formation of the magnetopause in the earth's magnetosphere and in the confinement of hot plasmas by magnetic fields. This magnetic barrier sheath[26] is mentioned here for completeness, although no analysis of the behavior of a radio-frequency source in such a sheath, nor of the appropriate boundary conditions, has been reported.

2.6 *Other Effects*

It should be kept in mind that in formulating the problem of a source in a plasma, any or all of the aforementioned sheath processes may be present in varying proportions. The extent of each and its importance must be assessed for the particular problem being attacked. The redistribution of charge due to the sheath effects will cause current flow in the material body and hence, in the case of an antenna, will modify its current distribution.

In the real situation, many other complicating processes affect the sheaths just discussed. These include the effect of the outgassing of

rockets and space vehicles,[27] ablation off the vehicle, contamination from rocket exhaust, photoionization by ultraviolet radiation, and meteor and cosmic-ray bombardment, all of which may alter the plasma surrounding the vehicle and, hence, in the vicinity of the source. It is little wonder that accurate formulation of the boundary conditions for a source in a plasma is one of the most formidable problems of physics.

III—Antennas in Plasmas

3.1 General Considerations

As indicated, a major difficulty in predicting the behavior of a source in a plasma is the determination of valid boundary conditions. Other difficulties include the determination of the correct current distribution in the source (which is not necessarily the same as the impressed current distribution) and specifications of the electromagnetic properties of the plasma (which cannot be done in terms of a simple plane wave, since a source in a plasma will excite a spectrum of plane waves). These compounded uncertainties make evaluation of source behavior in terms of a rigorous boundary-value problem quite impossible, even if the accompanying mathematical problems of specifying appropriate Green's functions, etc. were tractable and if physical processes such as Cerenkov radiation, moving charges, spectra of e-m waves guided by plasma layers, circulating waves, and acoustic sources are neglected.

The physics involved in a given theoretical situation, as well as the complexity, is strongly affected by the model used for the plasma. Most theoretical attacks have been based on the simple hydrodynamic or continuous-fluid model for a plasma. The properties of the plasma can then be derived from consideration of the behavior of "average" particles that constitute the plasma or from the moment equations resulting from averaging by integration of the Boltzmann equation over velocity space. The plasma can be isotropic or anisotropic (inclusion of magnetic-field effects), cold or warm (inclusion of thermal-pressure effects), and can consist of 1-fluid (electrons), 2-fluid (electrons and ions), or 3-fluid (electrons, ions, and neutrals) models. The number of wave modes that the various models of the plasma can support are summarized in Table I. The transverse electromagnetic wave modes depend upon whether the plasma is isotropic (one e-m mode) or anisotropic (two e-m modes), while the longitudinal electroacoustic wave modes depend upon the number of "warm" species that comprise the plasma—1-fluid (electron acoustic wave), 2-fluid (electron

acoustic and ion acoustic waves), 3-fluid (electron, ion, and neutral acoustic waves). Of course, coupling between all the various waves is possible and hybrid electromagnetic–electroacoustic modes occur. We have separated the various modes into purely transverse or purely longitudinal waves here only to simplify the discussion.

Further favorite assumptions often introduced into the hydrodynamic model for a plasma are (1) neglect of charged-particle coupling (i.e., collisions) with neutrals and (2) for the case of aniso-

Table I—Simplified Wave Modes Implied by Various Models of the Plasma

Plasma Model	Wave Modes	
	Transverse	Longitudinal
Isotropic, cold, 1 fluid	1	0
Isotropic, cold, 2 fluid	1	0
Isotropic, cold, 3 fluid	1	0
Isotropic, warm, 1 fluid	1	1
Isotropic, warm, 2 fluid	1	2
Isotropic, warm, 3 fluid	1	3
Anisotropic, cold, 1 fluid	2	0
Anisotropic, cold, 2 fluid	2	0
Anisotropic, cold, 3 fluid	2	0
Anisotropic, warm, 1 fluid	2	1
Anisotropic, warm, 2 fluid	2	2
Anisotropic, warm, 3 fluid	2	3

tropic plasmas, that the dielectric coefficient of the plasma is uniaxial. With the latter assumption and using scaling procedures, the resulting fields of the source (under simplified boundary conditions) can be directly related to the fields that would be generated in free space. Little work has been devoted to the limitations of the fluid model[28] as it applies to antenna source problems,[29] although it is clear that this description of the plasma fails to predict the plasma behavior in resonance regions (see Section 4.3). Source problems based on the Vlasov and Boltzmann-equation approaches have been limited, although such an approach is essential for resonance regions, prediction of Landau damping, determination of some aspects of anisotropic pressure effects, etc.

Approaches to the analysis of the behavior of sources in plasmas thus far have been based on (a) theoretical analyses using highly idealized simple models, (b) semi-rigorous theoretical analyses, and (c) an intermediate approach to theoretical analysis with as much experimental verification as is possible. With the first approach, an almost infinite number of problems can be formulated and consider-

able insight can be gained as to the various aspects of source behavior. The results may, however, not be applicable to any physical situation, since the validity of the approximations made is unknown.

The semi-rigorous analyses encounter considerable difficulties in physics and mathematics. When, and if, results are obtained, they are often not in a usable form and simplifications are necessary. They do, however, provide a better indication of the physics of the situation.

The third approach tends to lead to a gradual improvement in the rigor of understanding, since it is closely coupled to the physics of the situation, and invalid aspects are ascertained and discarded and new features are introduced. Experiments with sources in plasmas have limitations, however, whether in the laboratory or in a space environment, and the time needed to obtain results is lengthy.

Before proceeding to discuss recent developments, let us briefly review the concepts fundamental to the discussion of sources in plasmas.[30-34]

The behavior of the electric field of an e-m source in a plasma must, of course, obey the wave equation

$$\nabla \times \nabla \times \mathbf{E} - k_0^2 \, \mathbf{K} \cdot \mathbf{E} = j \, \omega \mu_0 \, \mathbf{J} \, \delta(r) \, ,$$

where $\mathbf{E}$ is the electric field, k_0 the free-space wave number, $\mathbf{K}$ the dielectric permittivity of the medium, ω the radian frequency, $\mathbf{J}$ the impressed current distribution, and $\delta(r)$ the delta function, where r represents space coordinates.

The electric field can then be obtained using Fourier transform techniques in terms of a superposition of wave modes of the form

$$E = \int F(r) \, dr \, .$$

The function $F(r)$ contains the dependence on both the current distribution and the medium properties. The exact form of $F(r)$ and the variables of integration, including the number required, depend on whether the field is expressed in terms of cavity modes or guided modes, the details of which are contained in the papers by Felsen.[32,34*] The medium properties are contained in the dispersion relation relating the refractive index of the medium to the frequency. In turn, the refractive index depends on the forces acting on each species of

* See also L. B. Felsen, "Radiation and Scattering in Bounded Plasmas," Invited paper, XVth General Assembly, Munich, 1966.

particles comprising the plasma and on the inter-relation between these forces. In loose physical terms, the electromagnetic properties of the plasma are determined by the effect of the force of the transverse electric field of the wave on the free charges of the plasma, while the electroacoustic properties arise from the thermal forces on the charges.

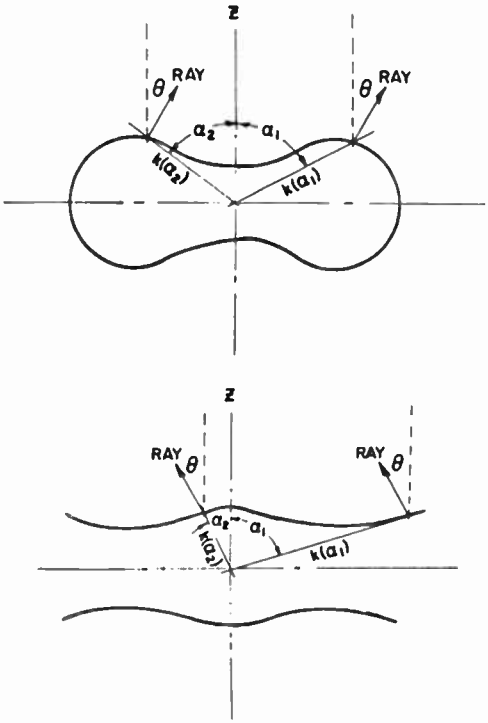

Fig. 4—Typical refractive index variation with angle relative to the magnetic field that can occur for an anistropic plasma.

Even for a cold anisotropic plasma, the situation can get quite complicated and interesting. Since an anisotropic plasma can support two wave modes, two refractive index surfaces[35,36] can be associated with the properties of the medium (corresponding to the ordinary and extraordinary wave modes). When the refractive index is real, the dependence of the refractive-index surfaces on the angle relative to the direction of the static magnetic field can take the form shown in Figure 4. The refractive index surface can be closed (it is circular for an isotropic plasma) or open, and either of the two modes can be of any one of these two types. The group velocity and hence the

direction of energy travel is perpendicular to the refractive-index curve. The open surface is particularly interesting, since no energy can propagate in the regions where the refractive index is infinite. These infinities, or resonance regions, also give rise to fields that diverge if excited by infinitesimal sources.[29,37] This problem is currently being investigated, and no degree of accord on the limitations of present theoretical treatments has yet been established.

Because of the complexity of the problem of a source in a plasma, simple antenna configurations have been principally considered. The following sections describe the effort since 1963 classed in terms of types of antennas.

3.2 Electric Dipoles

Electric dipoles in plasmas have received considerable attention. Seshadri[38] has calculated the radiation resistance for an electric dipole in a cold lossless plasma and found that at high frequencies when losses are low, it varies as the square of the frequency. Isotropic warm plasmas showing the coupling between electromagnetic and electroacoustic modes, radiation resistance, and far-zone fields have been determined for 1-fluid models (showing the electron acoustic wave) and the 2-fluid model (showing both the electron acoustic and the ion acoustic wave). In general, the radiation resistance is reduced with increasing temperature when the radian frequency ω is less than the plasma frequency ω_p. At the same time, the electromagnetic radiation is mostly broadside for $\omega > \omega_p$ and tends to shift to directions along the axis of the dipole for $\omega < \omega_p$. In an attempt to ascertain the effect of the ion sheath, a model in which an electric dipole is oriented within a column of insulation has been studied.[1] The insulation was found to affect principally the acoustic-wave mode, which varied as the inverse of the radius of the insulation. Also the power in the acoustic mode was found to decrease at high frequencies.

For cold anisotropic media, the radiation pattern of an electric dipole calculated for infinite magnetic field showed a special direction of principal radiation. For arbitrary orientation of the dipole to the direction of magnetic field, Mittra and Deschamps[39] arrived at the far-field pattern in terms of a finite range integral suitable for numerical computation. Numerous radiation patterns corresponding to various plasma regions of the CMA (phase-velocity) diagram were calculated.[36] Typical radiation patterns are shown in Figure 5. Weil and Walsh,[40] Staras,[37] and Seshadri[41] have evaluated the radiation resistance. Weil and Walsh present extensive numerical data for both the ordinary and extraordinary wave modes. Staras calculates values

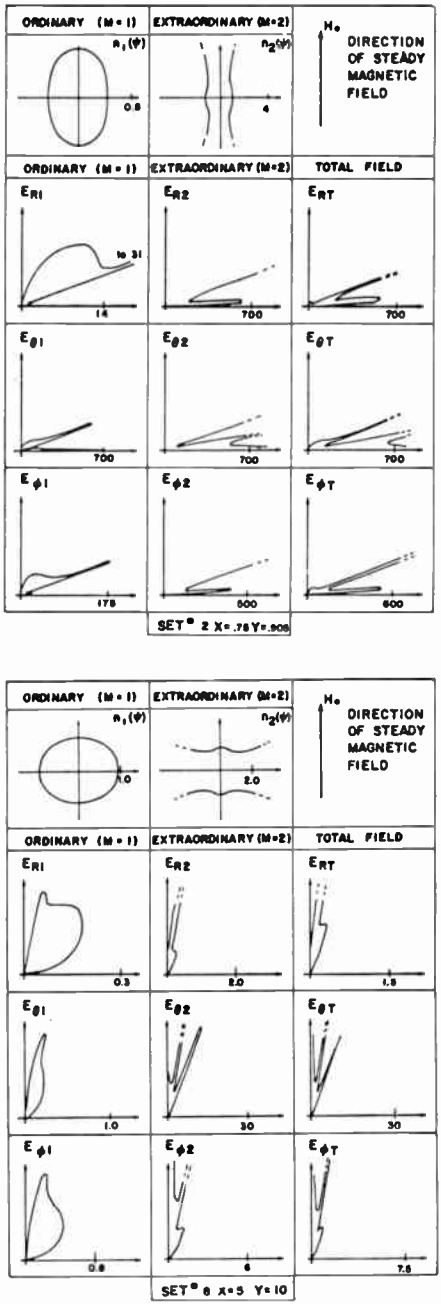

Fig. 5—Radiation patterns for an electric dipole in an infinite, anisotropic cold plasma.[36]

for the whistler region and illustrates the infinity "catastrophy" for the radiation resistance. Seshadri's calculations for a triangular current distribution for a dipole oriented parallel to the magnetic field confirms the infinities for a collisionless plasma at the resonance regions. Electric dipole antennas for operation in the ionosphere have also been reported.[42] The influence of magnetic field strength on a dipole located in an axially magnetic plasma column to improve communications has also been calculated[43] (the maximum radiation shifts from broadside ($\omega > \omega_p$) to end-fire ($\omega < \omega_p$)). Electric dipoles in a warm, anisotropic plasmas[44] showing the transfer of power between electromagnetic and electroacoustic wave modes have also been investigated.

3.3 Magnetic Dipoles

After the analysis of Wu[45] for the radiation patterns of both electric and magnetic dipoles, the magnetic dipole received little attention until 1965. Since then calculations have been published for the radiation resistance,[46] input impedance,[47] and admittance of a magnetic dipole in an infinite, cold, anisotropic plasma. An interesting result of Duff and Mittra's calculation[47] is that the effect on the input impedance of the plasma does not depend on the off-diagonal terms of the plasma permittivity tensor. Weil and Walsh[46] present radiation resistance calculations for magnetic dipoles similar to their analysis of the electric dipole.[40] A comparison between the electric and magnetic cases reveals that significant differences exist even for isotropic plasmas. For the isotropic plasma the radiation resistance varies as $(1 - X)^{1/2}$ for the electric dipole but as $(1 - X)^{3/2}$ for the magnetic field dipole in the region $X \leqslant 1$ (X is the plasma frequency normalized to the radian radio frequency.) Thus the derivative with respect to X for the electric dipole becomes infinite as X approaches unity for the electric dipole, while it is zero for the magnetic dipole.

Thus the electric dipole is much more sensitive than the magnetic dipole to the plasma properties in the region $X = 1$ and in cutoff regions, even in the presence of a magnetic field. It is thus apparent that the electric dipole is a preferable probe for sensing plasma parameters, while the magnetic dipole may have advantages as a transmitter or receiver of r-f energy, since it is less sensitive to the local plasma.

Some progress in studying the influence of inhomogeneous stratified media above a horizontal conducing plane has been reported. Formal expressions for Hertz potentials have been obtained in the form

of contour integrals that can be evaluated in terms of known functions.[48] In addition, the influence of a conical sheath has been investigated.[49]

3.4 *Short Antennas*

The initial attack on the input impedance of a short cylindrical antenna in a magnetoplasma, made by Balmain,[50] was based on quasi-static electromagnetic theory and included both collisions and dependence on orientation relative to the magnetic field. His analysis for a triangular current distribution could be put in terms of scaled coordinates and compared favorably with experiment. Seshadri[41] presented numerical values for the radiation resistance of a similar current distribution in a linear current filament parallel to the magnetic field in a uniaxial anisotropic medium. Mittra[51] subsequently showed that the above two results were compatible.

Balmain[52] extended his treatment to include thermal pressure effects and quoted the input impedance for a plane, a sphere, and a dipole. These are summarized in Table II. Further analyses of short antennas have been made by Fejer[53] for an antenna consisting of two spherical conductors and by Balmain[54] for a spherical probe.

The behavior of short antennas for the absorptive boundary conditions[55] (i.e. no sheath, but, of necessity, current drain) shows very little influence for frequencies above the plasma frequency. For frequencies below the plasma frequency, however, there is a significant change in input impedance, the input impedance turning resistive as the frequency approaches zero. The relevant formulas are compared in Table II.

3.5 *Magnetic Line Source*

Because of its two-dimensional symmetry, the magnetic line source has attracted much attention. Excitation of plasma waves for a line source parallel to a magnetic field has been analysed.[56] The power in the electromagnetic mode is always less than for free space. A magnetic line source located off-axis in a multi-layer concentric cylinder of plasma has been considered by Harris.[57] He was able to obtain an exact solution for a single plasma layer and evaluate the more general case asymptotically. The radiation pattern for a single layer shows narrow peaks and valleys at certain angles. Beyond the critical angle, the power is reduced in some directions. In addition there is a profound influence on the polarization; the emerging radiation is, in general, elliptically polarized.

The line source in a half-space[58] and in a cylindrical cavity[59] have also been analyzed to ascertain the coupling to electroacoustic waves at a boundary in a warm plasma. A two-fluid plasma[60] excited by a line source shows, as expected, the addition of an ion acoustic wave. Finally, a line source on a conducting plane covered with a warm plasma layer[61] has been treated to determine the far field, the surface-wave contribution, and the relative power in the various wave modes. The far fields are not affected by plasma temperature unless the frequency is in the vicinity of the plasma frequency or unless the electrons have relativistic energies. Energy transported by surface waves is often greater than the energy in the radiation field. The radiation field is dominant for a source of finite width. Some attention has also been paid to a magnetic line source in an inhomogeneous plasma.[62]

3.6 Cylindrical Antennas

Initial work on cylindrical antennas was devoted to analyzing the dispersion relation and wave modes along conducting wires in order to obtain some indication of how to determine the radiation pattern and radiation resistance. Since then, the radiation resistance for the acoustic and electromagnetic components and the far field patterns have been determined for both warm and cold isotropic and anisotropic plasmas. Both finite and infinite cylindrical antennas with sinusoidal current and δ-function distributions have been considered.

Hurd[63] has presented consderable numerical data on the admittance of a long, thin, center-fed, cylindrical antenna in a uniaxial medium oriented at an arbitrary angle to the medium axis. Aoki[64] has computed the input impedance by Fourier-transform techniques for both infinite and finite cylindrical antenna in a homogeneous anisotropic plasma (neglecting sheath effects). An interesting result is that a small circulation current (which is spatially modulated) is found in linear antennas in anisotropic media. An approximation to the influence of the ion sheath has been to investigate an infinite cylindrical antenna with concentric free-space gap[65] in a warm plasma. Calculations of the radiation resistance and radiation pattern shows that as in other antenna configurations, the free-space gay significantly affects only the electroacoustic mode.

3.7 Slot and Open-Waveguide Antennas

The slot type of antenna has been considered primarily because of its convenience as a radiator on a re-entry body, and the effect of the re-entry plasma sheath[66] on its impedance and radiation charac-

Table II—Comparison of Input Impedance for Rigid[7] and Absorptive[13] Boundary Conditions (for cold plasma, $\alpha \to \infty$).

Antenna Configuration	Z_{in} (Rigid Boundary Condition)	Z_{in} (Absorptive Boundary Condition)
Sphere	$\frac{1}{j\omega 4\pi\epsilon_o (1-X) R}\left[1-\frac{X}{1+\alpha R}\right]$	$\frac{1}{j\omega 4\pi r_o\epsilon_o K_o}\left[1-\frac{X}{U}(1+\alpha_1\rho_o - j\alpha X^{\frac{1}{2}}\alpha_1^2\rho_o)^{-1}\right]$
Plates	$\frac{2D}{j\omega\epsilon_o (1-X) S}\left[1-\frac{X\tanh\alpha D}{\alpha D}\right]$	$\frac{2L}{j\omega\epsilon_o K_o S}\left[1-\frac{X}{U}\left\{\alpha h(\coth\alpha h - j\sqrt{\frac{2}{\pi}}X^{\frac{1}{2}}\alpha)^{-1}\right\}\right]$
Dipole	$\frac{1}{j\omega\pi\epsilon_o (1-X) L}\left[\ln\frac{L}{\rho} - 1 - XI_o(\alpha\rho) K_o(\alpha\rho)\right]$	

Definitions:

R = sphere radius

S, D are parallel plate area and spacing

$\alpha = \frac{\omega}{V}(X-U)$

$V = \sqrt{\frac{\gamma\kappa T}{m}}$

$U = 1 - jZ$

L = dipole length

ρ = dipole radius

I_o = modified Bessel function of 1st kind

K_o = modified Bessel function of 2nd kind

ϵ_o = permittivity of free-space

$h = L/\lambda_D$

$\alpha^2 = 1 - UX^{-1}$

L = plate spacing

$K_o = 1 - XU^{-1}$

S = plate area

r_o = sphere radius

$\rho_o = r_o/\lambda_D$

$a = (2/\pi)^{\frac{1}{2}}$

$\lambda_D = \sqrt{\frac{\epsilon_o\kappa T}{ne^2}}$

teristics has been treated in detail. The re-entry sheath has a significant influence for frequencies below the plasma frequency; in practice, the source frequency is nearly always below the plasma frequency during some part of the re-entry trajectory. The admittance of slots radiating[67] into uniform plasmas and into stratified magnetoplasmas has been determined, as well as the susceptance and conductance of a slotted sphere[68,69] including the influence of electron–neutral collisions.

The use of a magnetic field to improve electromagnetic transmission through the plasma has been analyzed by Yeh[70] for an arbitrary slot on a conducting cylinder coated with a plasma layer having an axial magnetic field. Numerical results for an infinite axial slot show the expected improvement in transmission as well as an influence on the asymmetry of radiation pattern by the magnetic field. Hodara's analysis[71,72] for a sheath-covered rectangular slot with a magnetic field perpendicular to the slot shows that the radiation pattern is not greatly affected for $\omega_b^2 \gg (\omega^2 + \nu^2)$, where ω_b is the electron cyclotron frequency and ν the collision frequency. In addition, the match is improved by the plasma under some conditions.

3.8 Other Antenna Configurations

Other antenna investigations that have been considered include calculation of the capacitance of a biconical antenna of circular cross-section[73] (neglecting the ion sheath) and the admittance of a thin biconical antenna in an isotropic warm plasma.[74] In the latter case, an ion sheath in the form of a sphere of radius equal to the half-length of the antenna was assumed.

The behavior of both a spherical aperture antenna[75] in a warm plasma (using the absorptive boundary condition) and a prolate spheroidal antenna[76] in a warm plasma (with a confocal sheath) have been treated by Wait. Emphasis has been on the energy in the electroacoustic modes.

Flat strip antennas have also been of interest. Seshadri[77] determined the radiation resistance of a linear current strip of finite width in a uniaxial anisotropic plasma. The strip width is arranged parallel to the magnetic field, while the current path is along the length of the strip; the current is constant over the length and triangular in shape across the strip. The radiation resistance for frequencies below the plasma frequency is found to be several orders of magnitude greater than that for frequencies above the plasma frequency. The dependence on frequency and length of the current strips was determined. Galejs[78] has looked at the behavior of finite-width antennas

and insulated antennas located in several dielectric layers parallel to the surface. Calculations for an antenna in a dielectric layer and for an insulated antenna are presented.

3.9 *Plasma-Clad Cylinders*

The axially slotted plasma-clad cylinder is another favorite configuration for obtaining an assessment of the influence of the re-entry plasma sheath on an on-board source. Complete radiation patterns obtained[79] for isotropic plasma sheaths[79] show that the maximum of the radiation swings to broadside as the plasma frequency exceeds the

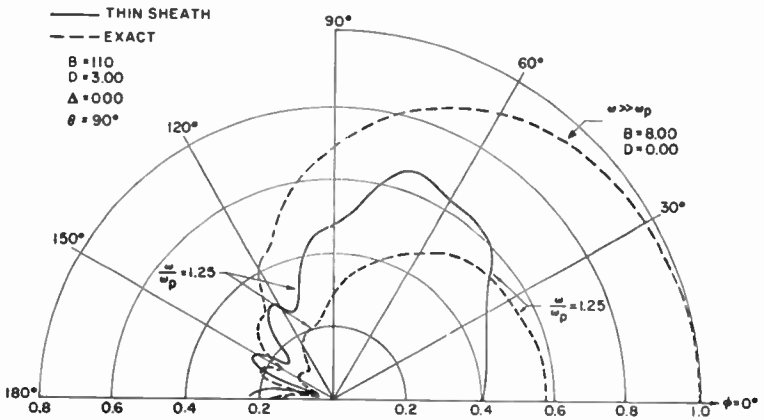

Fig. 6—Typical radiation pattern for an axially slotted cylinder covered with a plasma sheath.[79]

radio frequency, as shown in Figure 6. The analysis has been extended to an inhomogeneous plasma layer in the radial direction corresponding to a typical flow field.[80] Analysis of insulation of the plasma by a thin dielectric coating[81] shows numerous nulls in the radiation pattern and high attenuation (which is reduced by collisions) for frequencies below the plasma frequency. The influence of an axial magnetic field gives rise to rapid fluctuation with angle in the radiation characteristics.

The radiation from a magnetic dipole located within an isotropic plasma-clad cylinder shows numerous peaks at different angles and little radiation beyond the critical angle. The broadside transmission is less than predicted by plane-wave theory.

Ring sources in a plasma cylinder, which have also been of interest, show coupled electroacoustic and electromagnetic modes for a warm plasma. Consideration of a loop antenna with a cylindrical core

surrounded by an insulator in a cylindrically stratified plasma shows that the radiation pattern is more sensitive to the thickness of the plasma sheath than to the value of the dielectric constant in the sheath.

3.10 Laboratory Experiments

The literature on antennas in plasmas abounds with theoretical papers, while meaningful experimental results are relatively difficult to find. The reported experiments can be classified into two kinds: those using artificial dielectrics to simulate plasmas and those using actual gaseous plasmas.

Following the original work of Rotman[82] showing that dielectric constants of less than unity could be obtained using artificial dielectrics (and, hence, plasmas could be simulated in this way), Golden and Smith[83-86] have used wire-grid artificial dielectrics to experimentally study the influence of a plasma sheath on horn antennas,[85] slots,[84] and slotted cylindrical antennas[86] located an arbitrary distance from a thin sheath. In each case, good agreement between theory and experiment has been obtained. This is not surprising for carefully performed experiments, since the boundary conditions for artificial dielectrics are the same as those used in the theoretical analysis, i.e. ion-sheath effects are not of importance. Tyras et al[87] have used a different technique for simulating a plasma sheath, namely, the idea that the e-m interaction is governed by the relative dielectric constants between two different media, and propagation into a dielectric medium whose dielectric constant is less than one can be simulated by propagating from a medium of higher dielectric constant into one that is less. Again fair agreement between theory and experiment was obtained.

Experiments with gaseous plasmas are more difficult to perform because of the problems of generating a suitable plasma. They are more difficult to analyze, because of the problem of specifying the properties of the plasma. Experiments on the effect of a gaseous isotropic plasma sheath on a horn antenna[88] show the decrease in signal level at normal incidence due to the plasma, the redistribution of energy over angle, and the shift of the critical angle.[89] Typical experimental results[89] are shown in Figure 7. The broadside radiation pattern of a slot covered with an anisotropic plasma with the magnetic field perpendicular to the plasma layer shows an enhancement of transmission with magnetic field.[90] The results show that in order to make accurate predictions, it is necessary to also consider collisions, impedance variations, and electro-acoustic resonances. As can be seen from Figure 8, where only simple considerations are included, a fair agreement between theory and measurements[90] is obtained.

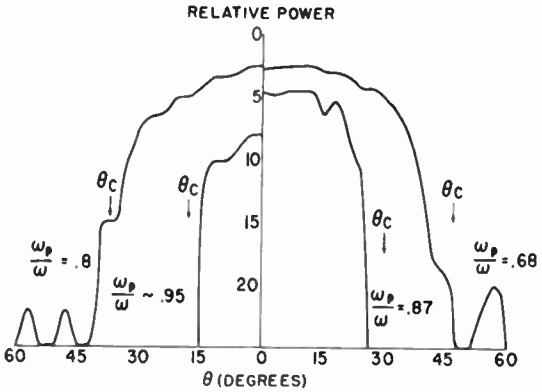

Fig. 7—Typical measured radiation patterns for plasma-covered horn antennas.[89]

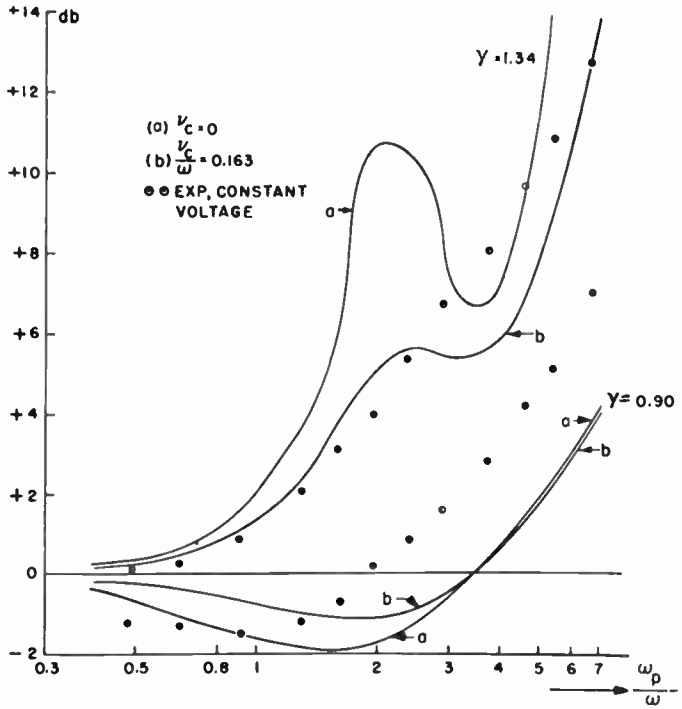

Fig. 8—Measured transmission enhancement due to a magnetic field for a plasma-sheath-covered slot and comparison with theory.[90]

3.11 Space Experiments and Considerations

Sheath effects are of prime concern in the evaluation of space measurements, and an attempt is made to include the ion sheath in most considerations of ionospheric experiments. To explain various space measurements, it has been necessary to invoke electroacoustic waves,[91] ∇E field effects,[8] etc.

Mlodnosky and Garriott[92] have determined the admittance of a dipole at very low frequency by considering the dipole as a symmetrical Langmuir probe whose ion sheaths follow the applied r-f fields. The ion sheath thus determines the admittance. Following this idea, Herman[93] has made an analysis of a capacitance impedance probe (built as part of the nose of a rocket) to measure electron density. The analysis includes collisions as well as the ion sheath.

Although the impedance probe is theoretically suited for space plasma measurement, space experimental results obtained[94] with it are not particularly satisfactory. Interpretation in terms of recent analysis,[95] which has included the effects of ion sheaths, collisions, and magnetic fields, may prove more useful.

Antenna performance in ionized media may also be of importance to the idea, recently suggested by Hall and Fooks,[96] of measuring electron densities in the D-region ($\sim 10^4$/cc) by sending up rockets to measure the received signal strength of a low-frequency signal sent from the ground.

IV—Nonlinear Phenomena

The near fields of antennas embedded in ionized media subject the surrounding plasma to very strong nonlinear electromagnetic fields. When considering sources in plasma, therefore, one must be very conscious of nonlinear effects that negate the assumption of small-signal theory upon which most theoretical investigations of antenna behavior have been based. Electric fields as low as 2×10^{-7} volts/cm can influence the properties of an isotropic plasma, and even lower field strengths can have an effect in the presence of magnetic fields. The strong electromagnetic fields can result in plasma heating, breakdown, various resonances, and stimulated phenomena, as well as collisionless dissipation of e-m energy. In this section, a brief account of some of these effects is given.

4.1 Plasma Heating

The linearized set of equations usually used to describe e-m wave interaction with plasma is valid provided the signal amplitudes are

sufficiently weak so as not to perturb the equation of state of the plasma. For sufficiently high r-f power levels, the electron velocity distribution, the electron temperature, electron density, electron-temperature gradients, and electron-density gradients may vary with the strength of the interacting wave. Thus, all the transport coefficients of the plasma become functions of the amplitude of the r-f field.

In recent investigations[97-99] concerned with the change in plasma properties in presence of a strong e-m field it has been found that the

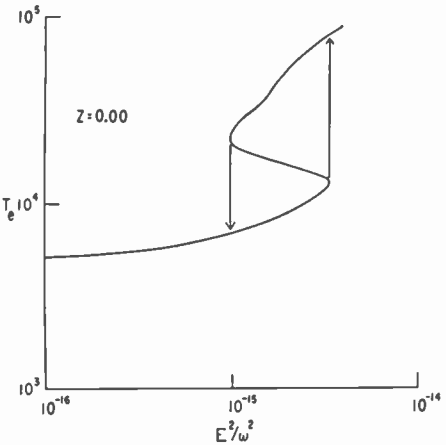

Fig. 9—Variation of electron temperature with square of ratio of electric field strength to radian frequency showing a "hysterisis" type effect.[100]

boundary separating the un-ionized from the ionized medium is perturbed by the strong r-f fields so that it appears to move towards the r-f source. The absorption of energy by the plasma results in electron temperature and electron density becoming functions both of time and space. In general, the time scale for temperature changes is much longer than the time scale for density changes.

An interesting factor appears in investigations[100,101] of the steady-state response of a plasma subjected to intense r-f fields, namely, a multivalued solution for the electron temperature exists for certain ranges of the square of the electric-field strength divided by the square of the frequency, as shown in Figure 9. Within a given range, the plasma can exist in three temperature states, two of which are stable and one unstable, thus forming a "hysteresis" in electron temperature.

In addition, the effect of the heating of the plasma by the wave can give rise to demodulation, generation of harmonics, frequency

mixing, and interactions with longitudinal waves; in the case of nonthermal particles, it can also result in enhanced plasma-density fluctuations, stationary fields and currents, and changes in the transmission, reflection, and absorption characteristics of the medium.

The effects of such factors as plasma compressibility due to an anisotropic pressure tensor, heat conduction within the electron gas, and the influence of strong fields on plasma boundary conditions have, as yet, received little attention.

4.2 Plasma Breakdown

High-power e-m signals in a plasma cause breakdown and, hence, no r-f power passes through the plasma medium. The breakdown condition thus places an upper limit on the useful signal strength that can be generated by a source in a plasma. Although r-f breakdown in plasmas is well understood under ideal situations, a number of complicating factors occur under real experimental conditions. Attention has recently been given to such complexities. Chief attention has been given to obtaining better values of breakdown conditions in air,[102] near atmospheric pressure,[103] in compressed air,[104] and in the constituents of air,[105] as well as to the breakdown conditions in the atmosphere.[106] Other factors that have been taken into account include the effects on breakdown of a hot surface, gases at high temperatures, nanosecond pulses, shocks, flow turbulence, and nonuniform electric fields.

4.3 Plasma Resonances

The launching of the first ionospheric topside sounder satellite[107] on September 29, 1962 (Alouette I) ushered in a new era of ionosphere exploration. The swept-frequency topside sounder operates much as a ground-based sounder, transmitting short pulses of radio energy and determining the delay time in receiving the return echo. By sweeping the frequency (in the region 0.5 — 12 MHz) a profile of electron density (reflection frequency) versus height (delay time) is obtained. Such a series of ionograms from a Topside Sounder satellite gives information on the region of the ionosphere above the region of maximum electron density, a region impossible to study adequately from the ground. (A second topside sounder satellite working at six successive fixed frequencies was launched August 25, 1964.)

In addition to the usual ionogram, the topside-sounder results have shown the surprising feature of immediately observed signals that decay slowly in time at certain critical sounder frequencies.

These have been termed spikes by Lockwood[108] and have proved to be dependent on the value of the magnetic field and on the electron density in the vicinity of the satellite. They were identified[109] as corresponding to harmonics of the electron cyclotron frequency (ω_h), the local plasma frequency (ω_p), and the local hybrid frequency ($\omega_T = (\omega_h^2 + \omega_p^2)^{1/2}$). Similar types of resonances are observed with fixed-frequency topside ionospheric sounders. Up to 19 cyclotron harmonics have been observed as well as occasional spikes at $2\omega_T$.

The frequencies of the harmonics are, within experimental accuracy, integral multiples of the cyclotron frequency derived from the best estimate of the earth's magnetic field at the satellite height.

A number of heavy-ion effects (e.g., proton gyro frequency and electron-ion hybrid extraordinary wave resonance[110]) have recently been observed in some of the Alouette data. Consequently the topside sounder appears to be a potentially very important technique for studying ion phenomena in the upper atmosphere.

All the above resonant frequencies appear naturally in plasma theory so that, in principle, determination of the details of excitation, manner of decay, and dependence on the various parameters could lead to a detailed knowledge of the ionospheric plasma. Such phenomena involve the behavior of an electromagnetic source in a plasma and the coupling between electrostatic waves and the normal transverse waves predicted by the Appleton–Hartree equation. Unfortunately, the theory for such interactions is not well developed, particularly in relation to predicting quantitative results, although Shkarofsky[111,112] has recently extended the theory for collisionless plasmas in the regions critical to such resonances.

Early qualitative explanations of the satellite experiments were offered in terms of bunching of electrons in phase with the electric field of the transmitter and in terms of the electron sheath surrounding the antenna. Recent approaches, each with simplifying assumptions, have been in terms of quasistatic waves[113] with the wave vector approximately normal to the direction of the magnetic field, in terms of cold plasma theory,[114] and considering both electromagnetic and thermal effects. Shkarofsky and Johnston[115] point out that the group velocity for some of the resonances may be of the same order as the satellite velocity. Thus it appears that for a complete theory, satellite motion as well as the electromagnetic and thermal effects should be included, although the above approaches may be valid in certain circumstances. Thus, the prediction of detailed quantitative results is yet to come.

Hagg[116] has demonstrated yet another resonance that he has termed

the remote resonance, since it occurs (unlike the aforementioned resonances) remote from the satellite. This resonance occurs at a height at which the frequency of the extraordinary wave is twice the electron gyrofrequency. Muldrew and Hagg[117] explain the remote resonance as due to a transmitted extraordinary wave pulse that travels down an ionospheric field-aligned duct into the region where the carrier frequency of the pulse train is twice the electron gyrofrequency, and a cyclotron resonance is excited in this region. At some later time, the phase distribution of the electrons as function of height is such that an extraordinary wave pulse at the transmitted frequency is generated in this region and propagates back up the duct to the satellite. This generated pulse is the remote resonance.

Considerable dependence must be placed on the results of laboratory experiments to guide the theoretical speculation. Laboratory experiments of direct application to the ionospheric phenomena are still at a relatively undeveloped stage.

Studies of such ionospheric phenomena involves a knowledge of the behavior of a radio frequency antenna or r-f source in a plasma. Experimental data on this subject are presently very scarce. Only Balmain[50] has reported laboratory measurements of the impedance of a monopole antenna immersed in a magnetoplasma. His experimental results are in broad agreement with analysis based on quasi-static electromagnetic theory. A laboratory experiment more closely related to the Alouette gyro-resonance phenomena has been conducted by Crawford, et al.[118] In this experiment direct measurement of transmission in a plasma across the magnetic field between two antennas was made using Langmuir probes located a few millimeters apart in a mercury-vapor discharge. The frequency was fixed at 400 MHz while the magnetic field was varied. A typical result is shown in Figure 10. The harmonics of the electron cyclotron frequency are clearly evident, although other measurements have shown multiple peaks in the resonances. These peaks have been attributed to two different wave modes trapped internally in the plasma due to the nonuniformity of the electron density[119] in the cylindrical plasma column used in the experiments. A direct comparison of these results with theory cannot as yet be made.

A considerable number of laboratory experiments have been performed on resonance phenomena pertaining particularly to noise emission and radio-wave absorption. These have been summarized by Crawford[120] and are not dealt with here. Recent theory and experiments on electromagnetic-wave interaction with thin cylindrical columns of magnetoplasmas have shown the influence of electroacoustic

waves in establishing longitudinal standing waves that determine the resonant scattering. In general, the laboratory techniques for studying resonances is better established than the ones used in space in which the source must be immersed in the plasma. To date, most laboratory results have not been directed towards the kind of phenomena of direct relevance to a source in a magnetoplasma, but the laboratory techniques are well enough advanced to enable many of the phenomena to be explored under controlled conditions.

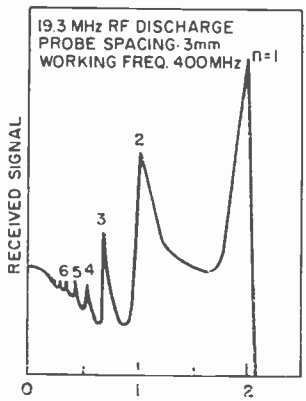

Fig. 10—Laboratory measurements of transmission between two Langmuir probes across the magnetic field direction in a plasma.[118]

4.4 Stimulated Echoes

Recent experiments[121] conducted in magnetized after-glow plasmas for various gases (Ar, Ne, He, N_2) show that echo trains can be stimulated by repeated r-f pulsing. In the experiment, pulses of $\sim 10^{-8}$ second duration, 15 watts of power at 10 GHz were applied with their electric vector perpendicular to the magnetic field. If the microwave frequency of the pulsed signal approached the electron cyclotron frequency of the anisotropic plasma, then two successive applied pulses separated by a given time could produce one or more stimulated echoes delayed by intervals that are the same as the time between applied pulses. A third pulse applied after some considerable time could also produce such stimulated echoes. This sequence of applied pulses and observed stimulated echoes is illustrated in Figure 11. The best results were obtained in neon at a pressure of 5×10^{-3} to 10×10^{-3} torr and an electron density $\sim 5 \times 10^{9}$ cm^{-3}. Echoes 50 db above the noise level could be observed.

Gould[122] was the first to offer an explanation of this phenomenon based on an ensemble of noninteracting charged particles being subjected to two or three successive delta-function excitation pulses. The result is that a periodic particle rebunching phenomenon can occur in velocity space after the second pulse. In addition to the periodic bunching some nonlinear mechanism is required in order to generate the echoes. The nonlinear mechanism currently suggested includes the velocity dependence of an oscillator frequency[123] (as for relativistic particles in a magnetic field or particles in certain inhomogeneous magnetic fields), the inhomogeneity of the driving electric field,[124] and the velocity dependence of the electron–neutral collision frequency.[125] The appropriate mechanism is yet to be delineated, but with increasing experimental data this should be forthcoming.

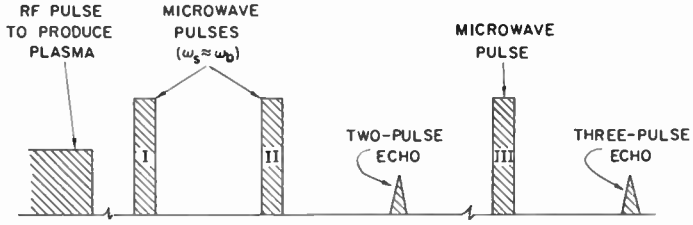

Fig. 11—Sequence of applied pulses and observed echoes in experiments on stimulated phenomena.[121,125]

4.5 Noncollisional Dissipation of E-M Energy

Collisionless dissipation of energy during magnetic field interaction with plasmas resulting in collisionless shock waves has been known for some time[126]; it has recently been invoked to explain the magnetically turbulent region found in the transition from the interplanetary solar-wind region to the magnetospheric windward boundary. As a result, intensive investigations of collisionless shock phenomena are being made by space and laboratory experiments and by theoretical studies.

Of concern in the behavior of a source in a plasma is the fact that a large-amplitude transverse e-m wave propagated across a magnetic field can give rise to a collisionless shock wave. Such an experiment has been performed in a rarefied plasma ($10^{11} - 10^{14}$ electrons per cm^3) by Zagorodnikov et al[127], and a steep shock front was observed. Observations of the rate of increase of the electron temperature show that a collisionless mechanism of energy dissipation is most probable. In the experiment, the contribution of ionization to electron-energy distribution is negligible. The time width of the observed

shock front was within a factor of two of the theoretical limit. Little work appears to have been done on the formation of collisionless shocks due to intense r-f sources.

Collisionless (Landau) damping for electrostatic waves due to interaction between particles of an appropriate velocity range with the wave is well known. Extensive theoretical work has been done on the subject and no mention of the theoretical work is made here. Recent experiments[128,129] have demonstrated this heavy damping of the longitudinal electron waves caused by electron traveling at the phase velocity of the wave. In the experiment, this damping is several orders of magnitude greater than that due to collisions. The dependence on wave-phase velocity and plasma temperature appears to agree with the Landau theory. Ion acoustic waves should be subject to Landau damping as well, and recent experiments have verified this.

Although the influence of longitudinal waves on a source in a plasma has been considered, most of the approaches have been based on the hydrodynamic model of a plasma that cannot predict Landau damping. The effect of Landau damping on longitudinal waves excited by an e-m source in terms of the source characteristics is unknown.

4.6 Other Phenomena

A number of other phenomena have been studied recently; the results have a bearing on the behavior of r-f sources in plasma media. These include *the properties of longitudinal waves* (their excitation and amplification). Considerable attention has been paid to both electron and ion thermal or acoustic waves and theoretical and experimental results have been obtained relating to the effects of collisions, their behavior at resonances, their velocities and attenuation.

A second relevant area is *the interaction between electroacoustic and electromagnetic waves*, such as scattering and modification of the plasma permittivity. Since an r-f source in a plasma can excite electroacoustic waves as well as electromagnetic waves, the behavior of the acoustic mode and their coupling to the e-m waves is essential in understanding the behavior of an r-f antenna in a plasma.

In many experiments, particularly those relating to space exploration, the r-f sources used in the experiment operate in a pulsed mode. Thus, this review would not be complete without mentioning that recent activities have been devoted to *the behavior of pulses in plasma media*.

Finally, one could mention the influence of multi-ion species in plasmas including the effect of heavy negative ions.

V—Unsolved Problems and Future Areas of Research

Based on this review of sources in unbounded plasmas, the following are considered the areas of research that require attention.

a. *Boundary Conditions*

Considerable effort is required on *all realistic aspects* of the conditions at the boundary between a source and a plasma. Attention must be paid to the plasma physics of the problem as well as the electromagnetics. Considerable progress could be made by applying what is already known on plasma sheaths to the electromagnetic problem. Due concern must be paid to the limitations of the various approximations that are being made, and their region of validity, if any, must be ascertained.

b. *Antennas in Plasmas*

Here, the major problem is to determine the correct current and voltage distribution in an antenna (which may be modified by, and changes with, the sheath conditions) and the correct impedance variation. Thus, future attacks should be concerned with more realistic theoretical models. It is doubtful whether consideration of further oversimplified cases will contribute much to what is already known. The behavior of sources near resonance-frequency regions and their time dependence requires additional attention. Estimates of the contribution due to Landau damping, anisotropic thermal pressure, turbulence, etc., under various conditions also require further elucidation.

c. *Nonlinear Phenomena*

The entire field of nonlinear phenomena generated by a source in a plasma is only beginning to be explored and considerable room exists for contributions in all aspects. Of great significance is finding the limiting field strengths required under various conditions in order (a) to be able to neglect nonlinear phenomena and (b) to be able to generate specific nonlinear phenomena. Considerable future effort should be devoted to plasma heating and the associated modulation, demodulation, harmonics, frequency-mixing, etc.; to ascertaining the boundary condition appropriate under strong-field circumstances; to the effect of heat conduction in the plasma; and to noncollisional dissipation, resonances, stimulated phenomena, electroacoustic coupling to electromagnetic waves, pulse propagation in plasma, and effect of multispecies of ions.

d. *Comparison of Theory and Experiment*

A most significant deficiency in the work on sources in plasma is the lack of comparison between theory and experiment. In many cases

the theory is too idealized to permit comparison with experiment; in other cases the experiments are "too messy" to be amenable to theoretical interpretation.

This is the area on which the greatest emphasis should be placed. Theories and experiments that can be compared as directly as possible must be developed. This is essential for all three of the aforementioned areas.

REFERENCES*

1. S. R. Seshadri, "Effect of Insulation on the Radiation Resistance of an Electric Dipole in a Simple Anisotropic Medium," *Proc. IEE*, Vol. 113, p. 593 (1966).
2. R. B. Hall, "Parallel Plate Plasma Problem," *Am. Jour. Phys.*, Vol. 31, p. 696 (1963).
3. E. Wasserstrom, C. H. Su, and R. F. Probstein, "Kinetic Theory Approach to Electrostatic Probes," *Phys. Fluids*, Vol. 8, p. 56 (1965).
4. M. H. Cohen, "Radiation in a Plasma—I Cerenkov Effect," *Phys. Rev.*, Vol. 123, p. 711 (1961); "—II Equivalent Sources," *Phys. Rev.*, Vol. 126, p. 389 (1962); and "—III Metal Boundaries," *Phys. Rev.*, Vol. 126, p. 398 (1962).
5. M. I. Sancer, "Boundary Conditions for a Unique Solution to the Linearized Warm Plasma Equations," *Radio Sc.*, Vol. 1, p. 1067 (1966).
6. T. W. Johnston, "Time-Averaged Effects on Charged Particles in A-C Fields," *RCA Review*, Vol. XXI, p. 570 (1960).
7. R. Bardet, T. Consoli, and R. Geller, "Theorie du 'Bouchon Dynamique' a Gradients de Champs H.F. et Magnetique Statique," *C.R. Acad. Sc.*, Vol. 259, p. 2388 (1964).
8. H. A. Whale, "Ion Sheath Effects Near Antennas Radiating Within the Ionosphere," *Jour. Geophys. Res.*, Vol. 69, p. 447 (1964).
9. C. G. Gotmantsev and N. G. Denisov, "Concerning an Effect During Measurement of Electron Concentration in the Ionosphere by the Antenna Probe Method," *Geomagnetism and Aeronomy*, Vol. II, p. 575 (1962).

* A limited number of copies of a more extensive bibliography on Sources in Plasmas covering primarily the period 1963-1966 is available from the author on request.

40. H. Weil and D. Walsh, "Radiation Resistance of an Electric Dipole in a Magnetoionic Medium," *IEEE Trans.*, Vol. AP-12, p. 297 (1964).
41. S. R. Seshadri, "Radiation Resistance of Elementary Electric-Current Sources in a Magnetoionic Medium," *Proc. IEE*, Vol. 112, p. 1856 (1965).
42. L.R.O. Storey, "Antenna Electrique Dipole pour Reception T.R.F. dans l'Ionoshere," *L'onde Electrique*, Vol. XLV, p. 1427 (1965).
43. S. R. Seshadri and G. L. Yip, "Radiation from an Electric Dipole in an Axially Magnetized Plasma Column," *Elect. Let.*, Vol. 2, p. 30 (1966).
44. H. H. Kuehl, "Excitation of Waves in a Warm Plasma by a Current Source," *Phys. Fluids*, Vol. 6, p. 1465 (1963).
45. C. P. Wu, "Radiation from Dipoles in a Magnetoionic Medium," *Trans. IEEE*, Vol. AP-11, p. 681 (1963).
46. H. Weil and D. Walsh, "Radiation Resistance of an Elementary Loop Antenna in a Magnetoionic Medium," *IEEE Trans.*, Vol. AP-13, p. 21 (1965).
47. G. L. Duff and R. Mittra, "Input Impedance of a Small Loop of Uniform Electric Current in an Anisotropic Cold Plasma," *Elect. Let.*, Vol. 1, p. 127 (1965).
48. S. N. Samaddar, "Radiation from a Vertical Magnetic Dipole in Inhomogeneous Stratified Media Above a Horizontal Conducting Plane," *Int. Jour. Electronics*, Vol. XIX, p. 279 (1965).
49. D. C. Pridmore-Brown, "The Radiation Field of a Magnetic Dipole in a Coaxial Sheath," *Jour. Math. & Phys.*, Vol. XLIII, p. 199 (1964).
50. K. G. Balmain, "The Impedance of a Short Dipole Antenna in a Magnetoplasma," *IEEE Trans*, Vol. AP-12, p. 605 (1964).
51. S. R. Seshadri, "Radiation from Electromagnetic Sources in Plasma," *Trans. IEEE*, Vol. AP-13, p. 79 (1965).
52. K. Balmain, "Impedance of a Short Dipole in a Compressible Plasma," *Radio Sc.*, Vol. 69D, p. 559 (1965).
53. J. A. Fejer, "Interaction of an Antenna with a Hot Plasma and the Theory of Resonance Probes," *Radio Sc.*, Vol. 69D, p. 1171 (1964).
54. K. G. Balmain, "Impedance of a Spherical Probe in a Magnetoplasma," *Trans. IEEE*, Vol. AP-14, p. 402 (1966).
55. K. Balmain, "Impedance of a Radio Frequency Probe with an Absorptive Surface," *Radio Sc.*, Vol. 1, p. 1 (1966).

56. S. R. Seshadri, "Excitation of Plasma Waves in an Unbounded Plasma by a Line Source," *Trans. IEEE*, Vol. MTT-11, p. 39 (1963).

57. J. H. Harris, "Radiation Through Cylindrical Plasma Sheaths," *Radio Sc.*, Vol. 67D, p. 717 (1963).

58. S. R. Seshadri, "Radiation from an Electromagnetic Source in a Half-Space of Compressible Plasma-Surface Waves," *Trans. IEEE*, Vol. AP-12, p. 240 (1964).

59. J. R. Wait, "Radiation from Sources Immersed in Compressible Plasma," *Can. Jour. Phys.*, Vol. 42, p. 1760 (1964).

60. S. R. Seshadri, "Radiation Propagating Transverse to the External Magnetic Field from an Electromagnetic Source in an Unbounded Plasma," *Trans. IEEE*, Vol. AP-13, p. 106 (1965).

61. G. E. Stewart and P. R. Caron, "Radiation from a Line Source in a Ground Plane Covered by a Warm Plasma Slab," *Trans. IEEE*, Vol. AP-13, p. 600 (1965).

62. F. H. Mitchell and F. J. Tischer, "Radiation Through Plasma Adjoining a Conducting Surface," *Trans. IEEE*, Vol. AP-13, p. 468 (1965).

63. R. A. Hurd, "The Admittance of a Linear Antenna in a Uniaxial Medium," *Can. Jour. Phys.*, Vol. 43, p. 2276 (1965).

64. K. Aoki, "On a Cylindrical Antenna in a Homogeneous Anisotropic Medium," *Can. Jour. Phys.*, Vol. 44, p. 1239 (1966).

65. S. R. Seshadri, "Infinite Cylindrical Antennas Immersed in a Warm Plasma," *Trans. IEEE*, Vol. AP-13, p. 789 (1965).

66. M. P. Bachynski, "Electromagnetic Wave Penetration of Reentry Plasma Sheaths," *Radio Sc.*, Vol. 69D, p. 147 (1965).

67. J. Galejs, "Admittance of Annular Slot Antennas Radiating into a Plasma Layer," *Radio Sc.*, Vol. 68D, p. 317 (1964).

68. J. R. Wait, "Theory of a Slotted-Sphere Antenna Immersed in a Compressible Plasma, Parts I & II," *Radio Sc.*, Vol. 68D, p. 1127 (1964).

69. J. R. Wait and K. P. Spies, "Theory of Slotted-Sphere Antenna Immersed in a Compressible Plasma, Part III," *Radio Sc.*, Vol. 1, p. 21 (1966).

70. C. Yeh, "Electromagnetic Radiation from an Arbitrary Slot on a Conducting Cylinder Coated with a Uniform Cold Plasma Sheath with an Axial Static Magnetic Field," *Can. Jour. Phys.*, Vol. 42, p. 1369 (1964).

71. H. Hodara, "Radiation from a Gyroplasma Sheathed Aperture," *Trans. IEEE*, Vol. AP-11, p. 2 (1963).

72. H. Hodara, "The Effect of a Static Magnetic Field on a Rectangular Slot Antenna Radiating Through the Plasma Sheath," *E-M Aspects of Hypersonic Flight, The 2nd Symp. on the Plasma Sheath—Its Effect Upon Re-entry Communication and Detection*, W. R. Rotman, H. K. Moore, R. Papa, editors, Spartan Books, Inc., Baltimore (1964), p. 118.

73. V. P. Pyati and H. Weil, "Capacitance Biconical Antenns in Magnetoionic Media; Elliptical Cone Capacitance," *Radio Sc.*, Vol. 69D, p. 291 (1965).

74. J. W. Carlin and R. Mittra, "Terminal Admittance of a Thin Biconical Antenna in an Isotropic Compressible Plasma," *Elect. Let.*, Vol. 2, p. 199 (1966).

75. J. R. Wait, "Radiation from a Spherical Aperture Antenna Immersed in a Compressible Plasma," *Trans. IEEE*, Vol. AP-14, p. 360 (1966).

76. J. R. Wait, "Theories of Prolate Spheroidal Antennas," *Radio Sc.*, Vol. 1, p. 475 (1966).

77. S. R. Seshadri, "Radiation from a Current Strip in a Uniaxially Anisotropic Plasma Medium," *Can. Jour. Phys.*, Vol. 44, p. 207 (1966).

78. J. Galejs, "Driving Point Impedance of Linear Antennas in Presence of Stratified Dielectric," *Trans. IEEE*, Vol. AP-13, p. 725 (1965).

79. W.W.T. Rusch, "Radiation from a Plasma-Clad Axially Slotted Cylinder," *Radio Sc.*, Vol. 67D, p. 203 (1963).

80. C. T. Swift, "Radiation Patterns of a Slotted-Cylinder Antenna in the Presence of an Inhomogeneous Lossy Plasma," *Trans. IEEE*, Vol. AP-12, p. 726 (1964).

81. D. L. Sengupta, "Radiation Field Produced by an Infinite Slot in an Infinite Cylinder Surrounded by a Homogeneous Plasma Sheath," *Proc. IEE*, Vol. 111, p. 950 (1964).

82. W. Rotman, "Plasma Simulation by Artificial Dielectric and Parallel Plate Media," *Trans. IEEE*, Vol. AP-10, p. 82 (1962).

83. K. E. Golden and T. M. Smith, "Simulation of a Thin Plasma Sheath by a Plane of Wires," *Trans. IEEE*, Vol. NS-11, p. 225 (1964).

84. T. M. Smith and K. E. Golden, "Radiation Pattern of a Slot Covered by a Simulated Plasma Sheath," *Trans. IEEE*, Vol. AP-13, p. 285 (1965).

85. K. E. Golden, "Plasma Simulation with an Artificial Dielectric in a Horn Geometry," *Trans. IEEE*, Vol. AP-13, p. 587 (1965).

86. T. M. Smith and K. E. Golden, "Radiation Patterns from a Slotted Cylinder Surrounded by a Plasma Sheath," *Trans. IEEE*, Vol. AP-13, p. 775 (1965).

87. G. Tyras, P. C. Bargeliotes, J. M. Hamm and R. R. Schell, "An Experimental Study of Plasma Sheath Effects on Antennas," *Radio Sc.*, Vol. 69D, p. 839 (1965).

88. G. G. Cloutier and M. P. Bachynski, "Antenna Radiation Patterns in the Presence of a Plasma Sheath," *E-M Aspects of Hypersonic Flight, The 2nd Symp. on the Plasma Sheath—Its Effect Upon Re-entry Communication and Detection*, W. Rotman, H. K. Moore, R. Papa, editors, Spartan Books, Inc., Baltimore, p. 169 (1964).

89. D. J. Jacavanco, "Electromagnetic Properties of Plasma Covered Antenna," *Radio Sc.*, Vol. 69D, p. 965 (1965).

90. G. Meltz and H. J. Schmidt, "Observation of the Effect of Magnetic Fields on the Radiation from Plasma Covered Antennas," *Trans. IEEE*, Vol. AP-13, p. 169 (1965).

91. H. A. Whale, "The Excitation of Electroacoustic Waves by Antennas in the Ionosphere," *Jour. Geophys. Res.*, Vol. 68, p. 415 (1963).

92. R. F. Mlodnosky and D. K. Garriott, "The VLF Admittance of a Dipole in the Lower Ionosphere," *Proc. Int. Conf. on the Ionosphere*, Inst. of Phys. & Phys. Soc., p. 484 (1963).

93. J. R. Herman, "A Method of Determining D-Region Electron Density Profile Utilizing a Capacitance Impedance Rocket Probe," *Jour. Geophys. Res.*, Vol. 69, p. 2329 (1964).

94. K. D. Baker, A. M. Despain, and J. C. Ulwick, "Simultaneous Comparison of R-F Probe Techniques for Determination of Ionospheric Electron Density," *Jour. Geophys. Res.*, Vol. 71, p. 935 (1966).

95. H. Oya, "Effects of Resonances on the Admittance of an R-F Plasma Probe Surrounded by an Ion Sheath," *Rep. of Ionosphere & Space Research in Japan*, Vol. 19, p. 243 (1965).

96. J. E. Hall and J. Fooks, "The Electron Distribution in the Quiet D-Region Derived from Rocket Measurements of Low-Frequency Propagation," *Planet. and Space Sc.*, Vol. 13, p. 1013 (1965).

97. K. T. Yen, "Non-Steady Response of Plasmas to Microwave Radiation," *Phys. Fluids*, Vol. 7, p. 1612 (1964).

98. R. J. Papa, "The Nonlinear Interaction of an E-M Wave with a Time Dependent Plasma Medium," *Can. Jour. Phys.*, Vol. 43, p. 38 (1965).

99. R. J. Papa, "Form of the Electron Distribution Function in a Time-Varying Plasma," *Phys. Fluids*, Vol. 8, p. 1408 (1965).

100. R. J. Papa and C. T. Case, "The Nonlinear Interaction of a Radio Frequency Wave with an Inhomogeneous Plasma Slab: I Kinetics of High Temperature Air in the Presence of an E-M Field," *Can. Jour. Phys.*, Vol. 43, p. 2021 (1965).

101. R. J. Papa and C. T. Case, "The Nonlinear Interaction of a Radio Frequency Wave with an Inhomogeneous Plasma Slab: II Nonlinear Reflection and Transmission Coefficient of an Inhomogeneous Plasma Slab," *Can. Jour. Phys.*, Vol. 43, p. 2021 (1965).

102. D. A. Clavier and L. Webb, "Electromagnetic Breakdown of Air," *Jour. Appl. Phys.*, Vol. 37, p. 742 (1966).

103. B. L. Ward, "Calculation of Electrical Breakdown in Air at Near Atmospheric Pressure," *Phys. Rev.*, Vol. 138, p. A1357 (1965).

104. A. G. James, H. McL. Ryan, and C. W. Powell, "Electric Strength of Compressed Air in a Uniform Field," *Nature*, Vol. 210, p. 403 (1966).

105. A. D. McDonald, D. V. Gaskell, and H. N. Gitterman, "Microwave Breakdown in Air, Oxygen, and Nitrogen," *Phys. Rev.*, Vol. 130, p. 1841 (1963).

106. W. E. Scharfman, W. C. Taylor, and T. Morita, "Breakdown Limitations on the Transmission of Microwave Power Through the Atmosphere," *Trans. IEEE*, Vol. AP-12, p. 709 (1964).

107. J. O. Thomas, "Canadian Satellite: The Topside Sounder Alouette," *Science*, Vol. 139, p. 229 (1963).

108. G.E.K. Lockwood, "Plasma and Cyclotron Spike Phenomena Observed in Topside Ionograms," *Can. Jour. Phys.*, Vol. 41, p. 190 (1963).

109. W. Calvert and G. B. Goe, "Plasma Resonances in the Upper Ionosphere," *Jour. Geophys. Res.*, Vol. 68, p. 6113 (1963).

110. R. L. Smith, N. M. Brice, J. Katsufrakis, D. A. Gurnett, S. D. Shawhan, J. S. Belrose, and R. E. Barrington, "An Ion Gyrofrequency Phenomenon Observed in Satellites," *Nature*, Vol. 204, p. 274 (1964).

111. I. P. Shkarofsky, "Dielectric Tensor in Vlasov Plasmas Near Cyclotron Harmonics," *Phys. Fluids*, Vol. 9, p. 561 (1966).

112. I. P. Shkarofsky, "Dispersion of Waves and Cyclotron Harmonic Resonance Regions in Plasmas," *Phys. Fluids*, Vol. 9, p. 570 (1966).

113. J. A. Fejer and W. Calvert, "Resonance Effects of Electrostatic Oscillations in the Ionosphere," *Jour. Geophys. Res.*, Vol. 69, p. 5049 (1964).

114. J. A. Fejer, "Excitation of the Lower Hybrid Resonance by an Antenna in the Ionosphere," *Radio Sc.*, Vol. 1, p. 447 (1960).

115. I. P. Shkarofsky and T. W. Johnston, "Cyclotron Harmonic Resonances Observed by Satellites," *Phys. Rev. Let.*, Vol. 15, p. 51 (1965).

116. E. L. Hagg, "Remote Cyclotron Resonance Phenomena Observed by the Alouette Satellite," *Nature*, Vol. 210, p. 927 (1966).

117. D. B. Muldrew and E. L. Hagg, "A Novel Ionospheric Cyclotron Resonance Phenomena Observed in Alouette I Data," *Can. Jour. Phys.*, Vol. 44, p. 925 (1966).

118. F. W. Crawford, G. S. Kino, and H. H. Weiss, "Excitation of Cyclotron Harmonic Resonances in a Mercury Vapour Discharge," *Phys. Rev. Let.*, Vol. 13, p. 229 (1964).

119. S. J. Buchsbaum and A. Hasegawa, "Excitation of Longitudinal Plasma Oscillations Near Electron Cyclotron Harmonics," *Phys. Rev. Let.*, Vol. 12, p. 685 (1964).

120. F. W. Crawford, "A Review of Cyclotron Harmonic Phenomena in Plasmas," *Nuclear Fusion*, Vol. 5, p. 73 (1965).

121. R. M. Hall and D. E. Kaplan, "Cyclotron Resonance Echo," *Phys. Rev. Let.*, Vol. 14, p. 1062 (1965).

122. R. W. Gould, "Echo Phenomena," *Phys. Let.*, Vol. 19, p. 477 (1965).

123. W. H. Kegel and R. W. Gould, "On the Theory of Pulse Stimulated Radiation From a Plasma," *Phys. Let.*, Vol. 19, p. 531 (1965).

124. G. F. Herrman and R. F. Whitmer, "Echo Processes in a Plasma," *Phys. Rev.*, Vol. 143, p. 122 (1966).

125. F. W. Crawford and R. S. Harp, "Electron Cyclotron Echoes from Plasmas," *Phys. Let.*, Vol. 21, p. 292 (1966).

126. S. A. Colgate, "Collisionless Plasma Shock," *Phys. Fluids*, Vol. 2, p. 485 (1959).

127. S. P. Zagorodnikov, L. I. Rudakov, G. E. Smolkin, and G. V. Sholin, "Observations of Shock Waves in a Collisionless Plasma," *Soviet Phys. JETP*, Vol. 20, p. 1154 (1965).

128. J. H. Malmberg and C. B. Wharton, "Collisionless Damping of Electrostatic Plasma Waves," *Phys. Rev. Let.*, Vol. 13, p. 184 (1964).

129. V. V. Dolgopolov, A. I. Ermakov, N. I. Nazarov, K. N. Stepanov, and V. T. Tocok, "Experimental Observation of Landau Damping in a Plasma," *Soviet Phys. JETP*, Vol. 18, p. 866 (1964).

THE DEPOSITION OF VITREOUS SILICON DIOXIDE FILMS FROM SILANE

By

N. Goldsmith* and W. Kern#

Summary—A process is described for depositing uniform films of vitreous silicon dioxide by the controlled oxidation of silane. Deposition rates of 1500 Å min^{-1} at 325°C, and of 2900 Å min^{-1} at 475°C can be obtained. The equipment used to deposit the films and typical deposition parameters are described. Some physical and chemical properties of the films, including dielectric constant, dielectric strength, refractive index, and dissolution rate are discussed and compared with silicon dioxide films obtained by the thermal oxidation of silicon. Changes in the film properties and various heat treatments are compared and evaluated. Units employing this process are currently in use in producing SiO_2 layers for solid-state devices.

Introduction

A NUMBER of methods of obtaining films of vitreous silicon dioxide have been reported. These include the pyrolysis of organo-oxysilanes,[1] the vapor phase hydrolysis of silicon tetrachloride,[2] the sputtering of fused silica,[3] and the glow-discharge decomposition of ethyl silicate.[4] The objective of the work described here was to develop a practical process, using a minimum of complex equipment, for the rapid deposition at low temperatures of uniform films of silicon dioxide free from organic impurities. The oxidation of silane was chosen as the basis for such a method:

$$SiH_{4\,(g)} + 2O_{2\,(g)} \rightarrow SiO_{2\,(s)} + 2H_2O_{(g)}$$

Silane is a gas that ignites spontaneously in air to form a powdery deposit of colloidal silica that, as such, is useless in semiconductor technology. However, it was found that the reaction can be readily controlled by diluting the silane with an inert gas prior to reacting

* RCA Electronic Components and Devices Division, Somerville, N.J.
# RCA Laboratories, Princeton, N.J.

[1] E. L. Jordan, "A Diffusion Mask for Germanium," *Jour. Electrochem. Soc.*, Vol. 108, p. 478, May 1961.

[2] T. L. Chu, J. R. Gavaler, G. A. Gruber, and Y. C. Kao, "A Novel Technique for the Deposition of Silica Films," *Jour. Electrochem. Soc.*, Vol. 111, p. 1433, Dec. 1964.

[3] P. D. Davidse and L. I. Maissel, "Dielectric Thin Films Through RF Sputtering," *Jour. Appl. Phys.*, Vol. 37, p. 574, Feb. 1966.

[4] L. L. Alt, S. W. Ing, Jr., and K. W. Laendle, "Low-Temperature Deposition of Silicon Oxide Films," *Jour. Electrochem. Soc.*, Vol 110, p. 465, May 1963.

it with oxygen. By a proper choice of reaction conditions, transparent and adherent films of silicon dioxide can be deposited on a wide variety of materials at temperatures ranging from 250° to 550°C. Three types of reaction apparatus were developed in the course of this project: (1) a small prototype, (2) an intermediate-size research/production unit, and (3) a large-scale unit that is currently being used as a pro-

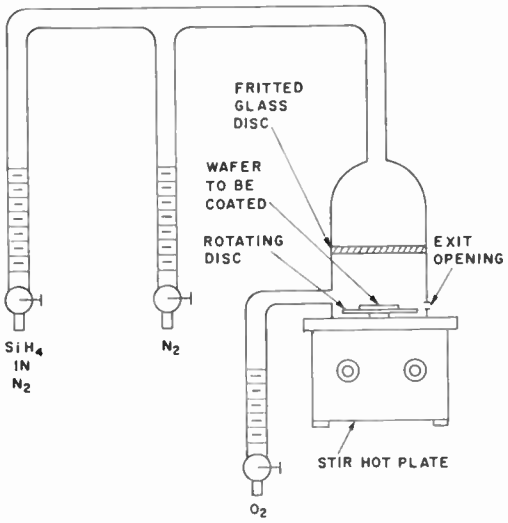

Fig. 1—Prototype deposition apparatus.

duction unit. While all three units are based on the same principle, each has specific features and capabilities. The principle of the technique and typical results for each of the systems are described here.

General Description Of Systems

The source material in all three systems was commercially available high-purity silane diluted with either nitrogen or argon. Dilutions of 3 and 10% by volume were used and these were further diluted with 99.998% pure nitrogen during metering into the system. Tank oxygen of 99.7% purity was used. The nitrogen and oxygen were passed through 0.22-micron Millipore* filters. In all three systems, the heat sources are hot plates on which the samples are rotated inside the glass reaction chamber.

A schematic drawing of the prototype apparatus of 50 cm^2 deposition area is shown in Figure 1. The diluted silane passes through a

* Registered trade name.

coarse-fritted Pyrex* disc sealed to the top of the deposition chamber. The frit serves to reduce the reaction of the silane in the upper part of the chamber and provide a uniform flow over the sample area. Oxygen is metered into the deposition chamber below the fritted disc. The gases exit from the deposition chamber through small slots cut into the rim. The samples are placed on an aluminum disc turning on a needle-point bearing. The disc is rotated by the eddy currents induced by a magnet rotating beneath the hot plate.

The intermediate-size research/production unit of 80 cm^2 deposition area was designed to permit the use of higher temperatures and variable modes and speeds of sample rotation. It uses reaction chambers having variable geometries for the gas inputs.

The factory unit was designed with a deposition area of 200 cm^2 and uses a standard hot plate modified to include a motor-driven rotating disc. This system represents a simple scale-up of the prototype system.

As will be shown, the deposition rate is temperature sensitive, requiring ±5°C control if reproducible deposition rates are to be obtained. The thermostatic element in the hot plate was adequate at the lower temperatures. However, when operating near the upper limit of the hot plate, an auxiliary temperature controller operating from a chromel–alumel thermocouple embedded in the plate was used. Temperatures on the wafer surface were checked using a surface thermometer calibrated at the melting point of cadmium. *The safety precautions followed are outlined in the Appendix.*

Deposition Parameters

Prototype System

The initial investigation of the effects of the several variables was made using the prototype system. The layers were deposited on polished silicon substrates that were cleaned prior to deposition by degreasing followed by etching in aqueous hydrofluoric acid. Since the growth of the silicon dioxide layer can be easily followed by observing the successive changes of interference color of the growing film, the growth rates were obtained by determining the time required to obtain a film thickness of approximately 6,000 Å. The exact thickness was then determined by multiple-beam interferometry or from the reflection spectrum of the film.[5]

---

* Registered trade name.

[5] N. Goldsmith and L. A. Murray, "Determination of Silicon Oxide Thickness," *Solid State Electronics*, Vol. 9, p. 331, April 1966.

The effect of varying the oxygen and silane flows is shown in Figure 2 for a deposition temperature of 330°C. The silane flow rates given are for the pure gas and were calculated from the actual flow rates and the tank concentration as specified by the supplier. For the range investigated, the rate of growth per cm^3 of silane was

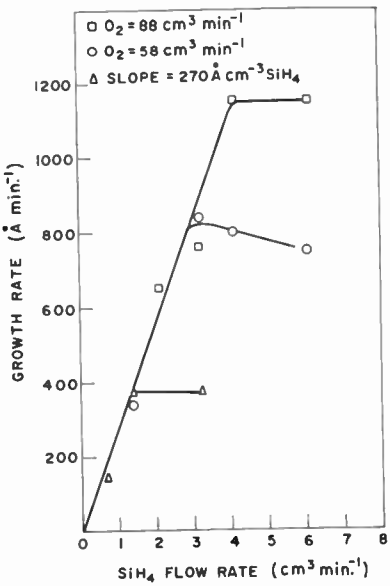

Fig. 2—Effect of oxygen on growth rate (prototype system, 50 cm^2 deposition area, nitrogen flow rate of 2420 cm^3 min^{-1}, substrate temperature 330°C).

essentially constant until an oxygen-to-silane ratio of about 20 was reached. At this point, the deposition rate became relatively independent of the silane flow rate. As the oxygen-to-silane ratio decreased, however, a fog of colloidal silica particles formed in the vapor phase. Since these particles settle on the growing film, depositions for the purposes of device fabrication are best carried out under conditions corresponding with those near the knee of the curves. The effect on the growth rate of lowering the temperature is illustrated by the results presented in Table I.

Intermediate-Size System

All work with the intermediate-size system was conducted at a constant carrier-gas (nitrogen) flow of 1900 cm^3 min^{-1}. The deposition rate as a function of substrate temperature and oxygen-to-silane ratio

Table I—Effect of Temperature on Growth Rate in Prototype System (50 cm^2 deposition area; N_2 flow rate 2420 cm^3 min^{-1}, SiH_4 3.4 cm^3 min^{-1}, O_2 88 cm^3 min^{-1})

Deposition Temperature, C°	Deposition Rate, Å min^{-1}
330	810
310	450
280	no deposit in 10 min

of 3.1 is shown in Figure 3 (data points given by circles). Deposition begins at about 250°C, and the rate rises steeply until, at 300°C a temperature-insensitive plateau occurs that extends to 350°C. With further increases in the temperature, the deposition rate again increases. Experiments to determine the optimum oxygen-to-silane ratios at various temperatures were also conducted. At 325°C the maximum

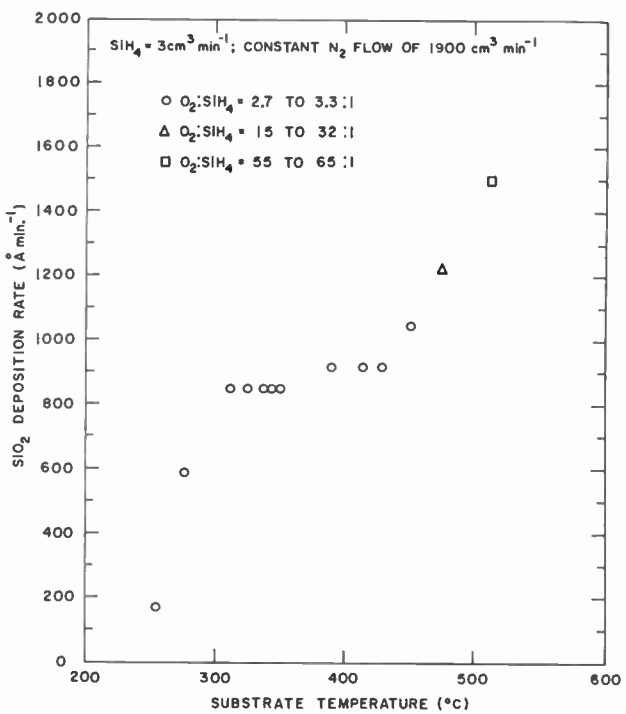

Fig. 3—SiO_2 deposition rate as a function of substrate temperature for different O_2:SiH_4 ratios.

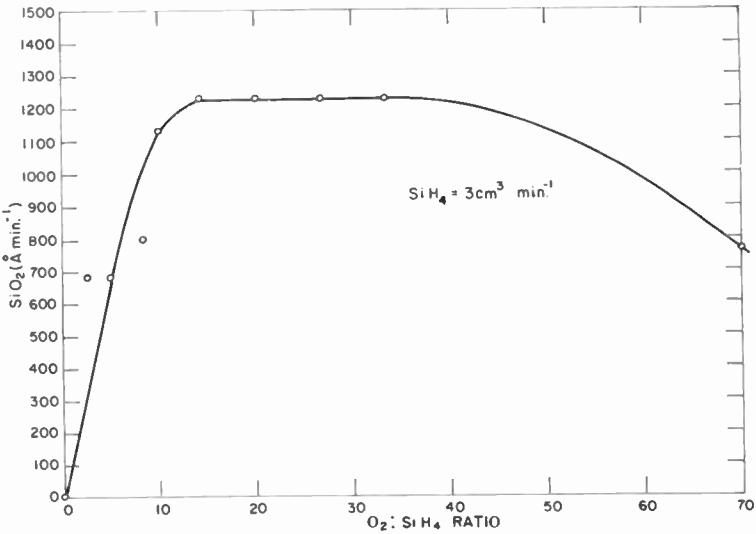

Fig. 4—SiO_2 deposition rate as a function of O_2:SiH_4 ratio at 475°C.

deposition rate was achieved with an oxygen-to-silane ratio of $(3 \pm 10\%)$: 1, the rate decreasing rapidly outside of these limits. This ratio gave optimum results up to 450°C. Figure 4 shows deposition rate as a function of oxygen-to-silane ratio at 475°C. Here, the optimum ratio occurs over the range 14:1 to 33:1. At 515°C, the

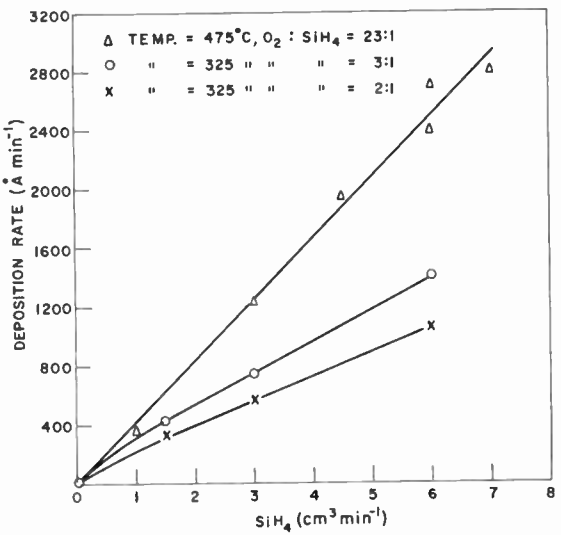

Fig. 5—Deposition rate as a function of SiH_4 flow (intermediate-size system).

optimum ratio rises to 55:1 to 65:1. Data points taken within the optimum ratio ranges for 475°C and 515°C are included in Figure 3.

Fixed substrate temperatures of 325°C and 475°C were chosen for an investigation of the effects of silane flow rate on the deposition rate. These isotherms are shown in Figure 5. Except for very low flow rates, the deposition rate increases in proportion to the silane flow. At 515°C, a deposition rate of 3,500 Å min^{-1} was obtained for a silane flow of 7 cm^3 min^{-1}.

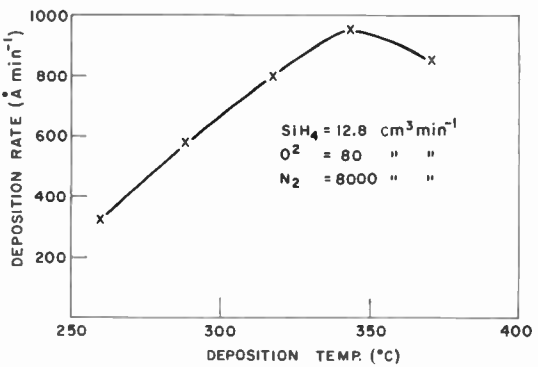

Fig. 6—Deposition rate as a function of substrate temperature (large system).

Large-Scale System

Although the full-size production system is modeled on the prototype, the results obtained differ in several respects. In particular, the growth rate is less temperature sensitive, and lower oxygen-to-silane ratios are required for optimum results. The effect of temperature on the deposition rate in this system is shown in Figure 6. Results obtained when the silane and oxygen flows are varied at constant temperature are shown in Figure 7, while Figure 8 shows the dependence of the rate on the nitrogen flow.

Film Properties

Some of the important properties of the films for application to semiconductor processing are uniformity, etch rate, "pin-hole" density, masking capability, and purity. In general, the films obtained on silicon wafers are comparable in all their properties, with the exception of etch rate, to those normally found for oxides thermally grown on silicon.

On properly cleaned wafers, the films are uniform across the entire

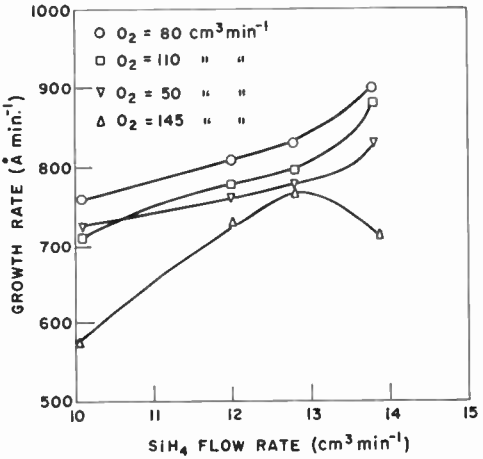

Fig. 7—Effect of silane and oxygen flow rates on growth rate (large system, 200 cm^2 deposition area, nitrogen flow rate of 8000 cm^3 min^{-1}, substrate temperature 320°C).

surface of a 1.5-inch-diameter silicon wafer, 6,000 ± 200 Å being a typical example of the capability of the method. The deposition of a 1,500 to 3,000 Å layer serves as a very sensitive test for the cleanliness of a wafer. The deposition rate and the quality of the films are strongly influenced by the nature of the substrate surface. Improper precleaning of the silicon results in film defects such as non-uniform thickness, pin holes, and particulate contaminants.

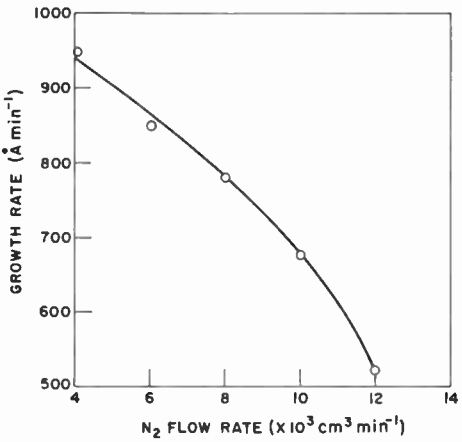

Fig. 8—Dependence of deposition rate on the nitrogen flow rate (large system, 200 cm^2 deposition area, oxygen flow rate of 80 cm^3 min^{-1}, silane flow rate of 12 cm^3 min^{-1}, temperature 320° C).

Despite the mismatch in thermal expansion coefficients, hard and adherent films of SiO_2 can be deposited on silicon. The maximum thickness of the film appears to depend upon the deposition rate and temperature. These observations are summarized in Table II.

The cracking of the film cited in Table II is not due solely to the difference in thermal coefficient of expansion of silica and silicon.

Table II—Effect of Temperature and Deposition Rate on Maximum Film Thickness

Temperature, C°	Deposition Rate, Å min^{-1}	Result
325	400	no cracking at 20,000 Å
325	1,000	cracks at 16,000 Å
475	2,500	no cracking at 16,000 Å
515	3,000	no cracking at 23,000 Å

First, the films crack before the wafer is cooled; second, deposition on quartz plates also results in cracking of the film. A possible explanation for the breakup of the film is that the film undergoes a densification during the time it takes for the deposition, and the shrinkage of the lower portion of the film causes the entire layer to crack.

The purity of the layers is adequate for most applications. The silane, as supplied, contains several parts per billion of boron. Conversion of n-type silicon of 10 ohm-cm resistivity was not observed except after 16 hours of heating at 1200°C in argon. An extensive study of the masking capabilities has not been made, but the films have been use successfully in place of thermally grown oxides as diffusion masks for boron and phosphorus.

Table III gives results of experiments to compare silicon dioxide films deposited at various temperatures with thermally grown oxides using one-micron-thick films on 0.01 and 10 ohm-cm silicon wafers. The samples were prepared under the following conditions:

(1) low temperature and low deposition rate (325°C, 760 Å min^{-1})

(2) high temperature and high deposition rate (475°C, 2,500 Å min^{-1})

(3) thermal oxidation (1250°C, steam)

The samples prepared on the low-resistivity wafers were used to determine the dielectric response of the films. The higher-resistivity silicon substrate was used for the optical analyses because of its higher infrared transmittance. Samples with deposited SiO_2 were

Table III—Comparative Data for Deposited and Grown One-Micron-Thick SiO_2 Films on Silicon

Method of oxide Preparation	Thickness by Tolansky Interferometry, Å	Refractive Index	Dissolution Rate* at 25°C, Å sec^{-1}	Dielectric Constant at 1 KHz	Dielectric Field Maximum Sustained, V_{DC} cm^{-1}
From SiH_4 325°C 760 Å min^{-1}					
(a) as deposited	10,190	1.43	87.3	5.73	2.5×10^6
(b) 12 min 770°C Argon	9,560	1.46	30.3	4.75	3.7×10^6
(c) 5 min 1250°C Argon	9,670	1.46	21.3		
From SiH_4 475°C 2500 Å min^{-1}					
(a) as deposited	10,340	1.46	81.1	4.30	3.4×10^6
(b) 12 min 770°C Argon	9,930	1.46	33.4	4.93	4.0×10^6
(c) 5 min 1250°C Argon	9,570	1.49	18.3		
Thermal Oxidation 1250°C., steam 10,000 Å $hour^{-1}$					
(a) As grown	9,650	1.47	18.1	3.83	5.2×10^6

* Standard $HF\text{-}NH_4F$ buffer solution: 454 g ammonium fluoride crystals; 654 cm^3 distilled water; 163 cm^3 hydrofluoric acid solution 49% (used after 24 hours of equilibration).

also heat treated for various periods of time in argon, air, or steam at 550°, 770°, and 1250°C.

Measurements of thickness and refractive index were carried out interferometrically. Infrared absorption and reflection spectra were taken on a Perkin–Elmer Model 137B double-beam spectrophotometer. One half of each high-resistivity sample wafer was stripped of its coating and used in the reference beam to cancel out absorptions introduced by the substrate. Reflection spectra were referred to an aluminum mirror. Parallel-plate capacitors were prepared using vacuum-evaporated, circular, palladium–gold counter electrodes of 1.06×10^{-2} cm^2 and 9.22×10^{-4} cm^2 areas. Determinations of the metal electrode area were made with a Nikon Model 6C Shadowgraph at an estimated precision of 1%. The capacitance values were obtained using a type 716-C General Radio Company capacitance bridge. The d-c dielectric strength of the films was tested with an oscilloscope by stepwise application of voltage until the breakdown voltage was reached. The values reported are the maximum fields the capacitors could take without breaking down. All dielectric tests were made with about eight capacitors per sample. The values listed in Table III are the weighted averages of the small- and the large-area capacitors.

From the results presented in Table III, it can be concluded that deposition at 475°C yields films with improved initial physical properties as indicated by a lower dielectric constant, a higher dielectric strength, a lower dissolution rate, a higher index of refraction and an infrared spectrum more closely resembling that of thermally grown SiO_2 films.

The etch rates of the deposited films, without subsequent heat treatment, are almost five times greater (in buffer etch) than those obtained for the thermally oxidized layers. The etch rates can be reduced by heat treating the deposited films, the etch rate decreasing with increasing temperature of heat treatment. However, prolonging the heat treatment at fixed temperatures above some saturation time has little effect on the etch rate. This can be seen in Figure 9, where the dissolution rates of oxides deposited at 475°C are plotted as a function of heating time. Heating at 1250°C yields an oxide with an etch rate indistinguishable from a thermal oxide within the first 2.5 minutes.

The infrared spectra for deposited silica films were similar to those noted by other workers.[6,7] Typical reflection spectra are shown in

[6] W. A. Pliskin and H. S. Lehman, "Structural Evaluation of Silicon Oxide Films," *Jour. Electrochem. Soc.*, Vol. 112, p. 1013, Oct. 1965.

[7] L. A. Murray and N. Goldsmith, "Nondestructive Determination of Thickness and Perfection of Silica Films," *Jour. Electrochem. Soc.*, Vol. 113, p. 1297, 1966.

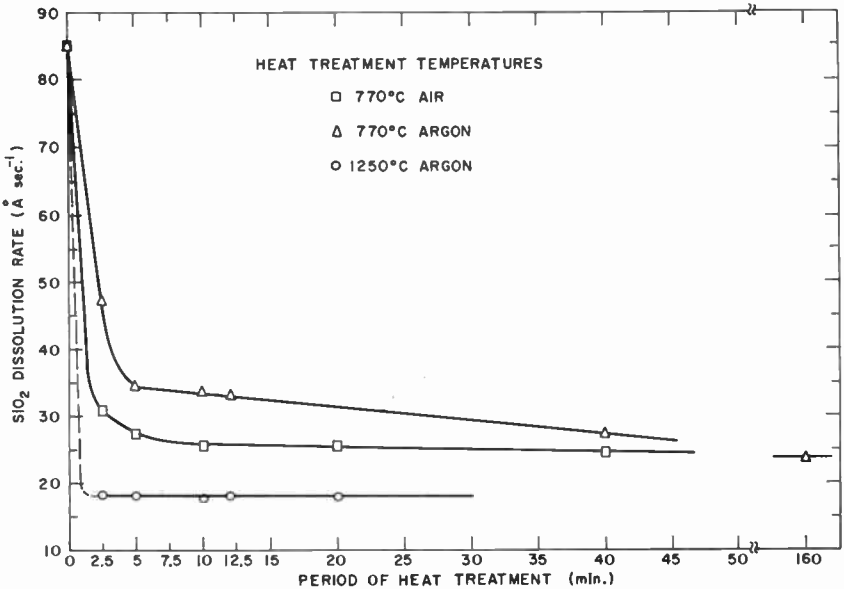

Fig. 9—SiO_2 dissolution rate in $HF\text{-}NH_4F$ buffer solution as a function of heating temperature, time, and ambient.

Figure 10 for oxides deposited at 350°C and for a densified oxide. The pronounced minimum near 8 microns is found to be a useful indicator of the densification process, since, unlike the peak at 9 microns, the wavelength position for the 8-micron minimum is invariant with thickness. Deposited oxides have this minimum shifted toward longer wavelengths. The intensity of this minimum was found to be sensitive to the angle of incidence. The spectra in Figure 10 were taken at a 30° incidence angle.

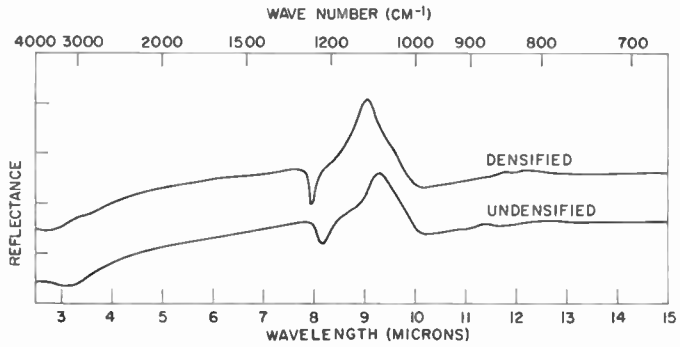

Fig. 10—Reflectivity of deposited silicon dioxide.

CONCLUSIONS

The oxidation of silane has been shown to yield high-quality films of silica. The process provides a new freedom in the fabrication of planar devices, since it is no longer necessary to make compromises in processing in order to obtain an oxide layer. Although most of the work has been directed toward the manufacture of silicon devices, it is obvious that the films can be applied to the fabrication of devices in other materials, such as germanium and gallium arsenide.

APPENDIX—SAFETY PRECAUTIONS

Silane is a toxic, poisonous gas requiring some care in handling. The entire gas system must be free of leaks to prevent decomposition from occurring within the gas lines. The deposition system should be installed in an exhaust hood and the tank storage should be exhausted to the outside. The systems installed at RCA's production facilities are designed with interlocked solenoid valves so that the flow of silane is automatically cut off if the hood door is opened or if the hood exhaust should fail. The pressure regulators should have metal diaphragms so that the connection between the tank and the pressure regulator can be evacuated. This is required to prevent the formation of particles of SiO_2 which may clog or otherwise interfere with the operation of the regulator.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the important contributions made by H. Hyman in the studies of the large-scale system, and wish to thank J. DePiazza, J. Velebir, and R. Heim for their competent technical assistance. A portion of this work was performed under Contract AF 33(657)-11615, sponsored by the Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Ohio.

RCA Technical Papers

Fourth Quarter, 1966

Any request for copies of papers listed herein should be addressed to the publication to which credited.

"Crystalline Solid Lasers," Z. J. Kiss and R. J. Pressley, *Proc. IEEE* (October) 1966

"Designing Silicon-Transistor Hi-Fi Amplifiers. Pt. 2. Circuit Configurations," R. D. Gold and J. C. Sondermeyer, *Electronics World* (October) 1966

"Discharge Dynamics in the Large Diameter Magnetically Confined Ion Gas Laser," S. A. Ahmed and T. J. Faith, Jr., *Proc. IEEE* (Letters) (October) 1966

"Physical Significance of the Power-Amplification/Bandwidth Product of an Active Device," E. O. Johnson, *Proc. IEEE* (Letters) (October) 1966

"Solid-State Inductive Element Using Magneto Resistance," M. Toda, *Proc. IEEE* (Letters) (October) 1966

"Some Performance Data on an Operational 4.2°K Parametric Amplifier," J. A. Stovman, *Proc. IEEE* (Letters) (October) 1966

"Structure of Superconductive Vortices Near a Metal-Air Interface," J. Pearl, *Jour. Appl. Phys.* (October) 1966

"Underwater Photography," L. E. Mertens, *Jour. SMPTE* (October) 1966

"Vehicular Interference Radiation Measurement Technique," J. R. Neubauer, *Trans. IEEE PTGVC* (October) 1966

"Critical Magnetic Field of Thin Superposed Films," R. Klein and G. Fischer, *Phys. Rev.* (7 October) 1966

"Demodulation of Low-Level Broad-Band Optical Signals with Semiconductors, Part III: Experimental Study of the Photoconductive Detector," H. S. Sommers, Jr., and E. K. Gatchell, *Proc. IEEE* (November) 1966

"Designing Silicon-Transistor Hi-Fi Amplifiers. Pt. 3. Performance of Practical Circuits," R. D. Gold and J. C. Sondermeyer, *Electronics World* (November) 1966

"Emitter Sidewall Junction Capacitance in Double-Diffused Transistors," N. Rumin, *Proc. IEEE* (Letters) (November) 1966

"Flux Jumps in a Hard Superconductor in Superimposed ac and dc Magnetic Fields," D. A. Gandolfo, C. M. Harper, and R. L. Hecht, *Jour. Appl. Phys.* (Communications) (November) 1966

"Intra-Cavity Color Selection in Ion Lasers," C. P. Wen, J. M. Hammer, I. Gorog, F. W. Spong, and J. A. vanRaalte, *IEEE Jour. Quantum Electronics* (November) 1966

"Linear Integrated Circuits," B. V. Vonderschmitt and R. L. Sanquini, *Electronics World* (November) 1966

"Man-In-Space Communications Subsystem," R. E. Davis, *Signal* (November) 1966

"Silicon Transistors in Audio Amplifiers," J. C. Sondermeyer, *Trans. IEEE PTGBTR* (November) 1966

"TV Applications of MOS Transistors," W. M. Austin, J. A. Dean, D. M. Griswold, and O. P. Hart, *Trans. IEEE PTGBTR* (November) 1966

"Tunneling Induced by Longitudinal Microwave Phonons in Al-Pb, Al-Sn, and Pb-Pb Superconducting Diodes," Y. Goldstein, B. Abeles, and R. W. Cohen, *Phys. Rev.* (4 November) 1966

"Exchange Interactions in Ferromagnetic Chromium Chalcogenide Spinels," P. K. Baltzer, P. J. Wojtowicz, M. Robbins, and E. Lopatin, *Phys. Rev.* (11 November) 1966

'Magnetoabsorption in Single-Crystal Semiconducting Ferromagnetic Spinels," G. Harbeke and H. Pinch, *Phys. Rev. Letters* (21 November) 1966

"The Acoustoelectric Effects and the Energy Losses by Hot Electrons—Part II," A. Rose, *RCA Review* (December) 1966

'Analysis of Parametric Action in Back-Biased P-N Junctions Carrying Injected Current," F. Sterzer, *RCA Review* (December) 1966

'Apertured Sheet Read-Only Memory," J. C. Miller and C. M. Wine, *Trans. IEEE PTGMAG* (December) 1966

"Data Display for Real Time Telemetry," A. P. Moll, *Trans. IEEE PTGCOM* (December) 1966

"Dielectric Resonance Probes and Double Probes Compared," F. H. C. Smith and T. W. Johnston, *Jour. Appl. Phys.* (Communications) (December) 1966

"The Frequency Modulation Feedback System for the Lunar-Orbiter Demodulator," F. Lefrak, H. Moore, A. Newton, and L. Ozolins, *RCA Review* (December) 1966

"A High-Frequency Representation of the MOS Transistor," H. Johnson, *Proc. IEEE* (Letters) (December) 1966

"The Hologram—Properties and Applications," E. G. Ramberg, *RCA Review* (December) 1966

"Infrared Reflectivity of Heavily Doped Low-Mobility Semiconductors. I. GaAs." L. A. Murray, J. J. Rivera, and P. A Hoss, *Jour. Appl. Phys.* (December) 1966

"Ionization and Displacement Damage in Silicon Transistors," G. J. Brucker, W. J. Dennehy, and A. G. Holmes-Siedle, *Trans. IEEE PTGNS* (December) 1966

"Irradiation of MIS Capacitors with High Energy Electrons," K. H. Zaininger, *Trans. IEEE PTGNS* (December) 1966

"Lithium-Doped Radiation-Resistant Silicon Solar Cells," J. J. Wysocki, *Trans. IEEE PTGNS* (December) 1966

"Lorentz Microscopy of Growing Domains in Permalloy Films," M. D. Coutts and H. Weinstein, *Jour. Appl. Phys.* (Communications) (December) 1966

"A Radiation-Induced Instability in Silicon MOS Transistors," W. J. Dennehy, G. J. Brucker, and A. G. Holmes-Siedle, *Trans. IEEE PTGNS* (December) 1966

"Saturation of Absorption and Fluorescence in Solutions of Phthalocyanines," W. F. Kosonocky and S. E. Harrison, *Jour. Appl. Phys.* (December) 1966

"Self-Pumped Parametric Amplification with an Avalanching Diode," A. S. Clorfeine, *Proc. IEEE* (Letters) (December) 1966

"Superconducting Properties of Nb_3Sn and of Multiphase Nb-Sn Alloys," R. E. Enstrom, *Jour. Appl. Phys.* (December) 1966

"Synchronization During Biased PCM Conditions," E. D. Bloedel, *RCA Review* (December) 1966

"Threshold Performance of Analog FM Demodulators," J. Frankle, *RCA Review* (December) 1966

"Use of Phase Subtraction to Extend the Range of a Phase-Locked Demodulator," A. Acampora and A. Newton, *RCA Review* (December) 1966

"X-Band Microstrip Circulator," B. Hershenov, *Proc. IEEE* (Letters) (December) 1966

"Large-Wave-Number, Almost Transverse Helicon Propagation, and the Specific Heat of Helicons," M. A. Lampert, J. J. Quinn, and S. Tosima, *Phys. Rev.* (9 December) 1966

"Photoemission of Electrons from n-Type Degenerate Silicon into Silicon Dioxide," A. M. Goodman, *Phys. Rev.* (9 December) .. 1966

"Photoemission of Holes from Silicon into Silicon Dioxide," A. M. Goodman, *Phys. Rev.* (9 December) 1966

"Application of Overlay Transistors to Solid State Mobile Equipment," N. G. Richards and F. A. Barton, *IEEE WESCON*, Part 3, *Electron Devices & Packaging* 1966

"High-Power Microwave Varactors and Varactor Multipliers," J. R. Collard, C. Sun, and T. E. Walsh, *IEEE WESCON*, Part 3, *Electron Devices & Packaging* 1966

"Microwave Power Transistors," H. C. Lee, *IEEE WESCON*, Part 3, *Electron Devices & Packaging* 1966

"Microwave Solid-State Multipliers for Space Systems," W. J. Dodds, *IEEE WESCON*, Part 3, *Electron Devices & Packaging* 1966

"A Novel Spacecraft Antenna Array," W. T. Patton, *IEEE WESCON*, Part 5, *Systems, Space Electronics* 1966

"RF Power Transistors in Vehicular Radio Communications Equipment," S. Matyckas, *IEEE WESCON*, Part 3, *Electron Devices & Packaging* 1966

"Semiconductor High-Frequency Power-Amplifier Design," R. Minton, *IEEE WESCON*, Part 6, *Instrumentation, Electronic Systems, Components* 1966

AUTHORS

MORREL P. BACHYNSKI received the B.E. degree in Engineering Physics in 1952 and the M.Sc. degree in 1953, both from the University of Saskatchewan, and the Ph.D. degree from McGill University in 1955. Until October 1955, as a member of the Eaton Electronics Research Laboratory, McGill University, he was engaged in investigations on microwave optics. Since that time, he has been with the RCA Victor Research Laboratories, Montreal, Canada, concerned primarily with electromagnetic wave propagation, microwaves and plasma physics. He is presently Director of the Microwave and Plasma Physics Laboratories. Dr. Bachynski is a senior member of the Institute of Electrical and Electronics Engineers, a member of the American Physical Society and of the Canadian Association of Physicists. He served on the Defence Research Board of Canada Advisory Committee on Gas Dynamics (1960-1963), is Chairman of Commission VI of the Canadian National Committee of the International Scientific Union (URSI) and is associated with the Graduate School of McGill University.

NORMAN GOLDSMITH received the B.A. degree with honors in chemistry form Hunter College in 1959 and the M.S. degree in physical chemistry from Stevens Institute of Technology in 1964. He joined the Electronic Components and Devices Division of RCA in 1959 working on elemental purification and the growth of bulk and epitaxial gallium arsenide. More recently he has been active in all phases of semiconductor materials and process technology including diffusion, epitaxial growth, vapor deposition, surface stabilization, and development of nondestructive measurement techniques. Mr. Goldsmith is a member of The American Chemical Society and The Electrochemical Society.

ISTVAN GOROG received the B.Sc. degree in Electrical Engineering in 1961 from the University of California, Berkeley, and subsequently, the M.Sc. (1962) and Ph.D. (1964) degrees from the same University. His doctoral studies concerned radiation losses from high-temperature plasmas as possible sources of light. Between 1957 and 1964, he was employed by the Wesix Company, San Francisco, studying the natural ion content of air; the E-H Laboratories, Oakland, developing microwave circuit components; and with the University of California Electronics Research Laboratories working on lasers and plasmas. He joined RCA Laboratories, Princeton, N. J., in 1964, where his research has centered on high-power high-efficiency gas lasers and laser components.

Dr. Gorog is a member of Eta Kappa Nu and the American Physical Society.

WERNER KERN received a certificate in chemistry in 1944 from the University of Basle, Switzerland, and a Swiss state diploma in chemical technology in 1946. His thesis work was on the chromatographic isolation of fluorescing polynuclear hydrocarbons, which he discovered existed in soil. He was employed as an analytical research chemist until 1948 with Hoffman-La Roche in Switzerland, and then joined the research department of the same company in New Jersey, specializing in the development of radiochemical methods. He received a BA degree in chemistry from Rutgers University in 1955. In 1958 he joined the research division of Nuclear Corporation of America as chief chemist directing research and development work in applied radiochemistry. He joined RCA Electronic Components and Devices in 1959 as radiochemist and radiological safety officer, conducting radiochemical research in semiconductor technology and process development. He is now at RCA Laboratories, Princeton, N. J., where he is engaged in the refinement of glassing and associated processes for the hermetic sealing of microelectronic solid-state devices.

Mr. Kern is a member of the American Chemical Society, the Electrochemical Society, the Research Society of America, the Franklin Mineralogical Society, and the Geological Society of New Jersey.

W. T. PATTON received his BSEE and MS degrees from the University of Tennessee in 1952 and 1958, respectively. He received his PhD in Electrical Engineering from the University of Illinois in 1963. While on the staff of the University of Tennessee Experiment Station, he did research on circular arrays. He was an instructor at the University of Illinois, and he conducted research on log-periodic and simply periodic antennas at the Antenna Laboratory. He was also a consultant with the Radio Direction Finding Laboratory. For three years he was Ship Superintendent for DER conversions at the Boston Naval Shipyard. Since joining the RCA Missile and Surface Radar Division, Moorestown, N. J., in 1962, he has done research on large phased arrays.

Dr. Patton is a member of the Institute of Electrical and Electronics Engineers, AAAS, Sigma Xi, Tau Beta Pi, and Eta Kappa Nu.

A. H. SOMMER received the Dr. Phil. degree in Physical Chemistry from Berlin University in 1934. From 1935 to 1945, he worked with Baird Television Company, London, in the field of photoelectric and secondary emission. In 1946 he joined EMI Research Laboratories in Hayes, England, to work on photomultipliers and television camera tubes. In 1953 he joined RCA Laboratories, where he is engaged in development and study of new photoemissive materials.

Dr. Sommer is a Fellow of the Institute of Electrical and Electronics Engineers and a member of the Physical Society and of Sigma Xi.

FRED W. SPONG received the B. S. degree in physics from the University of Utah, Salt Lake City, in 1958, and the Ph.D. degree in physics from the University of California at Berkeley. His dissertation involved an experimental investigation in Fermi surface physics, namely cyclotron resonance in aluminum. In 1964 he joined RCA Laboratories, Princeton, N. J., where his studies have centered on high power, high efficiency gas lasers.

Dr. Spong is a member of the American Physical Society, Sigma Ki, and Phi Beta Kappa.

MARTIN C. STEELE received the B.Ch.E. (Cum Laude) from Cooper Union Institute of Technology in 1940, and the M.S. and Ph.D. degrees in Physics from the University of Maryland in 1949 and 1952, respectively. He worked as an engineer for the Office of the Chief of Engineers, before the war, and as a research physicist in the Naval Research Laboratory from 1947 to 1955, where he was head of the Cryomagnetic Research Group in the Solid State Division. Since 1955 he has been with RCA, as head of the Solid State Electronics Research Group and for three years (1960-63) as Research Director of RCA's research laboratory in Tokyo, Japan. In August, 1963 he became head of the Plasma Physics Group of the Microwave Research Laboratory at RCA Laboratories, Princeton, New Jersey. His research has covered a wide range of areas in solid-state physics, including superconductivity, magnetic susceptibility, thermometry, theory of magnetic properties of metals, thermal, electric and magnetic properties of semiconductors, and the properties of plasmas in solids.

Dr. Steele is a fellow of the American Physical Society, a member of Sigma Xi and a former member of the Washington Academy of Science.

GEORGE A. SWARTZ received his B.S. degree from Massachusetts Institute of Technology in 1952 and the M.S. and Ph.D. degrees in physics from the University of Pennsylvania in 1954 and 1958, respectively. He joined the technical staff of RCA Laboratories in 1957. His work has been in the field of plasma propulsion, plasma stability, microwave phenomena in gaseous and solid-state plasmas, and solid-state physics.

Dr. Swartz is a member of the American Physical Society and Sigma Xi.