RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH • ENGINEERING

VOLUME XVIII SEPTEMBER 1957 NO. 3

PRINTED IN U.S.A.

RCA REVIEW, published quarterly in March, June, September, and December by RCA Laboratories, Radio Corporation of America, Princeton, New Jersey. Entered as second class matter July 3, 1950 at the Post Office at Princeton, New Jersey, under the act of March 3, 1879. Subscription price in the United States and Canada; one year $2.00, two years $3.50, three years $4.50; in other countries: one year $2.40, two years $4.30, three years $5.70. Single copies in the United States, $.75; in other countries, $.85.

RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH • ENGINEERING

Published quarterly by

RCA LABORATORIES

in cooperation with all subsidiaries and divisions of

RADIO CORPORATION OF AMERICA

VOLUME XVIII SEPTEMBER, 1957 NUMBER 3

CONTENTS

RCA REVIEW is regularly abstracted and indexed by *Industrial Arts Index Science Abstracts* (I.E.E.-Brit.), *Electronic Engineering Master Index, Chemical Abstracts, Proc. I.R.E.,* and *Electronic & Radio Engineer.*

RCA REVIEW

BOARD OF EDITORS

A TRANSISTORIZED HORIZONTAL-DEFLECTION SYSTEM

By

Hunter C. Goodrich

RCA Victor Television Division,
Camden, N. J.

Summary—*The good switching characteristics of junction transistors make them basically well suited to the generation of horizontal-deflection current in television receivers. A completely transistorized developmental horizontal-deflection system is described consisting of an oscillator, driver, output stage with ultor supply, and phase detector. Experimentally, 90-degree deflection at 10 kilovolts has been achieved with the use of selected audio-type power transistors in the output stage. To achieve this deflection it was necessary to operate these transistors beyond the manufacturer's current and voltage ratings. The results demonstrate the feasibility of the circuits and indicate the characteristics that would be desirable in an output transistor designed specifically for this application. In particular, expressions are given for the effect of output transistor switching power capability, saturation resistance, and cutoff time on the circuit performance.*

Introduction

THE developmental horizontal-deflection system described in this paper is shown in block diagram form in Figure 1. A pulse waveform at horizontal scanning frequency is developed in the horizontal oscillator, amplified by the driver, and applied to the deflection output stage. The output stage delivers a sawtooth current waveform to the horizontal winding of the deflection yoke and also supplies a high-amplitude flyback pulse which can be rectified to produce kinescope ultor voltage. A signal from the output stage is fed back to the phase detector where its phase is compared with that of incoming sync. The resulting error voltage is passed through an integrating circuit and used to control the oscillator frequency and phase.

Basic Output Circuit

In considering the system, it seems best to begin with the output stage since the reactive power required from it can be determined from measurements on conventional (vacuum-tube) television receivers. The output requirements of the preceding stages depend in turn on the drive requirements of the output stage. There are several methods by which a transistor might be used to develop a sawtooth deflection current. For example, it might be used, with suitable impedance

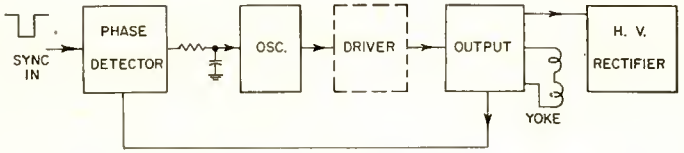


Fig. 1—Block diagram of horizontal-deflection system.

changes, in the same manner that a tube is normally used for deflection in a television receiver. In this system a damper diode conducts the current during the first portion of scan and the output tube is used as a controlled impedance element to control the current during the completion of scan.

Another method has been described by Guggi[1] in which energy is fed into the yoke during retrace and a large capacitance switched into the circuit to control the current during forward scan. The basic circuit to be discussed here represents a third method which is simple in structure and potentially very efficient with respect to the power dissipated in the output transistor.

It has been shown[2] that a linear sawtooth current can be achieved in an inductance shunted by a capacitance with the use of a voltage source and a bidirectional resistanceless or ideal switch as shown in Figure 2. If switch S is closed at time t_1, the current will increase linearly according to

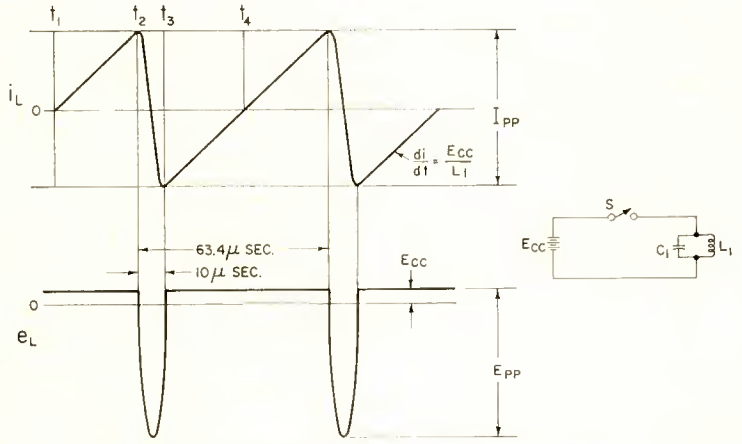


Fig. 2—Basic output circuit and waveforms.

[1] W. B. Guggi, "Retrace Driver Deflection Circuit," *I.R.E. Transactions on Broadcast and Television Receivers*, Vol. PGBTR-2, p. 65, October 1956.

[2] G. C. Sziklai, "Current Oscillator for Television Sweep," *Electronics*, Vol. 19, p. 120, September 1946.

$$\frac{di}{dt} = \frac{E_{cc}}{L_1} \tag{1}$$

until the switch is opened at time t_2. The current then flowing in L_1 will discharge through C_1 in an oscillatory manner. If S is then closed after one-half cycle of free oscillation, the reverse current flowing in L_1 will flow into the power supply, decreasing linearly to zero at t_4. The cycle then repeats.

In this analysis it was assumed that switch S has negligible resistance to currents flowing in either direction. Sziklai discovered that a junction transistor can be operated to provide a reasonable approxi-

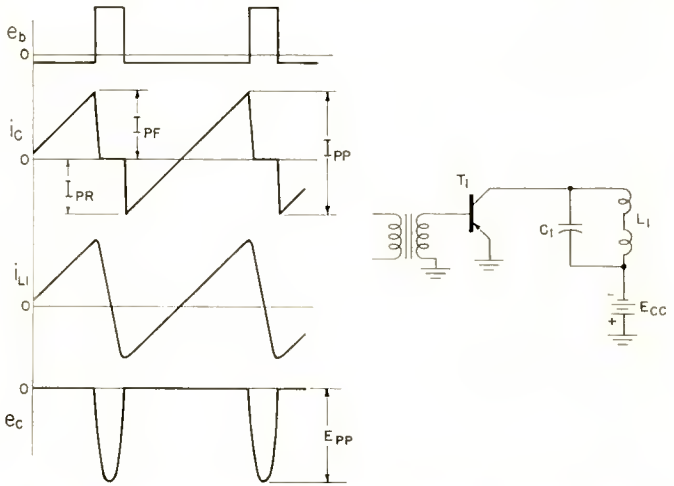


Fig. 3—Basic transistor output circuit and waveforms.

mation to the required switch.[3] It will pass current bi-directionally between emitter and collector. When the transistor is saturated, the emitter-to-collector resistance is a fraction of an ohm; when it is cut off, this resistance may be in megohms. The basic output circuit incorporating a transistor is shown in Figure 3, together with the corresponding waveforms. The operation is analogous to that of Figure 2. During scan, the base of T_1 is biased sufficiently in the forward direction to place the transistor in saturation. At the end of scan the base is driven in the reverse direction by the applied drive signal, cutting off T_1, while L_1C_1 goes through a half cycle of free oscillation. C_1 is chosen to give a desired return time with the particular yoke inductance employed.

[3] G. C. Sziklai, "Symmetrical Properties of Transistors and Their Applications," *Proc. I.R.E.*, Vol. 41, p. 717, June 1953.

During this flyback interval a relatively large voltage pulse appears on the collector. The yoke inductance must be adjusted so that the peak voltage does not exceed the collector breakdown voltage. At the end of the base pulse, which should coincide with the completion of flyback, T_1 conduction is restored for the next scanning period. The peak T_1 collector current reached prior to flyback is in the normal or forward direction and is designated I_{pf} as indicated in Figure 3. The reverse peak following flyback is termed I_{pr} and the peak-to-peak collector current, I_{pp}.

Figure 4 illustrates a method of reducing the flyback voltage pulse on the collector. L_2C_2 is tuned approximately to an odd harmonic of the flyback frequency—in this case the third harmonic. The voltage e_{L_2} across L_2C_2 is a damped sine wave of third-harmonic frequency which is initiated by the front edge of the flyback pulse. This voltage is 180 degrees out of phase with the voltage across C_1 at the center of the flyback pulse and thus reduces the peak collector voltage about

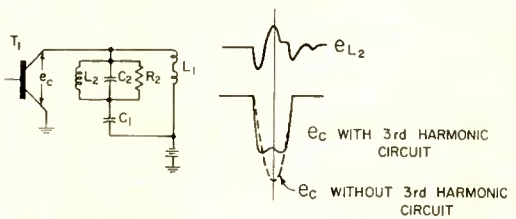


Fig. 4—Reduction of flyback peak with 3rd harmonic circuit.

30 per cent. The shape of the resulting collector pulse e_c is indicated by the solid-line waveform in Figure 4.

When a high-voltage winding is placed on the horizontal-output transformer, a similar result can be achieved with a design such that the distributed capacitance and leakage inductance resonate at an odd harmonic of the flyback frequency. This has become common practice in television receivers employing vacuum tubes.

Switching-Power Requirements

It is now of interest to determine the desired specifications on the output switching device, i.e., the output transistor. For this purpose the problem of deflecting a 90-degree kinescope with a 10-kilovolt ultor voltage will be considered. The results can be extrapolated for other deflection angles and ultor voltages.

It is evident that for a given yoke geometry and ultor voltage the deflection angle, D, will be proportional to the peak-to-peak ampere turns in the horizontal winding.

$$D \propto I_{pp} N \propto I_{pp} \sqrt{L}, \tag{2}$$

where I_{pp} = peak-to-peak yoke current,

N = number of turns in yoke winding,

L = yoke inductance.

Further, for a given return time and flyback pulse shape,

$$E_{pp} \propto L I_{pp}, \tag{3}$$

where E_{pp} = peak-to-peak yoke voltage during scanning cycle.

Then from Equations (2) and (3),

$$D \propto (I_{pp} E_{pp})^{1/2}. \tag{4}$$

Thus, the deflection that can be achieved with a particular transistor is determined by the product of its peak-to-peak current handling capability and the peak voltage it can withstand. This represents the bidirectional switching power capability. The bidirectional switching power employed in a vacuum-tube television receiver with a 90-degree kinescope and a 10-kilovolt ultor supply has been measured at 1,300 volt-amperes. This value has been confirmed with measurements on a transistor deflection stage. Selected transistors of commercially available types have been found which are capable of this switching power, although such operation is beyond the manufacturers' ratings on peak voltage and current.

For a given flyback waveshape the voltage pulse applied to the collector during flyback is proportional to the inductance seen from the collector. If the yoke is direct-coupled, this is the yoke inductance itself; if transformer coupling is used, the inductance reflected to the collector determines the voltage. To obtain full advantage of the switching power capabilities of a transistor, this inductance should be such that the maximum safe pulse voltage is developed on the collector when the desired deflection has been achieved. With a fixed retrace pulse shape and return time, there is a fixed ratio between the flyback pulse amplitude and the supply voltage, E_{cc}. With a 9-microsecond retrace time and third-harmonic cancellation, this relationship is

$$E_{pp} \approx 10\, E_{cc}. \tag{5}$$

Because the drive requirement for the output transistor decreases with

the collector current to be controlled, and because the power-supply filter cost tends to decrease for higher load impedances, it seems desirable to have transistors with higher breakdown voltages than are now available. The collector junction breakdown and punch-through voltages determine the maximum permissible collector peak voltage. There is little tendency for the output stage to "run away" because of the switch-type mode of operation. During forward scan, collector current is controlled by the load circuit; during retrace, the base is reverse biased so that the collector current is limited to the I_{co} value of the collector junction.

When the deflection power and maximum voltage are specified, the peak-to-peak current to be handled is automatically determined. Thus, if the required switching power is 1,300 volt-amperes, and the safe collector voltage is 130 volts, the peak-to-peak current is 10 amperes. In any practical circuit this current swing will be greater in one direction than in the other. The yoke energy following retrace is that which was in the system prior to retrace less the flyback losses and energy delivered to the ultor supply. The dissymmetry is thus a function of the ratio of dissipated to stored energy in the system. Typically, the current might be 60 per cent in the forward direction, 40 per cent in the reverse direction as illustrated in Figure 3.

Transistor Saturation Resistance

The series saturation resistance, R_s, that can be tolerated between emitter and collector during forward scan is a function of the resulting nonlinearity which is acceptable. In a purely inductive circuit, the current will increase linearly as indicated in Equation (1). Any resistance in the circuit will result in a sawtooth voltage component being added to the applied voltage, E_{cc}. The variation of the deflection current slope from edge to edge is in proportion to the variable component added to E_{cc}. Thus R_s results in linearity distortion from edge to edge, as follows:

$$\text{Per cent nonlinearity} = \frac{I_{pp} R_s}{E_{cc}} \times 100. \qquad (6)$$

This nonlinearity will be added to others that will arise from other resistances in the circuit, primarily in the yoke winding and in the transformer if a transformer is used. In the example where E_{cc} is 13 volts and I_{pp} is 10 amperes, if R_s is allowed to contribute a 5 per cent nonlinearity, Equation (6) indicates a maximum R_s of 0.065 ohm. Note that since E_{cc} is proportional to the allowable breakdown voltage

and I_{pp} is inversely proportional (for a given deflection power), the allowable R_s increases with the square of the permissible collector voltage.

The value of R_s also largely determines the power dissipated in the transistor during forward scan. In the example being considered, if R_s is 0.065 ohm, this dissipation is less than 1 watt.

Switching Time

With presently available transistors, a large portion of the dissipation in the output transistor occurs during collector current cutoff. The collector flyback pulse starts at the beginning of the collector-current cutoff interval. It is shown in Figure 5 that when third harmonic cancellation is used with a return time of 10 microseconds, the voltage increases approximately linearly to the value E_{pp} in 3 microseconds. To calculate the transistor dissipation during cutoff,

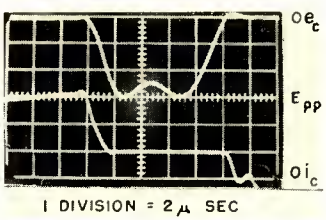


Fig. 5—Collector voltage and current waveforms during retrace.

it is assumed that the current cutoff and voltage build-up are linear. From the waveforms of Figure 5, this approximation seems reasonable. Thus during the cutoff interval,

$$e_c \approx \frac{t}{3} E_{pp} \qquad \text{when } t < 3 \text{ microseconds,} \qquad (7)$$

$$i_c \approx I_{pf}\left(1 - \frac{t}{T_{co}}\right) \qquad \text{when } t \leq T_{co} \text{ microseconds,} \qquad (8)$$

where $t =$ time from beginning of cutoff,

$T_{co} =$ time required for i_c cutoff.

Dividing the energy dissipated during cutoff by the duration of a horizontal period gives the average collector power dissipation due to cutoff, P_{co}.

$$P_{co} = \frac{1}{63.4} \int_{t=0}^{t=T_{co}} i_c e_c \, dt = \frac{T_{co}^2}{1140} E_{pp} I_{pf}, \qquad (9)$$

where I_{pf} = peak collector current prior to flyback.

This expression is approximately correct for cutoff times up to 3 microseconds. In the chosen example, if I_{pf} is 6 amperes, the average dissipation due to cutoff is 0.69 watt when T_{co} is 1 microsecond and varies with square of T_{co}. The desirability of minimizing cutoff time is obvious.

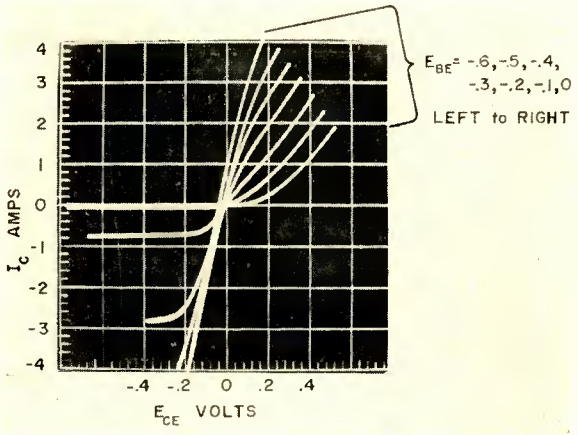


Fig. 6—Bilateral saturated collector characteristic.

Static Characteristics of Present Power Transistors

One pertinent characteristic that the manufacturers of power transistors usually fail to give is the bilateral collector–emitter characteristic in the saturation region. As an example, Figure 6 shows such a characteristic for an RCA type 2N301A power transistor. Collector–emitter voltage versus collector current is shown for several values of forward base-to-emitter voltage. It will be noted that the saturation resistance is lower and remains low over a wider range with higher values of base bias. This resistance reaches an average of 0.05 ohm over the 8-ampere peak-to-peak current range plotted. Constant base voltages were used in Figure 6 because in the operating circuit a low-impedance driving source is used. However, the forward driving power depends on the input resistance when the transistor is in the saturated state and on the base current required to maintain that state. The

minimum base current required to maintain saturation during the last half of scan is

$$I_{b\min} = \frac{I_{pf}}{\beta}, \tag{10}$$

where β = base–collector current gain at I_{pf}.

As noted, the saturation resistance will be less if a base current in excess of $I_{b\min}$ is used during scan.

The fact that the reverse collector current peak, I_{pr}, is approximately 40 per cent of the peak-to-peak current swing, I_{pp}, indicates the desirability of using an output transistor with approximately symmetrical current gains to maintain saturation during the first half of scan. In practice, the forward base voltage applied during scan may have some slope due to lack of low-frequency response in the driving transformer. This slope is in a direction to increase the relative base current during the first portion of scan and thus decreases to a degree the relative reverse direction current gain required. If a highly unsymmetrical transistor is used, a diode can be placed in shunt with the transistor to carry the current in the reverse direction.

Cutoff Behavior of Present Power Transistors

Most existing commercial power transistors have not been designed for rapid switching, and they turn off more slowly than is desirable in the horizontal-deflection application. Nevertheless, with proper drive, such units will function in the circuit described. If one were to attempt to cut off a collector current of several amperes in a typical power transistor by merely reducing the forward base bias to zero, the cutoff time would be many times the total permissible flyback time. However, if the cutoff pulse applied to the base is of sufficient amplitude to bias it in the reverse direction with respect to the emitter, a transient reverse base current will flow as stored minority carriers are removed from the base layer and the collector current cutoff time will be reduced. The cutoff time has been found to vary approximately in proportion to the peak collector current and in inverse proportion to the applied reverse base bias. The base and collector waveshapes during cutoff are shown in Figure 7. The interval between the initial application of the cutoff pulse to the base and the completion of collector current cutoff has two components—the *storage time*, T_s, during which the transistor remains saturated and continues to act as a closed switch, and the *cutoff time*, T_{co}, which represents the interval of collector current decay.

The duration of T_s is essentially zero when the base current during scan has the value I_{bmin} but may increase to several microseconds with higher levels of base current. In a sync-driven deflection system, T_s will introduce a time delay, but in the automatic-phase-control system to be described, the oscillator phasing will automatically shift to eliminate any phase discrepancy of the output deflection current with respect to sync. As long as the saturation resistance remains low during the storage interval, the duration of this interval has no serious effect on the system.

The collector flyback pulse builds up during the cutoff interval. The importance of minimizing T_{co} has already been discussed. To make presently available power transistors cut off as rapidly as practically

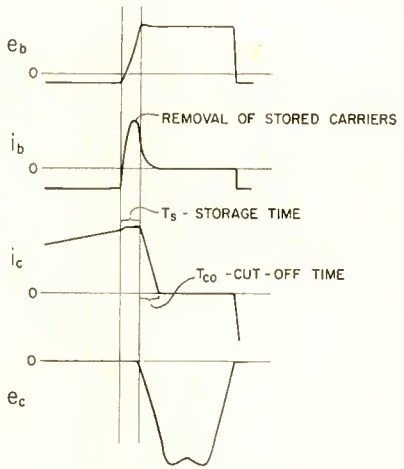


Fig. 7—Base and collector waveforms during cutoff.

possible, a reverse base pulse on the order of 10 volts has been employed. Depending on the output transistor used, up to several amperes of reverse base current may flow during the cutoff process and a peak drive power as high as 40 watts may be required to achieve a cutoff time of the order of 3 microseconds. When inherently faster transistors with the required switching capacity are developed, it is anticipated that the peak drive power required will be much less.

Developmental Output Stage

A developmental deflection and high-voltage output circuit is shown in Figure 8. A bootstrap output circuit is used so that the transistor collector and case can be grounded for best heat conduction. A 1-milli-

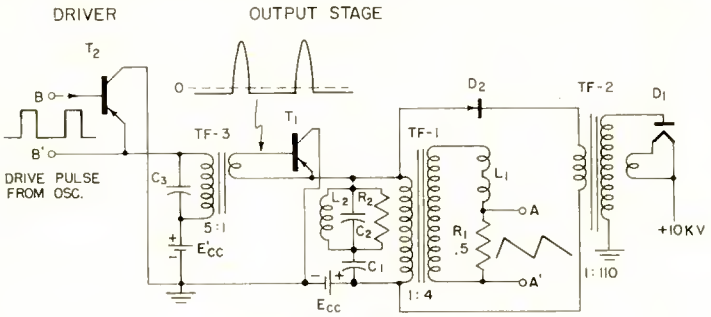


Fig. 8—Output stage and driver.

henry yoke is used with transformer TF1 which reflects an inductance of 62 microhenrys at the transistor T_1. High voltage is taken from transformer TF2 which reflects an inductance of 400 microhenrys at the transistor. C_1 tunes the reflected inductance to produce the desired flyback period. L_2C_2 provides third harmonic cancellation as previously described. Diode D_2 opens during scan and prevents damped oscillatory voltages in TF2 from causing "ripples" in the deflection current. An oscillogram of the output current in the deflection yoke is shown in Figure 9.

A voltage reference to the deflection current appears across sampling resistor R_1. This signal is applied to the phase detector which is discussed later. The 1.2-volt peak-to-peak sawtooth voltage developed across R_1 has about a 2 per cent effect on linearity. If a lower impedance yoke were used, this effect would become more serious because the sawtooth component would represent a greater proportion of the voltage across the yoke during scan.

Driver Stage

The relatively large amount of drive power needed to make present power transistors cutoff sufficiently fast makes the use of a driver stage, T_2, desirable, as shown in Figure 8. This stage is operated in

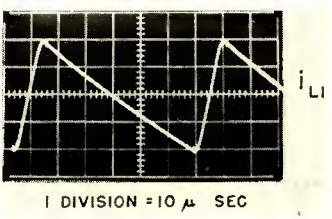


Fig. 9—Deflection yoke current.

the same manner as described for the basic output stage. It is saturated during scan and cut off by a pulse from the oscillator during retrace. During retrace a pulse analogous to the flyback pulse appears across the primary of TF3. This pulse, after being stepped down in voltage to reduce the circuit impedance, is used to cut off the output stage. The pulse has the shape of one half a sine-wave cycle. One might think that a steeper-front pulse would produce faster cutoff of the output current; however, this is not necessarily true. If the output stage is operated with sufficient forward bias to produce a storage time of a few microseconds, the cutoff pulse will be close to its full amplitude during the cutoff interval. It is during this interval that carriers should be removed as rapidly as possible to reduce the cutoff time.

A-C coupling the drive signal results in a d-c axis about 10 per cent above the signal level during scan, as indicated by the T_1 base waveform shown in Figure 8. Thus a forward bias equal to about one-tenth the peak-to-peak drive amplitude is applied to the T_1 base. This eliminates the need for an added bias supply and protects the output stage in case of drive failure. When faster deflection output transistors become available, it seems probable that the driver stage can be eliminated.

Oscillator Stage

The horizontal oscillator and phase detector circuits are shown in Figure 10. The oscillator is of the blocking type in which T_3 is saturated during flyback and cutoff during scan. This on–off mode of operation results in relatively high peak output compared with the transistor-dissipated power. Much of the dissipation occurs on the transitions between the off and saturated states, so that the transition intervals should be brief. Since this stage must supply its own drive and drive the following stage as well, the oscillator transistor needs more high-frequency gain than is possessed by the type used in the output stage. The maximum peak output from the oscillator is proportional to its switching power capability. The transistor used is an experimental type with a maximum oscillation frequency of the order of 10 megacycles, a maximum collector voltage of 35 volts, and a peak current rating of 0.5 ampere.

The oscillator collector voltage and current waveforms are shown in Figures 11a and 11b, respectively, with the oscillator loaded by the driver stage. The oscillator transition intervals between on and off states vary manyfold with loading, being longer when the load is heavy. The output waveform applied to the driver stage as shown in Figure 11c is about 6 volts in amplitude. The peak power delivered is about

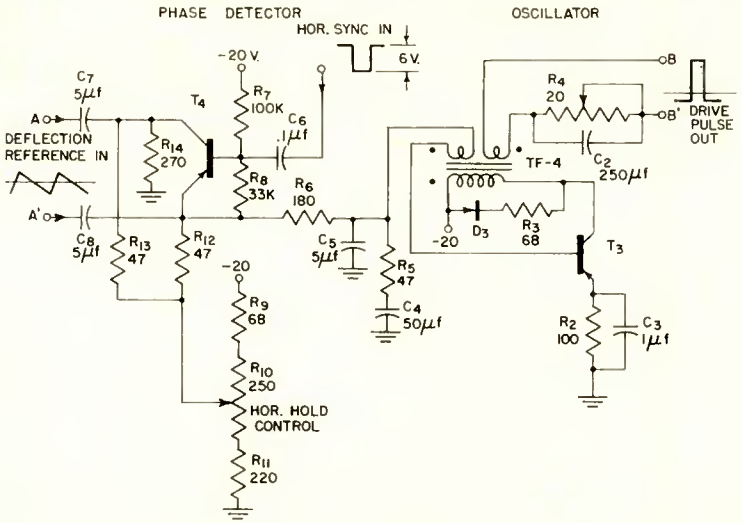


Fig. 10—Oscillator and phase detector.

2 watts. The fact that the load is nonlinear limits the power that can be delivered to the driver, because the impedance match must be compromised to accommodate the low base impedance offered during the forward-scan portion of the cycle and the higher base impedance existing during the retrace portion of the cycle. The oscillator repetition frequency and pulse width depend primarily on the load R_2, C_3 and

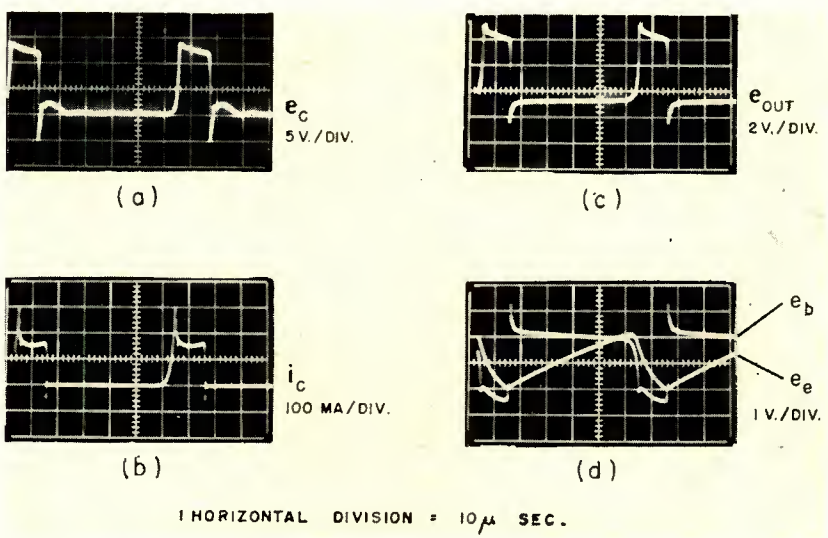


Fig. 11—Oscillator waveforms.

the ratio of T_3 base bias to collector supply voltage (Figure 10). For a given load, adjustment of the other three factors gives a wide range of control over the repetition frequency and pulse duty factor. The base and emitter voltage waveforms which determine the oscillator timing are shown in Figure 11d. In operation the oscillator frequency is controlled by the base bias, which is determined by the hold control setting and phase detector output.

Diode D_3 (Figure 10) serves to limit the transformer inductive overshoot during turn-off so that the collector breakdown voltage is not exceeded. The output signal to be applied to the following stage is taken from a third winding on the oscillator transformer. Forward bias for the driver is obtained by a-c coupling the input signal in the same manner as on the output stage. Since the forward bias tends to be greater than is required, the operating point is shifted somewhat in the reverse direction by bias resistor R_4 (Figure 10).

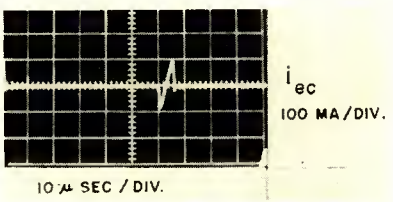


Fig. 12—Phase detector current.

Phase Detector

In Figure 10 transistor T_4, which is a symmetrical p-n-p type, functions as a horizontal phase detector.[3] The signal developed across R_1 (Figure 8), which corresponds to the deflection current sawtooth waveform, is applied between emitter and collector of T_4. Negative sync pulses applied to the base of T_4 permit emitter–collector current to flow during the sync interval, while self-bias developed by base current flow through R_8 keeps the transistor cut off between sync pulses. When the horizontal system is properly phased, the duration of retrace will overlap the sync interval. As indicated by the waveform shown in Figure 10, the collector-to-emitter voltage of T_4 is negative during the first portion of retrace and positive during the last portion. The instantaneous direction of emitter–collector current flow during the sync pulse is determined by this polarity.

This current is shown in Figure 12 for the case where retrace is approximately centered on sync. The integrated voltage output under this condition is essentially zero. If the phasing shifts, the current

flow becomes unsymmetrical and a d-c component is developed across R_{12} (Figure 10). This error signal is applied through the anti-hunt integrating circuit $R_6C_4C_5R_5$ to the oscillator base winding and acts to oppose the original phasing error. The frequency sensitivity of the oscillator to changes in base bias is on the order of 2,000 cycles per volt. The impedance of the phase detector is made relatively low to reduce the d-c internal error voltage drops due to oscillator base current flow. This current is about 20 milliamperes average. The pull-in range of the system is on the order of ±200 cycles.

Conclusions

A completely transistorized horizontal deflection system has been developed. The angular deflection that can be achieved increases with the square-root of the bidirectional switching power capability of the output transistor. The circuit is potentially very efficient with respect to the power dissipated in the output transistor. For example, the total power that would be dissipated in an output transistor generating 90-degree deflection at 10 kilovolts is calculated to be of the order of 2 watts if the transistor were capable of cutting off in 1 microsecond. This is less than one tenth the total power dissipated in the vacuum tubes normally used to perform this function.

Acknowledgment

The author wishes to acknowledge the contribution of I. G. Maloff of the RCA Victor Television Division to the work described. He is responsible for the specific driver and output circuits shown.

TRANSISTOR RECEIVER VIDEO AMPLIFIERS

By

MARSHALL C. KIDD

RCA Victor Television Division,
Camden, N. J.

Summary—*During the last two years the improvement in high-frequency transistors has been significant. Transistors suitable for video output stages have been made with large increases in collector breakdown voltage and power dissipation. While these units are now of an experimental nature, they represent only minor modifications of commercial transistors.*

The design of the receiver video amplifier is a function of the television system and the transistor. In the receiver, the video amplifier must drive the kinescope, supply the horizontal and vertical sync separators, sound i-f and a-g-c circuits with suitable signals, and provide an associated gain control for the signal to the kinescope. The transistor amplifier must fulfill these requirements. In this paper, the general design problems are discussed and a practical amplifier is described.

TRANSISTORS FOR VIDEO AMPLIFIERS

UNTIL recently, it was difficult to use transistors for driving a kinescope in a conventional television receiver. Special circuitry was needed to overcome the collector-voltage and power limitations of transistors to obtain the 80 to 150 volts normally available from tubes.

High-frequency transistors are normally made by decreasing the base width and the area of the collector and emitter dots or contacts. Also, low resistivity material is used in the base pellet to reduce the base-lead resistance. The low resistivity and narrow base widths lower the collector breakdown voltage as the frequency response is improved. Because of the small dimensions involved, low power dissipation seemed a limitation to high-frequency transistors.

The drift transistor[1, 2] and other transistors using diffusion techniques in their manufacture appear to represent a major breakthrough in the transistor art. It is now possible to improve the frequency response of a given transistor configuration by an order of magnitude by using diffusion techniques. This is achieved with an increase in

---

[1] A. L. Kestenbaum and N. H. Ditrick, "Design, Construction, and High-Frequency Performance of Drift Transistors," *RCA Review*, Vol. XVIII, p. 12, March, 1957.

[2] H. Kroemer, "The Drift Transistor," *Transistors I*, RCA Laboratories, Princeton, N. J., 1956, p. 202.

collector voltage breakdown. The diffusion process allows a variation of impurity between the base and collector so that low resistivity is achieved at the base and high resistivity at the collector side of the junction. Since the breakdown voltage is proportional to the resistivity of the material, collector breakdown voltages for drift transistors run two or three times those of normal high-frequency transistors such as the 2N139 type used in 455-kilocycle i-f amplifiers.

Figure 1 shows a comparison of a conventional 2N139 alloy junction high-frequency transistor with two drift transistors. The 2N247 drift transistor has geometry similar to that of the 2N139 except that diffusion techniques are used in its manufacture. The type A drift transistor has electrically similar characteristics to the 2N247 with the exception of the increased breakdown voltage and higher collector dissipation. The dissipation rating is increased by lowering the thermal resistance of the unit.

TYPE	DESCRIPTION	f_{max} MC	$f_{\alpha e}$ MC	$r_{bb'}$ OHMS	$C_{b'c}$ $\mu\mu f$	$V_{CE\ max.}$ VOLTS	COLLECTOR DISSIPATION AT 70°
2N139	ALLOY p-n-p	14	4.7	75	9.5	16	35 MW
2N247	DRIFT p-n-p	136	30	40	1.7	>40	35 MW
TYPE A	DRIFT p-n-p	136	30	40	1.7	>100	≈200 MW

Fig. 1—Transistor characteristics.

The drift transistors have reduced base resistance, $r_{bb'}$, in the order of 40 ohms compared to 75 ohms for the 2N139. The maximum frequency of oscillation, f_{max}, and the frequency at which the common base current gain falls to 70 per cent of its low-frequency value, $f_{\alpha b}$, are both increased by an order of magnitude. Since the depletion layer between the collector and base is much wider in the drift transistor, the depletion layer capacitance, $C_{b'c}$, is also greatly reduced.

The 2N247 transistor has frequency characteristics that are suitable for low-power video application. The type A units can dissipate over 200 milliwatts with a maximum peak-to-peak voltage of 100 volts with the same bandwidth as the 2N247.

Current and Power Gain

The performance as a video amplifier can be evaluated from the curves given in Figure 2. The common-emitter (beta) current gains and power gain for three typical type A units show a frequency response that is constant up to a given cutoff frequency. Above this frequency the curves roll off at a 6-decibels-per-octave rate. As a rough

approximation the 3-decibel cutoff frequencies of both curves will be the same. The beta cutoff frequency, $f_{\alpha c}$, is easily measured since its cutoff is in the video frequency range. The unilateralized power gain curve can be determined by two measurements (1) a low-frequency measurement at audio frequencies, and (2) a measurement in the order of ten times the 3-decibel cutoff frequency. In the curves given in Figure 2 a 40-megacycle measurement was used to determine the high-frequency portion of the curve. The 6-decibels-per-octave slope has been checked by a series of power-gain measurements including f_{max} data. The 2N247 transistors have beta cutoff frequencies in the order of 0.5 megacycle.

The power-gain curve is obtained by conjugately matching and unilateralizing the transistor. The power gain of the transistor as a video amplifier will be related to the curve of Figure 2.

Assume a given bandwidth such as 3.5 megacycles. From the power-gain curve the maximum power gain would be 28 decibels for a single conjugately matched stage. Generally, it is not possible to conjugately match the transistor over a wide frequency range, and the maximum power gain is not realized. However, using a low-impedance generator of the order of 50 ohms and shunt series peaking with a 5000-ohm load, gains of 22 decibels are possible with a 3.5-megacycle bandwidth.

The maximum power gain of available transistors varies greatly over the video spectrum. In order to overcome this limitation, gain must be traded for bandwidth. Feedback offers a solution to this problem and most transistor video amplifiers use feedback.

The Common-Emitter Amplifier

The common emitter is the most commonly used transistor configuration. As an amplifier it has the highest power gain of the three possible arrangements and gives both voltage and current gain. The amplifier and its hybrid pi equivalent circuit[3] are given in Figure 3. This equivalent circuit uses a current generator similar to the pentode, and for drift transistors using a low generator impedance, the low-frequency gain is approximately,

$$A_{LF} = g_m R_L. \tag{1}$$

It can be seen from the equivalent circuit that for small values of load

[3] L. J. Giacoletto, "Study of PNP Alloy Junction Transistor from DC Through Medium Frequencies," *RCA Review*, Vol. XV, p. 506, December, 1954.

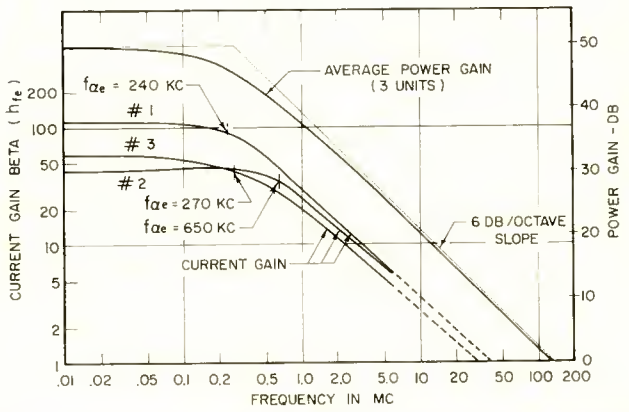


Fig. 2—Common-emitter current and power gain versus frequency.

resistance, the input circuit and the emitter capacitance are combined in $C_{b'c}$, and the equivalent shunt resistance, R_{eq}, across it determines the cutoff frequency;

$$R_{eq} = \frac{r_{b'e}\,(r_{bb'} + r_g)}{r_{b'e} + r_{bb'} + r_g}. \tag{2}$$

The 3-decibel cutoff frequency neglecting the feedback capacitance, $C_{b'c}$, is

$$f_{3db} = \frac{1}{2\pi\, R_{eq}\, C_{b'e}}. \tag{3}$$

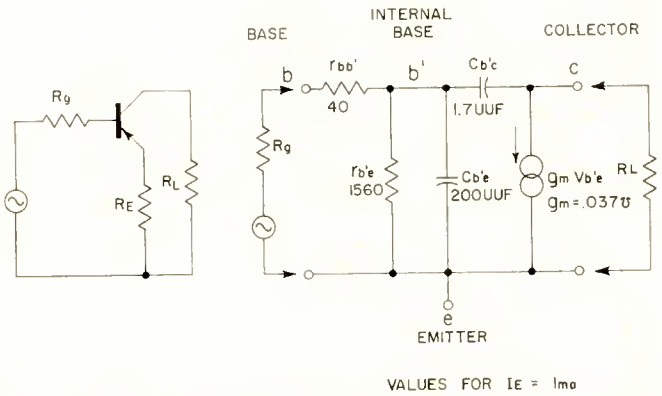


Fig. 3—Common-emitter amplifier and hybrid-pi equivalent circuit.

These equations show the importance of using a low source impedance for wideband amplifiers. Two examples will be given to show this relationship using the equivalent circuit values and again neglecting $C_{b'c}$.

(a) $$R_g = 10{,}000 \text{ ohms.}$$

Then, from Equation (2),

$$R_{eq} = \frac{1560\ (40 + 10{,}000)}{1560 + 40 + 10{,}000} = 1350 \text{ ohms,}$$

so that f_{3db} from Equation (3) is

$$f_{3db} = \frac{10^6}{2\pi\ 1350 \times 200} = 0.56 \text{ megacycle} \approx f_{ac},$$

(b) $$R_g = 40 \text{ ohms.}$$

$$R_{eg} = \frac{1560\ (40 + 40)}{1560 + 40 + 40} = 76.2 \text{ ohms,}$$

$$f_{3db} = \frac{10^6}{2\pi\ 76.2 \times 200} = 10.5 \text{ megacycles.}$$

The above relations are useful when small values of load resistance are used, making it possible to neglect $C_{b'c}$. For the more general case $C_{b'c}$ cannot be neglected and accounts for most of the bilateral nature of the transistor. This has been analysed by Bruun[4] and compared to the Miller effect in tubes. The feedback capacitance, $C_{b'c}$, is multiplied by the gain of the amplifier and added to $C_{b'c}$. This gives an equivalent capacitance

$$C_{eq} = C_{b'e} + g_m R_L C_{b'c}. \tag{4}$$

The cutoff frequency of the input then would be

$$f_{3db} = \frac{1}{2\pi\ R_{eq}\ C_{eq}}. \tag{5}$$

[4] G. Bruun, "Common-Emitter Transistor Video Amplifiers," *Proc. I.R.E.*, Vol. 44, p. 1561, November, 1956.

In a practical amplifier this feedback capacitance may reduce the frequency response to one half the value without feedback, depending on the load.

In the video output stage where the transistor has a large resistive load shunted by a capacitance of 20 or 30 micromicrofarads the output circuit may limit the frequency response and cannot be neglected in determining the amplifier bandwidth. The feedback capacitance becomes important since its effect is to reduce the output impedance at high frequencies. The load resistance, R_L, will then be shunted by a transistor impedance of the same order of magnitude. The cutoff frequency of the output circuit will then always be greater than the frequency at which the load resistance equals the shunt capacitance.

Video Amplifier Configurations

A transistor may be operated common emitter, common base, or common collector.[5] These arrangements correspond roughly to the vacuum tube grounded cathode, grounded grid, and grounded plate or the cathode follower. Each configuration has advantages and disadvantages for a particular application. The common-emitter amplifier is the most commonly used and is the only one suitable for a single-stage video amplifier in a receiver using a diode detector. This is because of the two configurations giving voltage gain—common emitter and common base—the input impedance at low frequencies is at least ten times higher for the common-emitter amplifier. At low frequencies the input impedance of the common-base amplifier is approximately

$$R_{IN_B} \approx 1/g_m. \tag{6}$$

For one milliampere this is roughly 27 ohms. This low input impedance limits the use of the amplifier considerably, but it may be used where no phase reversal is desired.

The common collector or emitter follower is used where high input and low output impedance is required. Its input and output impedances are approximately

$$Z_{IN_C} \approx h_{fe} R_L, \tag{7}$$

$$Z_{OUT} \approx 1/g_m + R_g/(h_{fe} + 1), \tag{8}$$

where h_{fe} is beta, the common-emitter current gain (Figure 2), g_m is

[5] R. F. Shea, *Principles of Transistor Circuits*, John Wiley & Sons, Inc., N. Y., 1953.

milliwatts to meet the dissipation requirement. If loads of 6600 or 5000 ohms are used, the minimum dissipation will be 375 and 500 milliwatts respectively. This assumes satisfactory stability factors for the amplifier. As the load resistance is reduced, the frequency response will be improved if the amplifier is limited by the output circuit.

The maximum and minimum voltage range can be seen from Figure 6. The maximum collector voltage is the breakdown voltage, BV_{CE}, which is shown at 106 volts. This is the voltage at which the collector current of the transistor begins to increase rapidly and is usually specified at a fixed value of current such as 50 microamperes.

The low-voltage limit can also be seen from the nearness of the knee of the curves to the current axis. The output voltage swing can come within a volt or so of the d-c emitter voltage.

Temperature Limitations

Another consideration in using a transistor as an output video amplifier is its power limitation with temperature. Normally a maximum safe junction temperature is established by the manufacturer based on life test data. This temperature is usually given at 85° C, though there are indications that a higher value may be practical. The 85° C value is used in Figure 7, and the relation between collector dissipation and the ambient operating temperature is shown. Less dissipation is allowed as the ambient temperature is increased. The five curves correspond to different values of thermal resistance. A thermal resistance of 50° C per watt means that the junction temperature would rise 50° above the ambient operating case temperature for 1 watt dissipation. Thus for a given maximum junction temperature the curves will appear as in Figure 7.

The case temperature of a transistor may be reduced with a heat sink. The high dissipation region shown in Figure 7 for high-frequency transistors can be realized only with a heat sink such as a metal clip on a chassis. The 2N139 and 2N247 transistors have thermal resistances in free air of 500° C per watt while for the type-A units it is 80° C per watt or less. Special heat-conducting materials and heat sinks make increased dissipation possible. Figure 7 shows that if a transistor must dissipate 250 milliwatts and operate at a case temperature as high as 55° C, a thermal resistance of 125° C per watt or less would be required. At a room temperature of 20° C, 520 milliwatts could be dissipated with this transistor. This assumes a maximum safe junction temperature of 85° C. If this value were increased to 100° C, then a thermal resistance of 200° C per watt would meet the original specification of 250 milliwatts at 55° C.

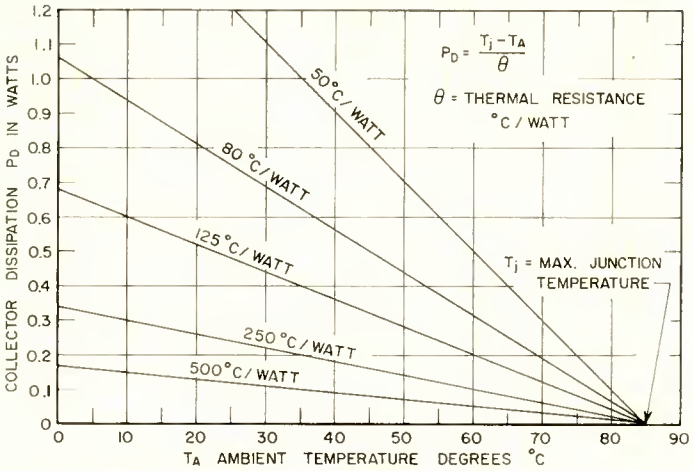


Fig. 7—Collector dissipation versus temperature.

Two-Stage Receiver Video Amplifier

The two-stage amplifier of Figure 8 shows a practical amplifier circuit that provides the basic requirements for a television receiver. The circuit uses a common-collector to common-emitter combination with two supply voltages, +12 and +300 volts. The common-collector input circuit allows a large detector load resistor to be used assuring maximum efficiency and highest detector output voltage. The low output impedance makes large voltage gains possible in the second stage with an over-all bandwidth of 3.5 megacycles; a-g-c voltage is supplied from the emitter of the first stage. The detector is direct coupled so that

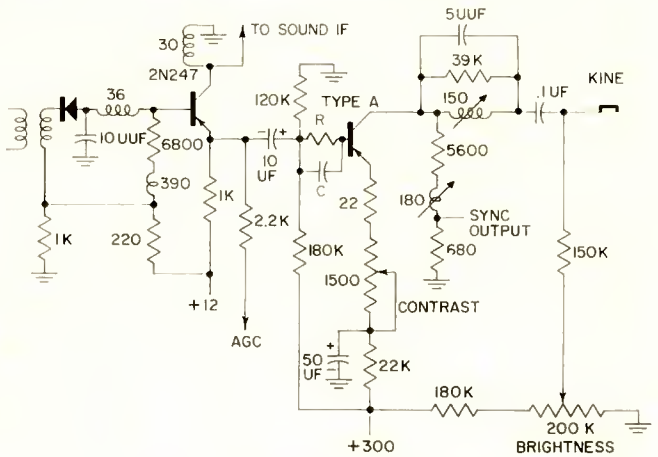


Fig. 8—Two-transistor amplifier.

no coupling capacitor is required, and the d-c component is carried through to the a-g-c circuit. The bias for the first stage is supplied from the 1000-ohm and 220-ohm voltage divider. Sound is taken from the collector circuit.

The inductance load in the collector of the first stage has the advantage of raising the input impedance due to the feedback provided by the capacitance $C_{b'c}$. The second stage is a-c coupled since it must operate at a 100-volt emitter voltage and would require a high breakdown voltage for the first transistor if it were d-c coupled. The power for the output stage is taken from the 300-volt kinescope screen supply. This large voltage allows an emitter resistance of 22,000 ohms to be

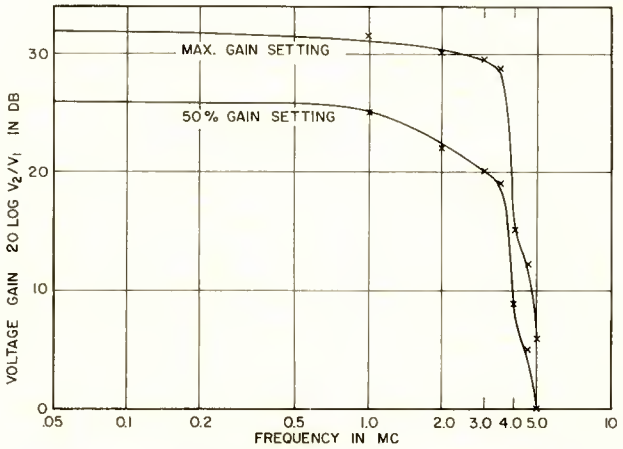


Fig. 9—Frequency response of two-stage amplifier.

used, giving an essentially constant current supply to the transistor and making the circuit relatively independent of temperature variations. The stability factor for the stage is three. Conventional shunt series peaking is used and an RC peaker is added between the stages to improve the frequency response. Sync is removed across a 680-ohm resistor in series with the load. The cathode of the kinescope is driven since 15 to 20 per cent less drive is required. When the cathode is driven, the signal is in the direction to cancel the screen voltage and thus increase the effective output since the drive required is proportional to the screen voltage.

The frequency response of the amplifier is shown in the curves given in Figure 9. The voltage gain for a 3.5-megacycle bandwidth is 40. The curves show the effect of the emitter gain control and show that the bandwidth is slightly increased for full gain indicating some

positive feedback in the output stage due to the indictive loading and internal capacity feedback. The square-wave response gives a large-signal rise time of 0.18 and a fall time of 0.21 microsecond. These results are comparable to those obtained with conventional tube amplifiers.

Conclusions

Experimental transistors can now meet the video amplifier receiver requirements. The drift transistor with improved thermal properties is capable of driving the conventional kinescope. Practical amplifiers have been made and found to give results comparable to tube performance. The circuit designer is confronted with the special problems of the transistor such as its bilateral nature and temperature sensitivity but can achieve improved efficiency in his circuits along with reduced heat, size, and weight.

AN IMAGE-CONVERTER TUBE FOR HIGH-SPEED PHOTOGRAPHIC SHUTTER SERVICE

By

R. G. Stoudenheimer and J. C. Moor

RCA Electron Tube Division,
Lancaster, Pa.

Summary—*A developmental image-converter tube having electrostatic focus, a shutter grid, and electrostatic deflection is described. This tube is intended for multiple-frame photography of high-speed events as short as ten millimicroseconds. Operating characteristics of the tube are described, with emphasis on the low gating-power and deflection-power requirements.*

Introduction

An image-converter tube is an electron device which provides a visible image of an irradiation pattern incident on its photocathode. Such a tube permits visual observation or photographic recording of optical images produced by nonvisible or relatively nonactinic radiation. In terms of radiation wavelength, the conversion is usually (but not necessarily) from longer to shorter, as, in "snooperscope" types which provide visible images of objects or terrain illuminated by "black" (infrared) light. Image-converter tubes now have many applications in science and industry as well as in the military field. A particularly important application is as shutters in very-high-speed photography.

A developmental image-converter tube designed specifically for high-speed photographic service and having low gating- and deflection-power requirements is described in this paper. The tube has a maximum ultor-voltage rating of 15 kilovolts and a deflection factor of only 75 volts per inch per kilovolt of ultor voltage, and requires a gating voltage of only 17 volts per kilovolt. It uses electrostatic focusing and deflection and is provided with a special P11 phosphor screen capable of resolving better than 22 line pairs per millimeter. Exposure times as short as 10^{-8} second can be achieved without loss of resolution.

Although the use of image-converter tubes as photographic shutters has been previously described,[1-3] the tubes used in this early work

[1] J. S. Courtney-Pratt, "A New Method for the Photographic Study of Fast Transient Phenomena," *Research*, Vol. 2, p. 287, June, 1949.

[2] A. W. Hogan, "Use of Image Converter Tube for High-Speed Shutter Action," *Proc. I.R.E.*, Vol. 39, p. 268, March, 1951.

[3] F. C. Gibson, M. L. Bowser, C. W. Bamaley, and F. H. Scott, "Image Converter Camera for Studies of Explosive Phenomena," *Rev. Sci. Instr.*, Vol. 25, p. 173, February, 1954.

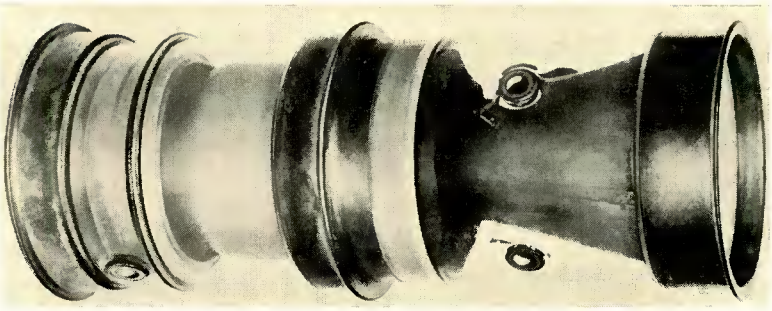


Fig. 1—Photograph of the developmental image-converter tube.

required substantial amounts of gating and deflection power, and, frequently, had been developed for purposes other than shutter service. Not until very recently have tubes designed specifically for shutter service and having low gating- and deflection-power requirements been described.[4]

Design Details

A photograph of the image-converter tube described in this paper is shown in Figure 1, and a cross-sectional drawing is shown in Figure 2. The tube is slightly less than four inches in diameter and is approximately 10 inches long. The circular photocathode is one inch in diameter and has a spherical radius of curvature (as proposed by Morton[5]) of four inches, to minimize curvature of the image field and pincushion distortion.

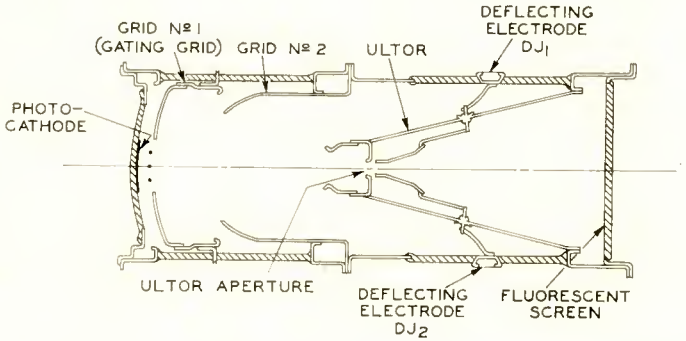


Fig. 2—Cross-sectional drawing of the developmental tube.

[4] B. R. Linden and P. A. Snell, "Shutter Image Converter Tubes," *Proc. I.R.E.*, Vol. 45, p. 513, April, 1957.

[5] G. A. Morton and E. G. Ramberg, "Electron Optics of an Image Tube," *Physics*, Vol. 7, p. 451, December, 1936.

Grid No. 1, grid No. 2, and the ultor comprise an electrostatic converging lens which produces on the fluorescent screen an inverted electron image of the photocathode illumination pattern. The magnification obtained is 0.75. The pair of electrostatic deflecting plates permits positioning of a series of displayed images so that none overlaps a preceding image.

Image Formation and Focusing

The position of the electron image on the tube axis is determined by the point at which the electron paths converge. These paths, in turn, are determined by the convergence of the internal electric field, or the related potential distribution along the axis. The normal potential distribution (approximate) along the axis for a focused electron image in the center of the screen is shown in Figure 3 by the solid

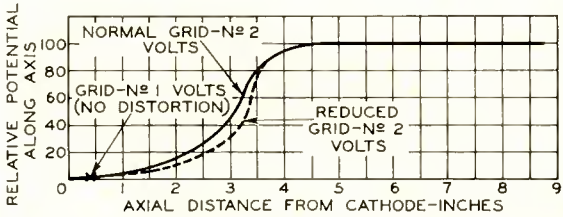


Fig. 3—Approximate potential distribution along the tube axis required to provide a focused electron image on the fluorescent screen. Solid curve: normal distribution; dashed curve: effect of a reduction in grid No. 2 voltage.

curve. The convergence of the electric field at a point of specified potential is proportional to the curvature of the potential-distribution plot (d^2V/dZ^2) at that point. As shown by the dashed curve in Figure 3, a reduction in the potential applied to grid No. 2 increases the curvature of the potential-distribution plot. Consequently, if the field produced by the normal potential distribution is not sufficiently convergent to provide a focused image at the screen, the image may be moved back towards the apex of the ultor by a reduction of the potential applied to grid No. 2.

The gating grid (grid No. 1) is placed close to the photocathode. It is normally operated at a small positive potential with respect to the photocathode (indicated in Figure 3 by an "X"). For complete suppression of cathode emission the gating grid must be supplied with a negative bias voltage sufficient to provide a negative voltage gradient at every point on the cathode. If the gating grid contained simply a single large aperture, a large bias voltage would be required for cutoff.

The cutoff bias can be substantially reduced, however, by the use of fine wires across the grid aperture, and, with a sufficiently fine mesh, can be made as small as 2 or 3 volts, i.e., equal to the maximum energy with which electrons leave the photocathode. A mesh of such fineness, however, degrades the electron image, and, therefore, is not desirable. In the developmental tube a cutoff voltage of only —4 volts per kilovolt of ultor voltage is achieved by the use of three wires across the gating-grid aperture. This voltage is satisfactorily small in relation to the normal gating-grid operating voltage (approximately 12 volts per kilovolt of ultor voltage), and is not substantially reduced by the use of closer grid-wire spacing.

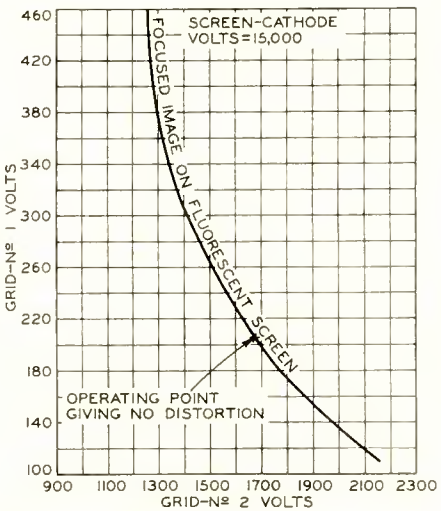


Fig. 4—Loci of grid No. 1 and grid No. 2 voltages which provide a focused electron image on the fluorescent screen. Point X indicates only combination of grid voltages for which there is no scattering of beam electrons by the cross wires of grid No. 1.

The gating grid voltage, like grid No. 2 voltage, is effective for controlling the convergence of the electron-lens system. As shown in Figure 4, for a wide range of grid No. 2 voltages, there are corresponding grid No. 1 voltages which will provide a focused image in the center of the screen. However, unless the potential applied to grid No. 1 is the same as the potential that would exist at the plane of the grid aperture in the absence of the grid cross wires, the resulting field will be distorted in the vicinity of the grid wires, and will not have the axial symmetry required for good imaging. There is, consequently, only one combination of grid No. 1 and grid No. 2 potentials which

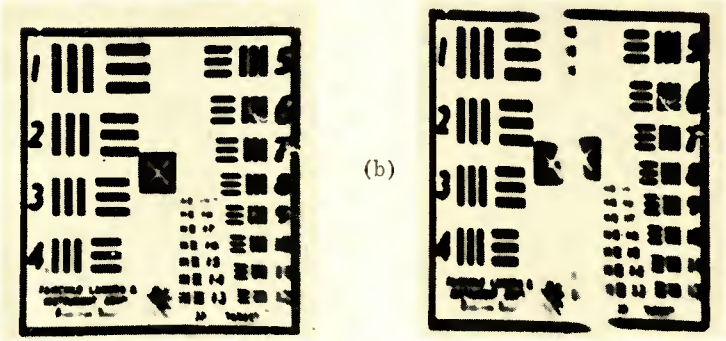


Fig. 5—(a) Focused, undistorted image produced when proper voltages are applied to grids No. 1 and No. 2; (b) segmented image produced when grid No. 1 and grid No. 2 voltages do not have the proper values.

provides an undistorted as well as focused image on the screen. At any other combination of grid voltages, the cross wires of grid No. 1 cause scattering of the electrons in their immediate vicinity and consequent distortion of the image. Figure 5a shows the appearance of the image when the proper voltages are applied to grids No. 1 and 2; Figure 5b shows how a focused image may be broken into segments when these voltages do not have the proper values.

Deflection

The electric field of the electrostatic lens used in the developmental tube is essentially a radial field which directs the electrons towards the apex of the ultor, where the beam diameter becomes very small. It is practical, therefore, to employ a simple electrostatic deflection system using internal deflecting plates. Because the diameter of the beam at the crossover point (the ultor aperture) is only about 0.125 inch at an ultor voltage of 10 kilovolts, and is even smaller at higher ultor voltages, the deflecting plates can be closely spaced so as to obtain high deflection sensitivity. The spacing used in the developmental tube

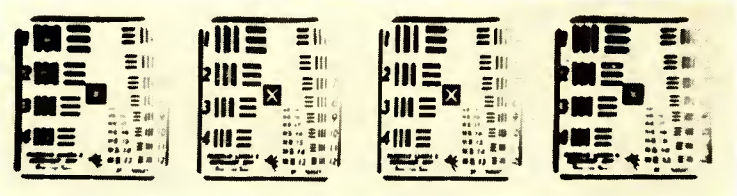


Fig. 6—Linear display of sequential frames of the same test pattern obtained by application of suitable potentials to the electrostatic deflecting plates.

is 0.15 inch, providing a deflection factor of only 75 volts per inch per kilovolt of ultor voltage.

Figure 6 shows four images of the same resolution chart displayed in a linear array by the application of suitable voltages to the deflecting plates. The quality of these images indicates that wide-angle deflection can be achieved without serious loss of resolution or distortion. Figures 7 and 8 show resolution and distortion (relative magnification), respectively, as functions of deflection.

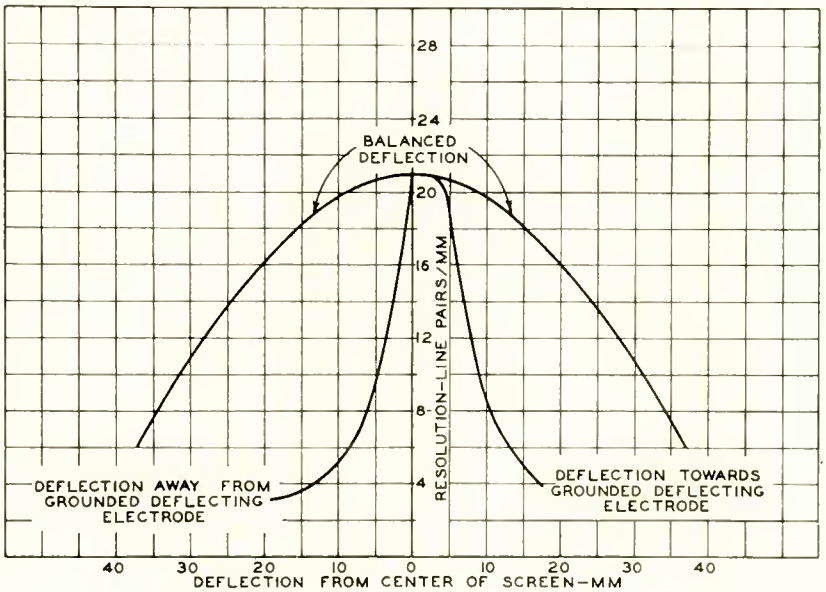


Fig. 7—Resolution of the developmental tube as a function of deflection distance.

Conversion Gain

The phosphor screen used in the new tube is a P11 type, specially prepared to have a very fine grain and high resolution capability. The screen is also very efficient, being capable of converting as much as 10 per cent of the incident beam energy into useful radiant energy. The aluminum film which backs the screen provides the following advantages:

1. It increases the useful light output of the screen by sending back through the screen light emitted towards the inside of the tube.

2. It prevents light feedback from the fluorescent screen to the photocathode, and thus prevents the loss of contrast which would otherwise result.

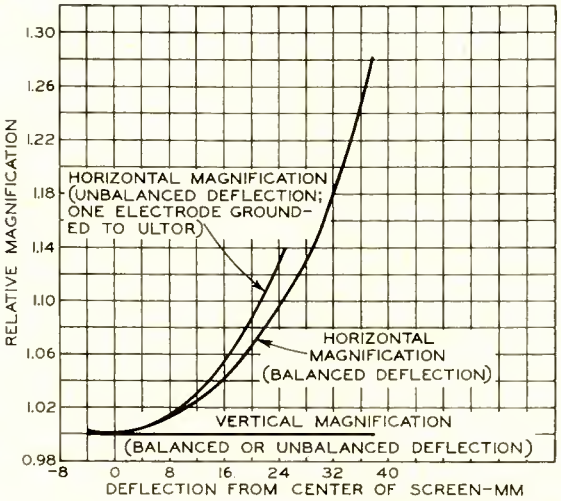


Fig. 8—Distortion (relative magnification) of the electron image as a function of deflection distance.

3. It prevents the illumination on the photocathode from being transmitted directly through the tube to the fluorescent screen and thus avoids possible exposure of photographic films during intervals when the gating grid is biased beyond cutoff.

Figure 9 shows the spectral emission characteristic of the P11 phosphor screen and the spectral sensitivity of the S11 photocathode. Effective application of the tube requires a knowledge of the ratio of output radiant energy to incident radiant energy. This ratio, commonly called the conversion gain, depends on the spectral distribution

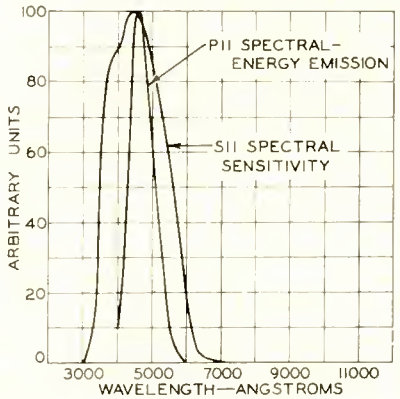


Fig. 9—Spectral characteristics of the photocathode and fluorescent screen used in the developmental image-converter tube.

CONVERSION GAIN OF TYPICAL IMAGE CONVERTER TUBE
HAVING S-11 CATHODE RESPONSE AND P11 PHOSPHOR

Radiation on Cathode (Wavelength = 4400 Å)	Cathode Emission (Sensitivity 0.024 μa/μw)
1 μwatt	0.024 μampere

Energy Delivered to Screen (Screen Voltage = 10 KV)	Screen Radiation (Screen Efficiency = 10 per cent)	Conversion Gain
240 μwatts	24 μwatts	24

Fig. 10—Calculation of conversion gain at the wavelength of maximum photocathode sensitivity.

of the incident light. The expected conversion gain for illumination having the wavelength of maximum photocathode sensitivity (4400 Å) can be easily calculated from the relations shown in Figure 10. A typical photocathode sensitivity of 30 microamperes per lumen (0.024 microamperes per microwatt at 4400 Å) is assumed. The conversion gain for illumination of any other wavelength can be calculated by multiplying the corresponding relative sensitivity on the S11 response curve by the conversion gain at 4400 Å.

Application

Figure 11 shows schematically the complete optical system of an image-converter camera, and the brightness gain achieved under typical operating conditions. If the scene observed is sufficiently distant so that the effective magnification (M) of the objective lens is very small, the brightness gain may be as much as four times that shown in Figure 11.

In most cases the loss of light in the objective lens used to form an

SCENE BRIGHTNESS = B

LENS F:No. (f/d) = 2.0 TRANSMISSION (T) = 80% MAGNIFICATION (M) = 1

IMAGE CONVERTER TUBE CONVERSION GAIN (G) = 24 MAGNIFICATION (m) = 0.75

CAMERA

ILLUMINATION ON TUBE	SCREEN BRIGHTNESS	BRIGHTNESS GAIN
$E = \frac{BT}{4(f/d)^2(1+M)^2} = \frac{B}{80}$	$B' = \frac{EG}{m^2} = \frac{43B}{80}$	$\frac{B'}{B} = \frac{TG}{4(f/d)^2(1+M)^2m^2} = 0.5$

Fig. 11—Block diagram of image-converter camera system, and calculation of brightness gain.

optical image on the photocathode of an image-converter tube is great enough to cancel most of the energy gain achieved in the tube itself. For applications requiring the highest possible image brightness the objective lens should, therefore, have the largest possible aperture preferably f:1.1 or larger. The conversion gain of the image-converter tube itself can be increased by operation at higher ultor voltage. Although the developmental tube has a tentative maximum ultor-voltage rating of 15 kilovolts continuous, it has improved design features which permit the use of ultor voltages up to 30 kilovolts where a pulse-type ultor-voltage supply is used.

The principal hazard involved in operation at an ultor voltage higher than 15 kilovolts is internal breakdown, with resulting discharges which may destroy the photocathode or scatter cesium through the interior of the envelope. The most common cause of internal breakdown is the accumulation of a high positive charge on the interior of the glass envelope in a region where the envelope is sealed to a negative electrode. The resulting field gradient is generally steep enough to produce field emission from the negative electrode, thereby initiating an internal discharge, which recurs periodically at a rate depending on the time required for the positive charge on the envelope to reach the field-emission value. The effects of such breakdowns may be minimized by the use of external impedances which will limit the resulting current through the tube to a safe value. However, even though the discharge current is limited in this manner the repeated dispersal of cesium throughout the tube increases the tendency towards breakdown.

Because a finite time is required for the positive charge on the envelope to reach a significant value, the danger of breakdown at higher-than-rated ultor voltages can be minimized by the use of a pulse-type ultor-supply circuit. With such an arrangement all adjustments may be made at a lower ultor voltage (15 kilovolts), and the voltage then raised to as much as 30 kilovolts for periods up to a second. It is important to limit the capacitance across the high-voltage supply circuit to assure that, if breakdown does occur, the energy dissipated within the tube during the resulting discharge is not great enough to cause damage.

The stiff-wire leads used in earlier tubes for the external connections to the gating grid and the deflecting plates were occasionally responsible for corona. This difficulty is avoided in the developmental tube by the use of recessed-cavity caps.

Exposure Time

In practice, the shutter speeds attainable with the new tube are limited by the ability of the external circuitry to supply good square-

wave pulses of sufficiently short duration. With perfect pulse-forming circuits, the minimum exposure time is limited by electron transit time. Electrons are defocused if they are not beyond the influence of the gating grid when its voltage returns to the cutoff value at the end of the gating pulse. The calculated transit time of electrons between the cathode and ultor for 15 kilovolts on the ultor is approximately 5×10^{-9} second. Loss of resolution will probably become noticeable, therefore, at exposure times in the order of 10^{-8} second, and become progressively worse as exposure time is decreased.

Acknowledgments

The development of the new tube was partly sponsored by the University of California under AEC Contract W7405-ENG-48. Many of the design features incorporated in the tube are the results of earlier research and development sponsored by the U. S. Army Engineer Research and Development Laboratories, Fort Belvoir, Virginia. Grateful appreciation is expressed for the enthusiastic support and encouragement given by R. W. Engstrom.

QUASI-ELECTRIC AND QUASI-MAGNETIC FIELDS IN NONUNIFORM SEMICONDUCTORS*

By

Herbert Kroemer

RCA Laboratories,
Princeton, N. J.

Summary—*In a chemically uniform semiconductor with a constant impurity concentration and without elastic strains, the forbidden band has a constant width, and there are no internal electric fields. A nonhomogeneous impurity distribution introduces electric fields, while nonuniform elastic strains introduce a nonuniform bandwidth. If the semiconductor is an alloy, such as a germanium–silicon alloy, a change in the alloy composition also changes the bandwidth.*

A change in bandwidth means gradients of the band edges which are different for the conduction and the valence bands. These gradients act upon the electron and hole movement as though they were electric fields, but, because the two slopes are different, these "quasi-electric" fields are not the same for electrons and holes, contrary to the case of real electric fields.

In addition, a type of "quasi-magnetic" field is produced when an inhomogeneity produces a shift of the location within the Brillouin zone of the energy minimum of the band. These quasi-magnetic fields not only are different for electrons and holes but they are also different for the electrons (holes) inside the various energy minima of that band.

Inhomogeneous Semiconductors

THIS paper deals with the electron movement in semiconductors in which the crystal potential is not perfectly periodic but where, instead, the *shape* of the atomic potential changes gradually from cell to cell if one proceeds through the crystal, as illustrated in Figure 1. It is this type of inhomogeneity to which we refer in this paper when we speak of "inhomogeneous" semiconductors. An example of an inhomogeneous semiconductor in this sense is a semiconductor under nonuniform elastical strains. Another example is an alloy of several semiconductors such that the composition of the alloy changes gradually throughout the crystal (e.g., a nonuniform germanium–silicon alloy). Actually, the atomic potential in such an alloy does not vary continuously from one cell to the next, but rather in a discontinuous and random fashion. In first-order considerations, however, it is permissible to neglect those statistical fluctuations by treating a

* This paper was presented at the Symposium on "The Role of Solid State Phenomena in Electric Circuits," in New York, April 24, 1957, and is included in the records of this symposium.

"local average" of the potential, which may be formed by averaging the actual potential over a certain number of cells adjoining the cell under consideration. In a nonuniform alloy this local average will then represent a slowly varying atomic potential as in the case of elastical deformation.

It is believed that nonuniform semiconductors in the sense used here are much more common than would be indicated by these two examples. Lattice discontinuities, like surfaces, are believed to lead to potential deformations in their neighborhood. Polarizations of the electronic orbits due to electric fields or due to high densities of injected carriers also belong in this class, and there may be other types. In many cases the inhomogeneity effects may be small compared to other effects so that they are overlooked in experiments not designed to show

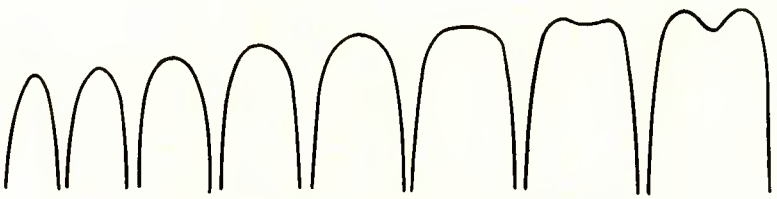


Fig. 1—An example of a nonuniform crystal potential.

them up. The purpose of this paper is to outline the general effects upon the electron movement of inhomogeneous potentials, regardless of their origin, and of the strength of the effects compared with other effects that might be present simultaneously.

Quasi-Electric Fields

In a nonuniform semiconductor one has essentially a different kind of semiconductor in every portion of the crystal. Different semiconductors, in general, have different widths of the forbidden energy band, so the nonuniform semiconductor leads to the concept of a nonuniform band gap. The question immediately arises as to whether this is a legitimate concept, i.e., whether the following three conditions are satisfied:

(1) A local density of states can still be defined, and the distribution of the states over the various energy levels can be described by *sharp* band edges and by effective mass tensors.

(2) The band gap and the mass tensors at each point are the same, as in a homogeneous crystal which has the same

atomic potential throughout as that which the inhomogeneous crystal has at the given point.

(3) The dynamics of the electron are still governed by Newton's Law, using as the mass of the electron the above effective mass and as the force upon the electron the slope of the edge of that band to which the electron belongs.

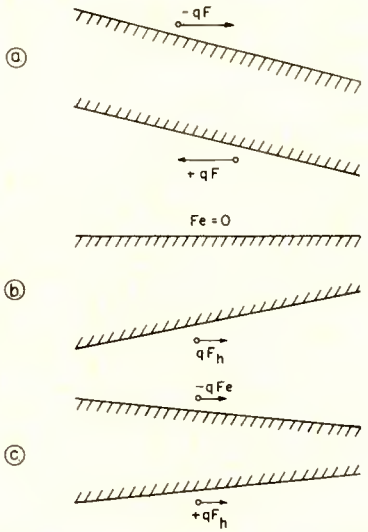


Fig. 2—(a): Effect of a true electric field; (b) and (c): Effects of quasi-electric fields.

The third of these conditions is the most interesting one. In a homogeneous semiconductor the band slope under an external field is the same for all bands and, as a result, the forces upon electrons and holes are equal in magnitude and opposite in direction. This is not the case with a varying band gap. If the concept of a varying band gap is legitimate, the forces would no longer be equal and opposite. It should, for example, be possible to have a force acting only upon one kind of the carriers, or to have a force which acts in the same direction for both (Figure 2). Electrical forces in uniform crystals can never do this. This is why we call these forces "quasi-electric." They present a new degree of freedom for the device designer to enable him to obtain effects with the quasi-electric fields that are basically impossible to obtain with ordinary circuit means involving only "real" electric fields. Two simple examples are given in the next section.

The principal question as to the correctness of the above three conditions and, therefore, the reality of the quasi-electric fields, has been answered in the affirmative; Bardeen and Shockley[1] have shown in their treatment of the electron-lattice scattering via deformation potentials that the concept of a variable band gap is legitimate in the above sense, provided the nonuniformity arises from nonuniform elastical deformation. The present author[2] has extended their proof and has shown that the concept holds true regardless of the shape or the origin of the variation of the atomic potential, provided this variation is a sufficiently gradual one.

Mathematically, this result can be expressed as an extension of the Wannier–Slater[3] Theorem: Assume that $E_{\vec{k}}(\vec{x})$ is the energy of an electron with the wave vector $\vec{k}$, which moves in an exactly periodic potential of such a shape as exists in the nonuniform crystal at the position $\vec{x}$ inside the crystal. Then an operator $E_{-i\nabla}(\vec{x})$ can be derived from $E_{\vec{k}}(\vec{x})$ by replacing $\vec{k}$ with the operator $-i\nabla$. The behavior of the electron can then be described by a Wannier–Slater wave equation:

$$E_{-i\nabla}(\vec{x})\Phi(\vec{x}) = E\Phi(\vec{x}), \tag{1}$$

where $\Phi(\vec{x})$ is the familiar Wannier–Slater amplitude function for the electron.†

In a semiconductor with a single energy minimum at $k = 0$,

$$E_{\vec{k}}(\vec{x}) = \frac{\hbar^2 k^2}{2m^*} + E_B(\vec{x}), \tag{2}$$

where $E_B(\vec{x})$ is the position-dependent‡ band edge. This leads to

---

[1] J. Bardeen and W. Shockley, "Deformation Potentials and Mobilities in Non-Polar Crystals," *Phys. Rev.*, Vol. 80, p. 72, October, 1950.

[2] H. Kroemer, "Band Structure of Semiconductor Alloys with Locally Varying Composition," *Bull. Amer. Phys. Soc.*, Vol. 1, p. 143, March, 1957.

[3] J. C. Slater, "Electrons in Perturbed Periodic Lattices," *Phys. Rev.*, Vol. 76, p. 1592, December, 1949.

† Equation (1) is identical with Equation (6) in Slater's paper. The operator $-i\nabla$ in $E_{-i\nabla}$ operates only upon $\Phi(\vec{x})$, not upon $\vec{x}$ inside $E_{-i\nabla}(\vec{x})$.

‡ It is now irrelevant whether this position dependence is due to external electric fields or changes in the band gap. Actually, the two effects cannot be separated from each other. Equation (3), therefore, covers as well the older case of an electric perturbation in a truly periodic potential.

$$\left[-\frac{\hbar^2}{2m^*} \nabla^2 + E_B(\vec{x}) \right] \Phi(\vec{x}) = E\Phi(\vec{x}), \tag{3}$$

the familiar wave equation for an electron in a potential, $E_B(\vec{x})$. Since the position dependence of $E_B(\vec{x})$ will be different, in a nonuniform semiconductor, for the conduction band and for the valence band, Equation (3) represents the mathematical expression for the existence of the quasi-electric fields.

Two Examples

In this section the potential usefulness of the quasi-electric fields will be illustrated by two examples of how transistor performance can be improved by the incorporation of quasi-electric fields.[4]

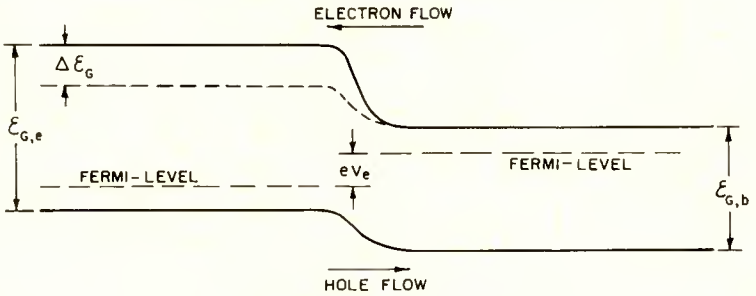


Fig. 3—Wide-gap emitter.

The Wide-Gap Emitter

Figure 3 shows a forward biased p-n junction where the p side has a wider band gap than the n side. In the transition region we then have a quasi-electric field opposing the hole flow to the right, and another quasi-electric field opposing the electron flow to the left. The electron field is stronger. This means that such a junction has a higher ratio of hole-to-electron current than a constant-gap junction with the same impurity distribution and the same mobilities. Used as an emitter in a p-n-p transistor, the transistor has a higher emitter efficiency and a higher current-amplification factor than an otherwise identical constant-gap transistor. This higher emitter efficiency may be utilized in either one of two ways:

[4] H. Kroemer, "The Theory of Diffusion and Drift Transistors, Part III—Dimensional Equations," *Archiv der Elektrischen Übertragung*, Vol. 8, p. 499, November, 1954.

(1) In low-frequency transistors the high emitter efficiency results in a reduction of the well-known and undesirable falloff of the current amplification factor with increasing current.

(2) In high-frequency transistors the intrinsically higher emitter efficiency of a wide-gap emitter may be utilized to decrease the doping of the emitter region without adversely affecting the current amplification factor. Such a decrease in doping results in a decrease of the emitter capacitance, which is one of the most seriously frequency-limiting quantities of modern high-frequency transistors.

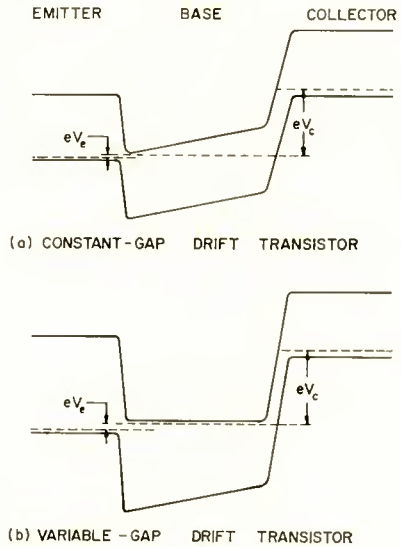


Fig. 4—Two drift-transistor types.

The wide-gap emitter will be described in more detail elsewhere.[5]

Graded-Gap Drift Transistor

In an ordinary drift transistor the drift field is generated by inhomogeneous doping in the base region (Figure 4a). Figure 4b shows a structure in which the quasi-electric field due to a decreasing band gap is used as a drift field. Such a structure has two advantages:[4]

1. Stronger drift fields are obtainable than by inhomogeneous doping.

[5] H. Kroemer, "Theory of a Wide-Gap Emitter for Transistors." To be published.

2. In an ordinary drift transistor a neutralizing electron accompanies each hole that travels through the base region. While the hole is aided by the drift field, the electron is opposed by it. It can be shown[4] that this effect wipes out the advantage of the drift field at high current densities. In a graded-base drift transistor, as shown in Figure 4b, there is no quasi-electric field opposing the electron flow, and the drift effect persists to much higher current densities.

Quasi-Magnetic Fields

Equation (3) was derived under the assumption of Equation (2), i.e., of spherical energy surfaces around $k = 0$. In many semiconductors known today this assumption is not fulfilled; instead, there are several nonspherical minima located at symmetrical points in k-space. If one of them is centered around, say, $\vec{k} = \vec{k}_o$ then, in the neighborhood of this minimum one has, instead of Equation (2),

$$E_{\vec{k}}(\vec{x}) = \frac{\hbar^2}{2m^*}(\vec{k} - \vec{k}_o)^2 + E_B(\vec{x}), \tag{4a}$$

where $1/m^*$ is now a tensor. If there is a minimum at $\vec{k} = \vec{k}_o$ there will also be an identical minimum at $\vec{k} = -\vec{k}_o$:

$$E_{\vec{k}}(\vec{x}) = \frac{\hbar^2}{2m^*}(\vec{k} + \vec{k}_o)^2 + E_B(\vec{x}). \tag{4b}$$

By substituting Equations (4a) and (4b) into Equation (1), one obtains:

$$\left[\frac{\hbar^2}{2m^*}(-i\nabla - \vec{k}_o)^2 + E_B(\vec{x})\right]\Phi = E\Phi \tag{5a}$$

$$\left[\frac{\hbar^2}{2m}(-i\nabla + \vec{k}_o)^2 + E_B(\vec{x})\right]\Phi = E\Phi. \tag{5b}$$

These are exactly the equations for an electron moving under the influence of a vector potential,

$$\vec{A} = \pm\frac{\hbar^2}{e}\vec{k}_o. \tag{6}$$

In a uniform semiconductor, $\vec{k}_o$ is constant and the vector potential has no physical meaning and can be transformed out of the equation by a simple gauge transformation. In a nonuniform semiconductor, however, $\vec{k}_o$ may vary with position. In general, then, curl $\vec{A}$ will no longer vanish. This means that the electron movement is as though a magnetic field

$$\vec{B} = \text{curl } \vec{A} = \pm \frac{\hbar c}{e} \cdot \text{curl } \vec{k}_o \tag{7}$$

were present. This is the quasi-magnetic field.

Thus far this result is a purely mathematical one. But the existence and the physical origin of the quasi-magnetic field can actually be understood qualitatively without resorting to the above mathematical derivation.

For simplicity we assume E_B to be constant, i.e., no quasi-electric fields are present. We wish to show, then, in a special example that the electron moves in an orbit if curl $\vec{k}_o \neq 0$. We assume that the crystal has a primitive cubic structure and that it is inhomogeneous along one of the [100] directions, which direction we call the x-direction. We further assume that the energy minima lie along the [100] directions of the k-space and that the minima move away from $k = 0$ if one proceeds in the positive-x direction. The k-space, then, is shown by Figures 5a and 5b. In such a case the direction of curl $\vec{A}$ for the six ellipsoids is shown in Figure 5c. Only for the two ellipsoids on the k_x axis is curl $\vec{A} = 0$.† The four other ellipsoids should see finite quasi-magnetic fields of the same magnitude, but of different directions.

An electron near the k_y minimum will be studied in more detail. Figure 6 shows the position of this minimum in k-space for three different positions inside the crystal, $x = 0$ and $x = \pm \Delta x$. Assume now that an electron is located at $x = 0$ with a k-vector corresponding to the point "A" in the diagram. Since the velocity, $\vec{v}$, of the electron is proportional to the gradient of the energy in k-space, the electron in that moment moves exactly in plus-x direction. After the time interval $\Delta x/v$ the electron has arrived at $x = +\Delta x$, still having the same absolute $\vec{k}$ vector. But since the ellipsoid itself has shifted, the position of the electron with respect to $\vec{k} = \vec{k}_o$ has changed. The gradient of the energy now has a component in the minus-k_y direction.

† For these two ellipsoids, div $\vec{A} \neq 0$, contrary to ordinary vector potentials.

Correspondingly, the velocity has a component in the minus-y direction, i.e., the electron is constantly deflected towards the right of its instantaneous direction of motion.

An electron at $x = 0$ and at point "B" in k-space has an initial velocity entirely in the plus-y direction. It therefore does not move into a region of different atomic potential. Nevertheless it, too, will be deflected towards the right. This is due to the fact that the distance of point "B" in k-space from $\vec{k}_o$ will vary for different values of x, and therefore the energy will change with x for fixed $\vec{k} = \vec{k}_B$. As shown in the bottom drawing of Figure 6, the energy decreases with increasing x. This gradient of the energy acts as a force upon the electron

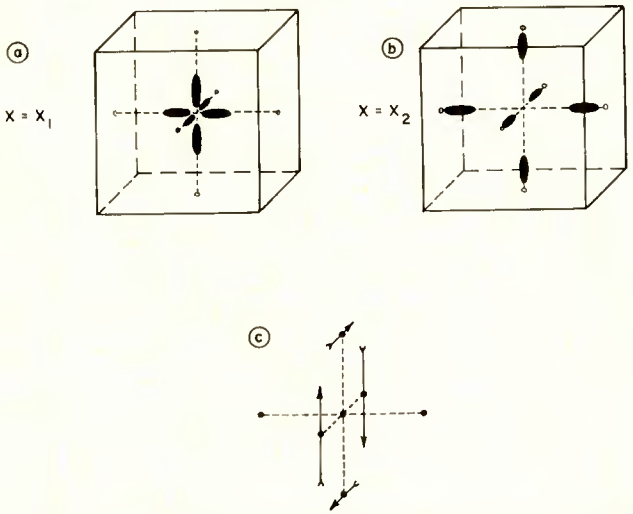


Fig. 5—(a) and (b): Examples of the location of the energy ellipsoids in the Brillouin zone at different positions inside the crystal; (c): Direction of the quasi-magnetic field for the different ellipsoids.

and also deflects it to its right.

For points intermediate between "A" and "B" on the energy ellipsoid, both of the two described effects are present and deflect the electron to the right. The net result is in all cases that the electron circles in k-space about $\vec{k} = \vec{k}_o$ along a path of constant energy, as in the presence of a true magnetic field.

The magnitude of the quasi-magnetic field is given by Equation (7). We assume a favorable case, namely that $\vec{k}_o$ wanders from the center of the Brillouin zone to its edge, and that the transition is rather steep, namely only about 2×10^4 atomic distances. For a lattice constant of 5×10^{-8} centimeter, then, $B = 625$ gausses.

This is not a negligible field. Of course, in practical cases the change of $\vec{k}_0$ may be smaller, but it is conceivable that the transition, on the other hand, can be made steeper. It appears reasonable, therefore, to expect magnetic fields of several hundred gausses.

The consequences of the quasi-magnetic fields are limited. As shown already in Figure 5, the fields for the two ellipsoids of each pair are in opposite directions. In thermal equilibrium, the two ellipsoids are equally populated. As a result, all those magnetic effects cancel which are linear in the magnetic field. There is, therefore, no self-Hall effect, no change in the low-field–high-temperature diamagnetic properties,

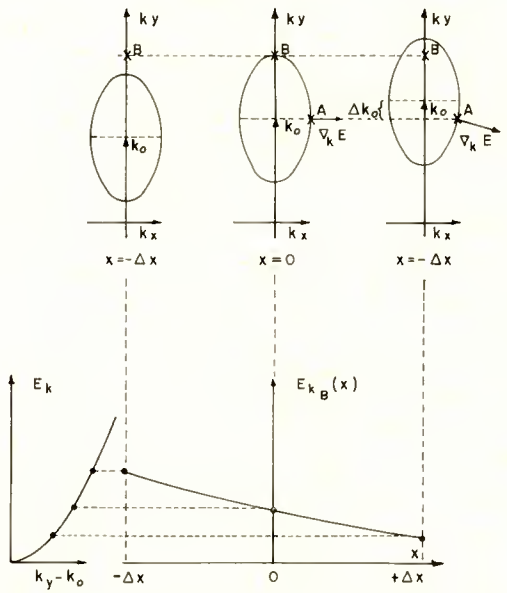


Fig. 6—Explanation of the origin of the quasi-magnetic field.

etc. There should be, however, an indication of effects that are nonlinear in the magnetic field, such as the magnetoresistance. For example, the mobility in a graded semiconductor alloy with noncentral energy bands will always be lower than in a uniform semiconductor alloy, due to the magnetoresistance effect caused by the quasi-magnetic field. It is possible, however, that this decrease of the mobility will not be observable, because already without the quasi-magnetic field the mobilities in a semiconductor alloy are lowered due to disorder scattering, and, the lower the original mobilities are the less they will be changed by a magnetic field. The situation might be different, how-

ever, if the nonuniform band structure were due to elastic deformation rather than to nonuniform composition, because then there is no disorder scattering. This should lead to a mobility reduction with nonuniform strains. One such type of strain is the thermoelastic lattice waves which are responsible for the lattice scattering of the electrons. Our theory, then, predicts that, in addition to the already-known scattering mechanism, these waves also reduce the lattice mobility due to the quasi-magnetic field they produce in multi-valley semiconductors. It is believed that this effect has not yet been studied.

Another effect that should be altered by quasi-magnetic fields is the cyclotron resonance. There the external magnetic field is added to the quasi field producing different net fields for different ellipsoids. If the line widths are small, the result should be a splitting of each cyclotron resonance line into a *multiplett*. The line width condition here means in particular, that no large change of the effective mass is associated with the change in the position of the ellipsoids.

For example, assume the external magnetic field to be parallel to the y-direction in Figure 5. The resonance corresponding to the four ellipsoids on the k_x and the k_z axes would then split up into a triplet, the center line corresponding to the two k_x ellipsoids, the two (symmetrically lying) side lines corresponding to the two k_z ellipsoids.

Other effects in which the quasi-magnetic field would show up could be expected if the equal population of the symmetrical minima could be disturbed.

A THIN-WINDOW CATHODE-RAY TUBE FOR HIGH-SPEED PRINTING WITH "ELECTROFAX"

By

ROGER G. OLDEN

RCA Laboratories,
Princeton, N. J.

Summary—A cathode-ray tube with a thin mica window for contact printing on "Electrofax" paper has been built. Tests were carried out which showed that good definition and very high printing rates can be obtained. Traces were recorded on the "Electrofax" paper at spot-speeds of 44,000 inches per second. The data obtained shows that character writing speeds of over 10,000 letters per second are possible.

INTRODUCTION

THE problem of photographing the transient image on the screen of a cathode-ray tube involves questions of brightness of the phosphor, available exposure time, optics with adequate speed and resolution, and sensitivity of the recording film or paper. Satisfactory solutions have been found for oscillographic work, as represented by a number of special cameras now on the market. More recently the problem has arisen in connection with output printing for electronic computers where the required printing speeds are thousands of characters per second. Mechanical printers can no longer keep up and attention is being directed to cathode-ray techniques. The oscilloscope cameras would be fast enough but the cost of silver halide emulsions and the time and awkwardness of liquid development make them impractical for this use.

These difficulties are all eliminated in the "Electrofax"[1] process. The paper is relatively inexpensive, and automatic dry development at web speeds of more than 120 feet per minute can be provided. However, difficulty has been experienced due to the fact that the present process is considerably less sensitive than conventional photography. At maximum phosphor brightness and with the image projected through an f:4 lens, a trace can be printed on the paper at a spot speed of not more than 500 inches per second, equivalent in facsimile-type operation to 200 to 300 characters per second.

The thin-window cathode-ray tube has been developed to overcome this sensitivity handicap by, in effect, permitting contact rather than

[1] C. J. Young and H. G. Greig, "'Electrofax' Direct Electrophotographic Printing on Paper," *RCA Review*, Vol. XV, p. 469, December, 1954.

projection printing from the face of the tube. The geometrical conditions are shown in Figure 1. Assuming, according to an actual case, a 4-inch scanning line on the tube and a lens with an opening of f:4 (Figure 1a), then the approximate diameter of the lens is 1 inch. Further assuming a maximum enclosed angle of 30° at the lens for flat field, the distance D of the lens from the light source becomes

$$\tan 15 = \frac{2}{D},$$

$$D = 7.46 \text{ inches.}$$

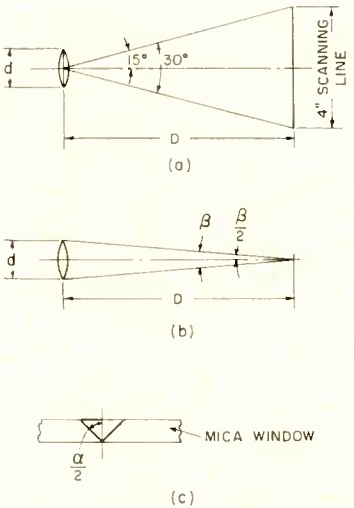


Fig. 1—Optical arrangement for projection printing (a and b) and for contact printing (c).

Half of the enclosed angle from each point of phosphor to the lens is (Figure 1b)

$$\tan \frac{\beta}{2} = \frac{.5}{7.46},$$

which gives, for the enclosed angle, $\beta = 7.666°$.

If, on the other hand, the recording paper is brought close to the phosphor and separated from it only by the thickness of the window, then all the light from the phosphor point, except that lost by total

reflection, will reach the paper. The angle, at which total reflection occurs is (Figure 1c)

$$\sin \frac{\alpha}{2} = \frac{1}{i}$$

in which i is the refractive index which, for Muscovite mica, is 1.56. Then

$$\sin \frac{\alpha}{2} = \frac{1}{1.56},$$

$$\alpha = 79.734°.$$

The ratio of the two angles is $79.734/7.666 = 10.4$ and the increase in light intensity therefore $10.4^2 = 108$. This gain factor is in good agreement with the results obtained in printing tests.

In order to preserve definition, the window separating the phosphor light source from the "Electrofax" paper should be as thin as possible but strong enough to hold a vacuum of the order of 10^{-7} mm Hg. The thickness of the membrane depends not only on the pressure but also on the width of the slot. For a window .080 inch wide a safe thickness for mica is .003 inch and for glass, due to the greater brittleness of the material, .005 inch.

Construction of the Tube

An experimental cathode-ray tube (Figure 2) for contact printing was built and tested. The window consisted of a clear mica strip ¾ inch wide and .003 inch thick which was sealed to the metal tube body by means of low-melting borosilicate glass. The slot in the metal was .080 inch wide and 8 inches long. The tube header was made semi-cylindrical to insure good contact between the window and the "Electrofax" paper.

After sealing, a target screen of P-11 phosphor was settled on the inside of the window, filmed and aluminized in the conventional manner. A 5W-type gun was sealed in the glass neck of the otherwise all-metal tube.

Along with the development of a mica thin-window tube, a parallel program was carried through which has established the technique for making a glass thin window. One tube was assembled which had a circular flat face plate made of Kovar into which a slot, 4 inches long by .080 inch wide, was cut. The window consisted of .005-inch-thick

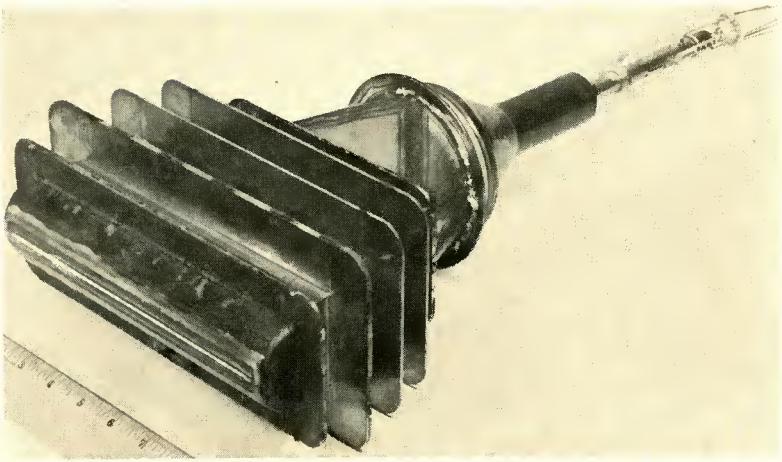


Fig. 2—Thin-window cathode-ray tube with 8-inch mica window.

glass which was stretched and sealed to a raised rim surrounding the slot. After sealing a screen of P-11 phosphor was settled on the glass, filmed and aluminized (Figure 3). Although there is no reason to believe that the glass-window and mica-window versions of the tube would differ in performance characteristics, it was felt that the investigation was desirable to determine which type was more easily fabricated.

Performance Tests

The results expected of the thin-window tube as a printing-out device were proven by a facsimile type test using the mica-window tube.

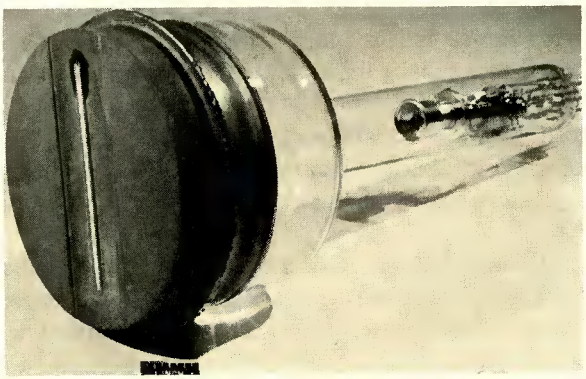


Fig. 3—Thin-window cathode-ray tube with 4-inch glass window.

To carry out the test, an existing flying spot scanning system was used. The scanning rate was 240 cycles per second with 120 lines per inch line frequency so that the linear speed of the recording paper web was 2 inches per second. In this system the beam from a 5ZP16 tube was focussed, by a lens, on a lantern slide containing the information to be recorded. The slide moved at the given rate of speed past the beam and, at the end of the travel, quickly returned to the initial position. A block diagram of the system is shown in Figure 4 and the arrangement of the thin-window recorder tube with the paper drive is illustrated in Figure 5.

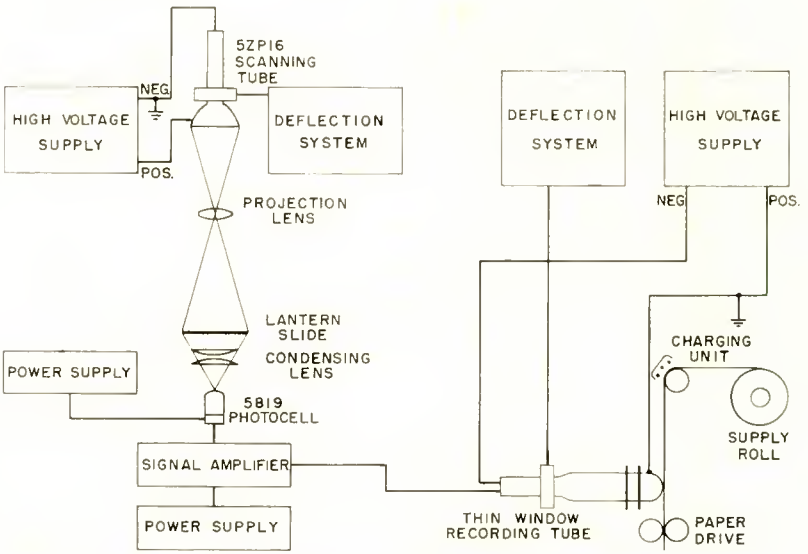


Fig. 4—Block diagram showing set-up for testing thin-window tube.

As can be seen from Figure 4 the tube, and in particular the cylindrical face plate, were at ground potential while the gun was at —25 to —27 kilovolts. This was necessary for safety reasons and also because such a high voltage on the metal part of the tube creates corona and stray electric fields which affect the charge on the "Electrofax" paper.

The paper used for the tests was standard "Electrofax" paper sensitized with .10 part of fluorescein sodium per 120 parts of zinc oxide. The paper was charged in the usual way by means of a corona unit. No attempt was made to include an automatic developer; the prints were developed manually using a magnetic brush as described

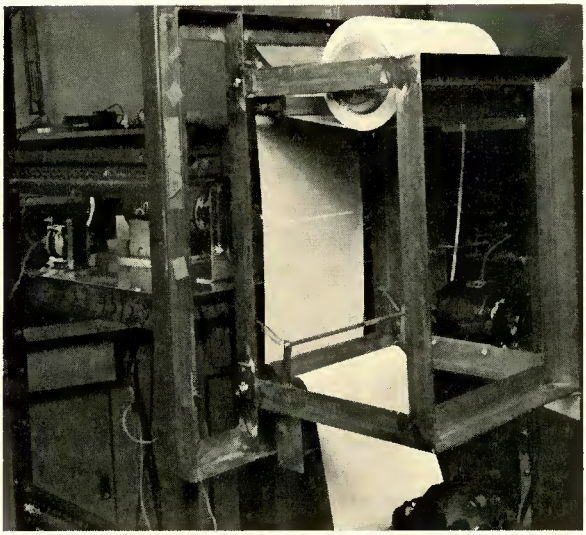
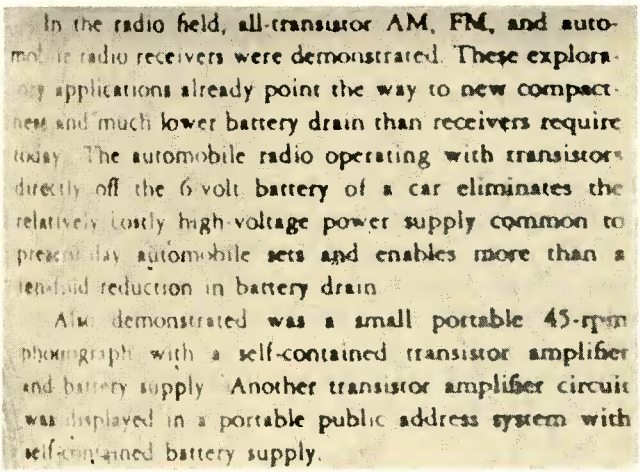


Fig. 5—Experimental printing arrangement showing "Electrofax" paper in contact with thin window.

by Young and Greig.[1] In view of the abrasiveness of the "Electrofax" paper and the softness of the mica window, a strip of clear Mylar .0005 inch thick was sandwiched between the window and the paper to prevent damage to the mica.

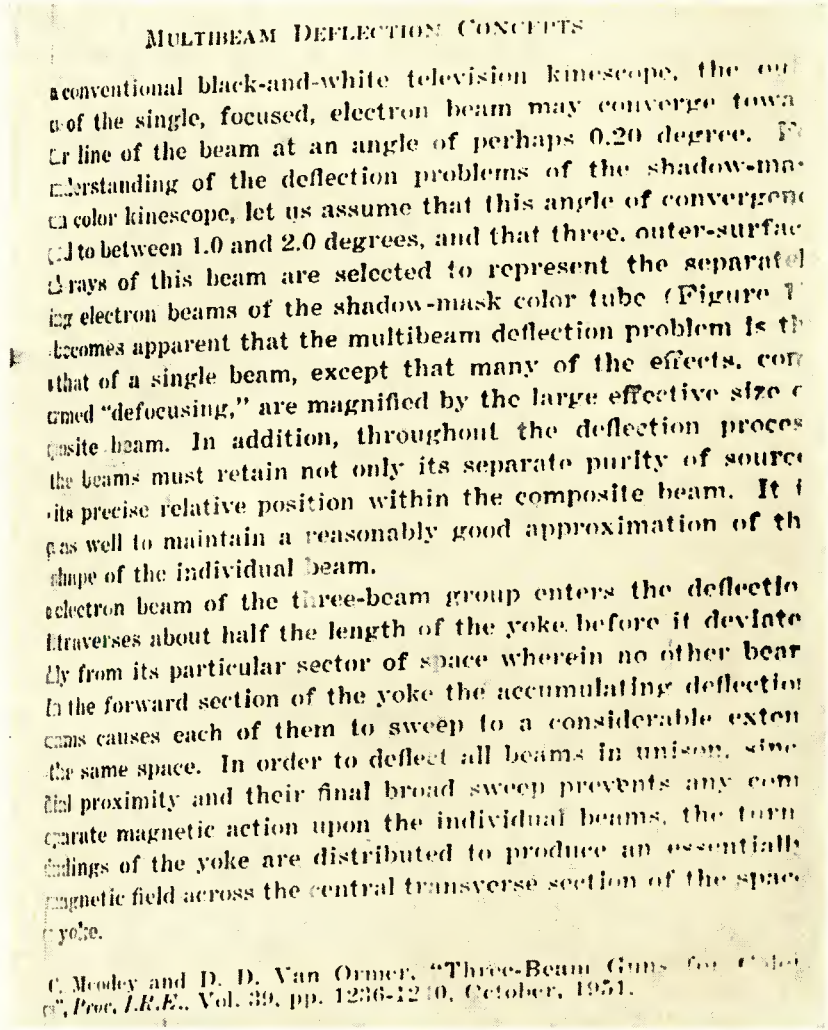
Samples of prints are shown in Figures 6 and 7. Figure 6 is a reproduction of a positive lantern slide (black letters on clear back-


In the radio field, all-transistor AM, FM, and auto-
mobile radio receivers were demonstrated. These explora-
tory applications already point the way to new compact-
ness and much lower battery drain than receivers require
today. The automobile radio operating with transistors
directly off the 6 volt battery of a car eliminates the
relatively costly high-voltage power supply common to
present day automobile sets and enables more than a
tenfold reduction in battery drain.

Also demonstrated was a small portable 45-rpm
phonograph with a self-contained transistor amplifier
and battery supply. Another transistor amplifier circuit
was displayed in a portable public address system with
self-contained battery supply.

Fig. 6—"Electrofax" print by direct development; cathode-ray beam off in letter areas.

ground) and was made with an average beam current of 3.5 microamperes. The print Figure 7 was made from a negative slide with an average beam current of .75 microampere. The distortion at both



MULTIBEAM DEFLECTION CONCEPTS

a conventional black-and-white television kinescope, the ou
n of the single, focused, electron beam may converge towa
er line of the beam at an angle of perhaps 0.20 degree. F
derstanding of the deflection problems of the shadow-ma
a color kinescope, let us assume that this angle of convergen
ed to between 1.0 and 2.0 degrees, and that three, outer-surfac
d rays of this beam are selected to represent the separatel
ing electron beams of the shadow-mask color tube (Figure 1
becomes apparent that the multibeam deflection problem is th
that of a single beam, except that many of the effects, com
monly termed "defocusing," are magnified by the large effective size o
composite beam. In addition, throughout the deflection proces
the beams must retain not only its separate purity of source
its precise relative position within the composite beam. It i
is well to maintain a reasonably good approximation of th
shape of the individual beam.
electron beam of the three-beam group enters the deflectio
traverses about half the length of the yoke before it deviate
ly from its particular sector of space wherein no other bear
In the forward section of the yoke the accumulating deflectio
coils causes each of them to sweep to a considerable exten
the same space. In order to deflect all beams in unison, sinc
their proximity and their final broad sweep prevents any com
separate magnetic action upon the individual beams, the turn
windings of the yoke are distributed to produce an essentially
magnetic field across the central transverse section of the space
the yoke.

C. Moodey and D. D. Van Ormer, "Three-Beam Guns for Color
e", *Proc. I.R.E.*, Vol. 39, pp. 1236-1240, October, 1951.

Fig. 7—"Electrofax" print by reverse development; cathode-ray beam on in letter areas.

ends of the line is due to nonlinearity of the saw tooth and can easily be corrected.

By scanning across a stationary bar pattern of graduated line width, dots of varying diameters were obtained. The diameter of the

smallest dots measured .010 inch, indicating that a 100 lines-per-inch definition is possible.

In order to get an indication of the writing speed obtainable, a sine-wave generator was connected to the thin-window tube. The paper was drawn manually across the window fast enough to separate the printed lines. At a beam current of 30 microamperes, a sweep frequency of 2,000 cycles per second, and an amplitude of 7 inches, the individual lines could clearly be distinguished. From this the writing speed, V_s, can be calculated.

$$V_s = \pi f A = 44{,}000 \text{ inches per second.}$$

Application

In an actual application it would be desirable to reduce the beam current to avoid browning of the phosphor. It would be reasonable to assume a current of 15 microamperes and a spot speed of 22,000 inches per second.

The tests which have been described do not demonstrate directly the speed capability of the combination of the thin-window cathode-ray tube with "Electrofax" printing. They do, however, supply data on the basis of which the high-speed possibilities of the system can be predicted. Among the possible applications are not only computer output printing but also fast facsimile and trace recording. For computers, it seems reasonable to predict 8½ inch webs of "Electrofax" paper running at two to four feet per second with continuous automatic development. This would mean the delivery of printed text at rates of 10,000 to 20,000 characters per second.

Acknowledgments

Many people have contributed to the success of the tubes described above. Particular credit belongs to J. E. Ruedy who was responsible for the parallel development of a glass version of the thin-window tube. Thanks are due to C. J. Young under whose guidance the project was carried out, H. C. Allen who designed the electronic equipment and cooperated in the tests, and P. G. Herkart for his help and advice in building the tube. Credit is also due to W. I. Rogers and W. H. Bleacher for the skill shown in carrying out the welding and sealing operations.

OPERATION AND PERFORMANCE OF THE 6866 DISPLAY STORAGE TUBE

BY

E. M. SMITH

RCA Electron Tube Division,
Lancaster, Pa.

Summary—*This paper describes a direct-view storage tube which provides an exceptionally bright visual presentation of radar-type information for relatively long periods. Principles of operation and important features of this type of tube are discussed.*

The design of the direct-view storage tube is explained to show how extended viewing duration, high brightness, half tones, fast writing and erasure, and good resolution are achieved. Performance characteristics are described, and the relation of these characteristics to the operating conditions is discussed.

INTRODUCTION

A MAJOR problem in the design of airborne fire-control radar equipment is the provision of information displays bright enough to be completely readable under the adverse viewing conditions frequently present in aircraft. A device designed to provide a solution for this problem is the 6866 display storage tube. This tube can display stored information with exceptional brightness (2750 foot-lamberts) for periods as long as a minute. It can also reproduce half tones with good resolution. It is capable of writing at speeds up to 300,000 inches per second, and has a nearly linear build-up writing characteristic which makes it possible to obtain very high signal-to-noise ratios in radar displays. Rapid erase may be achieved, and the rate of brightness decay, as well as the shape of the brightness-decay curve may be controlled by the erase process.

In addition to its usefulness in radar applications, the 6866 display storage tube is an extremely valuable tool for use in transient analysis, oscilloscope applications requiring slow-speed scanning, and other applications where cathode-ray tubes having long-persistence phosphor screens are normally employed.

DESCRIPTION AND OPERATING PRINCIPLES

Figure 1 is a photograph of the 6866, and Figure 2 shows the schematic arrangement of the tube. Table I shows operating voltages. The viewing section of the tube includes an axially centered viewing

gun provided with four grids, a storage unit consisting of a storage grid and a backing electrode, and an aluminized phosphor screen having very high visual efficiency. The writing section contains an axially offset writing gun which is provided with four grids and two pairs of electrostatic deflecting electrodes.

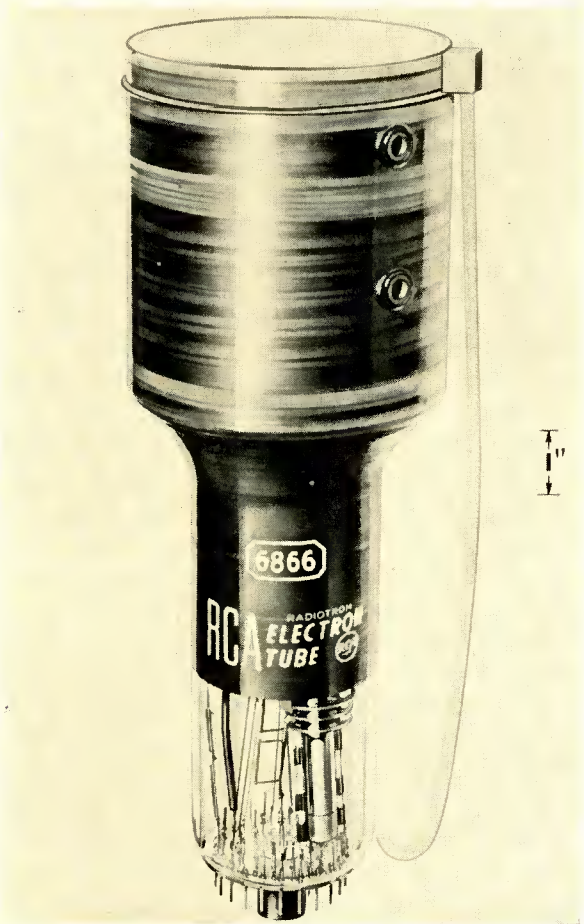


Fig. 1—Photograph of the 6866 display storage tube.

The viewing-gun cathode and the electrostatic lens formed by grids No. 1 and 2 produce a high-current, low-velocity stream of electrons which is collimated by grid No. 3, and floods grid No. 4, the storage grid, and the backing electrode. This stream determines the brightness of the display, and can have the high current density necessary for a very bright display because it is not focused, deflected, or modulated.

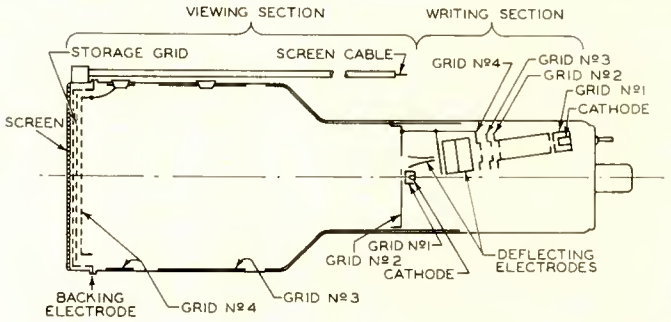


Fig 2.—Schematic arrangement of the 6866 display storage tube.

Collimation of the viewing-beam electrons is necessary so that they approach the storage grid along paths perpendicular to the plane of the grid and with uniform velocity, and can, therefore, be uniformly controlled by the storage-grid potential.

Grid No. 4 consists of a fine metal mesh and a band of conductive material on the interior of the glass envelope. It has four principal functions: (1) it provides an accelerating field for the collimated electrons of the viewing beam; (2) it repels positive ions produced by collisions of these electrons with molecules of residual gas in the region between the base and grid No. 4, so as to prevent these ions from landing on the storage grid and affecting a stored charge pattern; (3) it collects viewing-beam electrons which are turned back by the storage grid when the latter is at a sufficiently negative potential with respect to the viewing-gun cathode; (4) it collects secondary electrons displaced from the storage grid by the writing beam.

The storage unit consists of a very fine metal mesh (the backing electrode) coated on its gun side with a very thin deposit of high-quality insulating material (the storage grid). The coating does not obstruct the apertures of the backing electrode and, therefore, does

Table I—Typical Operating Conditions for the 6866 Display Storage Tube.

	Potentials, in volts, with respect to cathode of viewing gun	
	Viewing Section	Writing Section
Screen	10,000	—
Backing-Electrode (D.C.)	5	—
Grid—No. 4	210	70 to 105*
Grid—No. 3	50 to 150	—2100 to —1700
Grid—No. 2	70 to 105*	70 to 105*
Grid—No. 1	0 to—50	—2400 to —2515
Cathode	0	—2400

* Grid No. 2 of the Viewing Section is connected internally to Grid No. 4 of the writing section. Usually Grid No. 4 and Grid No. 2 of the writing section are connected externally.

not interfere mechanically with the flow of electrons to the screen. The storage unit represents a very great number of capacitive storage elements. One plate of a capacitive storage element is an elemental area of the gun-side surface of the storage grid, and the other plate, common to all elements, is the backing electrode. The dielectric is the intervening insulating material. The charges stored in these elements determine the number of viewing-beam electrons which reach the corresponding areas of the phosphor screen, and, therefore, control the brightness of the display.

Approximately 50 per cent of the electrons in the viewing beam are collected by grid No. 4, and the remainder are decelerated as they approach the storage grid to practically zero velocity. Initially, most of these electrons land on the storage grid, and, because of their very low velocity, produce a secondary-emission ratio of less than unity. These electrons, therefore, charge the storage grid to the potential of the viewing-gun cathode—i.e., to zero potential. As a result, viewing-beam electrons which subsequently approach the storage grid do not land there, but are, instead, accelerated by the positive high-voltage field of the phosphor screen through the apertures of the storage grid and backing electrode. These electrons then strike the phosphor screen and cause it to fluoresce over its entire area.

The storage grid can be used to control the screen current in the same manner as the grid of a triode is used to control the plate current — that is, when operated as a sufficiently negative voltage it can cut off the current to the screen. Under cutoff conditions, viewing-beam electrons which approach the storage grid are turned back and are collected by grid No. 4. For the typical operating conditions shown in Table I a storage-grid potential of —8 volts is sufficient to assure cutoff.

Because of the capacitive character of the storage elements the storage grid may be placed at cutoff potential by application of a positive pulse to the backing electrode. Assume that the storage grid is uniformly charged by the viewing beam to the potential of the viewing-gun cathode, i.e., to zero potential, and the backing electrode is at an applied potential of 5 volts. A charge corresponding to a potential difference of 5 volts is, therefore, stored in the dielectric between the gun sides of the storage grid and backing electrode. Assume, also, that a positive-going rectangular pulse having an amplitude of 8 volts and a duration of approximately 50 milliseconds is applied to the backing electrode. At the instant the pulse is applied, the potential of the storage-grid surface is shifted, by capacitive coupling to the backing electrode, to +8 volts. Viewing-beam electrons then land on the storage-grid surface and charge it to zero potential in

approximately 50 milliseconds. As a result, a charge corresponding to a potential difference of 8 + 5, or 13 volts is stored in the dielectric. At the end of the positive pulse the backing electrode returns to its initial applied potential of 5 volts, but, because of the charge stored in the dielectric, the potential of the storage-grid surface simultaneously changes from zero to a value of 5+ (—13), or —8 volts, the value required for cutoff.

Writing Section

The writing gun is similar to the guns used in oscillograph tubes using electrostatic focus and deflection. It produces a well-defined, focused, high-velocity beam of small area, which may be deflected and intensity modulated in the same manner as the beam of an oscillograph tube. This beam is used to scan the storage grid and to modify the charge distribution produced by the viewing beam and cutoff pulse in such a manner that the information to be displayed appears on the phosphor screen.

As shown in Table I, the writing-gun cathode is operated at a potential of approximately 2,400 volts negative with respect to the viewing-gun cathode and the storage grid. The electrons of the writing beam, therefore, land on the storage grid with sufficient energy to produce a secondary-emission ratio greater than unity. As a result, the written elements of the storage grid assume less-negative potentials proportional to the density and dwelling time of the writing beam. When these potentials are more positive than the cutoff value, viewing-beam electrons pass through the corresponding storage-grid apertures and strike the phosphor screen, where they produce luminous spots only slightly larger than the corresponding storage elements.

Because secondary electrons driven from the storage grid by the writing beam are attracted to grid No. 4, a written element of the storage grid may, theoretically, be driven to grid No. 4 potential — that is to a potential of approximately 210 volts. Under normal operating conditions, however, no storage-grid element will be written to a potential higher than that of the viewing-gun cathode — that is, zero volts — because of the continuous flooding action of the viewing beam. The brightness of the display at any point is determined by the number and energy of the viewing-beam electrons which strike the phosphor screen at that point as a result of the changes in storage-grid potential produced by the writing operation. The maximum brightness obtainable, therefore, is determined by the potentials applied to the various electrodes of the viewing section, and not by the characteristics of the writing beam. The latter, consequently, is primarily a control device which contributes little to the brightness of the display.

Half-Tone Displays

The 6866 can display half tones provided the writing and erasing processes are carefully controlled so that the potentials of written storage-grid elements remain in the linear region of the storage-grid control characteristics. Because of the narrowness of this linear region, extremely precise adjustments are necessary to provide even a limited number of half-tone steps. If the written storage-grid elements are overdriven, that is, charged to potentials equal to or more positive than that of the viewing-gun cathode, no half-tone steps can be displayed, and serious deterioration of resolution will result.

Duration of Visibility

In the absence of a cutoff pulse or the erasing process described below, the potentials of written storage elements are not affected by the flooding action of the viewing beam, and a visible pattern remains on the phosphor screen after writing has ceased. The duration of visibility is limited, however, by the effects of positive ions produced by collisions of viewing-beam electrons with molecules of residual gas in the region between grid No. 4 and the phosphor screen. The positive ions are attracted to the most negative (unwritten) areas of the storage grid, where they gradually neutralize the negative potentials, permitting flow of viewing-beam electrons to the phosphor screen. Since the effect of this action is a gradual increase in background brightness the duration of visibility is determined by loss of contrast rather than by the decay in the brightness of the display which occurs in conventional cathode-ray tubes.

Erasure

A display can be completely erased by application to the backing electrode of a positive pulse or series of pulses having sufficient amplitude and duration to restore the entire storage grid to cutoff potential. Under the typical operating conditions shown in Table I, the display can be erased by a single 10-volt rectangular pulse of approximately 40 milliseconds duration, or by a series of 10-volt pulses having a total duration of 40 milliseconds. For example, a series of 10-volt, 10-microsecond pulses having a repetition rate of 1,000 pulses per second will produce complete erasure in approximately four seconds.

A rectangular erasing pulse charges all elements of the storage grid to cutoff potential at substantially the same rate regardless of the charge on any element. Consequently, when such a pulse is used, the brightest areas of the displayed pattern remain visible for longer periods than do half tones. A sawtooth-type erasing pulse, on the other hand, charges the most positive elements of the storage grid

toward cutoff potential most rapidly because it permits the viewing-beam electrons to land on these elements for longer periods than on the other elements. Consequently, when this type of erasing pulse is used, half-tone areas of a display remain visible as long as the brightest areas.

Build-up Characteristic

A valuable feature of the 6866 display storage tube, particularly for radar applications, is the nearly linear build-up of display brightness obtainable by repeated addresses of the writing beam. Each address of the beam to a storage element increases the brightness of the corresponding point on the phosphor screen in approximately direct proportion to the beam density. Consequently, a periodic, consistent signal such as a radar target signal can be integrated by repeated addresses into an extremely bright indication, whereas the associated background noise, because of its random character, is not built up to the same degree. Radar displays which are substantially free of noise can, therefore, be obtained by adjustment of the erasing-pulse amplitude so that the storage grid is charged by the erasing process to a potential more negative than cutoff potential by an amount equal to the charge produced by the noise voltage. The noise voltage will, therefore, charge the storage grid to cutoff, and will not be displayed, whereas a useful signal superimposed on the noise will charge the storage elements to a potential more positive than cutoff and will be displayed.

Brightness Decay

Another advantage of this display storage tube is that the decay of brightness during or following application of the erasing pulse may be made approximately linear. For example, when conditions are adjusted so that 10 or more seconds are required for complete erasure, the brightness of the brightest area two seconds after writing has ceased is at least 80 per cent of the initial value. In contrast, the brightness of the display on a conventional cathode-ray tube using a long-persistence phosphor screen decays to less than one per cent of its initial value within two seconds after excitation is removed.

Performance Characteristics

The outstanding characteristic of the 6866 display storage tube is the exceptional brightness of the display. A typical value for light output of a display written to maximum brightness with the tube operated under the conditions shown in Table I is 2,750 foot-lamberts.

Maximum brightness is achieved when the phosphor screen and grid No. 4 of the viewing gun are operated at their maximum rated voltages. An increase in grid No. 4 voltage produces an approximately proportional increase in maximum phosphor-screen current and a corresponding increase in brightness. An increase in phosphor-screen voltage, however, is more effective because the maximum brightness varies approximately as the third power of the phosphor-screen voltage in the range between 5 and 10 kilovolts.

When the tube is operated under the conditions shown in Table I, the writing and viewing functions are substantially independent. The fact that the display brightness is independent of the writing operation is especially useful in application requiring both high deflection sensitivity and bright displays, because the writing gun may be operated at a relatively low voltage to obtain high deflection sensitivity, and high display brightness obtained by an increase in the phosphor screen voltage.

The potential at which the backing electrode is operated affects the erasing speed. Because a reduction in backing-electrode voltage reduces the storage-grid voltage required for cutoff, it also reduces the amplitude and duration of the pulse required to produce complete erasure within a given time.

The rate of erasure is also proportional to the number of viewing-beam electrons which land on the storage-grid surface. Because this number increases with the potential applied to grid No. 4 of the viewing gun, fastest erasure is obtained when grid No. 4 is operated at its maximum rated voltage.

The phosphor screen of the display storage tube reaches its maximum brightness during application of an erasing pulse. The background brightness during this interval is a function of the erasing rate and the backing-electrode potential, and is approximately equal to the saturated brightness multiplied by the ratio of erasing pulse duration to the time between the leading edges of successive erasing pulses.

Half-Tone Displays

The number of half-tone steps which can be displayed under optimum operating conditions is limited by nonuniformities in the control characteristics of various elements which tend to shift the cutoff potential of the storage grid. Such nonuniformities may be caused by defects in the backing-electrode mesh, by poor collimation of the viewing beam, by variations in the spacing between the storage grid and the phosphor screen and between the storage grid and grid No. 4, and by variations in the density of the viewing beam. Non-

uniformities due to mesh defects, poor collimation, and spacing variations can be detected if the screen is first written to uniform brightness over the entire viewing area, and the backing-electrode potential then slowly reduced until the brightness of the display is zero. The presence of such nonuniformities will be indicated by differences in the backing-electrode potential required for cutoff at different points in the display area.

Nonuniformities caused by variations in the secondary-emission ratio of the storage grid, in the density of the viewing beam, and in the capacitance between the corresponding elements of the storage grid and backing electrode are most easily detected during pulse-erase operation. Under these conditions the variations in cutoff voltage resulting from such nonuniformities will cause corresponding variations in the time required for all elements of the display to reach zero brightness.

Variations in the secondary-emission ratio of the storage grid and in the capacitance of the storage elements may also cause nonuniformity in writing speed. Deflection defocusing, however, may also have similar effects, because it changes the density of the writing beam and, therefore, the writing speed in particular areas of the screen. In most applications, satisfactory compensation for the effects of deflection defocusing may be achieved by a compromise adjustment which provides the same degree of focus at the center and edges of the screen.

The effects of many of the nonuniformities described above can be reduced by proper adjustment of tube operating conditions. For example, the additive effects of certain nonuniformities in the writing and erasing processes may be minimized if all old information is completely erased before new information is written. In other words, most uniform writing speed and erasure are achieved when the storage grid is uniformly charged to the viewing-beam cutoff voltage before writing.

Best resolution, on the other hand, is achieved when the storage grid is charged by the writing beam to a potential less positive than that required for maximum brightness, i.e., viewing-gun cathode potential. In other words, the resolution capabilities of the tube decrease as the stored charges on written areas approach the value required for maximum brightness. In this respect the display storage tube is similar to a cathode-ray tube, i.e., a decrease in grid bias produces an increase in brightness, but a relatively greater loss in resolution.

Writing

As stated earlier, the writing beam of the display storage tube is primarily a control mechanism, and does not contribute appreciably to the maximum brightness of the resulting display, although the relative brightness is determined by the density and dwelling time of the writing-beam current. Once a written element has been charged by the writing beam to zero potential, repeated writing of that element will not increase brightness, and, in fact, will cause serious deterioration of the resolution in that area of the display.

Except in applications requiring extremely high writing speeds, relatively little writing beam current and, therefore, low driving voltage are required to produce displays of maximum brightness. In applications using very slow deflection of the writing beam or where certain storage elements are written repeatedly, a driving voltage of one or two volts is usually adequate. While such low driving voltages simplify the signal-source requirements and make it possible to obtain very high resolution, they also make it difficult to maintain precise control of the resulting charges developed in the storage elements. Under low-drive conditions, therefore, it is necessary to use electrode-voltage supplies which are free from ripple.

When the writing-gun cathode is between 1,500 and 2,500 volts negative with respect to the viewing-gun cathode, the collimating electrodes of the viewing section have little effect on the high-velocity writing beam. When the potential difference between the two cathodes is only a few hundred volts, however, the collimating electrodes cause some deflection of the writing beam. Since the writing gun is mounted off-axis, the resulting distortion may appear as "keystoning," "barrel," or "pincushion" in the display. The severity of these effects will depend upon the collimating electrode voltages.

Display Duration

A display containing half tones will remain visible for at least 20 seconds when the 6866 is operated with the typical voltages shown in Table I, and for approximately 40 seconds when the viewing-gun grid No. 4 voltage is reduced from 210 volts to 150 volts. The duration of a display may be extended to two minutes or more if half tones are not essential. The increased duration is achieved by use of an erasing pulse having an amplitude several volts higher than that required for complete erasure. Such a pulse charges the unwritten storage elements more negative than cutoff voltage and consequently increases the time necessary to discharge these elements to cutoff potential.

THE EQUIVALENT CIRCUIT OF THE DRIFT TRANSISTOR

By

J. Almond and R. J. McIntyre

RCA Victor Company, Ltd.,
Montreal, Que.

Summary—*The four-terminal admittances for the drift transistor, derived from the admittance equations worked out by Kroemer,[2] have been examined as functions of frequency and approximated by an equivalent circuit. The equivalent circuit is in two forms, one suitable for the common-base connection and the second suitable for the common-emitter connection. The frequency response of each equivalent circuit admittance is plotted and compared with that of the device admittance function which it represents. The differences in the behavior of the phase of α at high frequencies are discussed qualitatively for the drift and diffusion transistors.*

Introduction

MUCH has been written about equivalent circuits for junction transistors.[1] Most of the development to date, however, has been concerned with equivalent circuits for the ordinary diffusion transistor. Since the drift transistor employs a new principle not utilized in the diffusion transistor, it is to be expected that equivalent circuits for the diffusion transistor will not be applicable in all respects to the drift transistor. Kroemer[2] in 1954 derived expressions for the four-terminal admittance parameters for the drift transistor and gave a low-frequency equivalent circuit for the common-base connection. These admittance expressions are re-examined with the object of deriving equivalent circuits for both the common-base and the common-emitter connection. The results of this investigation are reported.

In the first section, the four-terminal admittances as functions of frequency are examined and approximations necessary to reduce them to usable forms are introduced. The functions are then plotted for typical values of device parameters. A π configuration is assumed for the equivalent circuit and the elements of the π network are written in terms of the admittance functions. In the next section the elements

---

[1] R. L. Pritchard, "Electrical Network Representation of Transistors—A Survey," *Trans. I.R.E. PGCT*, p. 5, March, 1956.

[2] H. Kroemer, "The Theory of Diffusion and Drift Transistors, Part I—The Quadripole Matrix and its Low-Frequency Behavior," *Archiv Der Elektrischen Ubertragung*, Vol. 8, p. 223, May, 1954.

of the π network are approximated by equivalent two-terminal electrical networks. The procedure in arriving at the equivalent networks is to expand the combinations of the admittance functions, which make up a given π element, into a power series in a normalized frequency parameter x, where the frequency corresponding to $x = 1$ is approximately the α-cutoff frequency. The admittance of a suitably chosen two-terminal electrical network is also expanded in powers of a normalized frequency and its elements then evaluated by equating coefficients.

Values of the circuit elements are calculated for typical values of the device parameters. In the last section, the diffusion and the drift transistor are discussed qualitatively in an attempt to explain the difference in behavior of the phase of α at high frequencies for the drift transistor.

The physical principles and methods of construction of the drift transistor are not discussed in detail in this paper. A thorough treatment of these aspects may be found in the literature.[2-6]

The Four-Terminal Admittance Functions

The circuit equations for the three-terminal box shown in Figure 1 may be written:

$$\begin{vmatrix} i_1 \\ \\ i_2 \end{vmatrix} = \begin{vmatrix} Y_{11} & Y_{12} \\ \\ Y_{21} & Y_{22} \end{vmatrix} \begin{vmatrix} V_1 \\ \\ V_2 \end{vmatrix} \tag{1}$$

When the equations are written in this form, terminal 3 is taken to be at the reference or ground potential. It is often more convenient

---

[3] H. Kroemer, "The Theory of Diffusion and Drift Transistors, Part II — Frequency Dependence," *Archiv Der Elektrischen Ubertragung*, Vol. 8, p. 363, August, 1954.

[4] H. Kroemer, "The Theory of Diffusion and Drift Transistors, Part III — Dimensional Questions," *Archiv Der Elektrischen Ubertragung*, Vol. 8, p. 499, November, 1954.

[5] H. Kroemer, "The Drift Transistor," *Transistors I*, RCA Laboratories, Princeton, N. J., 1956, p. 202.

[6] A. L. Kestenbaum and N. H. Ditrick, "Design, Construction and High-Frequency Performance of Drift Transistors," *RCA Review*, Vol. XVIII, p. 12, March, 1957.

to leave the reference potential unspecified. We can do this by forming the indefinite admittance matrix;[7]

$$\begin{vmatrix} i_1 \\ i_2 \\ i_3 \end{vmatrix} = \begin{vmatrix} Y_{11} & Y_{12} & -Y_{11}-Y_{12} \\ Y_{21} & Y_{22} & -Y_{21}-Y_{22} \\ -Y_{11}-Y_{21} & -Y_{12}-Y_{22} & Y_{11}+Y_{21}+Y_{12}+Y_{22} \end{vmatrix} \begin{vmatrix} V_1 \\ V_2 \\ V_3 \end{vmatrix} \tag{2}$$

To form the admittance equations with any of the three terminals as the reference, the appropriate row and column of the indefinite matrix are crossed out. From this it is seen that, provided the admittance matrix is known for one terminal common, the admittance matrix can easily be obtained if we wish to consider either of the other two terminals common.

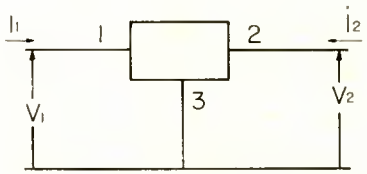


Fig. 1—Three-terminal box representation.

The four admittance functions for the drift transistor with the base lead as the common terminal have been worked out by Kroemer[2] for a p-n-p transistor and are repeated here for convenience.

$$(Y_{11})_b = (i_e^{(p)} - i_s)\,\frac{q}{kT}\,\frac{1+\dfrac{f}{\delta}\coth\dfrac{w}{\delta}}{1+\dfrac{f}{\delta_o}\coth\dfrac{w}{\delta_o}}, \tag{3}$$

$$(Y_{12})_b = i_e^{(p)}\,\frac{dw}{dV_c}\,\frac{e^{-\frac{w}{f}}}{\delta\sinh\dfrac{w}{\delta}}, \tag{4}$$

[7] J. Shekel, "Matrix Representation of Transistor Circuits," *Proc. I.R.E.*, Vol. 40, p. 1493, November, 1952.

$$(Y_{21})_b = -(i_e^{(p)} - i_s)\frac{q}{kT}\frac{e^{\frac{w}{f}}}{\delta \sinh\frac{w}{\delta}\left(\frac{1}{f}+\frac{1}{\delta_o}\coth\frac{w}{\delta_o}\right)}, \tag{5}$$

$$(Y_{22})_b = -\left[i_c^{(p)}\left(\frac{1}{\delta\tanh\frac{w}{\delta}} - \frac{1}{f}\right) - i_D\frac{1}{L}\right]\frac{dw}{dV_c}, \tag{6}$$

where
$$i_s = \frac{qAD_p p_o}{\delta_o}\left(\frac{1}{\tanh\frac{w}{\delta_o}} + \frac{\delta_o}{f} - \frac{e^{\frac{w}{f}}}{\sinh\frac{w}{\delta_o}}\right) \tag{7}$$

$= $ the emitter saturation current which flows when a reverse voltage $\gg kT/q$ is applied at the emitter.

$$i_d = \frac{qAD_p}{L}p_o(w) \tag{8}$$

$= $ the hole saturation current of a p-n diode where the impurity concentration on the n side is constant.

$i_e^{(p)} = $ the emitter hole current,

$i_c^{(p)} = $ the collector hole current,

$D_p = $ Diffusion constant for holes,

$\tau_p = $ recombination lifetime for holes,

$q = $ charge on an electron,

$A = $ junction area,

$p_o = $ equilibrium hole concentration at the base side of the emitter–base junction,

$\frac{dw}{dV_c} = $ change of base-width with collector voltage,

$w = $ the average base-width,

$\Delta V = $ potential-energy difference in the base region due to the impurity distribution $= qFw$,

F = drift field in the base region,

$p_o(w)$ = equilibrium hole concentration at the collector end of the effective base,

and the following notations and relations are used:

$$L = \sqrt{D_p \tau_p}, \quad f = \frac{2kT}{qF}, \quad \frac{1}{\delta_o^2} = \frac{1}{f^2} + \frac{1}{L^2},$$
$$\frac{1}{\lambda^2} = \frac{1 + j\omega\tau_p}{D_p\tau_p} = \frac{1}{L^2} + \frac{j\omega}{D_p}, \quad \frac{1}{\delta^2} = \frac{1}{f^2} + \frac{1}{\lambda^2}. \tag{9}$$

Equations (3) to (7) consider only the effect of the hole current (minority carriers). Kroemer points out[2] that the electron currents may be neglected in all of the parameters except Y_{11}. The effect of the electron current can be approximated in the admittance Y_{11} if we include the current gain, α, the emitter efficiency, γ, and the transport factor, β.

$$\alpha = -\frac{Y_{21}}{Y_{11}} = -\frac{Y_{21}^{(p)}}{Y_{11}^{(p)} + Y_{11}^{(n)}} = -\frac{Y_{21}^{(p)}}{Y_{11}^{(p)}} \cdot \frac{Y_{11}^{(p)}}{Y_{11}^{(p)} + Y_{11}^{(n)}} = \beta\gamma. \tag{10}$$

The emitter efficiency, γ, may be written

$$\gamma = \frac{1}{1 + \dfrac{Y_{11}^{(n)}}{Y_{11}^{(p)}}} \approx 1 - \frac{Y_{11}^{(n)}}{Y_{11}^{(p)}}. \tag{11}$$

The only time that this effect will be of interest will be when it is necessary to subtract from the term Y_{11} another admittance Y_{21}. These admittances are of the same order of magnitude. The important term under these conditions is $(1-\alpha)$ which may be written $(1-\alpha) = 1 - \beta\gamma$. At low frequencies the value of this term is much less than unity as both β and γ are close to unity. Consequently, the low frequency value of γ becomes quite critical. At higher frequencies both β and γ decrease with frequency, but β is a much stronger function of frequency than γ, since the only frequency dependent term in γ is the ratio $Y_{11}^{(n)}/Y_{11}^{(p)}$ which is small compared to unity (Equation (11)). We therefore can neglect the frequency variation of γ and use only the low-frequency value, γ_o, throughout. To use the complete admittance

$(Y_{11})_b$ we then simply substitute $i_e^{(p)}/\gamma_o$ for $i_e^{(p)}$. In this discussion we have neglected the effect of the emitter and collector transition capacitances which arise from the fact that with a change in voltage there is a change in depletion layer thickness and thus in the total amount of charge in the depletion layer. These capacitances could be added to the admittance functions[2] at this point and an equivalent circuit found which included them. However, another, perhaps simpler approach is to obtain an equivalent circuit for the intrinsic transistor and then to add the transition capacitances as external elements.

A few new terms will now be defined and Equations (3) to (7) rewritten. Also, the saturation currents i_s and i_d compared to i_e and i_c will be neglected.

$$\frac{w}{f} = \epsilon, \quad \frac{w}{\delta_o} = \epsilon' \quad \text{and} \quad \frac{f}{\delta_o} = \frac{\epsilon'}{\epsilon} = \sqrt{1 + \frac{w^2}{L^2\epsilon^2}}. \tag{12}$$

$\epsilon = \Delta V/(2kT)$ and is a measure of the strength of the drift field. If $\epsilon > 1$, $w^2/L^2\epsilon^2$ may be neglected compared to unity in most cases, and $\epsilon' = \epsilon$. However, there is one occasion when this slight difference is important and this difference is preserved for that reason. One further change is made in notation. Let

$$\Omega = \sqrt{1 + j\omega\delta_o^2/D_p}.$$

Equation (3) then becomes*

$$Y_{11} = \frac{i_c^{(p)}}{\gamma_o} \cdot \frac{q}{kT} \cdot \frac{1 + \frac{\epsilon'}{\epsilon}\Omega \frac{1 + e^{-2\epsilon'\Omega}}{1 - e^{-2\epsilon'\Omega}}}{1 + \frac{\epsilon'}{\epsilon} \frac{1 + e^{-2\epsilon'}}{1 - e^{-2\epsilon'}}}. \tag{13}$$

The hyperbolic cotangents have been written in their exponential forms. For $\epsilon' \gg 1$ the negative exponents may be neglected since for $\epsilon' = 2$ the error in neglecting the exponents would be of the order of 5 per cent, while for $\epsilon' = 3$, the error would be less than 1 per cent. It is assumed that the drift field is strong enough so that $e^{-\epsilon'}$ is

* The subscript (b) on the four-terminal admittances has been dropped for convenience. When these admittances are used in the remainder of the paper, they refer to the common-base connection.

negligible compared to unity. It is also assumed for this purpose that $\epsilon/\epsilon' = 1$.

Equation (13) then reduces to

$$Y_{11} = \frac{i_e^{(p)}}{\gamma_o} \frac{q}{kT} \frac{1+\Omega}{2}. \tag{14}$$

Making the same approximations as before, Equation (4) reduces to

$$Y_{12} = 2i_c^{(p)} \frac{\epsilon'}{w} \frac{dw}{dV_c} \Omega e^{-(\epsilon+\epsilon'\Omega)}. \tag{15}$$

Equation (5) reduces to

$$Y_{21} = -i_c^{(p)} \frac{q}{kT} \Omega e^{(\epsilon-\epsilon'\Omega)} \tag{16}$$

and Equation (6) reduces to

$$Y_{22} = -\frac{i_c^{(p)}}{w} \frac{dw}{dV_c} [2\epsilon'\Omega e^{-2\epsilon'\Omega} + \epsilon'\Omega - \epsilon]. \tag{17}$$

In this last equation one more term has to be used in the expansion of the hyperbolic tangent, since the earlier terms almost cancel. For convenience in calculation it is simpler to put the admittance functions in terms of their low frequency values. Then,

$$\frac{Y_{11}}{Y_{11}^{(o)}} = \frac{1+\Omega}{2}, \quad \frac{Y_{12}}{Y_{12}^{(o)}} = \Omega e^{\epsilon'(1-\Omega)}$$

$$\frac{Y_{21}}{Y_{21}^{(o)}} = \Omega e^{\epsilon'(1-\Omega)}, \quad \frac{Y_{22}}{Y_{22}^{(o)}} = \frac{\Omega e^{2\epsilon'(1-\Omega)} + \frac{1}{2} e^{2\epsilon'} \left(\Omega - \frac{\epsilon}{\epsilon'} \right)}{1 + \frac{1}{2} \left(1 - \frac{\epsilon}{\epsilon'} \right) e^{2\epsilon'}}. \tag{18}$$

These functions have been plotted for a value of $\epsilon' = 4$ and in the case of Y_{22} we have assumed that $\epsilon/\epsilon' = 0.999688$, which is equivalent to assuming $w/L = .1$. The functions are plotted in terms of x, where $x = \omega\delta_o^2/D_p$ and $\Omega = \sqrt{1+jx}$. These functions are shown in Figures 2 to 4. The shape of the curve for Y_{22} (Figure 4) depends quite critically on the assumed value of ϵ/ϵ'. The shape of the curve at low frequencies depends on the relative low frequency values of the two terms in

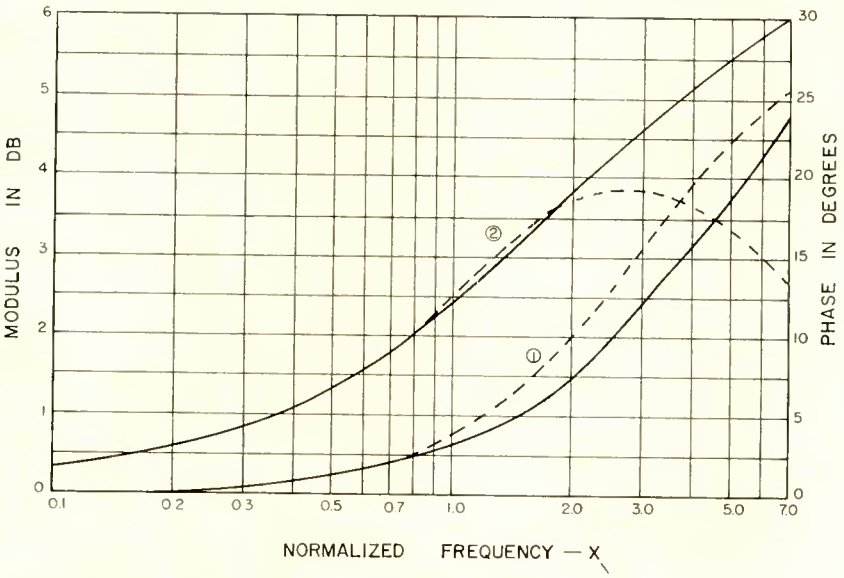


Fig. 2—The variation of $\frac{Y_{11}}{Y_{11}(0)}$ with frequency. Curve 1 represents the modulus and curve 2 the phase. The dotted curves show the equivalent circuit approximations.

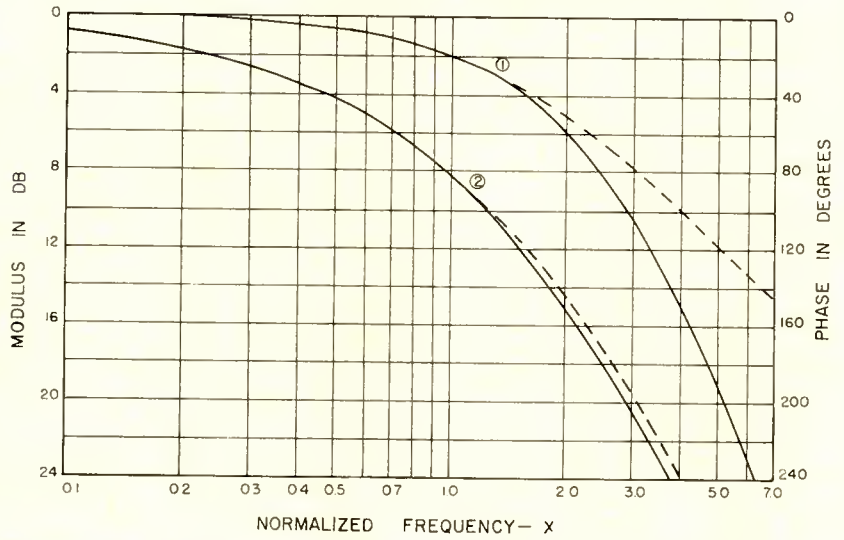


Fig. 3—The variation of $\frac{Y_{12}}{Y_{12}(0)}$ and $\frac{Y_{21}}{Y_{21}(0)}$ with frequency. Curve 1 represents the modulus and curve 2 the phase. The dotted curves show the equivalent circuit approximations.

Equation (17). At higher frequencies the second term dominates and the phase asymptotically approaches 45 degrees.

When considering the representation of the transistor by means of an equivalent circuit, the equivalent circuit parameters will be combinations of the four admittance parameters discussed above. To obtain the desired accuracy in some of these combinations, it may be necessary to obtain more accurate expressions for the admittance parameters than those used to plot the functions in Figures 2 to 4. It is useful at this point to consider just what combinations of the original admittance functions will arise in the equivalent circuit representation. The form of equivalent circuit chosen is a π network with a current generator in the output leg. The form of the circuit is shown in Figure 5.

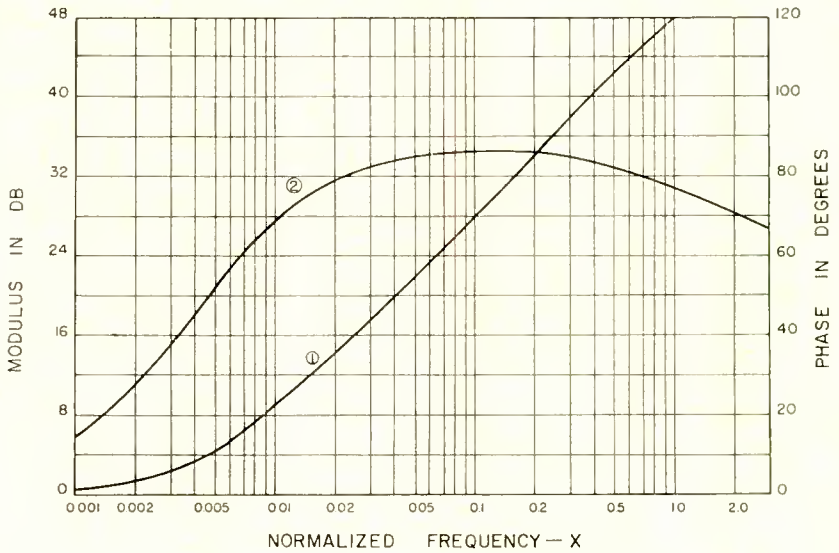


Fig. 4—The variation of $\frac{Y_{22}}{Y_{22^{(0)}}}$ with frequency. Curve 1 represents the modulus and curve 2 the phase.

The admittance $Y_{12} \ll Y_{11}$ or Y_{21} and therefore it may be neglected when added to these terms. However, Y_{22} and Y_{12} are of the same order of magnitude over part of the frequency range of interest and, as they are opposite in sign, the term $Y_{22} + Y_{12}$ should be considered. An equivalent circuit of the form shown in Figure 5 is quite useful for the common-base connection; however, for the common-emitter, it would be more suitable in the form shown in Figure 6.

This equivalent circuit can easily be formed from Figure 5 by using the indefinite admittance matrix previously described. The only element which appears different in this equivalent circuit is the input branch $Y_{11} + Y_{21} - Y_{12}$. Again, Y_{12} may be neglected compared to Y_{11} and Y_{21}. The two combinations of admittances which are of interest in the equivalent circuits of Figures 5 and 6 are $Y_{11} + Y_{21} = Y_s$ and $Y_{22} + Y_{12} = Y_c$. The expressions for these two combined admittances can be developed from Equations (4) to (7) and the results are given below.

$$Y_c = Y_{22} + Y_{12} = -\frac{i_c}{w}\frac{dw}{dV_c}\left[\epsilon'\left(\Omega - \frac{\epsilon}{\epsilon'}\right)(1 - 2\epsilon'\Omega e^{-2\epsilon\Omega})\right], \quad (19)$$

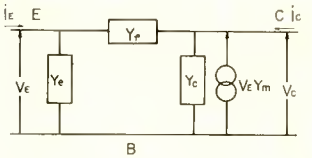


Fig. 5—π network for common-base connection, where $Y_r = -Y_{12}$, $Y_e = Y_{11} + Y_{12}$, $Y_c = Y_{22} + Y_{12}$ and $Y_m = Y_{21} + Y_{12}$.

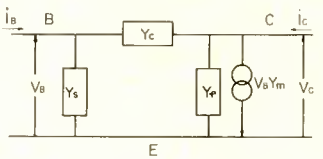


Fig. 6—π network for common-emitter connection, where $Y_s = Y_{11} + Y_{21} - Y_{12}$. The other elements are the same as in Figure 5.

$$Y_s = Y_{11} + Y_{21} = i_e \frac{q}{kT}\frac{1}{\gamma_o\left(1 + \frac{\epsilon'}{\epsilon}\right)}\left[\left(1 + \frac{\epsilon'}{\epsilon}\Omega\right) - 2\gamma_o\Omega e^{-\epsilon'\left(\Omega - \frac{\epsilon}{\epsilon'}\right)}\right]. \quad (20)$$

The effect of the exponential term in Equation (19) can be neglected apart from the slight modification it will make at low frequencies. At high frequencies the expression for Y_c will be of the same form as the expression for Y_{22}. In Figures 5 and 6, "V_c" refers to the output voltage. Distinguishing subscripts have been omitted.

The low-frequency value of Y_c will depend quite critically on the ratio of ϵ/ϵ'. For an assumed value of $\epsilon/\epsilon' = .999688$ and $\epsilon' = 4$, the function is plotted in Figure 9. The value of ϵ/ϵ' will also affect the value of $Y_{11} + Y_{21}$, but this effect is by no means as strong and can largely be accounted for in the low frequency value of γ_o. The function is plotted in Figure 10 for $\epsilon' = \epsilon$ and $\gamma_o = 0.975$.

Equivalent Circuit Representation

In this section equivalent circuit representations of the admittance functions Y_e, Y_c, Y_s, Y_ρ, and Y_m, described in the previous section, are derived. The method of derivation is essentially the same as that used by Kroemer.[3] The admittance functions are expanded to the third term in a power series in a normalized frequency parameter x, where $x = \omega\,\delta_o^2/D_p$, and $x < 1$. The frequency corresponding to $x = 1$ is of the order of the α-cutoff frequency of the intrinsic transistor. Three-element networks are used to represent the functions. The values of the elements are derived by writing the admittance of a network in powers of normalized frequency and equating coefficients to the coefficients in the expansion of the corresponding admittance function.

In this derivation the transition capacitances have been omitted as well as the external base-lead resistance. These elements will be added to the equivalent circuit of the intrinsic transistor in a later section.

Admittance Y_e

Substituting $\Omega = \sqrt{1 + jx}$ into Equation (14),

$$Y_e = \frac{i_e^{(p)}}{\gamma_o}\,\frac{q}{kT}\,\frac{1+\sqrt{1+jx}}{2}. \tag{21}$$

For $x < 1$ we may write $\sqrt{1+jx} = 1 + j\frac{x}{2} + \frac{x^2}{8}$ - - - -

Keeping only the first three terms in the expansion and substituting in Equation (21),

$$Y_e \approx \frac{i_e^{(p)}}{\gamma_o}\,\frac{q}{kT}\left(1 + j\frac{x}{4} + \frac{x^2}{16}\right) \text{ for } x^2 \ll 1. \tag{22}$$

The admittance of the circuit shown in Figure 7 may be written

$$Y = g_1 + \frac{j\omega C}{1+\dfrac{j\omega C}{g_2}} \approx g_1\left[1 + \frac{j\omega C}{g_1} + \frac{\omega^2 C^2}{g_2 g_1}\right] \text{ for } \frac{\omega C}{g_2} < 1. \tag{23}$$

If we write $x = \omega a$, where $a = \delta_o^2/D_p$, we may equate coefficients of ω in Equations (22) and (23) and solve for C, g_1, and g_2. Then

$$g_1 = \frac{i_e^{(p)}}{\gamma_0}\,\frac{q}{kT}, \quad C = \frac{a}{4}\,g_1, \text{ and } g_1 = g_2. \tag{24}$$

The variation of the admittance of the circuit shown in Figure 7 is plotted in Figure 2 (dotted) for the ratios of the elements given by Equations (24). It can be seen from this curve where the approximations break down above $x = 1$.

Admittance Y_ρ

This term may be dealt with in the same manner as that used for Y_{11}; the most suitable three-element circuit is the one shown in Figure 8. This circuit equivalent has however two disadvantages. The first is that the capacitance C turns out to be negative. The second is that at high frequencies the circuit admittance begins to increase with frequency whereas the admittance function Y_ρ keeps on decreasing with frequency. A more useful approximation would be simply the g and L elements in series, as in most circuit applications it is not necessary to know the exact variation of Y_ρ with frequency, but only to know that its admittance decreases with increasing frequency.*

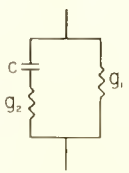


Fig. 7—Circuit representation of Y_t.

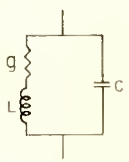


Fig. 8—Circuit representation of Y_ρ.

Admittance Y_m

In the equivalent circuits of Figures 5 and 6, Y_m appears as the multiplier of the input voltage to make up the current generator in the output leg. Although an equivalent circuit may be formed to represent its frequency variation, it would need to be somewhat complex and its usefulness would be limited. It was thought preferable therefore to obtain a simplified expression for the frequency variations of this function that could be used simply as a multiplier for the current generator.

Using Equation (16)

$$Y_m = i_e^{(p)} \frac{q}{kT} \Omega e^{(\epsilon - \epsilon'\Omega)}. \tag{25}$$

* A slightly better approximation may be obtained for Y_ρ at the lower frequencies by placing a small capacitance across the conductance in Figure 8.

Letting g_m equal the low frequency value of Y_m, Equation (25) may be written in the form

$$Y_m = g_m \Omega e^{\epsilon'(1-\Omega)}. \tag{26}$$

If Ω is expanded in powers of x and only terms up to x^2 retained, Y_m may be written

$$Y_m \approx g_m \frac{e^{-j\omega A}}{1 + j\omega B}, \tag{27}$$

where

$$A = \frac{a}{2}\left[\epsilon' - 1 - (\epsilon' - 2)^{1/2}\right],$$

$$B = \frac{a}{2}(\epsilon' - 2)^{1/2},$$

for $\epsilon' = 4$, $A \approx 0.8a$, and $B \approx 0.7a$. A good approximation to the original function can be obtained if A and B are made equal. Equation (27) is plotted on Figure 3 (dotted) as a function of the frequency parameter $x = \omega a$, for $A = B = 0.74a$. Comparing these curves with the original curves it can be seen that the modulus curve is a good approximation up to $1.5x$ while the phase curve gives a good approximation somewhat higher.

Admittance Y_c

The next element to consider is Y_c. Using Equation (19) and neglecting the last term,

$$Y_c = Y_{22} + Y_{12} \approx -\frac{i_c}{w}\frac{dw}{dV_c}\epsilon'\left[1 - \frac{\epsilon}{\epsilon'} + j\frac{x}{2} + \frac{x^2}{8}\right]$$

$$= -\frac{i_c}{w}\frac{dw}{dV_c}(\epsilon' - \epsilon)\left[1 + j\frac{x}{2}\frac{\epsilon'}{\epsilon' - \epsilon} + \frac{x^2}{8}\frac{\epsilon'}{\epsilon' - \epsilon}\right] \tag{28}$$

An equivalent circuit may be derived for this admittance function similar to the one used for Y_e and is illustrated in Figure 7. The two conductances and the capacitance will then be given by

$$g_1 = -\frac{i_c}{w}\frac{dw}{dV_c}(\epsilon' - \epsilon),$$

$$C = g_1 \frac{a}{2}\frac{\epsilon'}{\epsilon' - \epsilon},$$

$$g_2 = 2\left(\frac{\epsilon'}{\epsilon' - \epsilon}\right) g_1. \tag{29}$$

The admittance of this equivalent circuit is plotted in Figure 9 (dotted) and compared to the admittance function. It can be seen from this curve that the approximation is very good up to $x = 1$. At frequencies much above this point, however, the approximations break down and there is a large difference in the two curves, particularly in the phase curves.

Admittance Y_s

The last admittance function that is considered is $Y_s = Y_{11} + Y_{21}$.

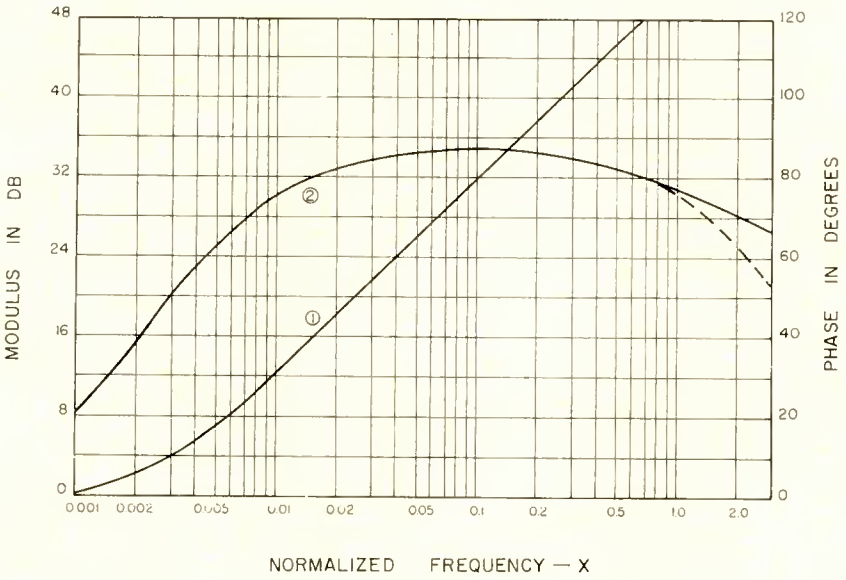


Fig. 9—The variation of $\dfrac{Y_{22} + Y_{12}}{(Y_{22} + Y_{12})^{(0)}} = \dfrac{Y_e}{Y_e^{(0)}}$ with frequency. Curve 1 represents the modulus and curve 2 the phase. The dotted curve shows the equivalent circuit approximation.

For this case we assume $\epsilon/\epsilon' = 1$ and use Equation (20). Again expanding in powers of x and keeping only the first three terms,

$$Y_s = Y_{11} + Y_{21} \approx i_e \frac{q}{kT} \frac{(1-\gamma_o)}{\gamma_o} \left[1 + j \frac{x}{4} \left(\frac{1 + 2\gamma_o (\epsilon - 1)}{1 - \gamma_o} \right) + \frac{x^2}{16} \frac{(1 + 2\gamma_o (\epsilon^2 - \epsilon - 1))}{1 - \gamma_o} \right]. \tag{30}$$

If we use the same type of circuit to approximate this function as the one shown in Figure 7, we get by equating coefficients,

$$g_1 = i_c^{(p)} \frac{q}{kT} \frac{(1-\gamma_o)}{\gamma_o},$$

$$C = \frac{a}{4} \left(\frac{1 + 2\gamma_o (\epsilon - 1)}{1 - \gamma_o} \right) g_1,$$

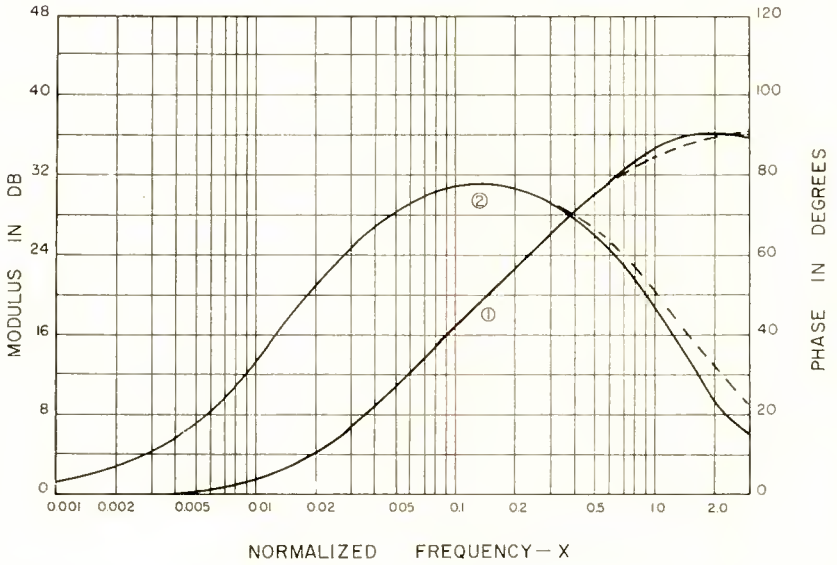


Fig. 10—The variation of $\frac{Y_{11} + Y_{21}}{(Y_{11} + Y_{21})^{(0)}} = \frac{Y_s}{Y_s^{(0)}}$ with frequency. Curve 1 represents the modulus and curve 2 the phase. The dotted curves show the equivalent circuit approximation.

$$g_2 = \frac{[1 + 2\gamma_o (\epsilon - 1)]^2 g_1}{[1 + 2\gamma_o (\epsilon^2 - \epsilon - 1)] [1 - \gamma_o]}. \tag{31}$$

The admittance of this network is plotted in Figure 10 (dotted) and may be compared to the function it is to represent for $\gamma_o = 0.975$ and $\epsilon = 4$. It can be seen that the equivalent circuit representation only differs from the admittance function appreciably for $x > 1$.

The admittance functions that have been represented in this section by two-terminal networks can be combined to give the complete equiva-

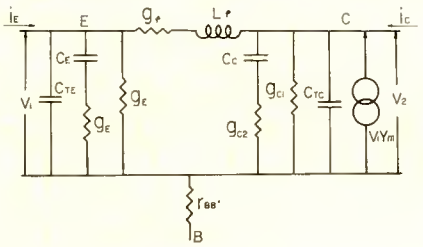


Fig. 11—Equivalent circuit for the common-base connection.

lent circuits of the transistor in the form of the circuits in Figures 5 and 6. This is done in Figures 11 and 12 for the common-base and common-emitter circuits respectively. In these circuits the collector and emitter transition capacitances as well as the external base resistance have been included.

In Table II the expressions, in terms of the device parameters, are summarized together with the calculated values using the device constants shown in Table I.

Current-Amplification Factor

An important transistor parameter which so far has not been considered explicitly is the current-amplification factor, α, for the grounded-base connection. One of the interesting features of the drift transistor is the manner in which the phase of α increases with increasing frequency. As was pointed out earlier, the controlling factor for the intrinsic transistor is the transport factor, β, the injection efficiency, γ, of the intrinsic transistor being relatively insensitive to frequency ($\alpha = \beta\gamma$). Therefore, we shall consider only β here. Curve No. 1 of Figure 13 is a plot of the modulus and phase of β for the drift transistor. Curve 2 represents β for the diffusion transistor and for comparison curve 3 represents the modulus and phase of the

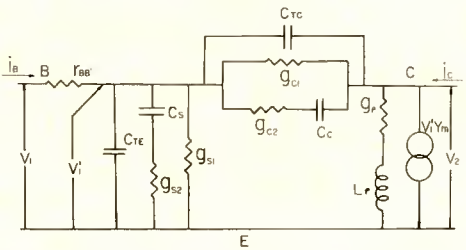


Fig. 12—Equivalent circuit for the common-emitter connection.

impedance of a simple parallel R-C network. Thus, at the 3-decibel point, the phase of β for a typical drift transistor is 105°, that for the diffusion transistor is 57°, and that of the R-C network is 45°.

This does not mean, of course, that at any given frequency the phase of β is larger for the drift transistor than for a diffusion transistor having the same device parameters. On the contrary, the drift field considerably reduces the transit time, and therefore the phase of β. The drift field has an even stronger effect on the modulus, however, extending its cutoff frequency sufficiently so that the phase

Table I—Estimated Typical Values for the Device Parameters of a Drift Transistor

$w = 10^{-3}$ cm,

$\frac{dw}{dV_c} = 1.2 \times 10^{-5}$ cm/volt,

$w/L = 0.1$,

$D_p = 44$ cm^2/sec,

$i_e = 1$ ma,

$q/kT = 40$ volt^{-1},

$\epsilon' = 4$,

$\gamma_o = 0.975$,

$\alpha = \frac{w^2}{D_p}\frac{1}{\epsilon^2} = 1.4 \times 10^{-9}$ sec (implying an α-cutoff of the order of 100 megacycles),

$\epsilon/\epsilon' = 0.999688$.

of β at that frequency at which the modulus is reduced by 3 decibels has grown comparatively large.

It is of interest to investigate the roles that diffusion and the drift field play in the mechanisms of carrier transit and of β fall-off in order to obtain a better understanding of just why the drift field improves the β fall-off situation more than it does the phase of β. It is found, contrary to the case of the diffusion transistor, where both processes depend on diffusion in an inseparable fashion, that in the drift transistor carrier transit depends primarily on the drift field, while β fall-off depends only on diffusion. If one could eliminate diffusion while keeping the drift field, the signal would be propagated across the base with a velocity proportional to the drift field and with no decrease in amplitude. Then the phase of β at the cutoff frequency would be infinite, a meaningless statement since the cutoff frequency

Table II—Equivalent Circuit Elements

Equivalent Circuit Element	Corresponding Term in Approximate Device Equations	Calculated Values Using Parameter Values Given in Table I
C_e	$\frac{a}{4} g_e$	14 $\mu\mu$f
g_e	$\frac{i_e q}{\gamma_o kT}$	41,000 μmho
g_p	$-2 \frac{i_e \epsilon e^{-2\epsilon}}{w} \frac{dw}{dV_e}$	0.032 μmho
L_p	$\frac{a}{2} \frac{(\epsilon - 1)}{g_p}$	0.066 henry
g_m	$i_e \frac{q}{kT} e^{(\epsilon' - \epsilon)}$	40,000 μmho
g_{c1}	$-\frac{i_e}{w} \frac{dw}{dV_e} (\epsilon' - \epsilon)$	0.015 μmho
g_{c2}	$2\left(\frac{\epsilon'}{\epsilon' - \epsilon}\right) g_{c1}$	96 μmho
C_c	$g_{c1} \frac{a}{4} \frac{\epsilon'}{\epsilon' - \epsilon}$	0.017 $\mu\mu$f
g_{s1}	$i_e \frac{q}{kT} \frac{(1 - \gamma_o)}{\gamma_o}$	1,025 μmho
g_{s2}	$\frac{[1 + 2\gamma_o(\epsilon - 1)]^2 g_{s1}}{[1 + 2\gamma_o(\epsilon^2 - \epsilon - 1)][1 - \gamma_o]}$	86,000 μmho
C_s	$\frac{a}{4}\left[\frac{1 + 2\gamma_o(\epsilon - 1)}{1 - \gamma_o}\right] g_{s1}$	100 $\mu\mu$f

It should be emphasized that the calculated values given above are only representative, but should correspond with the actual values of a drift transistor having the device parameter values given in Table I.

itself would be infinite (neglecting other limiting effects). With diffusion, fall-off in amplitude is roughly proportional to the product of the square of the average concentration gradient in the base and the

length of time the signal remains in the base (transit time). Thus, if the drift field is doubled, the transit time is approximately halved, the cutoff frequency is increased by approximately $2^{3/2}$, and the phase of β at the cutoff frequency is increased by approximately $2^{1/2}$.

In order to properly understand the physical processes which relate the phase and modulus of β, it is helpful to examine (1) the concentration of minority carriers in the base and (2) the current at any point in the base as functions of frequency. For a p-n-p transistor, if $p(s,t)$ is the concentration of excess minority carriers (holes) in the base, then, from Equation (19) of Kroemer[2]

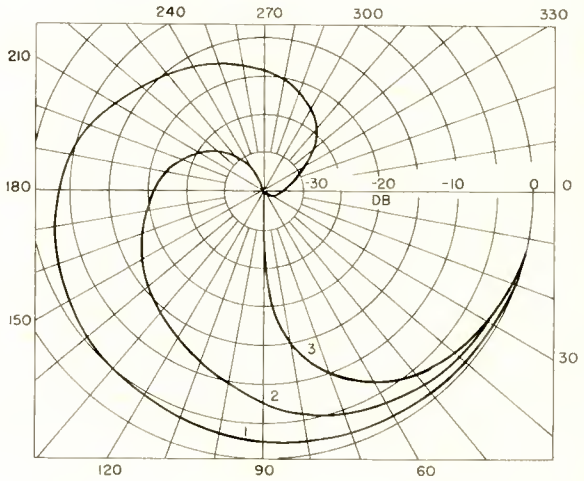


Fig. 13—Polar diagram showing the phase change of the current transfer ratio with its reduction in modulus. Curve 1 is for the drift transistor. Curve 2 is for the diffusion transistor. Curve 3 is for a simple R-C representation.

$$p(s,t) = \frac{q\,U_e}{kT}\, e^{j\omega t}\, p_0 e^{\frac{q\,V_e}{kT}}\, e^{-\frac{s}{f}}\, \frac{\sinh \frac{w-s}{\delta}}{\sinh \frac{w}{\delta}}, \qquad (34)$$

where $U_e\, e^{j\omega t}$ is the applied signal at the emitter and p_o is the equilibrium hole concentration at the emitter. s represents distance, measured from the emitter end of the base.

Thus
$$p(s,t) = p(s,\omega)\, e^{j\omega t} = p(o)\, e^{\frac{s}{f}} \frac{\sinh \dfrac{w-s}{\delta}}{\sinh \dfrac{w}{\delta}}\, e^{j\omega t}. \tag{35}$$

The a-c current $i(s,t)$ is given by

$$i(s,t) = A\mu_p kT \left(\frac{2}{f} p(s,t) - \frac{dp(s,t)}{ds} \right), \tag{36}$$

so that

$$\beta(s,\omega) = \frac{i(s,t)}{i(o,t)} = \frac{e^{\frac{s}{f}} \left[\sinh \dfrac{w-s}{\delta} + \dfrac{f}{\delta} \cosh \dfrac{w-s}{\delta} \right]}{\sinh \dfrac{w}{\delta} + \dfrac{f}{\delta} \cosh \dfrac{w}{\delta}}. \tag{37}$$

Here we have extended the normal definition of β to have meaning at any point in the base.

The Diffusion Transistor

For the diffusion transistor, Equation (35) reduces to

$$p(s,\omega) = p(o) \frac{\sinh \dfrac{w-s}{\delta}}{\sinh \dfrac{w}{\delta}} \approx p(o) \frac{\sinh \sigma z(1+j)}{\sinh \sigma(1+j)}$$

where
$$\sigma^2 = \frac{\omega w^2}{2D_p} = \frac{\omega}{\omega_o}, \quad z = \frac{w-s}{w}.$$

The approximation is the neglect of recombination in the base. Written in terms of modulus and phase, Equation (35) becomes

$$p(z,\omega) = p(o) \left[\frac{\sinh^2 \sigma z + \sin^2 \sigma z}{\sinh^2 \sigma + \sin^2 \sigma} \right]^{1/2} \exp -j \left[\tan^{-1} \frac{\tan \sigma}{\tanh \sigma} - \tan^{-1} \frac{\tan \sigma z}{\tanh \sigma z} \right] \tag{35a}$$

Similarly the hole current ratio (transport factor) becomes

$$\beta(z,\omega) = \frac{\cosh z\sigma(1+j)}{\cosh \sigma(1+j)} = \left[\frac{\cosh^2 \sigma z - \sin^2 \sigma z}{\cosh^2 \sigma - \sin^2 \sigma}\right]^{1/2} \exp - j$$
$$[\tan^{-1} \tan\sigma \tanh\sigma - \tan^{-1} \tan\sigma z \tanh\sigma z]. \qquad (37a)$$

At low frequencies ($\sigma \to o$),

$$p(z,\omega) = p(o)\, z\left[1 - \frac{\sigma^4}{45}(1-z^4)\right] \exp - j\frac{\sigma^2}{3}(1-z^2) \qquad (35a,1)$$

$$\beta(z,\omega) = \left[1 - \frac{\sigma^4}{3}(1-z^4)\right] \exp - j\sigma^2(1-z^2). \qquad (37a,1)$$

The Drift Transistor

For the drift transistor, Equations (35) and (36) may be written

$$p(z,\omega) = p(o)\, e^{\epsilon(1-z)} \frac{\sinh z\epsilon\Omega}{\sinh \epsilon\Omega} \qquad (35b)$$

$$= p(o)\, e^{\epsilon(1-z)} \left[\frac{\sinh^2 z\epsilon c + \sin^2 z\epsilon d}{\sinh^2 \epsilon c + \sin^2 \epsilon d}\right]^{1/2} \exp - j$$
$$\left[\tan^{-1} \frac{\tan \epsilon d}{\tanh \epsilon d} - \tan^{-1} \frac{\tan z\epsilon d}{\tanh z\epsilon d}\right]$$

where $$\Omega = \left[1 + \frac{j\omega\delta_0^2}{D_p}\right]^{1/2} = c + jd,$$

and $$\beta(z,\omega) = e^{\epsilon(1-z)} \left[\frac{\sinh \epsilon z\Omega + \cosh \epsilon z\Omega}{\sinh \epsilon\Omega + \cosh \epsilon\Omega}\right] \qquad (37b)$$

Discussion

The moduli of the hole concentrations $p(z,\omega)$, as given in Equations (35a) and (36a) are plotted in Figures 14 and 15 for the diffusion transistor and for the drift transistor respectively. The solid lines AO and BO represent the envelope of the concentration at very low frequencies. At higher frequencies the modulus of the concentration is less than the low-frequency limit at every point in the base excepting the emitter. The broken lines ACO and BDO represent the envelope of the concentration for that frequency for which the phase of $\beta(w,\omega)$

is —180°. For the diffusion transistor this occurs (see Equation (37a)) at $\sigma = \pi$, i.e., at approximately 10 times the cutoff frequency. For the drift transistor, the phase of β is —180° at approximately twice the cutoff frequency.

The solid line EO depicts the actual concentration at one instant in time at the frequency described above. The modulus of β may be obtained from this curve by inspection. The instant in time chosen is that at which the current at the emitter and at the collector is maximum in magnitude (and separated in phase by 180°). Remem-

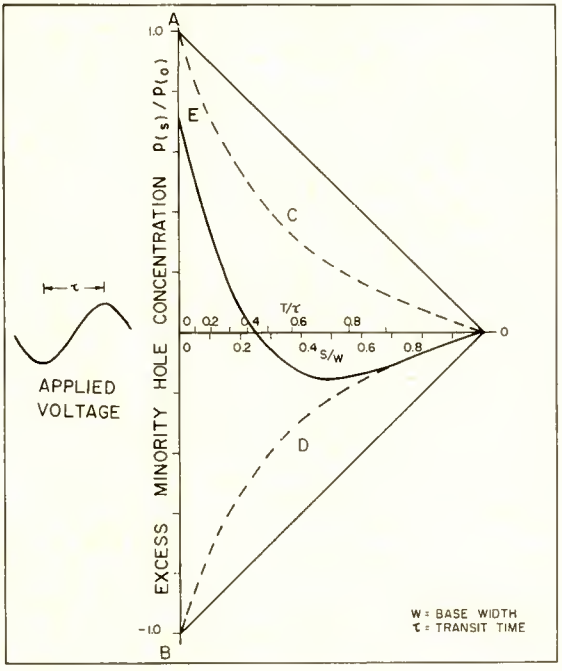


Fig. 14—Excess minority hole concentration and time delay in the base of a diffusion transistor.

bering that the current in a diffusion transistor is proportional to the concentration gradient, the modulus of β is merely the negative of the ratio of the concentration gradient at the collector to that at the emitter. For the drift transistor, from Equation (36)

$$\beta(w,\omega) = \frac{-p'(w,\omega)}{\frac{2}{f}p(o,\omega) - p'(o,\omega)}.$$

It is interesting to compare the emitter and collector currents of the two transistors at the frequencies giving a change in phase of 180° for β. For the diffusion transistor, β has fallen off 21.3 decibels, the emitter current increasing by 13 decibels and the collector current decreasing by 8.3 decibels. In the case of the drift transistor, β has fallen off only 8.5 decibels, the emitter current increasing 1.7 decibels and the collector current decreasing 6.8 decibels. It thus appears that one effect of the drift field is to keep the emitter current almost constant as a function of frequency.

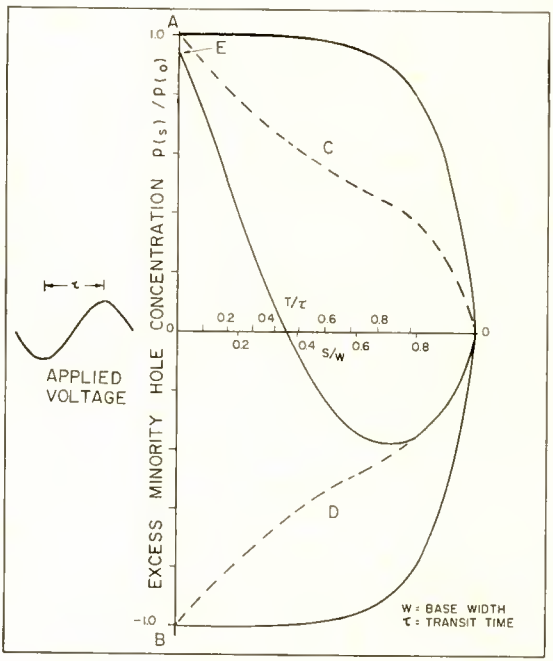


Fig. 15—Excess minority hole concentration and time delay in the base of a drift transistor.

The reason for the large phase change of β of the drift transistor at the 3 decibel point, compared to that of the diffusion transistor, now becomes clear. In the diffusion transistor, diffusion is the only mechanism working. The same concentration gradients on which hole transport depends are also responsible for the fall-off of β, since diffusion tries to make uniform a nonuniform hole concentration. In the drift transistor, on the other hand, diffusion is only a hindrance. It does not come into play until the frequency is high enough so that

there are appreciable concentration gradients in the base. If the diffusion constant were small, one could picture a signal in the base as an exponentially damped sine wave, the "damping" being due to the diffusion of the concentration "hills" into the concentration "valleys." This mechanism decreases the collector current and thus the modulus of β.

Conclusion

The four-terminal admittance equations derived by Kroemer[2] have been used to form an equivalent circuit for the drift transistor. The equivalent circuit is in two forms, one suitable for the common-base connection and the second suitable for the common-emitter connection. The equivalent circuit is rather complex for circuit analysis, but it can probably be simplified for any particular type of drift transistor.

ACTIVATION OF AN OXIDE CATHODE BY DEPOSITION OF ALKALINE EARTH METAL IONS VIA A MASS SPECTROMETER

By

R. M. Matheson,* L. S. Nergaard,* and R. H. Plumlee†

RCA Laboratories,
Princeton, N. J.

Summary—*A graphical display is given of thermionic activation effects produced by deposition of mass spectrometrically resolved ion beams of the alkaline earth metals on an oxide cathode. The maximum ion beam current density deposited was 7×10^{-8} ampere per square centimeter, corresponding to a deposition rate of the order of 0.02 monolayer per minute. Emission activations in excess of two orders of magnitude (referred to the thoroughly deactivated state) were produced. Maximum current densities of approximately 30 milliamperes per square centimeter (corrected to 1,000°K) were measured. These experiments, devised to facilitate the simplest and purest study feasible of the kinetic processes of cathode activation, show that cathode electrical behavior is dependent in a very complex way on stoichiometry and emission duty. As is to be expected from recent interpretations of cathode chemistry, all the alkaline earth metals produced essentially identical emission activation effects. Emission activation–decay cycles produced by metal depositions vary with cathode temperature, environment, and previous history. Plots of emission versus time during and following depositions at cathode temperatures in the range 765-910°K show some resolution of the rate processes in time and temperature. To describe the gross features of the build-up or recovery branches of these emission cycles analytically, as many as three terms exponential in q (the time-integrated electron charge passed) are required, and one term exponential in time. Series–parallel and opposing rate processes are indicated. A tentative and incomplete physical model of the emission–deposition behavior is given; however, a satisfactory analysis in terms of a physical-chemical model has not yet been derived.*

Introduction

WHEN these deposition experiments were undertaken in 1951, there was considerable question as to the nature of the polarization mechanism responsible for the pulse-emission decay of the oxide cathode. It was expected that, by depositing the alkaline earth atoms[1] of radii ranging from 1.60 angstroms for Mg to

* RCA Laboratories, Princeton, N. J.

† Formerly RCA Laboratories, Princeton, N. J.; now with Sandia Corp., Albuquerque, N. M.

[1] It seems likely that incident ions arriving with velocities not exceeding the thermal velocities of conduction electrons or the orbital velocities of electrons attached to donors become neutralized to the atomic state before entering the solid. As evidence concerning this and related behavior, see, for instance, the series of studies of Auger processes by H. D. Hagstrum, one of the latest being, "Auger Ejection of Electrons from Molybdenum by Noble Gas Ions," *Phys. Rev.*, Vol. 104, p. 672, November, 1956.

2.17 angstroms for Ba, the substantial variation in the rates of diffusion of these atoms or their ions into the coating should produce greatly different rates of electron-emission activation if the transport of cationic constituents rather than the transport of anionic constituents were responsible for pulse-emission decay phenomena. Alternatively, if the electron donors were indeed F-centers as they were then usually believed to be, transport of donors would most probably occur via an oxygen vacancy mechanism and activation would be rather independent of the precise identity of the "activating agent" or "reducing agent" employed.

Earlier experiments performed by Matheson and Nergaard[2] had shown that slight warming of getter patches and envelopes of receiving tubes almost invariably increased cathode emission. From the temperatures of the envelopes and from vapor pressure data for Ba, it was concluded that 10^{-12} mm Hg vapor pressure of Ba produced enough change in cathode emission to be easily detectable. This vapor pressure was so low that the question immediately arose as to whether Ba vapor really was responsible for the activation observed or whether other species derived from the tube envelopes were causing the activation. The rate of incidence of Ba atoms from the vapor density estimated to be effective in producing detectable activation of oxide cathodes was also well within the range attainable by deposition of the resolved ion currents of laboratory mass spectrometers.

The desirability of using the mass spectrometer to produce beams of doping elements for studies of cathode activation rests on the following features. The mass spectrometer can provide a rather pure beam of known chemical composition. In addition, it accurately measures the rate of deposition of the species; it provides a means for initiating and interrupting the deposition rapidly and precisely; it acts as a "cold" source of doping agents because the test cathode is far removed from hot filaments and can be examined in an environment practically unchanged by the process of turning the beam on and off target.

Probably the only a priori basic objection to the mass spectrometric deposition method of studying cathode activation is that surface-sputtering reactions or bombardment damage reactions might in some way nullify the intrinsic activating properties of the materials deposited. Literature on sputtering[3-6] suggests that sputtering yields

---

[2] L. S. Nergaard, "Studies of the Oxide Cathode," *RCA Review*, Vol. XIII, p. 464, December, 1952.

[3] Gottfried K. Wehner, "Controlled Sputtering of Metals by Low-energy Hg Ions," *Phys. Rev.*, Vol. 102, p. 690, May, 1956.

from metals with bombarding particle energies of the order of 1,000 electron volts do exceed unity[7] but that the yields from ionic compounds[5] and covalent substances[6] are much less than those from metals.

Work by Koch[8] and Young[9] and also theoretical analyses[10,11] indicate that the penetration of heavy particles into solids is somewhat less than ten lattice spacings per kilovolt of incident particle energy at energies under 10^5 electron volts.

Whatever deleterious reactions might actually proceed concurrently, any positive activating effect produced by deposition of particles with greater than thermal energies would indicate an over-all net gain and would constitute a worthwhile valid experiment; however, if the high-energy deposition produced no effect or produced a de-activating effect, the experiment could reasonably be classified as inconclusive rather than indicative that alkaline earth metals were not activating agents.

It was recognized that, chemically, the target cathode cannot distinguish the arrival of positive ions from the arrival of the parent atoms or molecules provided the work function of the cathode is less than the first ionization energy of the parent (neutral) particles and provided the electronic conductivity of the cathode is sufficient to allow charge neutralization of the surface. Chemically, the deposition on the cathode of any species less electronegative than oxygen constitutes a partial reduction of the oxide and hence should raise the Fermi

---

[4] Frank Keywell, "Measurements and Collision-Radiation Damage Theory of High-Vacuum Sputtering," *Phys. Rev.*, Vol. 97, p. 1611, March, 1955.

[5] Richard C. Bradley, "Sputtering of Alkali Atoms by Inert Gas Ions of Low Energy," *Phys. Rev.*, Vol. 93, p. 719, February, 1954.

[6] M. L. Smith, "Collection Problems in the Large Electromagnetic Separator," Chapter 7 of *Electromagnetically Enriched Isotopes and Mass Spectrometry*, p. 53, Academic Press, N. Y., 1956.

[7] These yield values may be erroneously high because the ions responsible for most of the yield may have had energies higher than those estimated previously. As shown in Reference (17), the numbers of multiply charged ions may be comparable with the number of singly charged ions in circumstances which allow ion trapping. Neither the results of Wehner nor those of Keywell are exempt from this source of error. Obviously multiply charged ions have impact energies which are multiples of the values for the singly charged ions assumed to be overwhelmingly dominant in most sputtering experiments.

[8] J. Koch, "Mass Spectrographic Separation of Isotopes of Gaseous Elements," *Nature*, Vol. 161, p. 566, April, 1948.

[9] J. R. Young, "Deterioration of Luminescent Phosphors under Positive Ion Bombardment," *Jour. Appl. Phys.*, Vol. 26, p. 1302, November, 1955.

[10] Niels Bohr, "The Penetration of Atomic Particles through Matter," *Kgl. Danske Vid. Selsk. Matfys. Medd.*, Vol. 18, No. 8, 1948.

[11] K. O. Nielsen, "The Range of Atomic Particles with Energies about 50 Kev," Chapter 9, *Electromagnetically Enriched Isotopes and Mass Spectrometry*, p. 68, Academic Press, N. Y., 1956.

level above that of the intrinsic oxide material.[12] Whether the species arrives with substantial velocity as an ion or as a neutral particle or arrives with very low (thermal) velocity is unimportant, in a chemical sense, under the conditions prescribed above. As long as the velocity of an incident ion does not exceed that of the thermally emitted or field-extracted electrons, it will be neutralized before it arrives within a few lattice spacings of the surface.[10,11] The incidence and attachment of any *electropositive* species having thus increased the reducing chemical potential of the oxide coating, the cathode is again made reactive toward any other incident residual *electronegative* gas. Such ensuing reactions obviously cause the oxide crystals to grow and also to become less active thermionically. The kinetics of cathode activation by deposition of ions via the mass spectrometer must be expected to proceed, therefore, as part of the kinetics of reactions of the coating with the remainder of its environment.

It is the purpose of the present paper to present a rather complete listing of the deposition data in graphical form and a tentative and incomplete interpretation. Abbreviated descriptions of these results have been presented previously.[13-15] Empirically, it was found possible to represent the electron current during and following depositions by formulas resembling the Sproull formula[16] for pulse-emission decay. It has also been attempted to derive these empirical formulas from a model which will be outlined later in the discussion of "typical" deposition behavior; however, the model in its present form has proved to be mathematically intractable and this more refined analysis has not yet been consummated and hence will not be discussed. Although in 1951 the thermochemical analysis of cathode phenomena and the $OH^{-}\cdot e$ donor proposal[12] had not then been conceived, they have not yet been found to be incompatible with the results of these deposition experiments.

---

[12] R. H. Plumlee, "The Electron Donor Centers in the Oxide Cathode," *RCA Review*, Vol. 17, p. 231, June, 1956.

[13] L. S. Nergaard, "Studies of the Oxide Cathode," *Report on the Thirteenth Annual Conference on Physical Electronics*, p. 21, Mass. Inst. of Tech., Cambridge, Mass., 1953.

[14] R. M. Matheson and R. H. Plumlee, "Activation of an Oxide Cathode by Ba, Sr, Ca, and Mg Deposited via a Mass Spectrometer," *Report on the Fourteenth Annual Conference on Physical Electronics*, Mass. Inst. of Tech., Cambridge, Mass., 1954.

[15] L. S. Nergaard, "Experiments Bearing on the Mobile-Donor Hypothesis in Oxide Cathodes," *Le Vide*, No. 52-53, p. 171, September, 1954.

[16] R. L. Sproull, "An Investigation of Short-Time Thermionic Emission from Oxide-Coated Cathodes," *Phys. Rev.*, Vol. 67, p. 166, March, 1945.

Equipment

The Vacuum System and Mass Spectrometer

Figure 1 gives the schematic arrangement of the vacuum system and the analyzer tube of the mass spectrometer. Two pumping lines were provided, one for the ion source region of the analyzer tube and one for the remainder of the analyzer tube. This differential pumping arrangement, with the restricted conductance path between the two sections of the analyzer, produced about a ten-fold lower partial pressure in the main part of the analyzer than in the ion source region for gases originating in the ion source region.

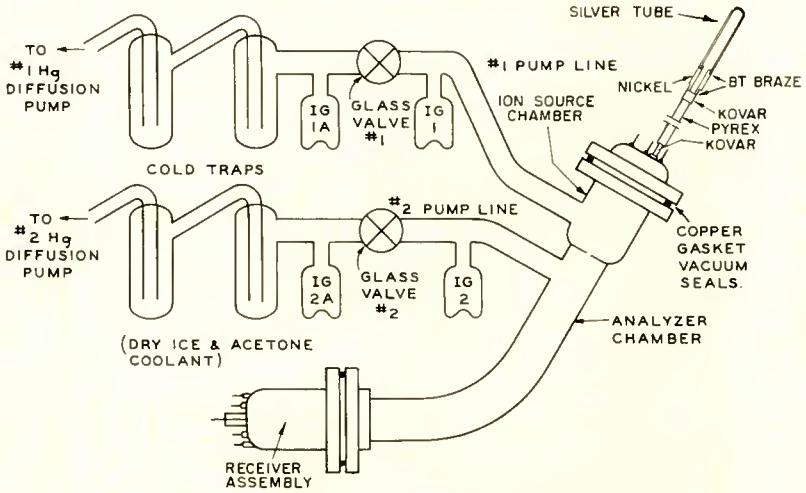


Fig. 1—Vacuum system schematic.

The mass spectrometer, constructed here, used a 60° sector magnetic field and an ion beam trajectory of approximately 5 inches radius of curvature. The ion source is shown in Figure 2. Details of operation of the ionizing chamber are described elsewhere.[17]

Metal Vapor Sources

Thirteen metal "evaporators" were contained in the upper deck of the ion source. Commercially available batalum getters were used to supply the Ba vapor. These consist of a reaction mixture of $BaBeO_3$ plus Ti powder contained in a molybdenum channel. The reaction mixtures for generating the other alkaline earth metal vapors were alkaline earth titanate powders plus aluminum powder mixed with a

[17] R. H. Plumlee, "Space-Charge Neutralization in the Ionizing Beam of a Mass Spectrometer," *Rev. Sci. Instr.*, in press.

nitrocellulose-solvent binder and packed into molybdenum channels similar to those used for the batalum getters. Each of these canoe-shaped channels, holding about 15 milligrams of material, was essentially three-fourths of a cylinder with crimped ends and was made from .002 × .100 × 1 inch molybdenum ribbon.

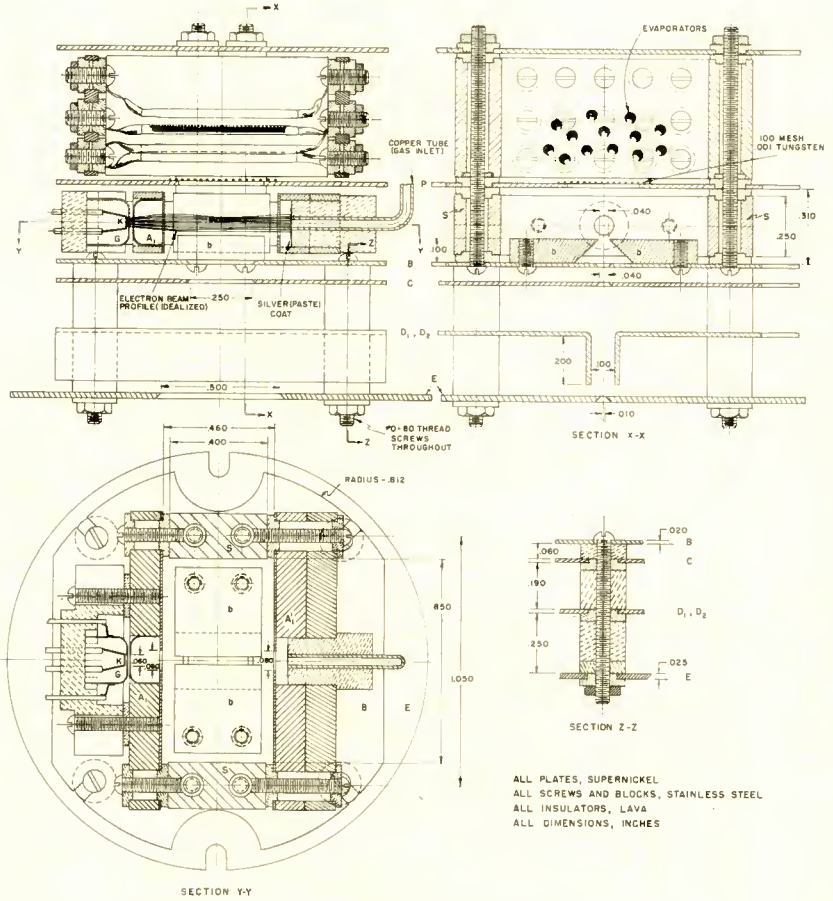


Fig. 2—Ion source assembly.

Incipient liberation of alkaline earth metal occurred from most of these reaction boats at temperatures around 1,000°C with a power input of approximately 30 watts (10 amperes heating current). Partial pressures of the order of 10^{-5} mm Hg of the principal alkaline earth metal were maintained for many hours from each of these evaporators by gradually increasing the boat temperature. Each boat was usually

operated at increasingly higher temperatures until it burned through. Thus, burnout commonly occurred at a power input around 300 watts, hence any alkaline earth metal previously generated and deposited on ion source surfaces gradually distilled out of the source region and condensed on the walls of the analyzer envelope. Calcium was produced with rather poor yield from this reaction; strontium with better yield; and barium and magnesium were produced quite readily. The vapor pressure of Al from the reaction mixture was never found to be more than about one-thousandth of that of the principal alkaline earth metal being generated.

Receiver Assembly and Target Cathode

The receiver assembly of the spectrometer is shown in Figure 3. The slit in plate A was about three times the expected width of the resolved ion beams. A half-plate, C, covered half the slit of plate A and was used as the ion-collecting electrode for the electrometer amplifier. The resolved ion currents could thus be brought to fall either on the ion-collecting electrode, C, or on the target cathode, K_1, by making only a small change in the ion-accelerating voltage or in the magnetic deflection field. Half-plates C′, F, and F′ were used as electrostatic deflectors to produce optimum efficiency for collecting ions on the cathode. Plates B, D, and E were biased negatively to suppress the ejection of secondary electrons from collector C. Plates G, H, and I formed the anode for the target cathode, K_1. The slit of plate I was covered with tungsten mesh, $.010 \times .010 \times .001$ inch. The probe wire (.001-inch platinum) embedded in the target coating was supported on lava insulators by plate S which also supported the cathode sleeve. Plate S was electrically connected directly to the cathode sleeve. The cathode sleeve and heater were of standard receiving-tube design intended to operate at about 1,000°K with 6.3 volts across the heater and 300 milliamperes of heater current.

The four-sided cathode sleeve was formed from an active nickel alloy (roughly 0.2 per cent each of silicon, manganese, and iron) and was spray-coated on two sides with triple alkaline-earth carbonates. Each coated area was about .05 square centimeter, .06 centimeter wide by .71 centimeter long, and had a coating weight of 12 milligrams per square centimeter. One of these coated areas, the No. 1 or front cathode, K_1, faced the slit-plate array and constituted the target cathode. The other coated area, the No. 2 or rear cathode, K_2, could not receive the ion beam (except by surface migration). This No. 2 cathode and its anode were used as a reference diode for measuring any activation effects due to surface migration of the deposited metal

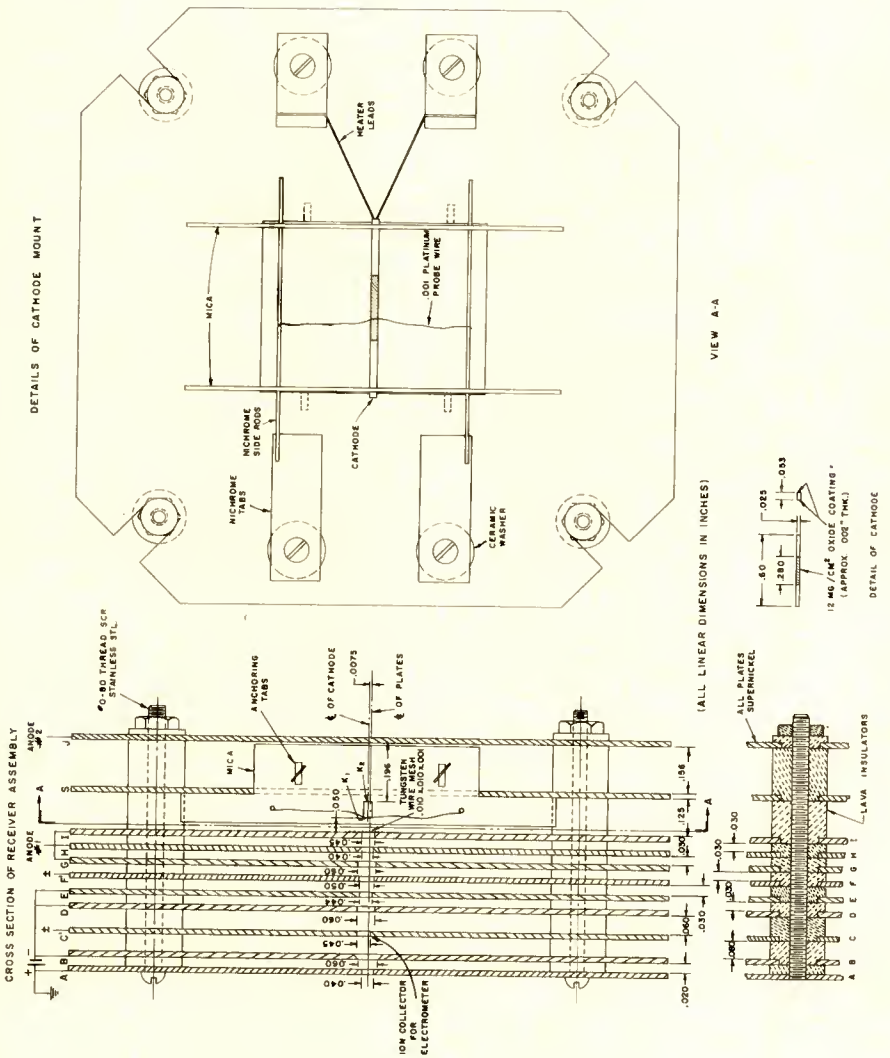


Fig. 3—Receiver assembly.

or changes in residual gas composition. This diode was also useful for estimating long-term drift changes in emission from the target cathode. During most of the deposition and recovery periods, the drift or change in the emission from K_2 was an insignificant fraction (0.1 to 1 per cent) of the change in emission from K_1, hence the changes in K_1 due to surface migration of deposited metal and to changes in gas atmosphere were clearly negligible. The perveance p_1 of the front diode, containing the target cathode, was

$$p_1 \simeq 4.3 \times 10^{-5} \text{ amp}/(\text{volt})^{3/2}.$$

Beam Currents and Deposition Rates Employed

The maximum resolved ion currents deposited by the mass spectrometer on the target cathode were about 3.5×10^{-9} ampere, corresponding to a current density of 7×10^{-8} ampere per square centimeter of oxide coating.[18] This is the rate at which Ba atoms would be incident (4.4×10^{11} per square centimeter per second) on a cathode immersed in Ba vapor at a partial pressure (at 1,000°K) of 4.6×10^{-9} mm Hg. Assuming the cathode surface to be a perfectly smooth (100) plane, the density of surface atom sites is 1.3×10^{15} per square centimeter, hence the maximum rate of deposition amounted to 3.4×10^{-4} monolayer per second or 0.02 monolayer per minute. The coating weight of 12 mg/cm^2, if pure BaO, would have contained about 4.6×10^{19} BaO molecules; hence in one minute the maximum quantity of deposited barium amounted to a mole fraction excess of roughly 6×10^{-7}. It has been estimated[2] that the density of donors in an active oxide cathode is 10^{15} to 10^{16} per cubic centimeter, which is equivalent to a concentration mole fraction around 4×10^{-8} to 4×10^{-7} (considering the coating as pure BaO having a concentration of 2.2×10^{22} BaO molecules per cubic centimeter.

Incidental Experimental Results

In addition to the effects on electron emission of deposits of alkaline earth metals on the target cathode, a few observations were made of the effects of deposits of ions of residual gases and of the effects of residual gas molecules.

[18] All such ion beam current density values listed in this paper have absolute errors perhaps as large as ±50 per cent; however, for comparison purposes, the relative values of the beam current densities are probably internally consistent within about 20 per cent for all operations for which normal high energy (1,200 electron volts) beams were used. The errors in relative magnitudes of beam current densities are always such that the beam current densities on the target may be less than those expected from the values measured by the electrometer.

Effects of Ions of Residual Gases on Electron Emission

The results of the efforts to measure the effects of ions *other* than those of the alkaline earth metals deposited on the target cathode suggest that these effects vary considerably with ionic species, with cathode condition, and with rate of deposition of the species. At least one stage more of differential pumping, and perhaps two stages more than was used would be required for making a thoroughly satisfactory study of the effects of ions of gases not condensable at room temperature. With the system used, it was apparent that the rates of incidence of ions of residual gases or of introduced gases could not be made to exceed the rates of incidence of the parent neutral molecules on the target cathode.

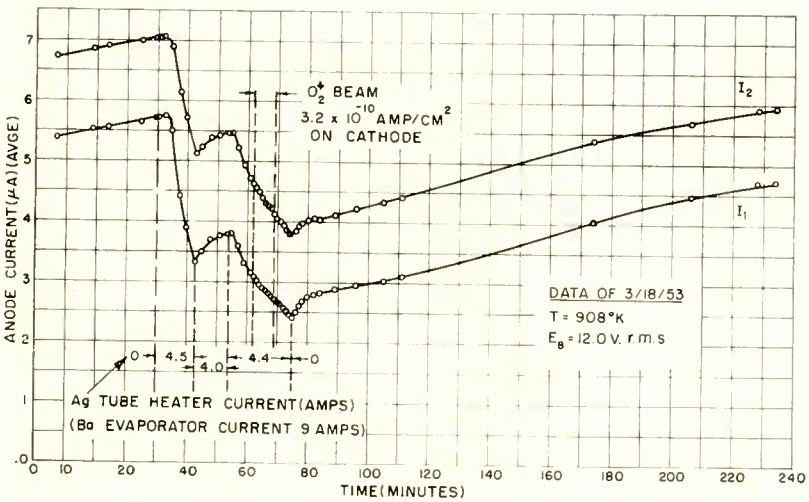


Fig. 4—Cathode deactivation by exposure to O_2 (and O_2+).

The effects of ions of residual gases were studied mostly during the early stages of the experiments before the system had become as thoroughly degassed as it eventually did during the metal depositions; however, at all stages, the change in electron emission produced by depositing the O_2+ beam (3.2×10^{-10} ampere per square centimeter) as in Figure 4 were so slight as to be of indeterminate sign. In contrast, a definite activation effect was produced in the early stages of operations by deposition of the mass 28+ ions beam (composed of N_2+ about 75 per cent, CO+ about 25 per cent, 3×10^{-9} ampere per square centimeter) as shown in Figure 5. In later operations, Figures 6 and 7, the effect of 28+ ions appeared quite different from that in Figure 4; nevertheless, even in Figures 6 and 7 the action of the 28+

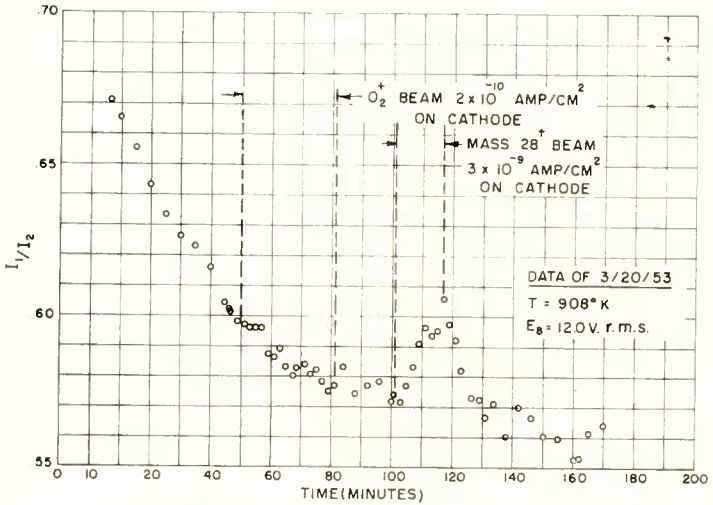


Fig. 5—Cathode activation by deposition of mass 28+ ions.

beam must be interpreted as a slight activation effect if the completely inactive condition of the cathode is taken as the reference state. From Figure 6, a threshold appears to have existed in the onset of activation by deposition of Ba+ ions. Below a certain rate of deposition of Ba+, the effect of the Ba+ beam was small and quite similar to that of the 28+ beam; nevertheless, for equal currents (Figure 6) slightly greater than the threshold value for activation by Ba+ ions, the Ba+ ions

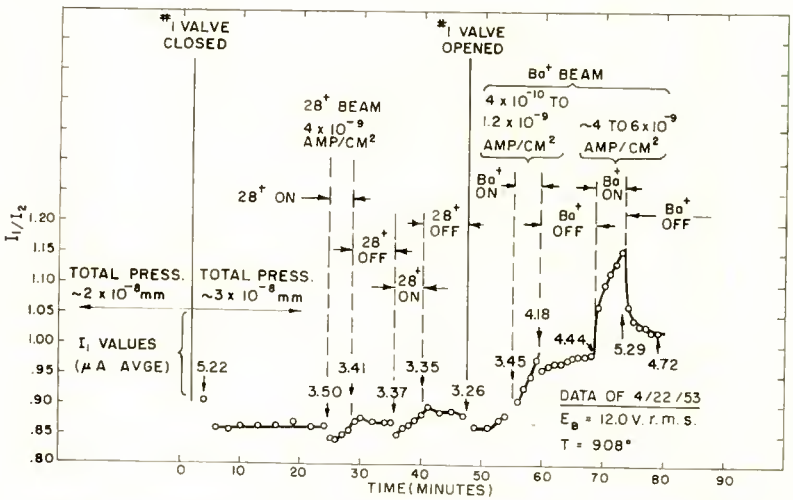


Fig. 6—Comparison of threshold activation effects of depositions of mass 28+ ions and Ba+ ions.

produced a clear-cut activation effect substantially larger than that produced by the 28+ beam. It is feasible to suppose that, because of electronegativity differences on other variations, the activation thresholds should lie at higher current densities for the CO^+ and N_2^+ ions than for Ba^+, Mg^+, etc.

The data for Figures 4 and 5 were taken while O_2 was being admitted to the spectrometer through a heated silver thimble attached to the gas inlet tube, *T*, of the ion source (Figure 2). This arrangement introduced the O_2 co-axially into the ionizing electron beam, the copper inlet tube being used also as the electron collector for monitoring the electron beam current. Even despite this artifice, together with the moderately high current density electron beam (50 milliamperes

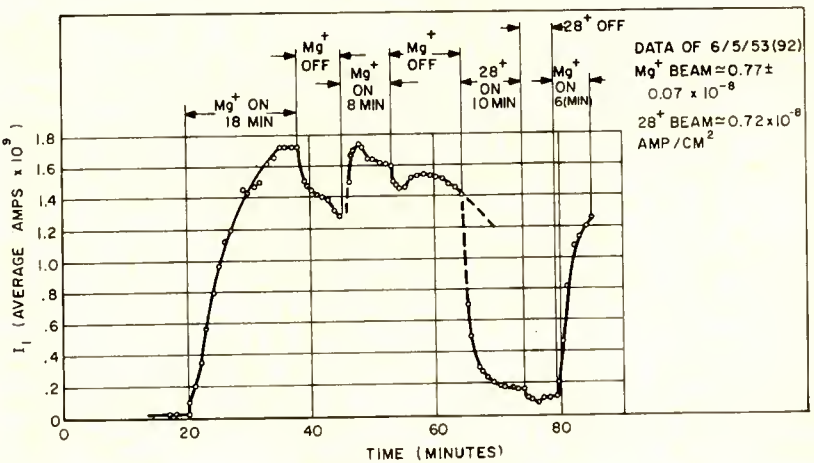


Fig. 7—Comparison of activation effects of depositions of Mg^+ and mass 28+ ions.

per square centimeter) and one stage of differential pumping, the rate of incidence of O_2 molecules on the target cathode exceeded the rate of deposition of O_2^+ ions by two or three orders of magnitude. The obtainable O^+ current was even smaller and deposition of O^+ was likewise ineffective. The same unfavorable situation existed for deposition of mass 28+ ions from residual N_2 and CO, assuming these two gases originated rather uniformly throughout the analyzer tube.

Table I gives the gas spectrum as detected in the ionizing chamber of the mass spectrometer for the two conditions, (a) the silver thimble heated for admitting O_2 and (b) the silver thimble at room temperature. (The silver thimble was obviously not thoroughly degassed at this point.) If for condition (b) one assumes the residual gas com-

position as detected in the ionizing chamber is roughly representative also of the gas composition over the target cathode, and if for condition (a) one assumes that an attenuation factor of 10 exists at the target cathode because of the differential pumping, it can then be estimated that for gases entering the system at the ion source, heating the silver tube increases the O_2 partial pressure at the cathode more than three decades above the residual O_2 pressure present when the silver tube is cold. For most of the other species the increase is trivial.

Table I—Measurements of Residual Gas Spectra Taken with Silver Thimble Cold and with Silver Thimble Heated to Admit O_2.

M/ne Numbers of Principal Peaks	Probable Identity	Ion Currents (amps $\times 10^{13}$) Thimble Hot	Thimble Cold
2	H_2^+	2.	0.5
14	N_2^{++}, CO^{++}, CH_2^+	47.	24.
15	CH_3^+	1.8	0.3
16	O^+, CH_4^+	20.5	0.3
17	OH^+	0.1	<0.1
18	H_2O^+	0.2	<0.1
20	A^{++}, HF^+	1.5	0.7
23	Na^+	0.5	0.5
28	N_2^+, CO^+	456.	117.
29	N_2^+, CO^+	2.6	0.83
32	O_2^+	170.	<0.1
34	H_2S^+	0.36	0.00
40	A^+	5.5	2.7
44	CO_2^+	0.15	<0.1
101	Hg^{++}	0.3	0.15
202	Hg^+	1.5	0.8
Total Pressure in Ion Source		$p \simeq 1 \times 10^{-7}$ mm Hg	$p \simeq 2 \times 10^{-8}$ mm Hg

Electrostatic scan; values of ion current not corrected for instrumental discrimination effects; one Ba evaporator at about 900°C, producing no Ba.

Effects of Residual Gases on Electron Emission

In the early stages of the deposition experiments, a threshold was found in the deactivating effect produced on electron emission from the test cathode by leakage of air into the system. During this interval it was apparent from the size of the O_2^+ peak in the mass spectrum that a leak, later discovered to be in the gasket seal adjacent to the target cathode, was contributing a partial pressure of O_2 at the cathode estimated as about 10^{-9} mm Hg. In this period, the cathode was

always found to deactivate slowly. When the leak was reduced so that the O_2 partial pressure was diminished by about a factor of 50, the cathode emission promptly reversed its direction of drift and slowly increased.

Gases evolved from the metal evaporators always produced some activation of the cathode whether or not detectable alkaline earth metal evolution was produced simultaneously from them. Gases from the spectrometer filament produced a much smaller activating effect. The general behavior of the emission during evolution of these gases is shown in Figure 8. The increase in total pressure produced *over the cathode* by one heated ribbon during the interval recorded in Figure 8 was probably less than 30 per cent of the background with ribbons

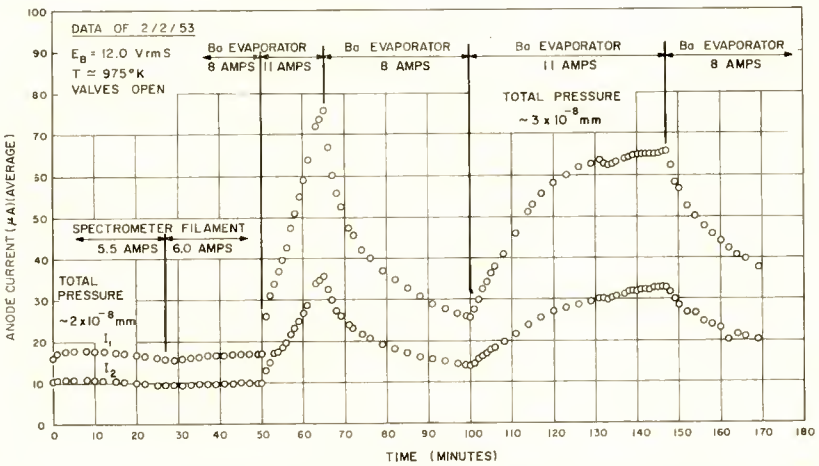


Fig. 8—Cathode activation by gases evolved from metal evaporators.

cold, the increase *at the ion source* being about a factor of three. The data of Figure 8 were taken with the valves in the pumping lines open.

Accumulation of residual gases upon closing the valves in the pumping lines always caused activation of the cathodes. This closure brought an increase in total pressure over the cathodes from about 2×10^{-8} mm Hg (with an evaporator ribbon at about 1,000° C) to about 2×10^{-7} mm Hg. Table II gives the gas composition computed for a typical accumulation of gases in the system with the pump lines closed. From consideration of observations given in the following paragraph, it is believed that the partial pressures accumulated are true equilibrium values for the materials evolving the gases rather than a representation of the steady-state balance between rates of evolution and the leakage conductance of the valves in the pumping lines.

Figures 9a and 9b show the time dependence of a similar accumulation of common gases in the presence of a barium evaporator which had been less completely degassed. The maximum accumulated total pressure in these figures is about 6×10^{-7} mm Hg. It is of interest that the increase in the H_2^+ and CH_4^+ ion currents in Figure 9b is somewhat more than two orders of magnitude while the increase in the 28^+ ($CO^+ + N_2^+$) peak, which is by far the largest resolved ion beam in the mass spectrum, is but a factor of thirty. This suggests that the closed-valve values for the partial pressures of at least those particular species whose concentrations increased less than two decades after valve closure are real equilibrium pressures rather than steady-state values determined by the leakage of the valves. It is estimated that the H_2^+ ion current plotted in these figures should be multiplied by a factor of approximately nine (three for instrumental discrimination, three for ionization cross section) to allow comparison of the partial pressure of H_2 with those of other species by direct comparison of ion currents. With this correction, it is apparent that at the maximum accumulation, the H_2 concentration equalled that of the $CO + N_2$ combined. The mass numbers 12, 14, 15, 16, and 26 presumably are due largely to ionization fragmentation of CH_4 and C_2H_2. Numbers 2, 3, 17, 18, are due to H_2^+, H_3^+, OH^+, and H_2O^+, respectively. Number 20 may be either A^{++} or HF^+; because it varied with ribbon temperature, it is probably HF^+, the HF being derived from the use of a mixture of HF and HNO_3 for the cleaning of tantalum strips which provided support and electrical connections to the evaporators. The mass numbers 40, 32, 29, 28, and 14 seem rather independent of evaporator temperature, hence they are mainly supplied by the remainder of the system.

Table II—Composition of Residual Gas Mixture Accumulated in System with Pumplines Closed and an Evaporator Ribbon at 1,000°C.

Constituent	Per Cent by Volume
H_2	45 ± 10
He	0.01
CH_4	1.8 ± 0.2
H_2O	1.5 ± 0.2
C_2H_2	0.04
$CO + N_2$	45 ∓ 10
O_2	0.03
A	4.0 ± 0.5
Total pressure $\simeq 2 \times 10^{-7}$ mm Hg. Evaporator Temperature $\simeq$ 1,000°C.	

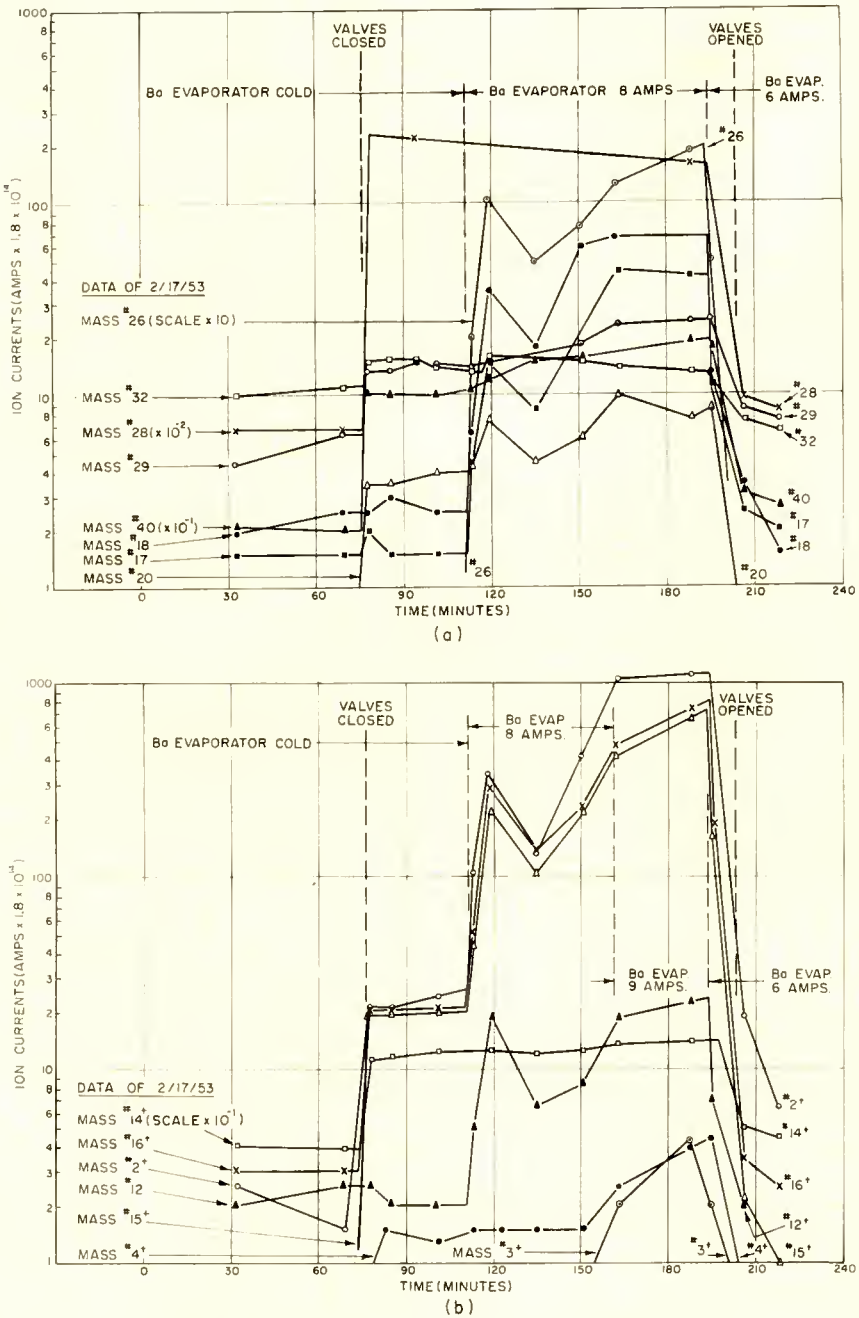


Fig. 9—Gas evolution from metal evaporators.

At somewhat higher evaporator temperatures, very small amounts of higher hydrocarbon fragments were discernible along with increasing amounts of inorganic fragments, such as Be^+, Na^+, K^+, Ti^+, TiO^+, Mn^+, Fe^+, Rb^+, Cs^+, and the alkaline earth metals. At the highest metal evaporation rates, CO and CH_4 constituted the bulk of the noncondensable gas evolved, and the total of the carboniferous ion currents was probably less than one tenth the magnitude of the ion current of the principal alkaline earth metal. H_2^+ was not apparent in the spectrum under these conditions; however, because the instrument was then adjusted to optimize the Ba^+ peak, the discrimination against H_2^+ may have been much more severe than the factor of nine applicable for normal operations. The possible importance of H_2 in cathode reactions was not appreciated at that time and no special effort was made to keep an account of the H_2^+ behavior. No H_2O^+ was detectable in the spectrum while appreciable alkaline earth metal partial pressures were present in the ion source.

Because of the differential pumping arrangement (which of course operated in both directions) it is quite likely that during deposition operations the residual gas composition around the target cathode was roughly that indicated by the final entry in Figure 9 and by the "cold thimble" values of Table I. On this basis, considering the correction factor of 9 applicable to these data for the H_2^+ current, one concludes that the H_2 partial pressure in the receiver chamber was probably the order of 10^{-8} mm Hg, CO was the order of 10^{-8} mm Hg, O_2 and CO_2 the order of 10^{-11} mm Hg. The rates of incidence of molecular species on the cathode at partial pressures of 10^{-8} mm Hg are approximately 3.5×10^{12} per square centimeter per second. This is roughly ten times the maximum employed rate of incidence of deposited Ba^+ ions.

It is important that this condition be kept in mind when one attempts to interpret the kinetics of the activation produced by metal deposition. The gross features of the emission response at low temperatures to metal deposition (e.g., Figure 22d) can be matched by assuming a bimolecular process for the rate of formation of donors, dN/dt, of the form,

$$\frac{dN}{dt} = \alpha \left(\frac{I_{Ba} t}{e} - N \right) (N' - N).$$

Here I_{Ba} = barium deposition current,

α = a rate constant,

N = concentration of donors all of which are formed by the reaction,

N' = initial concentration of a second species.

However, such a model does not encompass the behavior subsequent to deposition, because the peak in electron emission must be generated by consumption of the reservoir of the second species.

It should be noted that, in most "high-vacuum" environments, the partial pressures of relatively electronegative residual gases exceed the partial pressures of electropositive metals. Without substantial reducing action by the cathode base metal, this condition also exists in vacuum tubes and the cathode surface becomes somewhat electronegatively loaded.[19] With rates of incidence of electronegative molecules equalling or exceeding the metal deposition rates employed, the electronegative loading cannot be completely destroyed. Though some of the deposited metal atoms are stored as atoms, perhaps transiently, and some subsequently evaporate, most of the deposited metal atoms are expected to combine quickly with the electronegative surface. Of these latter, perhaps only a small fraction reacts in such a way as to form *effective* electron donors—the remainder may yield their electrons for local charge balance or for association with various other less effective ionized-donor materials. The chemical composition of the surface may thus be as important in the formation of donors as is the injection of reducing chemical potential. The kinetics of the reactions themselves, or of prerequisite surface diffusion processes may be involved in the long decay periods observed.

Studies of the Activation Phenomenon

Introductory Notes

Five separate series of deposition experiments were attempted; however, because of various mechanical and electrical failures or mishaps only two of these series, the first and the last, proceeded to the point that effective depositions could be made. Results of the first series have been described previously.[2,13,15] This series was superior to the final series in that the platinum probe wire embedded in the coating was usable during depositions; however, in the first series, only Ba evaporators were employed and these were exhausted before all the desired operations were completed. All the results described

[19] The principles behind this observation are quite analogous to those operative in the well known use of adsorption indicators for detecting end-points in analytical titrations which produce insoluble reaction products.

in the present paper were obtained from the final series of deposition experiments for which the platinum probe was not operable but for which adequate metal evaporators were employed.

In this final series, the following depositions were made: Ca^+, one; Sr^+, one; Mg^+, five; O_2^+, five; 28^+, five; Ba^+, fifty-five.

The series of measurements extended over a period of nearly six months and was terminated by mechanical failures of the vacuum system—cracks developed in glass around the leads to evaporators in the ion source and the cold traps became clogged with mercury from the diffusion pumps. Extensive though the number of Ba depositions is, it is not sufficient to allow determining how the deposition activation of the cathode varied with other environmental factors. Secondary differences in activations by the four metals were of course not detected.

The emission activity of the cathodes was measured by observing the changes in anode currents with a fixed 60-cycle anode voltage applied (usually 12 volts r-m-s). The cathodes were deactivated prior to each deposit on the front cathode, K_1, by heating them to 1,150°K for ten minutes without drawing anode current. This treatment always produced the result that a substantial fraction (as deduced from measurements of the probe potential) of the diode voltage drop appeared across the cathode coatings.[20] The current from the second cathode, K_2 (Figure 3), was always monitored to determine incidental emission drifts and effects of background gas composition changes.

General Features of Deposition Activation

As will be observed in Figure 10 and others following,[21] the cathode activation produced by depositions of alkaline earth metals is a complex function of each of several variables—residual gas environment, duration of deposition, rate of deposition, total amount of metal deposited, temperature of the cathode during deposition, the elapsed time after initiation and termination of deposition, and perhaps of other variables not clearly discernible. For this reason the results cannot be described briefly, either verbally or mathematically, and hence a large number of plots of the data must be presented to convey the nature of the behavior.

From these plots, it is apparent that a description of the "state of activation" of a cathode requires much more than simply specifying

[20] During these experiments, the electron currents drawn from the cathodes were always essentially resistance limited and space-charge limited in the diode rather than emission limited.

[21] All data plots carry the dates on which data were recorded. Unless otherwise specified, the kinetic energies of the ions deposited were in the range 1,200 to 1,400 electron volts.

the emission level attained at one point in the measurements. The thermionically active state of a cathode is usually a thermodynamically unstable state maintained only by the consumption of externally supplied energy. Accordingly, any change imposed on the state of the system (the cathode and its environment) generates a complex time dependence of the emission behavior such as that shown by the emission build-up during ion deposition and the emission decay during the recovery period following termination of the deposition.

Figure 10 shows one of the more simple curves of emission response to metal deposition. The major features of Figure 10 are:

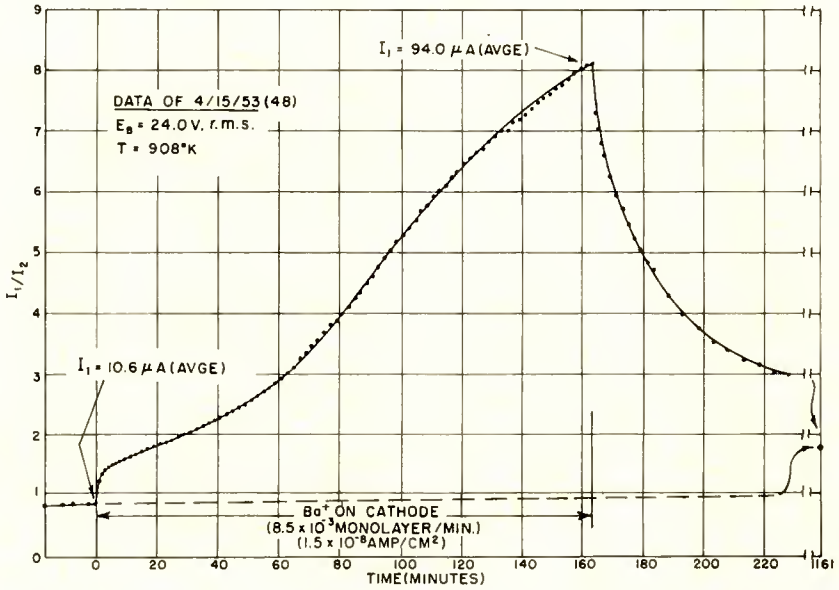


Fig. 10—Deposition activation by long duration medium current Ba+ beam.

(1) A small stable electron current prior to deposition on the deactivated cathode.

(2) A slow rise of electron current toward an asymptotic value during deposition. During this slow rise, the emission current, i, can be represented by the equation,

$$i = i_o + i_1 (1 - e^{-\alpha_1 q})$$

in which i_o = the initial current,

α_1 = a constant,

$i_o + i_1 =$ the asymptotic current value (which is proportional to the rate of metal deposition), and

$$q = \int i\,dt.$$

(3) A relatively short transient response when deposition starts. When the slow transient of (2) above is subtracted from the complete response, it is found that the *short transient* emission component, i_s,

$$i_s = i_3\,(e^{-\alpha_2 q} - e^{-\alpha_3 q}),$$

can be represented by

where α_2 and α_3 are constants, and
i_3 is the $q = 0$ intercept value of either exponential.

(4) A slow decay in electron current to an asymptotic value, i_∞, subsequent to termination of deposition. The emission current i_d during this slow decay may be fitted by the equation,

$$i_d = i_\infty + i_4\,e^{-\alpha_3 q},$$

where $i_\infty + i_4 =$ electron current value at termination of deposition. It is found that $(i_\infty + i_o)$ is proportional to the total amount of metal deposited.

(5) A fast transient whose component, i_f, subsequent to termination of deposition may be represented by

$$i_f = i_5\,(e^{-\alpha_4 q} - e^{-\alpha_5 q}),$$

where i_5 is the $q = 0$ intercept value of either exponential.

These features are interpreted as follows:

The slow rise during deposition is attributed to donor formation at the surface and to the electrolysis and diffusion of donors through the donor depletion layer. The donors moving through the depletion layer act as minority carriers of low mobility and give rise to a large increment in electron current. The slow rise to the asymptotic value indicates either that (a) the donor formation is immediate and the donor current into the cathode is proportional to the density of donors at the surface or (b) the donor current is limited by the rate of the reaction through which donors form, diffusion in this case carrying the donors into the coating as fast as they form. The data are inadequate for determining which mechanism is dominant.

The slow decay subsequent to termination of deposition is again

interpreted as due to diminishing enhancement of the electron current by the dwindling donor current through the depletion layer as the excess deposited barium on the surface is consumed.

The rapid transient when deposition starts is attributed to the spatial redistribution of the donor depletion layer. This is in part due to a direct effect of the deposition current. The electron current across the vacuum is space-charge limited. The barium-ion beam crosses the vacuum to reach the cathode surface, partially neutralizes the electronic space charge in the diode space, and increases the voltage across the cathode. The total current across the vacuum space is the sum of the measured electron current, i, and the barium ion current i_{Ba}. Thus the electron current through the *cathode* must be the sum $(i + i_{Ba})$. The excess current i_{Ba} through the cathode and the change in voltage across the cathode due to space-charge neutralization in the vacuum cause a redistribution of the donors and give rise to a transient in the electron current. This transient is closely related to the more familiar decay in pulse emission. The increase in donor density at the surface as the Ba is deposited also contributes to the relaxation. This increase probably contributes in a minor way at the higher cathode temperatures where donor mobility is relatively high. The possible effect of this increase at lower cathode temperatures will be discussed later.

The short transient in the decay subsequent to deposition is again attributed to donor-depletion layer relaxation as it was for the short transient in emission build-up.

Thus to describe the behavior of the emission during the deposition and recovery periods, three terms exponential in q are required. In some other plots, an additional term exponential in t is needed.

Figure 11 shows the similarity in rate and level of activation produced by comparable beams of Ba^+, Sr^+, Ca^+, and Mg^+ on the cathode surface. The similarity in the shapes of the emission versus time plots is quite evident.

In Figure 12 the fractional change in the emission current $(i_\infty - i_o)/i_o$, attained in the decay periods following depositions of various metals, is plotted as a function of the number of monolayers of metal deposited. The unlabeled points are those obtained by depositing Ba^+ on the cathode. This plot, too, indicates that the variation in activation produced by the different metals is not greater than the variation between successive activations with the same metal. Not enough depositions were made to discover and establish definitely any "second-order" differences in the effects of the different metals; however, as will be discussed in detail later, such differences are

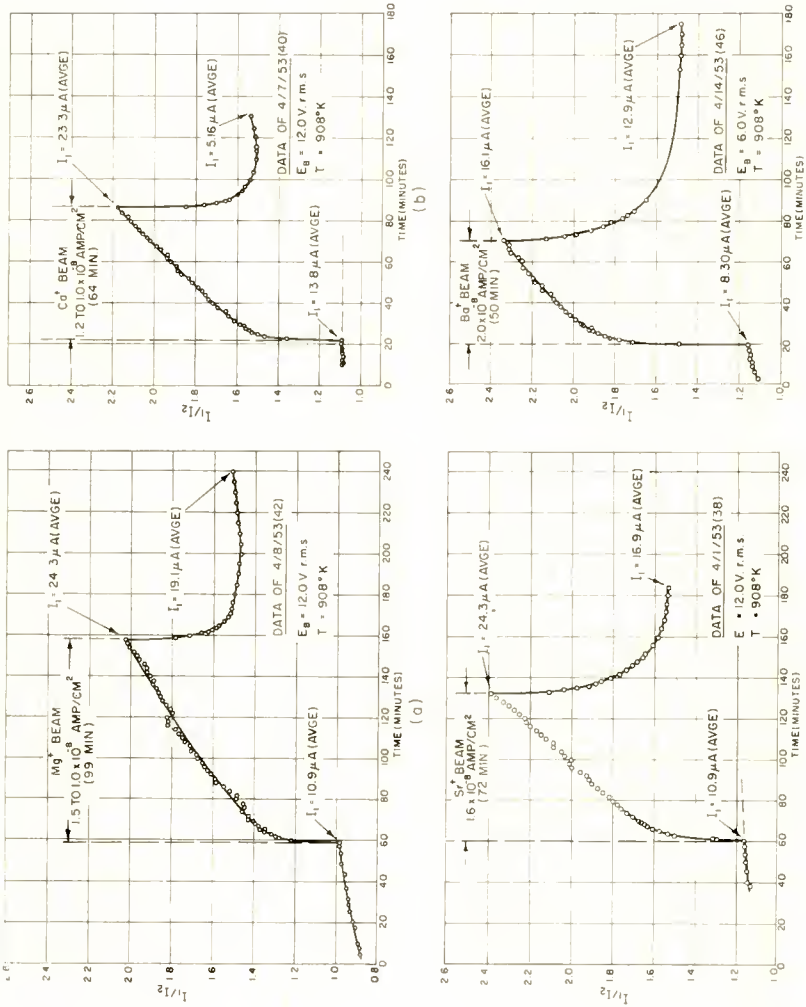


Fig. 11—Comparison of activations produced by depositions of Mg^+, Ca^+, Sr^+, and Ba^+.

expected to be trivial or nonexistent on thermochemical grounds. Any apparent differences are assumed to be due to (a) minor instrumental variables and (b) variation in cathode environment or previous history from one deposition to another.

The more complicated behavior which occurs at lower temperatures is shown in Figures 13, 17, and others. The curves which show maxima both during and after deposition are particularly striking. They strongly resemble the "short transients" observed in the higher temperature measurements (Figure 10). This resemblance suggests that the "slow" response of the higher temperature measurements is suppressed at lower temperatures and that the fast transients dominate the response.

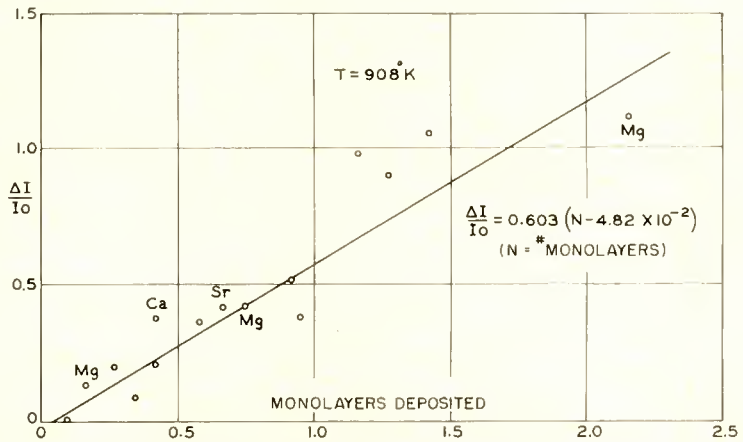


Fig. 12—Net activation produced by depositions of alkaline earth metals (unlabelled points are those due to Ba depositions).

A search for correlation between the appearance of maxima and various parameters, such as temperature and deposition current, has thus far disclosed a discernible correlation with but one parameter, I_{Ba}/i_o, where I_{Ba} is the deposition current, and i_o is the initial electron current. An analysis of 48 deposition runs shows that when $I_{Ba}/i_o < 2$, (19 runs) maxima do not appear, when $I_{Ba}/i_o > 10$, (10 runs) maxima always appear, and when $2 < I_{Ba}/i_o < 10$ maxima may (11 runs) or may not (8 runs) appear. The remaining 16 metal depositions were omitted from this consideration because they were too short or were otherwise inappropriate for determining whether they should show maxima.

The observation that maxima occur only when the deposition current exceeds the initial electron current, so that the electron current

through the cathode at least doubles at the inception of deposition, supports the view that the "short transient" attributed to donor-depletion-layer relaxation dominates the behavior under these circumstances. A simple model of a cathode which exhibits donor depletion is explored in the Appendix and, in particular, its behavior when a

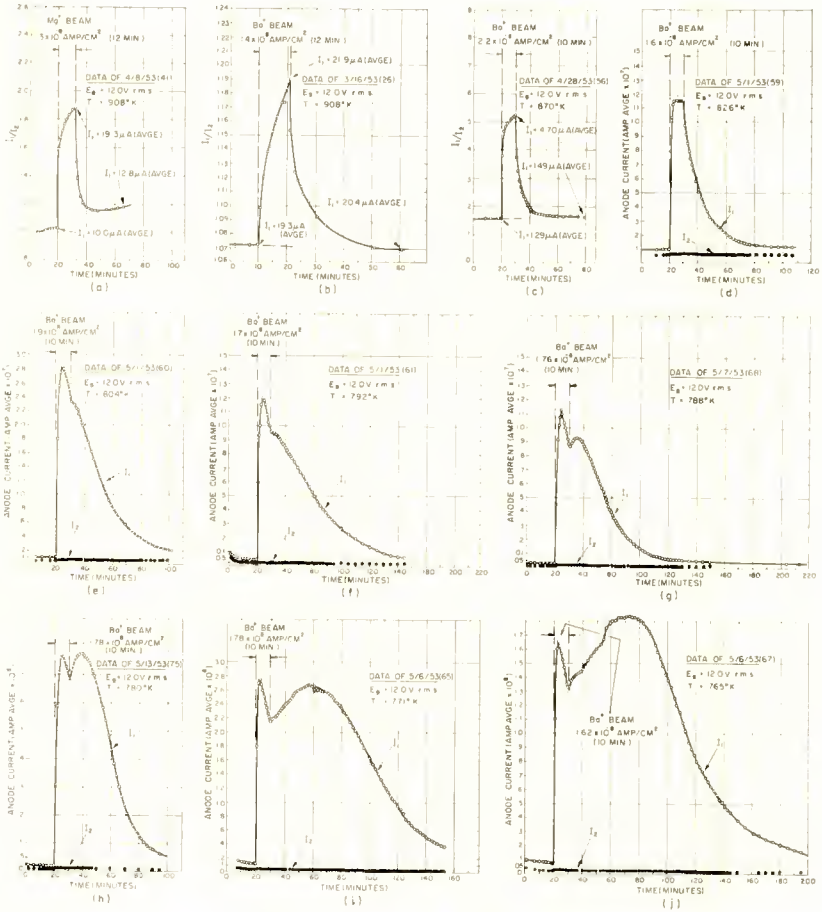


Fig. 13—Variation of deposition activation with cathode temperature.

donor current flows into the front surface is examined. It is found that when a donor current enters the surface, the donor density at the surface increases. Also, if the ratio of donor current I_o to initial electron current, i_o, is large enough, the donor-depletion layer moves into the cathode and, further, if

$$\frac{I_o}{i_o} > 12\,\frac{\mu^+}{\mu^-}$$

where μ^+ and μ^- are the donor and electron mobilities, respectively, the donor density at the point of minimum density in the depletion layer is less than the donor density which previously existed at the surface when no donor current was flowing. The donor mobility is of

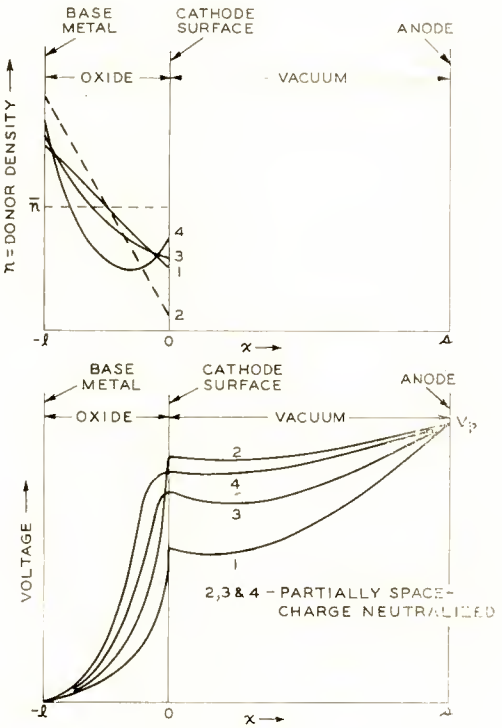


Fig. 14—Donor and voltage distributions prior to and during deposition.

the order of 10^{-5} and the electron mobility is of the order of 10. Hence, the relation cited above is approximately

$$\frac{I_o}{i_o} > 10^{-5}. \qquad (1)$$

As noted above, maxima in the experimental deposition curves appear when

$$\frac{I_{Ba}}{i_o} > 10, \tag{2}$$

approximately. Hence, if the donor current exceeds 10^{-6} times the deposited barium current, relation (1) is satisfied when relation (2) is satisfied. Thus a relatively long reaction time for the formation of donors should lead to the donor current required for a donor density minimum within the cathode.

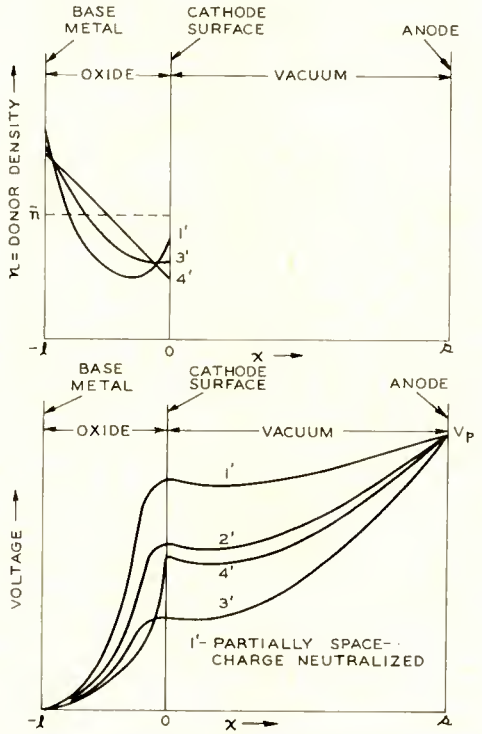


Fig. 15—Donor and voltage distributions subsequent to deposition.

With these considerations in mind a tentative explanation of the appearance of maxima shown in Figures 13, 17, 22, et al is offered. In Figures 14 and 15, the donor distribution is shown in the upper plots and the voltage distribution throughout the diode in the lower. In Figure 14, the initial donor distribution and voltage distributions are as shown by Curves 1. When deposition starts, the voltage across the vacuum drops abruptly because of space-charge neutralization; at the same time the electron current through the cathode increases by an amount I_{Ba}, so that the cathode voltage drop increases markedly with little change in the electron current to the anode. The voltage

is then as shown by Curve 2, lower plot. If there were no donor current into the cathode, the donor distribution would tend toward that shown by Curve 2, upper plot; however, the deposition increases the density of donors at the surface and produces a donor current into the cathode so that the donor distribution progresses to the distribution shown by Curve 4, upper plot, passing through the distribution of Curve 3 in the process. The corresponding voltage distributions are shown by the similarly numbered curves in the lower plots. The currents corresponding to these donor and voltage distributions are shown in Figure 16.

The behavior subsequent to deposition is shown in Figure 15. When deposition ceases, the voltage across the vacuum increases abruptly

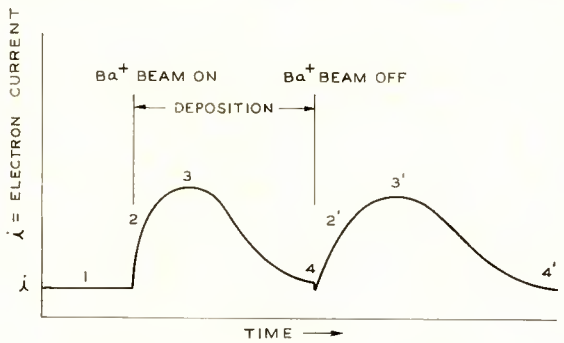


Fig. 16—Diode currents corresponding to the donor and voltage distributions of Figures 14 and 15.

because of the loss of space-charge neutralization in the vacuum. At the same time, the electron current diminishes by an amount I_{Ba}, thus reducing the voltage drop through the cathode. This process may result in a discontinuity in current such as is observed in Figures 23a, 23b, and others. The resulting voltage distribution is shown by Curve 2′, lower plot. The donor distribution then progresses through distributions such as 3′, upper plot, to the final distributions 4′. The corresponding changes in voltage distribution are shown by the similarly numbered curves in the lower half of Figure 15 and the corresponding currents are shown in Figure 16.

Although many models have been examined since the first observation of deposition curves exhibiting maxima, none has been found that accounts for as much of the observed detail as that outlined above. It is unfortunate that the model is so mathematically intractable, particularly in the transient case, that a quantitative comparison of the model with the experimental results has not been possible.

Variation of Deposition Activation with Impact Energy of Incident Ions

It has not been possible to evaluate unambiguously the nature and extent of variations in activation with kinetic energy of the incident ions. The comparative measurements are given in Figures 17 and 18. The similarities in curve shapes produced in emission build-up during depositions and in emission decay upon termination of depositions rather strongly indicate that no significant variations exist with beam energies in the range employed, 70 to 1,400 electron volts.[22] However, data can be extracted from these plots to suggest that the magnitude of activation increases slightly with decreasing energy of the incident ions. Such a set of comparative data is given in Table III. By other suitable cross comparisons of the same data, one can perceive ambiguities in the data.

Although in general the emission magnitudes and curve shapes were roughly reproducible from week to week with appropriately selected and adjusted rates of metal deposition, they were not reproducible in detail. This nonreproducibility was due in part to the variation in time elapsed between the initial "firing-up" of cathode, ion source, and metal evaporators, and the taking of data. The nonreproducibility was also due in part to the diminution in pumping speed of the vacuum lines because of the accumulation of mercury "trees" in the cold traps.

In addition to these environmental variables, instrumental variables were strongly involved in measurements of activation as a function of incident ion energy. Two methods were employed to produce resolved ion beams with low kinetic energies: (1) The use of small ion-accelerating voltages and appropriately small magnetic deflecting fields in the analyzer tube. With this method of operation, the intensity losses in the analyzer itself were so great for the low ion energies desired that the obtainable currents of ions incident on the cathode were in the threshold range (or lower) where the activation effects are strongly sensitive to incident ion current magnitudes. (2) The use of normally large ion-accelerating voltages at the ion source together with decelerating[23] voltages applied to the diode structure (target cathode plus plate S, anode #1 and anode #2 of Figure 3).

[22] These are kinetic energies of the ions before they entered the target diode space in which they were further accelerated or decelerated by the 60-cycle diode voltage (usually ± 17 peak volts) before striking the cathode.

[23] For Figures 17 and 18, the label "decelerated beam" is applied to distinguish this second type of deposition from the first or "normal beam" type which employed no significant post-resolution deceleration. When not otherwise specified, "normal beam" deposition of ions having impact energies 1,200-1,400 electron volts was employed.

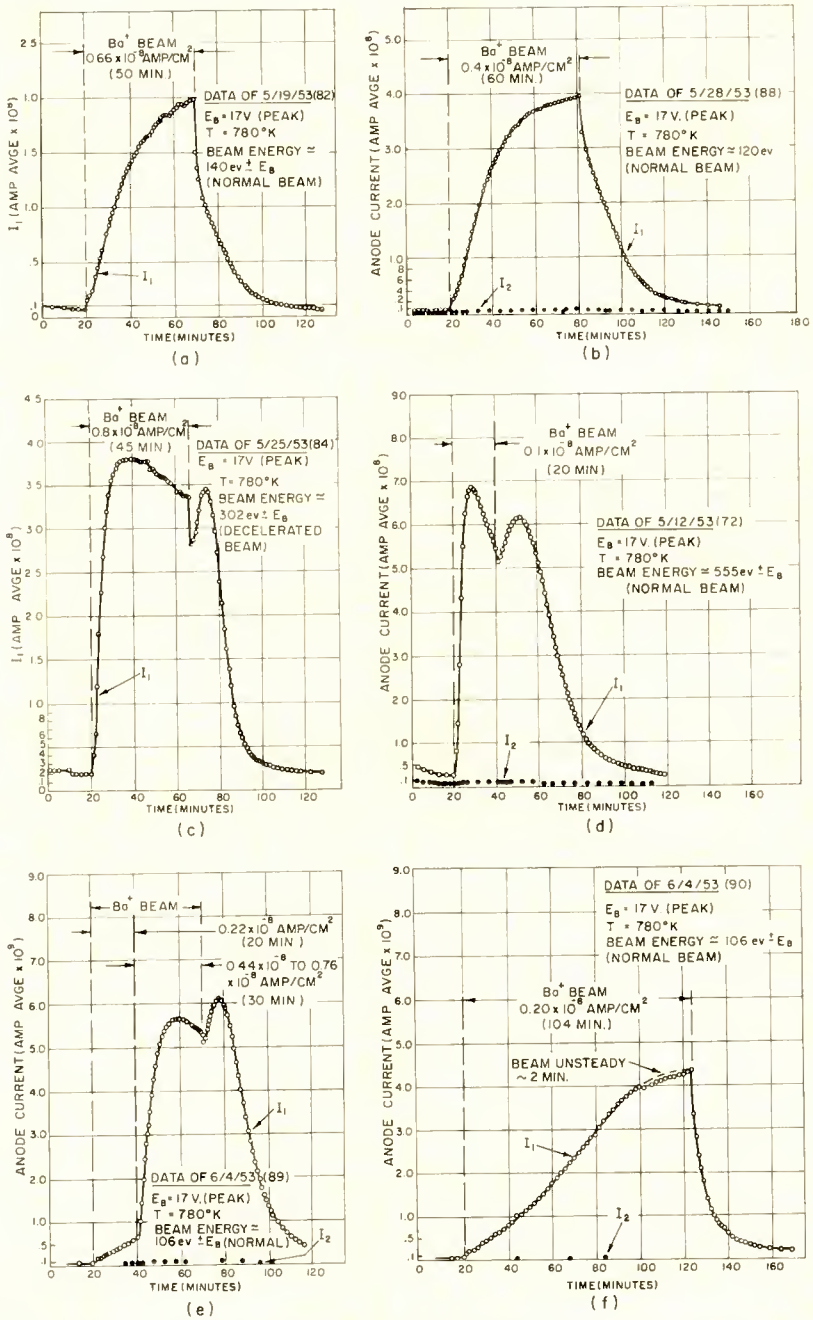


Fig. 17—Variation of deposition activation with beam energy and beam current density at low temperature.

Unfortunately, this method very probably produced defocusing of the beam after it passed the monitoring electrode (C of Figure 3) but before it struck the target cathode. For this reason, the Ba^+ beam current densities actually incident on the cathode during the use of this "decelerated beam" mode of operation were probably smaller than those measured by the electrometer and listed on the figures. From

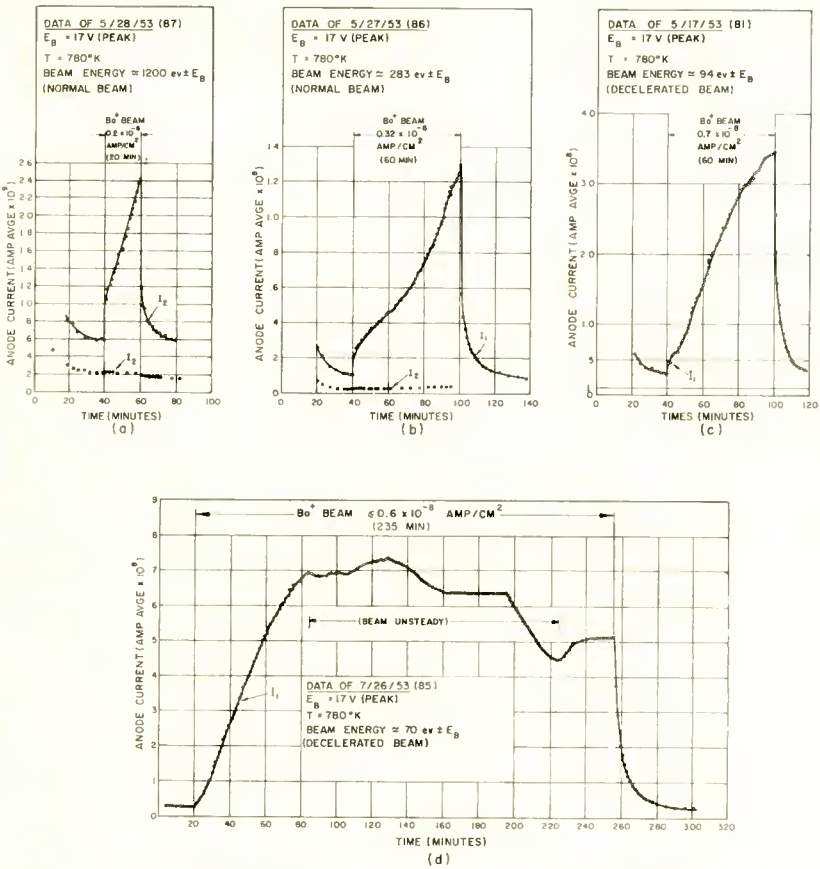


Fig. 18—Comparison of activation curves generated by deposition of beams having different kinetic energies.

comparisons of the curve shapes (a), (b), (f) in Figure 17, (a), (b), (c), (d) in Figure 18, it is seen that similar curve shapes can be generated by known low-current high-energy normal beams and (nominally) higher current low-energy decelerated beams. If, indeed, the real current densities of the decelerated beams are less than the apparent values, then the relative effectiveness of low-energy beams

for activating the cathode is greater than the data suggest when taken at face value. Considering these experimental factors, it is believed, therefore, that the activation produced by these depositions of metal is primarily a real chemical reaction rather than a "radiation" damage phenomenon. In fact, the bombardment damage appears to counteract, to some extent, the inherent reducing chemical potential of the metal deposited.

Table III—Summary of Data Concerning the Variation of Deposition Activation with Kinetic Energy of Incident Ions (Ba+ Beam Cathode Temperature 780°K)

Entry No.	Figure No.	Current Density (Amp × 10^8 cm^2)	Impact Energy (ev)	Elapsed Deposition Time for Which Emission Comparison is Made (minutes)	Electron Emission Obtained (amp)
1	17b	0.40 (normal)	1,200 ± 17	50	3.8×10^{-8}
2	17a	0.66 (decelerated)	140 ± 17	50	2.0×10^{-7}
3	17e	0.44 to 0.76 (normal)	1,066 ± 17	20	5.6×10^{-9}
4	17c	0.8 (decelerated)	302 ± 17	20	3.8×10^{-8}
5	18c	0.56 to 0.7 (decelerated)	94 ± 17	20	1.6×10^{-8}
6	18d	0.6 (decelerated)	70 ± 17	20	2.7×10^{-8}
7	23b	0.9 to 1.0 (normal)	1,140 ± 17	15	6.5×10^{-8}
8	17d	0.9 (normal)	555 ± 17	9	6.8×10^{-8}
9	18a	0.20 (normal)	1,200 ± 17	20	2.4×10^{-9}
10	17e	0.22 (normal)	1,066 ± 17	20	6.0×10^{-10}
11	17f	0.20 (normal)	1,066 ± 17	20	8.0×10^{-10}
12	18b	0.32 (normal)	283 ± 17	20	4.5×10^{-8}

Variation of Activation with Cathode Temperature, Rate of Deposition, and Duration of Deposition

Figures 19 through 23 comprise a rather complete set of plots of deposition activation and decay at three different temperatures with rate of deposition and duration of deposition as the variable parameters. From these and previous plots, it is seen that the threshold deposition current density is about 0.3×10^{-8} ampere per square centimeter, corresponding to about 10^{-3} nominal monolayer per minute. The maximum current densities employed are about twenty times larger. Still larger current densities would be desirable because they would probably yield equally useful results with the cathode at higher temperature and would also generate higher levels of activation.

Assuming ideal solution of the deposited metal in the oxide coating, it can be estimated that the highest deposition current employed, 7×10^{-8} ampere per square centimeter, producing a metal atom inci-

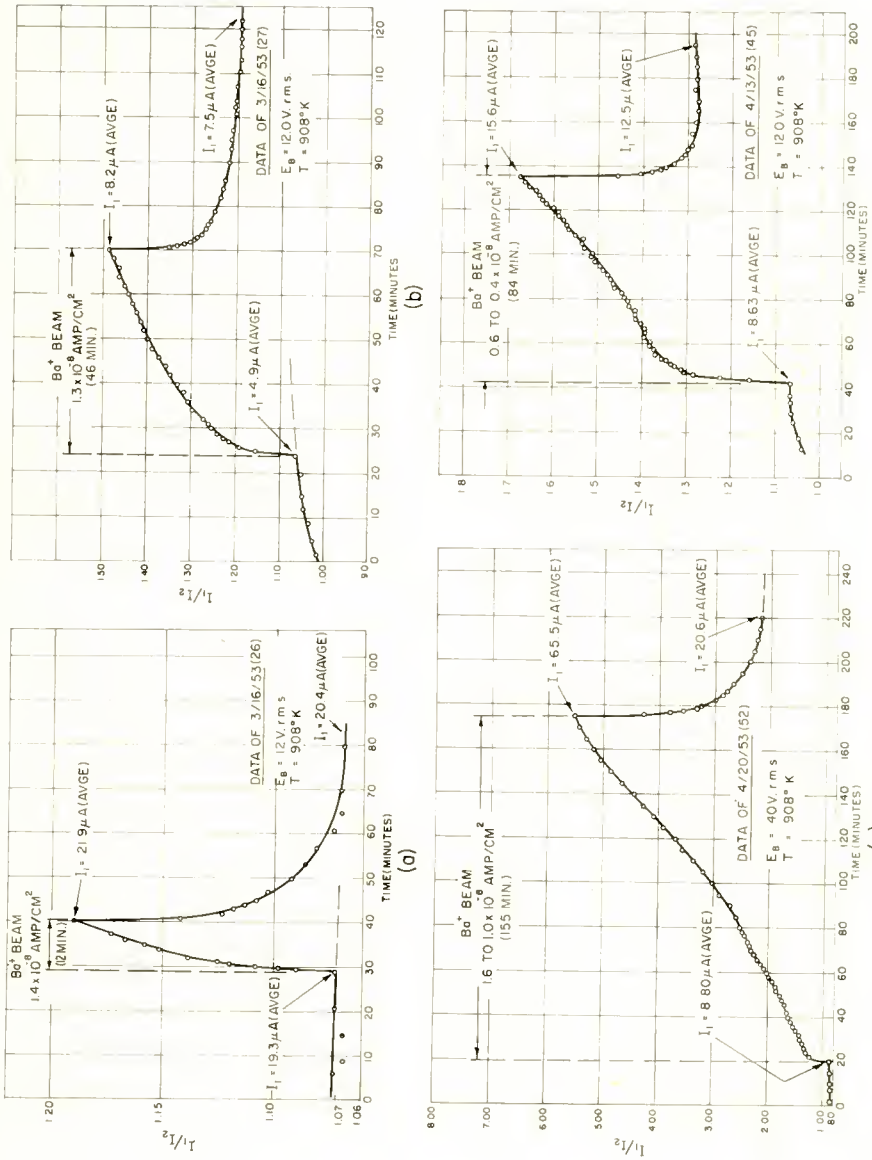


Fig. 19—Variation of deposition activation with duration and rate of deposition (cathode at high temperature, 908° K).

dence rate equivalent to that from a Ba partial pressure of 4.6×10^{-9} mm Hg, could have supported an equilibrium excess metal concentration of 3×10^{-7} mole fraction at 908°K and 2×10^{-5} mole fraction at 765°K, the extremes of cathode temperature used in the depositions. The partial pressures of various residual gases around the cathode were previously estimated at approximately 10^{-8} mm Hg for H_2, 10^{-10} mm Hg for H_2O, 10^{-8} mm Hg for CO, and 10^{-11} mm Hg for O_2 and CO_2. From simple rate of incidence and fractional surface coverage

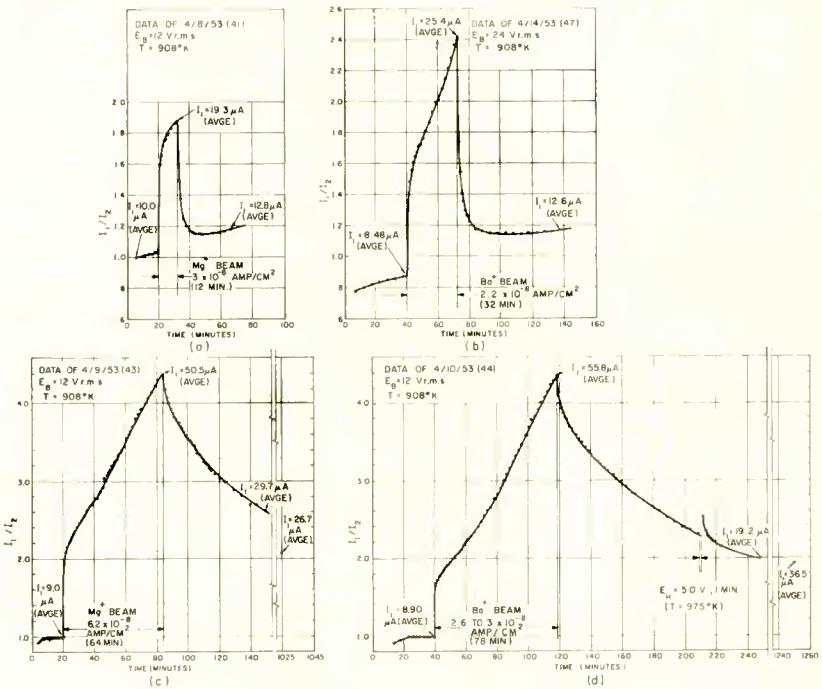


Fig. 20—(1) Comparison of activations by Mg+ and Ba+ at high deposition rates; (2) Variation of deposition activation with duration of deposition (cathode at high temperature, 908°K).

considerations, it is unlikely that the equilibrium mole fractions of excess metal available for donor formation can have been diminished by reaction with residual gases to less than a few per cent of the values given above.

Furthermore, with the range of deposition rates employed, the greatest amount of structure in the activation and decay curves was found with total deposits in the range .01 to 1.5 nominal monolayer put down, corresponding to a mole fraction excess of metal, assuming no reaction with gases, of 3×10^{-7} to 4×10^{-5} (about 7×10^{15} and

10^{18} atoms per cubic centimeter of oxide coating, respectively). Because the range of deposition currents employed varied only a factor of twenty from minimum to maximum, rate of incidence and fractional coverage considerations suggest that at least a few tenths of one per cent of these quantities should have escaped complete oxidation by incident gases at the front surface under the lowest rates of deposition

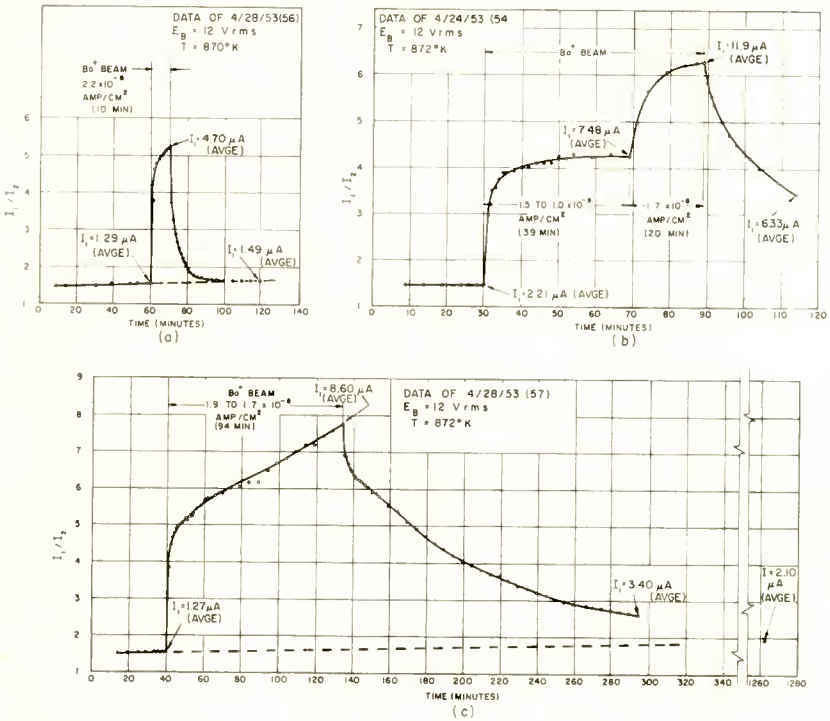


Fig. 21—Variation of deposition activation with duration of deposition at a fixed rate (cathode at moderate temperature, 872°K).

employed and should have penetrated the coating as active reducing chemical potential.

Thus, had each atom of these minimal quantities produced an electron donor 1.4 electron volts below the conduction band,[24] fairly high states of cathode activation should have been attained at the highest rate of deposition. At the lowest temperatures, increases in electron currents in excess of two orders of magnitude (e.g., Figures

[24] H. B. DeVore, "Photoconductivity Study of the Activation of Barium Oxide," *RCA Review*, Vol. XIII, p. 453. December, 1952.

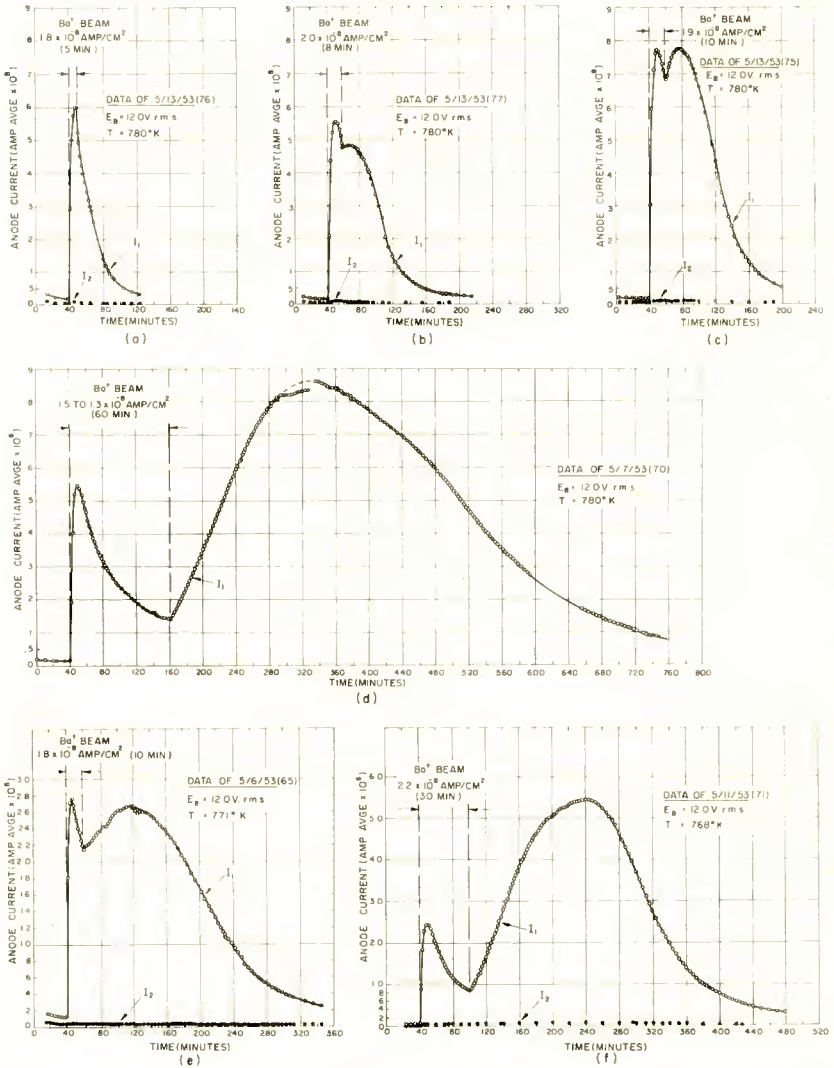


Fig. 22—Variation of activation deposition with duration of deposition at a relatively high rate of deposition (cathode at low temperature, 780°K, 770°K).

22d and 22f) referred to the inactive reference condition were produced; however, much smaller relative increases were measured at the higher temperatures, the maximum measured current density being about 6 milliamperes per square centimeter (peak) at 908°K (c.f. Figures 10, 19c, 20c, 20d), the highest cathode temperature at which

depositions were made. Corrected to 1,000°K, this would correspond to an emission density around 30 milliamperes per square centimeter. From the geometrical perveance value of the #1 diode, 4.3×10^{-5} amp/(volt)$^{3/2}$ it is apparent that the space-charge-limited emission value would not exceed 60 milliamperes per square centimeter with the anode voltage (17 volts peak) usually employed during depositions. Thus had the cathode coating shown zero resistance, 60 milliamperes per square centimeter electron current might have been drawn; to have a space-charge limited current of 6 milliamperes per square

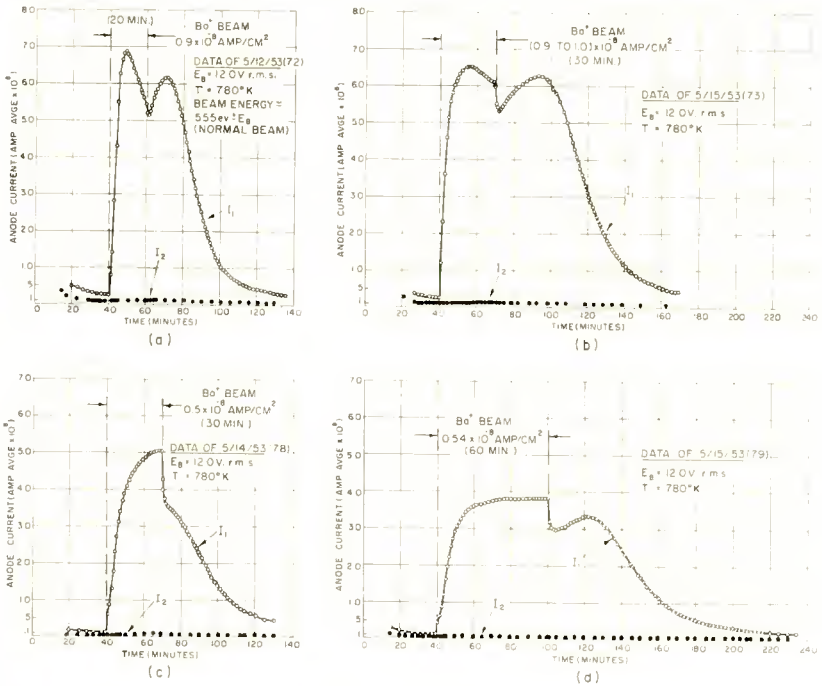


Fig. 23—Variation of deposition activation with duration of deposition. Comparison of effects of two low rates of deposition (cathode at low temperature, 780°K).

centimeter, 3.6 volts was required between anode and front surface of the cathode, or 13.4 volts across the cathode coating. With this current and voltage across the coating, the coating resistance at 908°K must have been

$$R_k \simeq 2{,}200 \text{ ohm-cm}^2.$$

In our experience, well-activated cathodes show coating resistances at 1,000°K of the order of

$$R_k \simeq 10 \text{ ohm-cm}^2.$$

Assuming an activation energy of 1.4 electron volts for conduction, one estimates from this that at 908°K,

$$R_k \simeq 44 \text{ ohm-cm}^2$$

as a *maximum* value which the cathode should have shown if it were well-activated.

From these considerations, it seems likely that although the maximum currents measured were space-charge limited values because of the voltage drop in the coating, the level of activation as indicated by the attendant coating resistance was about a factor of ten poorer than that present in a well-activated cathode. One might conclude one of the following:

(1) Only a very small fraction of the Ba deposited was utilized in donor formation, or

(2) excess Ba does not constitute the effective donor material and, alone, does not guarantee a high level of activation, or

(3) application of auxiliary processing is needed in addition to the provision of an adequate amount of excess barium. (High-field electrolysis might constitute an example of auxiliary processing.)

The general conclusion which encompasses all of these specific alternatives is that although excess Ba may constitute a necessary condition for producing an active cathode, it does not constitute a sufficient condition.

Miscellaneous Experiments

Deposition on a Cold Cathode

Figure 24 shows the effect of depositions made on a cathode at room temperature. A notable feature of this experiment is the negligible final activation caused by the deposition; Figure 24a shows that the final asymptotic valve of plate current is almost identical with that which would have been expected in the absence of the deposition. This suggests the possibility that the deposited metal is rapidly evaporated from the cathode during the period when the cathode was heated from room temperature to the measuring temperature of 908°K. Presumably most of the metal can be retained only if it reacts with the surface of the cathode. The data of Figure 24 suggest that the rate of reaction is much lower than the rate of evaporation at the lower temperatures.

The initial high plate current immediately following the deposition indicates that the number of donors created by the deposition is suffi-

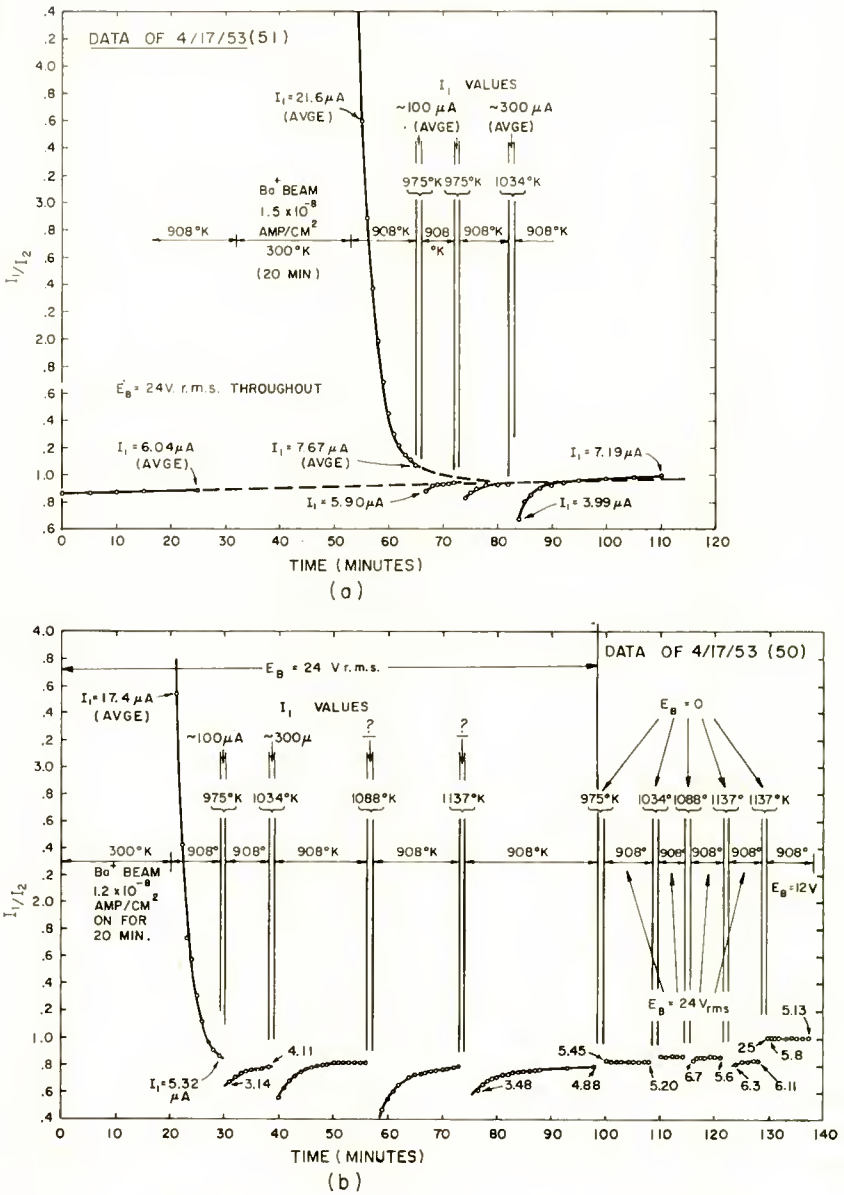


Fig. 24—Activation by deposition of Ba on cathode at room temperature; depletion layer formation by drawing anode current during "hot shot."

dent to eliminate the donor depletion layer at the surface of the cathode but not sufficient to affect the over-all activity of the cathode appreciably.

Both plots of Figure 24 also show the effect of enhancing the depletion layer by brief "hot-shots" with plate voltage applied. Unfortunately these "hot-shots" were applied too late in the decay following deposition to indicate clearly whether the plate current approaches the original decay curve or the final equilibrium asympote. If it does approach the decay curve, this would strongly suggest that potential donors can be "stored" at or near the surface.

Successive Depositions

Figure 25 shows one of the few studies made of cathode activation through repeated depositions interspersed with recovery intervals. The most notable feature of this plot is the very rapid rise in emission upon initiating the second and subsequent depositions. In the last of these, the emission values were read at 15-second intervals during the

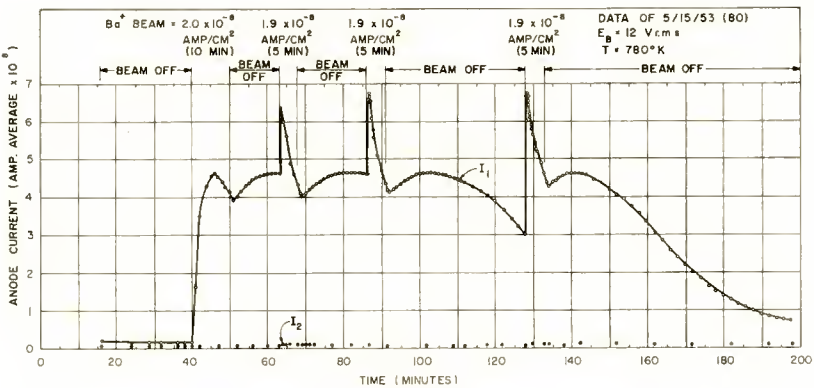


Fig. 25—Activation by successive depositions.

early part of the deposition. The maximum in emission during this deposition was attained in about 30 seconds, corresponding to a deposition increment of approximately 3×10^{-3} nominal monolayer.

Unfortunately similar measurements were not made at higher temperatures for which the emission maximum is absent from the build-up and recovery branches of the activation plot.

Discrepancies in Activations by the Four Metals

As pointed out earlier, not enough depositions were made with the different metals to determine whether secondary differences exist in their activation effects. The magnitudes and nature of such minor differences, which one should contemplate detecting, are indicated by the Mg deposition activations plotted in Figure 26. Whether fortuitous or significant, three features are present in this plot which were not found in the 780°K curves for Ba activations. They are

(1) the lack of sharpness in the maximum occurring in the deposition branch,

(2) the relatively greater sharpness of the maximum in the recovery branch (compared with that in the deposition branch), and

(3) the emission in the recovery branch exceeding that in the deposition branch.

With respect to the differences in electrical behavior one should expect to be produced by these different "doping" agents, certain idiosyncrasies of the operation of the mass-action principle should be considered. Too frequently, the doping of *compounds* is undertaken

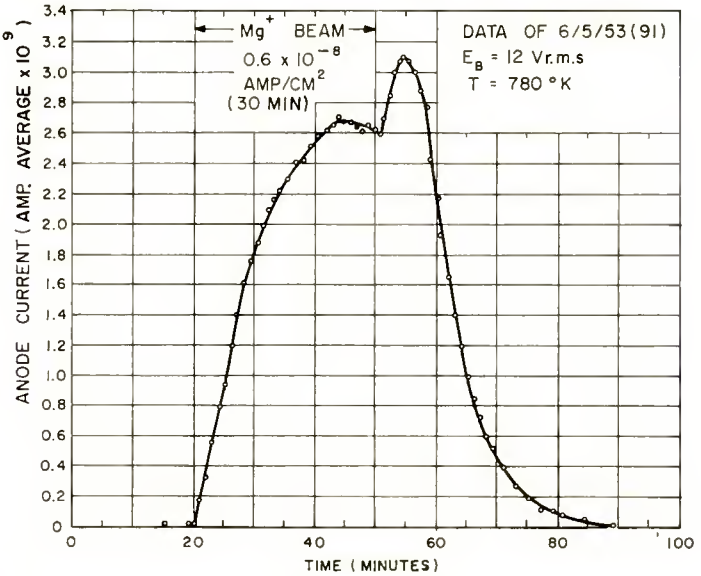


Fig. 26—Deposition activation by Mg+ beam. An example of curve shape which does not conform to pattern of activation recovery curves produced by Ba depositions.

as a process completely analogous with the doping of Ge and other Group IV B elements. In these elements, the covalent bond completely dominates the chemical behavior of the solvent, and hence doping amounts largely to a simple substitution of covalently bound doping atoms for solvent atoms. The electrical behavior of the solution is then rather uniquely related to the more flexible valence characteristics of the solute. In compounds, the situation is quite different. Often both the electropositive and the electronegative constituent of the solvent compound can exist in two or more valence states having but relatively little differences in energy content. This means that the

equilibrium constants for transformations between these states of oxidation are finite, and hence that significant activities of each of the oxidation states are always present at equilibrium. They all enter into the determination of the Fermi level of the material. Alteration of the concentration of any of these species alters the Fermi level; likewise, substitution of any species having an electronegativity different from that of the species it replaces alters the Fermi level and the distribution of concentrations of all other species.

Let us consider cathode "doping" with the aid of the reaction below.

$$\mathrm{Mg(s) + BaO(s) = MgO(s) + Ba(s).}$$

Initially, the MgO present as an impurity in the triple oxide coating may be very low, perhaps only a mole fraction of 10^{-4}. Assuming ideal solutions and activities equal to mole fraction concentrations and taking the equilibrium constant as near unity,

$$\frac{a_{\mathrm{Ba}}\ a_{\mathrm{MgO}}}{a_{\mathrm{Mg}}\ a_{\mathrm{BaO}}} \simeq 1.$$

At the end of the first Mg^+ deposition, the value

$$a_{\mathrm{Mg}} \leqq 10^{-5}$$

might have been attained. Taking BaO as the solvent, for which $a_{\mathrm{BaO}} \simeq 1$, and taking $a_{\mathrm{MgO}} \simeq 10^{-4}$, it is seen that

$$\frac{a_{\mathrm{Ba}} \times 10^{-4}}{10^{-5} \times 1} \simeq 1,$$

whence $a_{\mathrm{Ba}} \simeq 0.1$.

Thus, to satisfy the mass-action principle, a *very large* Ba activity would be generated in reaching equilibrium upon deposition of a *very small* quantity of Mg in a coating having a very low MgO content. The charge neutrality requirement, the volatilities of the metals, and diffusion limitations prevent such drastic behavior, but the implications inherent in the process are clear, nevertheless. As a consequence of the *limiting* factors, the result must be in practice that almost every Mg incident is transformed into a crystalline Mg^{++} site and one Ba atom is formed which evaporates or goes into solid solution.

In short, the foreign doping agent, Mg, is quantitatively removed from circulation as a potential electron donor, and only the reduction

product impurity of the solvent remains (the Ba atom) to behave as a donor or to react further to form a donor.

In the present experiment, because a triple oxide coating was employed such that

$$a_{CaO} \simeq a_{SrO} \simeq a_{BaO},$$

the nonforeign doping agents Ca, Sr, and Ba are *not* quantitatively removed by the operation of mass action. Of the three solvent cations, the most easily reduced (and/or most easily soluble in the oxide) will be the predominant reaction product, regardless of which of the four doping metal atoms is deposited by the spectrometer.

Thermochemically, therefore, only trivial differences in cathode activation can be expected from deposition of small amounts of the various electropositive elements on the oxide cathode. The activation effects of small quantities (10^{-3} mole fraction) of Mg, Al, C, Si, W, etc., should be essentially equivalent, regardless of the precise identity of the most important electron donor species. This, of course, is an agreement with general cathode experience.

Conclusions

(1) Activation measurements were successfully performed with metal deposition rates in the range 10^{-3} to 2×10^{-2} monolayer per minute, 2×10^{-2} monolayer per minute being the highest rate employed compatible with adequate life of metal evaporators used. Higher deposition rates would be desirable for making measurements of still higher levels of emission activation.

(2) The levels of emission activation produced by the quantities of metal deposited in the various deposition cycles, 2×10^{-7} to 2×10^{-5} mole fraction of the oxide coating, are less than should be expected if a substantial fraction of the "excess" metal deposited forms electron donors. The equilibrium metal vapor pressures from the cathode sustainable by the rates of deposition employed were presumably high enough to maintain substantial F-center concentrations. The absence of a very high emission level indicates either that (a) only a small fraction of the deposited Ba atoms forms donors, or (b) other additional impurity species are needed in the donor formation, or (c) other additional processing is needed.

(3) An abundance of emission measurements was obtained which showed the existence of much detail previously unknown and unsuspected in the kinetics of emission activation and decay. A tentative analysis of this behavior in terms of a physical model is presented. A complete physical-chemical model has not yet been derived.

(4) Supplementary data are given showing the behavior of the oxide cathode in the presence of small partial pressures of deliberately added O_2 and of gases evolved from metal evaporators. Accumulations of gases from the evaporators always produced activation. The activating gas species appear to be one or more of the following: H_2, CH_4, C_2H_2.

Appendix

To acquire some notions about the effect of a donor current on the donor depletion layer in a cathode, consider the following over-simplified model. Assume that all donors are ionized and that there is no space charge so that the free electron and ionized donor densities are equal. Then the electron current, i, and ionized donor current, I, are given by the equations,

$$i = ne\mu^- E + D^- e \frac{\partial n}{\partial x}, \tag{1}$$

$$I = ne\mu^+ E - D^+ e \frac{\partial n}{\partial x}. \tag{2}$$

Here, n = electron density = ionized donor density,

μ^- = electron mobility,

D^- = electron diffusion constant,

μ^+ = (ionized) donor mobility,

D^+ = (ionized) donor diffusion constant.

The electron current is continuous throughout the circuit so that

$$\frac{\partial i}{\partial x} = 0. \tag{3}$$

The donor current is discontinuous because of the low solubility of the donors in the base metal. Therefore,

$$\frac{\partial I}{\partial x} = -e \frac{\partial n}{\partial t} \neq 0. \tag{4}$$

E may be eliminated between Equations (1) and (2) yielding

$$I = \frac{\mu^+}{\mu^-} i - 2D^+ e \frac{\partial n}{\partial x}. \tag{5}$$

Substituting Equation (5) into Equation (4) gives

$$\frac{\partial^2 n}{\partial x^2} = \frac{1}{2D^+}\frac{\partial n}{\partial t}. \tag{6}$$

A quasi-stationary solution is

$$n = n_o + a_1 x + a_2 \frac{x^2}{2} + 2D^+ a_2 t. \tag{7}$$

The integration constants n_o, a_1, a_2 are determined from the boundary conditions

$$I = I_o \text{ at } x = 0 \text{ (at front surface)},$$

$$I = 0 \text{ at } x = l = \text{thickness of cathode.}$$

This is tantamount to assuming zero solubility for donors in the base metal.

$$\int_o^l n dx = \bar{n} l + \frac{I_o t}{e},$$

where $\bar{n}$ = the average density of donors initially. With the constants so determined, Equation (7) becomes

$$n = \bar{n} - \frac{1}{2D^- e}\left[i - \frac{\mu^-}{\mu^+} I_o\right]\frac{l}{2} - \frac{I_o}{2D^+ e}\frac{l}{6} + \frac{I_o t}{el}$$

$$+ \frac{1}{2D^- e}\left(i - \frac{\mu^-}{\mu^+} I_o\right) x + \frac{I_o}{2D^+ e}\frac{x^2}{2l}. \tag{8}$$

and the donor current becomes

$$I = I_o\left(1 - \frac{x}{l}\right). \tag{9}$$

The donor density has a minimum when

$$\frac{\partial n}{\partial x} = 0,$$

i.e., when

$$x = \left(1 - \frac{\mu^+}{\mu^-}\frac{i}{I_o}\right) l. \tag{10}$$

This minimum falls within the cathode when

$$\frac{\mu^- I_o}{\mu^+ i} > 1.$$

The minimum density is

$$\bar{n}_m = \bar{n} + \frac{I_o t}{el} - \frac{I_o}{2D^+ e}\left[\frac{1}{3} - \frac{\mu^+ i}{\mu^- I_o} + \left(\frac{\mu^+}{\mu^-}\frac{i}{I_o}\right)^2\right]\frac{l}{2}. \tag{11}$$

The question pertinent to the present problem is: Given an electron current i_o with the donor current at the surface, I_o, equal to zero so that the surface donor density is n_o, can the minimum donor density drop below n_o for some electron current i_i when $I_o \neq 0$? An alternative question is: for what values of i_o and I_o is the following relation satisfied?

$$\bar{n} + \frac{I_o t}{el} - \frac{I_o}{2D^+ e}\left[\frac{1}{3} - \frac{\mu^+ i_1}{\mu^- I_o} + \left(\frac{\mu^+ i_1}{\mu^- I_o}\right)^2\right]\frac{l}{2} < \bar{n} + \frac{I_o t}{el} - \frac{i_o}{2D^+ e}\frac{l}{2}.$$

In more compact form, this is

$$\xi_2^2 - \xi_2 + \frac{1}{3} > \xi_1 \tag{12}$$

where

$$\xi_1 = \frac{\mu^+ i_o}{\mu^- I_o},$$

$$\xi_2 = \frac{\mu^+ i_1}{\mu^- I_o}.$$

For this minimum to lie within the cathode $(0 \leq x \leq l)$,

$$0 \leq \xi_2 \leq 1$$

from Equation (10). The minimum value of the quadratic on the right of Equation (12) is 1/12. Therefore, if

$$\xi_1 < \frac{1}{12}, \tag{13}$$

the inequality is certainly satisfied. In terms of the currents, condition (13) becomes

$$\frac{I_o}{i_o} > 12\,\frac{\mu^+}{\mu^-}.$$

In terms pertinent to the deposition experiments, this relation states that if the ratio of donor current to initial electron current is great enough, a donor density minimum as deep as the original donor density minimum at the surface will form within the cathode.

The above model is an inadequate description of a cathode in many respects.

1. It assumes that all of the donors are ionized. This is certainly not true as is apparent from the dependence of cathode conductivity on temperature.

2. It tacitly assumes that the electron current is small enough so that the donor density does not drop to zero within the cathode. For large currents, depletion must be complete and the current must become space-charge limited to account for the large voltage drops within cathodes observed experimentally.

3. Space-charge neutrality has been assumed. This assumption requires the internal electric field to be constant. This is at variance with the voltage distributions observed within cathodes by probe measurements.

When these limitations are not imposed, the algebra becomes unwieldy, to say the least, and it has not yet been carried through. The transient case is even more forbidding.

RCA TECHNICAL PAPERS†

Second Quarter, 1957

Any request for copies of papers listed herein should be addressed to the publication to which credited.

"Alloying Properties of Germanium Free of Edge Dislocations," C. W. Mueller, *RCA Review* (June) 1957

"Analysis of Post Detection Integration Systems by Monte Carlo Methods," E. Ackerland and Co-author, *I.R.E. Convention Record, Part 2, Circuit Theory and Information Theory* 1957

"Anodized Aluminum or Mica Washers for Insulated Mounting of RCA-2N301 and 2N301-A Transistors," *RCA Application Note AN-171*, Semiconductor Division, Radio Corporation of America, Somerville, N. J. (June) 1957

"Antenna-to-Medium Coupling Loss," H. Staras, *Trans. I.R.E. PGAP* (April) 1957

"Automatic Gain Controlled Audio Program Amplifier," G. A. Singer, *Broadcast News* (April) 1957

"The 'Bizmac' Digital Data Processing System," J. C. Hammerton, *Electronic Engineering* (April) 1957

"Circuit Considerations for Audio-Output Stages Using Power Transistors," R. Minton, *I.R.E. Convention Record, Part 7, Audio and Broadcast Transmission Systems* 1957

"Circuit Considerations for High-Frequency Amplifiers Using Drift Transistors," J. W. Englund and A. L. Kestenbaum, *I.R.E. Convention Record, Part 3, Broadcast and Television Receivers and Electron Devices* 1957

"Clusters with Replaceable Pole Caps," E. Stanko, *Inter. Project.* (June) 1957

"Colour TV on Tape," H. R. L. Lamont, *Wireless World* (April) 1957

"Concentric-Shear-Mode 455-Kilocycle Electromechanical Filter," R. W. George, *RCA Review* (June) 1957

"Design Considerations in the First Stage of Transistor Receivers," L. A. Freedman, *RCA Review* (June) 1957

"Design Considerations in the First Stage of Transistor Receivers," L. A. Freedman, *I.R.E. Convention Record, Part 3, Broadcast and Television Receivers and Electron Devices* 1957

"Determination of Transistor Performance Characteristics at VHF," G. E. Theriauet and H. M. Wasson, *Trans. I.R.E. PGBTR* (June) 1957

"Dilemma of Engineers in Management," A. N. Curtiss, *Trans. I.R.E. PGEM* (June) 1957

"Dissociation Pressure and Cohesive Energy of Indium Phosphide," K. Weiser, *Jour. Phys. Chem.* (May) 1957

"The Dissolution of Germanium by Molten Indium," B. Goldstein, *RCA Review* (June) 1957

"A Dynamic Standard Signal for Black-and-White and Color Television Systems," *I.R.E. Convention Record, Part 7, Audio and Broadcast Transmission Systems* 1957

"Flattening Response of Crystal Pickups," A. L. Cleland, *Electronics* (May) 1957

† Report all corrections or additions to *RCA Review*, RCA Laboratories, Princeton, N. J.

"Frequency Conversion at VHF and a New Converter Design," D. J. Carlson, Thesis for Master's Degree, University of Pennsylvania (May) 1957

"Heater Surge Chart," M. P. Feyerherm, *Electronics* (June) 1957

"A High Input Impedance Transistor Circuit," P. J. Anzalone, *Electronic Design* (June 1) 1957

"Home Measurement of Phonograph System Performance," W. H. Erikson, *I.R.E. Convention Record, Part 7, Audio and Broadcast Transmission Systems* 1957

"How To Get Top Performance From The TK-21 Vidicon Film Chain," T. J. Shipferling and H. N. Kozanowski, *Broadcast News* (April) 1957

"Influence of Surface Oxidation on Alpha$_n$ of Germanium P-N-P Transistors," J. T. Wallmark, *RCA Review* (June) 1957

"The Influence of Threshold Action on the RMS Value of Input Gaussian Noise," A. H. Benner (coauthor), *Proc. I.R.E.* (May) (Letter to the Editor) 1957

"Installing Antenna Systems for AM Operations," J. Novik, *Broadcast News* (June) 1957

"Internal Friction and Defect Interaction in Germanium: Experimental," J. O. Kessler, *Phys. Rev.* (May 15) 1957

"Internal Friction and Defect Interaction in Germanium: Theoretical," J. O. Kessler, *Phys. Rev.* (May 15) 1957

"Low Noise Transistor Microphone Amplifier," J. J. Davidson, *I.R.E. Convention Record, Part 7, Audio and Broadcast Transmission Systems* 1957

"Minimizing the Effects of Ambient Light on Image Reproduction," G. L. Beers, *Jour. S.M.P.T.E.* (June) 1957

"A New Semiconductor Photocell Using Lateral Photoeffect," J. T. Wallmark, *Proc. I.R.E.* (April) 1957

"New Stabilizing Amplifier for Color and Monochrome," C. P. Corey, *Broadcast News* (June) 1957

"A New 2 x 2 Slide Projector For Television," A. E. Jackson, R. D. Houck, and R. F. Roundy, *Broadcast News* (April) 1957

"On the Nonlinear Behavior of Electron-Beam Devices," F. Paschke, *RCA Review* (June) 1957

"Optimizing the Dynamic Parameters of a Track-While-Scan System," J. Sklansky, *RCA Review* (June) 1957

"Opto-Electronic Properties of Mercuric Iodide," R. H. Bube, *Phys. Rev.* (May 15) 1957

"Precision Lapping Device," B. Goldstein, *Rev. Sci. Instr.* (April) (Notes) 1957

"Professional Projectors Give Top Picture Quality," R. F. Roundy, *Broadcast News* (June) 1957

"A Proposed Reference Signal for Broadcast Television Transmissions," J. W. Wentworth, *I.R.E. Convention Record, Part 7, Audio and Broadcast Transmission Systems* 1957

"Pulse-Firing Time and Recovery Time of the 2D21 Thyratron," J. A. Olmstead and M. Roth, *RCA Review* (June) 1957

"Radio Free Europe's Broadcast Operation," P. A. Greenmeyer, *Broadcast News* (June) 1957

"The RCA Ampliphase 50-KW AM Transmitter," J. Q. Lawson, *Broadcast News* (June) 1957

"An RCA High-Performance Tape Transport Equipment," S. Baybick and R. E. Montijo, *I.R.E. Convention Record, Part 4, Automatic Control, Electronic Computers and Medical Electronics* 1957

"Reduction of Snivets Interference in Television Receivers," *RCA Application Note AN-170*, Electron Tube Division, Radio Corporation of America, Harrison, N. J. (June) 1957

"Scientific Education . . . the New Challenge," D. Sarnoff, *Signal* (May) 1957

"Signal Mutilation and Error Prevention on Short-Wave Radio-Teleprinter Services," J. B. Moore, *I.R.E. Convention Record, Part 8, Aeronautical and Navigational Electronics, Communications Systems, Military Electronics and Vehicular Communications* 1957

"Simple Geometrically Tapered Ladder Matching Networks Exhibiting Reciprocal Symmetry," H. M. Wasson, Thesis for Master's Degree, University of Pennsylvania (May) 1957

"Simultaneous Color-TV Test Signal," R. C. Kennedy, *Electronics* (May) 1957

"A Six-Transistor Portable Receiver Employing a Complementary Symmetry Output Stage," D. D. Holmes, *I.R.E. Convention Record, Part 3, Broadcast and Television Receivers and Electron Devices* 1957

"Spectrochemical Determination of Ferric Oxide and Silicon Dioxide in Alumina," Doris Conrad, *Applied Spectroscopy* (May) 1957

"A Survey of Methods Used to Determine Contact Potentials in Receiving Tubes," E. R. Schrader, *RCA Review* (June) 1957

"Transistorized RC Phase-Shift Power Oscillator," L. J. Giacoletto, *Trans. I.R.E. PGA* (May-June) 1957

"A Transistorized Seven-Position Portable Mixer," K. Singer, *Jour. S.M.P.T.E.* (June) 1957

"Traveling Wave Antenna," M. S. Siukola and G. A. Kumpf, *Broadcast News* (April) 1957

"Tri-Partition of a Sound Stage," M. Rettinger (coauthor), *Jour. S.M.P.T.E.* (May) 1957

"Vapor Pressure Data for the More Common Elements," R. E. Honig, *RCA Review* (June) 1957

AUTHORS

J. ALMOND received the B.Sc. degree in Engineering Physics from the University of Saskatchewan in 1952 and the M.Sc. degree in Electrical Engineering at Imperial College, London, in 1954. From 1954 to 1956 he was employed by the Canadian Defence Research Board at the Radio Physics Laboratory in Ottawa. He joined the research staff of the Electronics Section of the RCA Victor Research Laboratories in Montreal in July 1956. Mr. Almond is a graduate member of the Institution of Electrical Engineers.

HUNTER C. GOODRICH received the B.A. degree in Physics from Wayne University in 1942. From 1942 to 1945 he was a civilian engineer with the Signal Corps Engineering Laboratory in Detroit working on radio interference measuring equipment. In 1945 he joined the Radio Corporation of America where he has worked on advanced development problems in automobile radio, color television, and transistorized television circuits. Mr. Goodrich is a member of Sigma Pi Sigma and a Senior Member of the Institute of Radio Engineers.

MARSHALL C. KIDD received the B.Ch.E. Degree from Ohio State University in 1944. From 1944 to 1945 he was associated with Bakelite Corporation in Bound Brook, New Jersey. In 1945 he joined the Allen B. DuMont Laboratories in Passaic, N. J., where he worked with cathode-ray tubes. In 1946 he returned to Ohio State University and received the B.E.E. degree in 1948. He then joined the Advanced Development Section of the RCA Victor Television Division, where he worked on loud speakers and television. His present work is with transistor circuits. Mr. Kidd is a Senior Member of the I.R.E.

HERBERT KROEMER received the degree of a Dr. rer. nat. from the University of Goettingen, Germany, in 1952. From 1952 to 1954 he worked in the Semiconductor Research Laboratory of the German Post Office. Since 1954 he has been with RCA Laboratories in Princeton, N. J. Dr. Kroemer is a member of the American Physical Society and the Institute of Radio Engineers, and Sigma Xi.

R. M. MATHESON received the A.B. degree from Cornell University in 1938 and the M.A. degree the following year. He attended graduate school at Princeton University in 1939 and 1940, holding a research assistantship during the latter year. From 1941 to 1946 he served in the U. S. Army Air Force, being released from active duty as a Captain. Prior to joining RCA in 1947, Mr. Matheson served as secretary to the Panel on Electron Tubes of the Joint Research and Development Board.

R. J. MCINTYRE received the B.Sc. degree in Physics from St. Francis Xavier University in Nova Scotia in 1950, and was an instructor in physics there in 1950-51. He received the M.Sc. degree in Physics from Dalhousie University in 1953 and the Ph.D. degree from the University of Virginia in 1956. Since then he has been a member of research staff of the Solid-State Section of the RCA Victor Research Laboratories in Montreal. Dr. McIntyre is a member of the Canadian Association of Physicists.

JOSEPH C. MOOR received the B.S. degree in Electrical Engineering from Texas Technological College in 1950. After teaching high-school science for two years, he joined the Electron Tube Division of RCA as a specialized engineering trainee in 1952. On completion of his training period, he was assigned to his present position in the Photo and Image Tube Design activity at Lancaster, Pa.

LEON S. NERGAARD received the B.S. degree in Electrical Engineering from the University of Minnesota in 1927, the M.S. degree from Union College in 1930, and the Ph.D. degree from the University of Minnesota in 1935. From 1927 to 1930, he was in the research laboratory and vacuum-tube engineering department of the General Electric Company; from 1930 to 1933 a teaching assistant in the department of physics at the University of Minnesota; from 1933 to 1942 in the research and development laboratory of the RCA Manufacturing Company; and since 1942 at the RCA Laboratories Division in Princeton, N. J. Dr. Nergaard is a Member of Sigma Xi and the American Association for the Advancement of Science, and a Fellow of the Institute of Radio Engineers and the American Physical Society.

ROGER G. CLDEN received his education in Austria where he graduated as a Mechanical Engineer from the Technical University of Vienna. He then worked for ten years in the mechanical engineering department of the city government of Vienna until the Nazi occupation of Austria forced him to leave. He joined RCA Laboratories in 1942 and has since worked in the fields of facsimile and electrophotography.

RALPH H. PLUMLEE received the B.S. degree in Chemistry from the University of Illinois in 1938, and the Ph.D. degree in Physical Chemistry from the Ohio State University in 1951. From 1938-1942 he was a teaching assistant at Ohio State University. In 1942 he joined RCA Laboratories in phosphor research. In 1944 he worked in the Manhattan Project under the auspices of the Ohio State University Research Foundation. In 1945 he returned to the RCA Laboratories in Princeton where he worked on mass spectrometric and other studies of oxide cathodes. In May 1957, he left to join the Sandia Corporation, Albuquerque, N. M. Dr. Plumlee is a member of the American Chemical Society, the American Physical Society, and Sigma Xi.

E. M. SMITH received the B.S. degree in Electrical Engineering in 1948 and the M.S. degree in Physics in 1950 from Ohio State University. During his graduate work, he was employed at the Research Foundation of Ohio State University in the development of infrared filters. He joined the Electron Tube Division of RCA in Lancaster, Penna. in 1950, and has been active in the design and development of cathode-ray devices such as oscillograph tubes and computer storage tubes of the oscillograph type. His latest assignment in the Oscillograph and Storage Tube Design activity was the design and development of a five-inch commercial display storage tube, the 6866. Mr. Smith is a member of the Institute of Radio Engineers.

RICHARD G. STOUDENHEIMER received the B.A. degree in Physics from the College of Wooster in 1940 and the M.S. degree in Physics from Syracuse University in 1942. He joined the Electron Tube Division of RCA in Lancaster, Penna. in 1942, and has since been engaged in the design and development of phototubes and image tubes. He is presently an Engineering Leader in the Photo and Image Tube Design activity. Mr. Stoudenheimer is a Senior Member of the Institute of Radio Engineers and a member of the American Physical Society.