RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH • ENGINEERING

VOLUME XVI MARCH 1955 NO. 1

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 XVI | MARCH, 1955 | NUMBER 1

CONTENTS

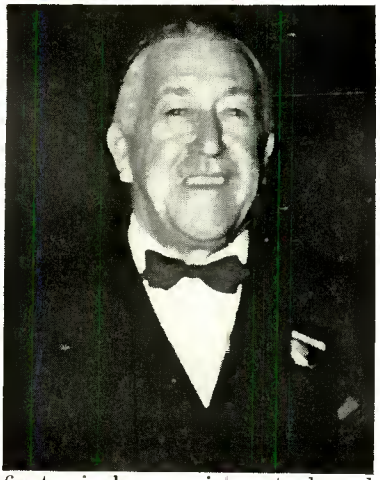
RCA REVIEW is regularly abstracted and indexed by *Industrial Arts Index. Science Abstracts* (I.E.E.-Brit.), *Electronic Engineering Master Index, Chemical Abstracts, Proc. I.R.E., and Wireless Engineer.*

RCA REVIEW

BOARD OF EDITORS

IN MEMORIAM
CHARLES J. PANNILL

It is with deep regret that the publishers of *RCA Review* announce the passing of Mr. Charles J. Pannill on February 7, 1955 in New York City. Mr. Pannill was one of the prime movers in establishing *RCA Review*, and served as the first Chairman of the Board of Editors. He continued to serve in this capacity until April 1942 when publication was suspended for the duration of the War.



Mr. Pannill was truly a pioneer in radio. He enlisted in the U. S. Navy in 1898 and served as a telegrapher during the Spanish-American War. He was associated with Professor Reginald A. Fessenden in his wireless experimentation from 1902-1909. During this period he installed the first wireless equipment aboard a United States battleship.

In 1912, Mr. Pannill received U. S. Radio Operator's License No. 1, issued under the enactment requiring licensing of Radio Operators. In 1914 he re-entered the Naval Service, this time as an expert Radio Aide to Rear Admiral W. H. G. Bullard. In 1917 he became Assistant to the Director of Naval Communications, a post which he held until 1919 when he resigned to return to civilian life.

During the period 1919-1927, Mr. Pannill was Vice President and General Manager, and subsequently President, of the Independent Wireless Telegraph Company. In 1928 he first became associated with RCA as Vice President and General Manager of Radiomarine Corporation of America. He was President of that company from 1935 until his retirement in 1947. He also served as President of RCA Institutes from 1932-1947.

The honors accorded Mr. Pannill are far too numerous to be listed here — suffice it to say that RCA, the Radio Industry, and the Country have lost a good friend and champion.

AN ELECTRON TUBE FOR HIGH-SPEED TELEPRINTING*

By

Warren H. Bliss and John E. Ruedy

Research Laboratory, RCA Laboratories,
Princeton, N. J.

Summary—*A new type of image-selection and display tube is being developed to reproduce letters and numbers in luminous form, which may then be recorded photographically. Selection from the compound beam produced by photoemission is accomplished by magnetic deflection from coded input control. A second deflection system positions the characters into any desired format on a fluorescent screen. A major field of application is in business machines where high-speed operation is required. Speeds of several thousand letters per second are achieved with a format of 4000 well-formed characters on a 5-inch screen. Precision electronic control is required for the deflection and focusing currents.*

Introduction

WITHIN the past few years the development of electronic applications in various fields has created a need for high-speed printing devices which operate directly from data in coded form. Since the speeds required begin to tax the upper limits of electromechanical devices, it appears natural to turn to electronics for a new method.

This paper introduces a new electron-image tube for selecting and displaying characters in a luminous form, which may then be photographically recorded. The operation of this tube simulates typesetting in the steps of selecting letters one-by-one from a "font" and placing them in a line or format. The "font" consists of a chart of characters on a lantern slide which is projected onto a photoemissive cathode. Selections are made from the resulting electron-image stream and directed to the desired position on a phosphor screen. The resulting luminous images may then be projected onto a photographic film as a means of recording or of making a master for printing.

The input control may come from any of the usual sources but preferably in a sequential pulse code. Perforated paper tape or magnetic tape are typical examples. The control can also come directly from a keyboard or from electrical storage. Similar signals received via wire or radio could also be used.

* Decimal Classification: R138.31.

The Image Selection and Display Tube

This new tube is fundamentally an electron-image device and consequently differs in several respects from the more common forms of beam-type cathode-ray tubes such as those used in oscilloscopes and television sets. The essential parts of the tube with its associated magnetic coils are shown in Figure 1. The photoelectrons liberated from the cathode when a letter chart is projected on it are accelerated by a potential of about 100 volts applied to the conductive wall coating. The resulting electron stream is focussed at essentially unity magnification upon the plane of a selecting aperture by a uniform axial magnetic field of about 100 gausses from the first focus coil. The aperture size is made such that one letter only from the array can pass through it at a time. In present models the aperture is round with a diameter of 0.04 inch. A set of magnetic deflection coils located inside the first focus coil deflects the entire electron-image stream for letter-by-letter selection.

The single-letter portion of the electron stream emerging from the aperture enters the positioning and reproducing end of the tube. Here it passes axially through a polished metal cylinder, the potential of which is maintained a few volts negative with respect to the aperture in order to suppress secondary emission. Next it enters the high-potential region where the wall coating is operated at 20 to 30 kilovolts. The letter is finally reproduced on a 5-inch-diameter aluminized phosphor screen. Focusing and positioning on the screen are accomplished by means of a second focus coil and a second set of deflection coils, also shown in Figure 1. The reproduced letters may be displayed in lines and columns as desired. The high-voltage limitation of the tube is set by field emission between the polished electrode and the wall coating. The image definition and brightness are both best at the maximum usable voltage. Present models of the tube are 25 inches long from photocathode to phosphor screen.

Adjustment of the second-focus-coil current and position may be used to vary the size of the letters on the screen from 1.5 to 12 times that at the selecting aperture. When properly adjusted the tube shows no interaction between the letter-selecting and letter-positioning operations. The reproduction and alignment of the letters is good from any location on the photocathode chart to any position of the image on the phosphor screen. It is essential that there be accurate alignment of the coils with respect to the tube and that the plane of the photocathode and of the selecting aperture be simultaneously in focus upon the screen. Optimum performance is obtained by a series of small trial adjustments of electrode potentials and focus-coil currents.

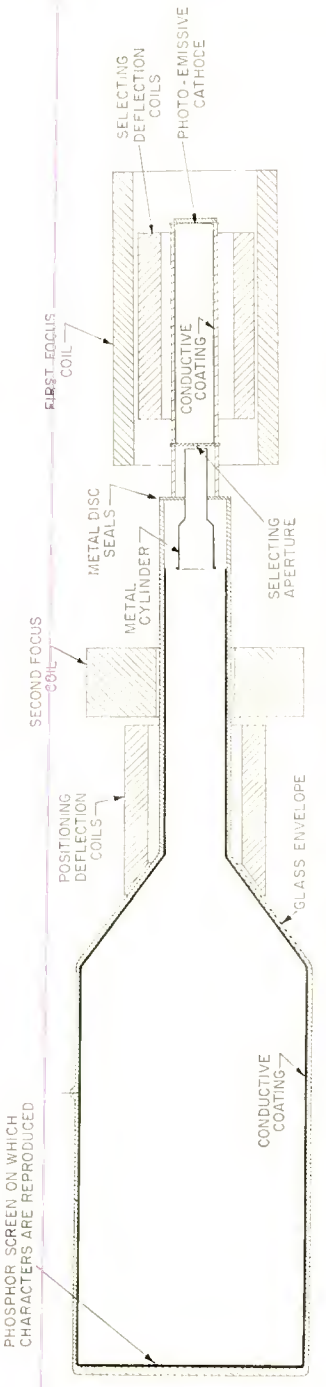


Fig. 1—Sectional view of image selection and display tube with focus and deflection coils.

System Operation

Before discussing performance characteristics and other features of the image selection and display tube, information will be given in connection with an experimental operating system as depicted in Figure 2. In this setup of equipment, which has been extensively used to evaluate the tube performance, standard 5-unit perforated teleprinter tape serves as a source of coded input information. A motor-driven loop of this tape with a repeated test combination of characters is convenient since the observed result appears to be continuously

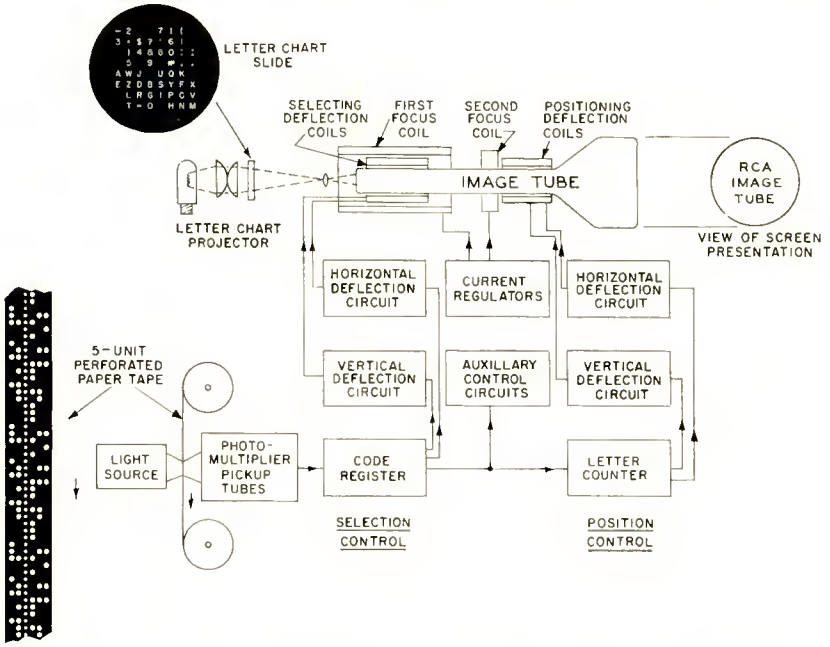


Fig. 2—Experimental operating system for testing image selection and display tubes.

displayed on the screen. To obtain high-speed reading of the perforated tape, the usual mechanical sensing is replaced by a light-beam method. The lateral rows of tape code holes are flooded with light and their images are projected onto a series of secondary-emission photomultiplier tubes. Each code combination then exictes a different combination of phototubes according to the letter to be selected. A set of locking-type multivibrators, actuated by the phototube outputs, acts as a code register to hold each selection until the next one arrives.

The letter chart to be used is projected steadily on the photocathode to produce the compound electron stream from which selections are

made. Since this image-type stream is subjected to twist or rotation as it passes through the focus-coil fields, the letter-chart slide in the projector must be rotated to produce proper character orientation on the screen. The selecting deflection yoke must also be appropriately positioned to make its deflection axes orthogonal to the chart.

Two letter-chart slides mounted in ring gears are depicted in Figure 3. The left-hand one comprises a set of conventional block-type letters, numerals and punctuation marks while the right-hand one consists of upper and lower case letters of Bodoni type face.

The arrangement of the characters in the make-up of the chart follows in this case from the standard teletype code and the use of binary stepped increments of horizontal and vertical deflection cur-

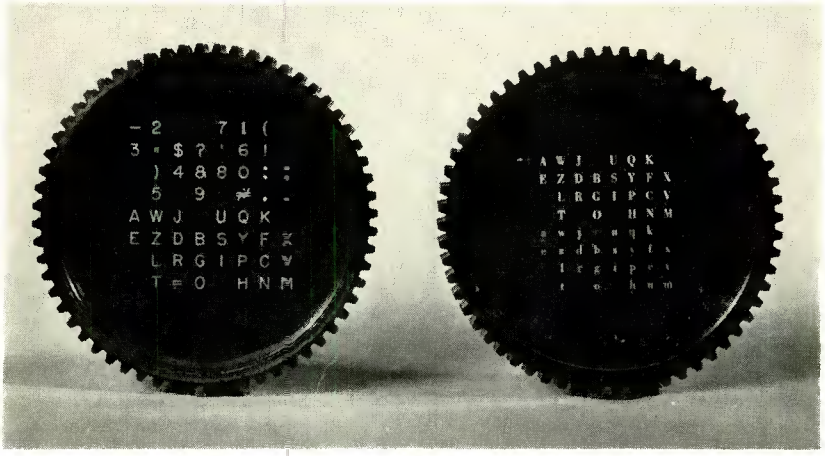


Fig. 3—Photograph of letter-chart slides mounted in ring gears for rotational adjustment.

rents. Each code hole is translated into a vertical or horizontal component of deflection current having a value of 1, 2, or 4 units. The general details of one of the deflection circuits are shown in Figure 4 in order to illustrate how this binary decoding scheme functions.

Deflection- or yoke-coil currents ranging from positive to negative values are supplied in a differential manner from a pair of tandem-connected triode tubes V1 and V2. With no input from the code register, the centering control is set to give a value of deflection current corresponding to zero or no letter selection (lower left-hand corner of chart). In this condition, as well as for all other selected steps, the yoke current is accurately stabilized by the feedback action of the circuit. Any stray variation in this current causes a change in the sensing potential across resistor R1. This changes the feedback ampli-

fier input which drives the grids of supply tubes V1 and V2 to bring the current back to the proper value.

A second and opposing component of the feedback amplifier input comes from a second sensing resistor R2 which is many times larger than R1. The passage of small control currents through R2 results in the production of relatively large currents in the yoke coil since the algebraic sum of sensing potentials must remain constant. The circuit may be regarded as self-balancing since it tends to hold the amplifier input potential at a constant value.

A series of switch tubes V3, V4, and V5 are arranged to draw binary-weighted control currents through resistors 4R, 2R and R, respectively. These resistors are large compared to the tube-plate impedances so that precise steps of current are drawn through R2. As one or more of these switch tubes are turned on by a combination

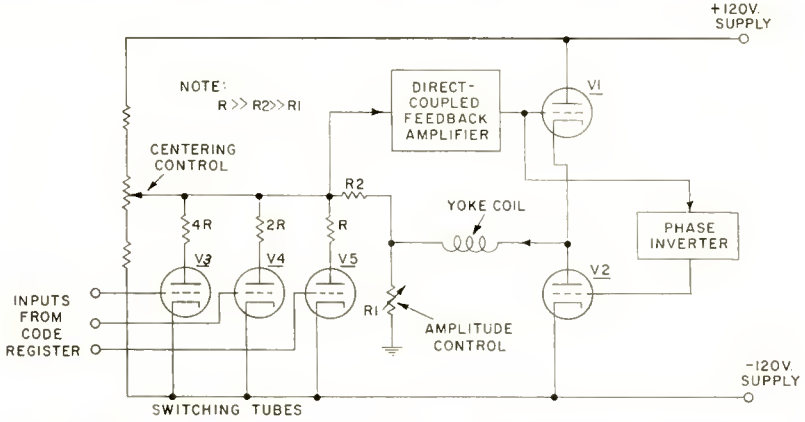


Fig. 4—Precision deflection circuit for producing binary controlled steps of current.

in the code register, the potential across R2 changes to a different value. This causes the feedback loop to stabilize on a different value of yoke current corresponding to the demand. With this triple binary arrangement, eight steps of current can be produced for the eight lines or columns on the chart.

If the selecting deflection current steps are accurately adjusted, each selected letter image in the cross section of the electron stream will pass precisely through the center of the aperture and good character alignment on the screen will be obtained.

Certain other circuit functions pertinent to the operation of the selecting end of the image tube are indicated by a block in Figure 2 designated Auxiliary Control Circuits. One of these is a blanking device which momentarily stops electrons from reaching the screen

during the interval of transfer from one letter to the next. Another circuit portion operates as a special code detector and vertical-deflection-shift means so that selections may be made from the figures (FIGS) section of the chart as well as from the letters (LTRS) section. This utilization of two code combinations for the FIGS–LTRS shift operation allows an increase from the 32 five-unit code combinations to 60 character selections. This is needed for the alphabet plus numerals plus punctuation marks and other symbols.

In the experimental test setup, the positioning of selected characters is controlled by a letter counter. With no positioning deflection current, all letters would be reproduced in a common position at the center of the phosphor screen. The letter counter receives one pulse for each letter selected and controls the horizontal and vertical positioning deflection circuits to put these letters uniformly across the screen in one or more rows. Since the counter is a binary-chain device, it allows use of a binary-step deflection system similar to that just described except that more steps are used.

Performance and Applications

The required quality of reproduced characters for a typical machine printing application is comparable to that of teletype printing. This level of quality lies somewhere above that required for accurate readability. In other words, each reproduced character should have not only sufficient sharpness to be unmistakably identified but must also be easy to read. It has been observed with block-type characters that a limiting condition has been reached in accurate identification when a letter B can just be distinguished from a figure 8. A practical printing system for the above type of application must do much better than this.

The original goal picked for the image tube was to produce 48 lines of 80 characters each on its screen. This represents a square raster since the final output printed copy was to have 10 characters per inch horizontally and 6 lines per inch vertically. On a round phosphor screen the equivalent number of letters across the diameter (corresponds to a diagonal of the square) is $80\sqrt{2} = 113$. For a nominal 5-inch tube having a 4.25-inch useful screen diameter, a requirement of 26.6 letters per inch is indicated.

It is pertinent to try to evaluate this requirement in terms of the television and facsimile terminology for definition. In this case definition is expressed as the maximum number of alternate black plus white lines that may be resolved across the full raster or line length. Various experimenters have observed that the so-called matrix type of printing requires a grid of at least 5 points by 7 points to produce letters and

figures that can be accurately read. Translating this directly into lines of resolution or definition and allowing 20 per cent for the space between characters gives 500 lines for the 80-letter line. Since this is the minimum requirement for accurate readability, the number should probably be doubled. Thus in common parlance, something in the order of at least 1000-line definition is required.

In addition to the requirement for definition there are the matters of speed and format composition. Operation at several thousand characters per second appears to be a practical speed and can readily be attained with an electronic device. In many applications there is a

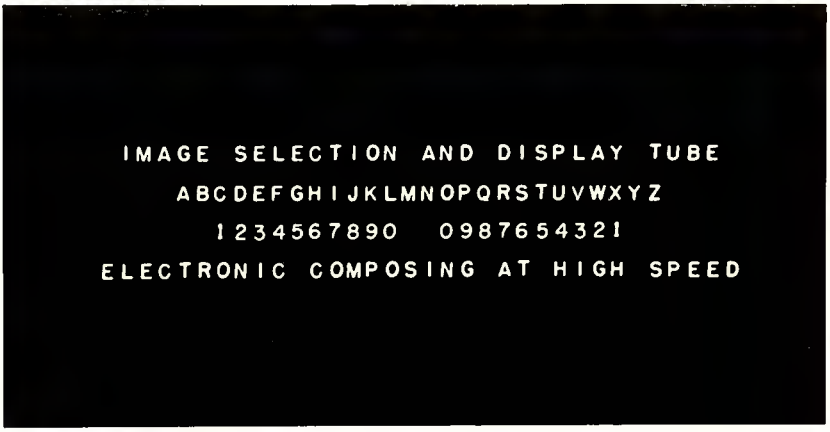


Fig. 5—(Upper) Photograph of image tube screen operating at 1100 letters per second spaced 12.5 per inch on the screen.
(Lower) Screen photographs of reproduced Bodoni type-face letters.

need to control the composition of output characters to fit into spaces on a prearranged form. The new tube offers wide flexibility in this matter since it is not only possible but entirely practical to include positioning information in the input code. For example in producing a financial statement, groups of words and figures must be placed in various positions on the form. A portion of the input coded data may be directed to control the positioning deflection circuit so as to accomplish this result. The tube and the system have been developed to meet these requirements of definition and format composition.

Figures 5 and 6 are photographic reproductions from the 5-inch screen of an experimental tube. The upper part of Figure 5 shows typical operation at 1100 letters per second of suitable size to give 12.5 characters per inch on the tube screen. Figure 6 is a raster of 48 lines of 80 characters each. Readable letters have been observed on a 5-inch screen with more than 35 per inch. Good reproduction with no defocusing occurs over the entire screen.

The present model of the image tube is the result of a series of several experimental tubes. The long spacing from positioning deflec-

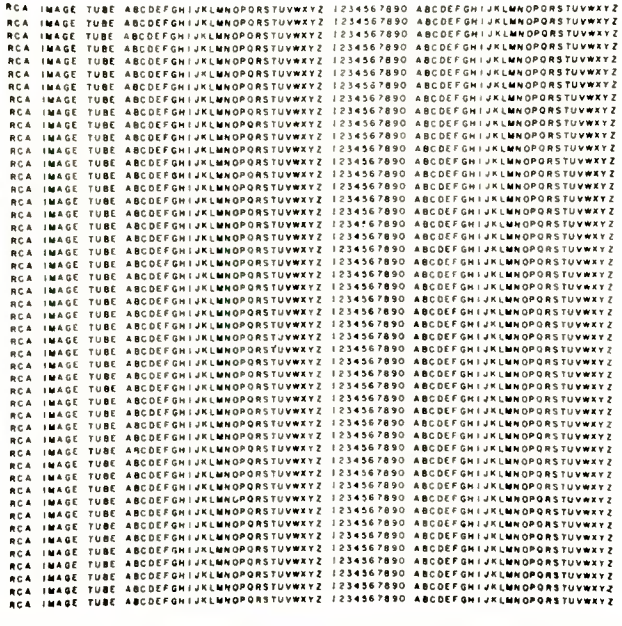


Fig. 6—Photograph of image-tube screen with an 80-letter loop of input tape repeated 48 times to fill the raster. Negative reproduction.

tion yoke to screen permits a low value of deflection angle, which not only lowers the deflection power requirements but decreases the distortion difficulties. Cathode-ray tubes usually experience a certain amount of defocusing at wide deflection angles but this has not been observed in any of the models. However, distortion in the form of letter twist or rotation has been observed in some cases at maximum deflection. This can be corrected readily by circuit means.

Power supplies for the system must have good regulation since variations in applied currents and voltages can affect both focusing and letter alignment. It has been found desirable to use regulators on the focus-coil currents to compensate for temperature changes. As

indicated previously, the deflection circuits require considerable precision and stability.

This discussion would not be complete without some reference to light levels and letter sizes in the system. The present practice is to project the letters onto the photocathode with the over-all chart size reduced to 0.25 by 0.25 inch. This results in individual letter heights of 0.018 inch. The initial end of the tube operates with a slight magnification (about 1.0 to 1.2) so that the letters are readily selected by the 0.04-inch round aperture. Without the aid of the second focus coil there is a magnification up to about 0.25 inch at the screen. For 26.6 letters per inch the focus coil current is increased sufficiently to reduce the letter height to 0.035 inch.

A high level of screen brightness is required for photographic recording at the speeds previously indicated. In one case ample light was available from the screen to expose 35-mm film at 2,000 characters per second. With high sensitivity film, speeds of 10,000 letters per second seem to be within reach before serious screen browning occurs. With photocathode sensitivities of 10 microamperes per lumen or better, these levels of output can be obtained with an input light value of thirty thousand foot-candles. This is provided by a slightly modified standard slide projector with a 300-watt lamp and a short-focal-length lens.

Although the immediate application of this new tube is in the field of high-speed record or message printing, there is a definite possibility of future use in the general printing of such material as books and magazines. A considerable amount of development work by the graphic arts industry has been done in the photographic reproduction of characters for typesetting by optical means. Although the image tube offers a means of more rapid selection than conventional methods and a readily controlled presentation of optical letter images in a given format, the high standard of quality now required in the printing industry is difficult to attain. This offers a challenge for future development.

In this connection some letter chart slides were made up with three commonly used type faces (one of these was mentioned before and is shown in Figure 3). The letters in the lower part of Figure 5 are the result of this type of reproduction. It is apparent that the effects of limiting resolution on the phosphor screen can be overcome by producing the characters in large size and optically reducing them.

Conclusions

The continued development of the image selection and display tube should produce a practical design which will fill the need for a high-

speed printer to operate directly from coded input information. Improvements in screen quality and size give promise of attaining at least 150 well-formed characters across the screen. This is more than ample for most printing where type lines seldom exceed 70 letters in length.

The outstanding features of this new electronic device are presented again in conclusion. As with all electronic devices, high speed of operation is inherent. Handling several thousand characters per second is practical in this case. Of particular interest is the feature that allows rapid change of letter style. The letter chart can be readily and quickly replaced by any one of a series kept in stock. For business machine printing, the matter of format or positioning control is important. The new device is extremely flexible in this respect and permits the composition of material to fit standardized forms. The feature of offering a visual display of information has merit for some applications.

Acknowledgments

The authors wish to acknowledge the advice and encouragement which have been received during the course of this project from C. J. Young and G. A. Morton. Credit is due also to P. G. Herkart for contributions to the tube design.

and (2) secondary emission of electrons by ions. In the gas, a few electrons are normally present, perhaps as a result of cosmic rays or even room-temperature thermionic emission from the cathode. When these electrons are accelerated by the applied voltage, some of them acquire enough energy to ionize gas atoms. The positive ions formed in this manner are attracted to the cathode, where they release electrons. These electrons, in turn, produce more ions. If, on the average, more electrons are released than those initially present in the tube, the process is regenerative and breakdown ensues.

Impact ionization can also occur in the depletion layer at the collector of a transistor. In this process, either an electron or hole which has acquired enough energy may produce a hole–electron pair by direct excitation across the "forbidden gap."[1, 2] This phenomenon may produce an avalanche of holes or electrons and cause electrical breakdown of the material. Depending on the type of transistor used (n-p-n or p-n-p) either holes or electrons produced by this impact ionization flow to the collector. The carriers having the opposite sign of charge flow back into the base region and add to the concentration of majority carriers. This additional charge biases the emitter junction in the conducting direction. The emitter, therefore, becomes analagous to the secondary-emitting cathode of the gas tube, and has a "secondary-emission ratio" equal to the base-to-collector current amplification factor, α_{cb}. Because α_{cb} is much greater than unity, it is only necessary that a few per cent of the carriers passing through the depletion layer cause ionization for the process to become highly regenerative. The static characteristics of a transistor operating in the DCC mode can be predicted by the use of ordinary junction-transistor theory provided avalanche multiplication is included. This paper presents an analysis of the DCC effect and discusses some of its practical applications.

Theory of Delayed Collector Conduction

Figure 2 is a schematic diagram of the hole-and-electron flow in a p-n-p transistor operating in the DCC mode. The following analysis is restricted to the steady-state operation of p-n-p junction transistors. However, the analysis also is applicable to n-p-n transistors provided the roles of holes and electrons are interchanged. The dashed lines in Figure 2 indicate hole-flow paths, and the solid lines electron-flow

[1] McKay and McAffee, "Electron Multiplication in Silicon and Germanium," *Phys. Rev.*, Vol. 91, p. 1079, 1953.

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

paths. The arrows indicate the direction of particle flow. The magnitudes of all currents are considered to be positive to simplify the equations.

Hole Flow

The emitter current, I_E, consists of two components: I_p, the current due to holes injected into the base, and I_e, the current due to electrons moving from the base into the emitter. These two components are related through the emitter efficiency, γ

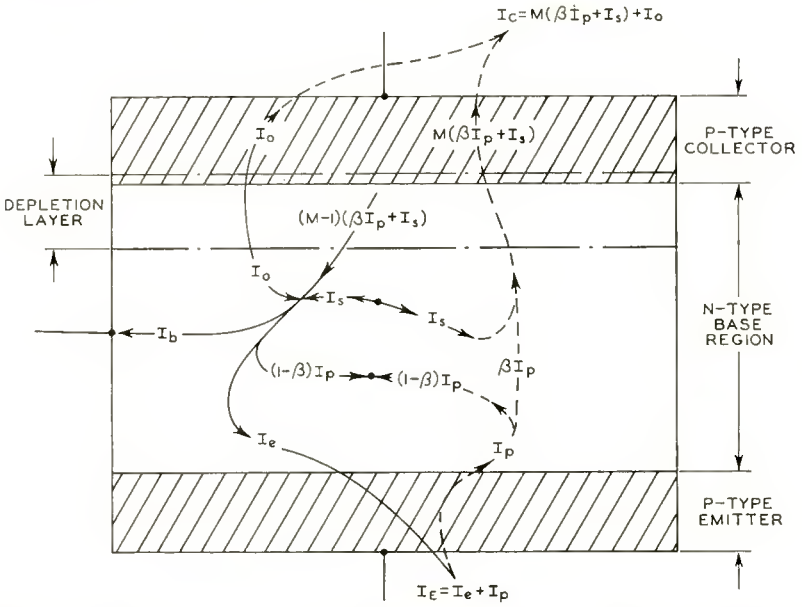


Fig. 2—Schematic representation of hole flow (dashed line) and electron flow (solid line) in a junction transistor.

$$\gamma = \frac{1}{1 + I_e/I_p} \approx 1 - \frac{I_e}{I_p}. \qquad (1)$$

since $I_p \gg I_e$ in practice.

The fraction of the injected holes which survive the trip across the base region is represented by β, as in conventional junction-transistor theory. The fraction of the hole current lost through recombination is equal to $(1 - \beta)$. In addition to the current flow to the collector due to injected holes, there is a component called the saturation current, I_s, which is due to thermal generation of hole–electron pairs. Although Figure 2 shows I_s originating in the base region, as

is essentially the case in alloy-junction devices, the analysis applies no matter where the thermal pair generation occurs.

The current $(\beta I_p + I_s)$ may be referred to as the "primary" collector current. In high-field regions such as the depletion layer at the collector junction, internal ionization multiplies this primary current by a factor, M, which depends on the field and the thickness of the depletion layer.[1, 2] Ebers and Miller[3] have shown that in alloy junctions this factor M may be approximately written as

$$M = \frac{1}{1 - (V/V_b)^n}. \qquad (2)$$

where V is the collector–base voltage, V_b is the ultimate avalanche breakdown voltage of the collector junction (which depends on the impurity density in the base material),* and n is a constant equal to about 3 for p-n-p transistors and ranging from 4.6 to 6.6 for n-p-n transistors. (This difference must be remembered if the analysis is applied to n-p-n types.)

In addition to the internal flow of holes and electrons across the collector junction, there may be a leakage current, I_0, (due, for example, to surface conduction) which is not multiplied by M. Even when I_0 does not exist within the transistor, it may sometimes be desirable to provide such a current by external circuitry. This leakage current will be included to make the analysis as general as possible.

Electron Flow

There are three sources of electrons in the transistor: (1) the leakage current corresponding to electron flow into the base lead; (2) an electron flow resulting from thermal generation (which also has the magnitude I_s because electrons and holes are produced in pairs); and (3) the multiplication mechanism in the junction. If continuity of current is to be maintained in the depletion layer, an electron current equal to $(M-1)$ times the primary current must flow back into the base region. There are only three ways in which electrons may leave the base region: through the base lead (I_b), through the emitter lead (I_e), or by recombination. The recombination rate for

[3] J. Ebers and S. Miller, "Avalanche Breakdown in Germanium Transistors," 1954 Conference on Semiconductor Devices Research.

* V_b is not always the observed breakdown voltage of the collector–base diode. Breakdown mechanisms other than the avalanche type exist and may occur at a lower voltage (e.g., Zener breakdown and, more often, highly nonlinear surface conduction).

electrons, however, is necessarily equal to that for holes, $(1-\beta)\ I_p$ in the steady state. The total flow of electrons into the base, therefore, may be equated to the flow out of the base:

$$I_o + I_s + (M-1)\ (\beta I_p + I_s) = I_b + (1-\beta)\ I_p + I_c\ . \tag{3}$$

As shown in Equation (1),

$$I_c = I_p\ (1-\gamma).$$

Then, by collection and rearrangement of terms,

$$I_p = \frac{I_o + MI_s - I_b}{(1-\beta) + (1-\gamma) - \beta\ (M-1)}\ . \tag{4}$$

The sum of the first two terms in the denominator, $(1-\beta) + (1-\gamma)$ is, in the first approximation, equal to $1/\alpha_{cb}$, when α_{cb} is the base-to-collector current amplification factor of the transistor. The factor α_{cb} varies in practice[4] with I_p and also, usually to a smaller extent, with the collector voltage.[5] For the moment, however, it will be considered a constant. Equation (4) may be further simplified by setting β equal to unity in the term $\beta(M-1)$. This simplification is permissible because β should be at least 0.97 in a practical transistor and because the equation is relatively insensitive to the accuracy of this particular approximation. Equation (4) then becomes

$$I_p = \frac{I_o + MI_s - I_b}{1/\alpha_{cb} - (M-1)}\ . \tag{5}$$

The factor M increases with collector voltage from a value of 1 at low voltages to infinity at the breakdown voltage of the collector junction. When the value of M is equal to $(1 + 1/\alpha_{cb})$, I_p becomes infinite, provided the numerator of Equation (5) is not zero. Because the value of α_{cb} may range from ten to several hundred, M need be only slightly greater than unity for I_p to become infinite.

When the numerator of Equation (5) is negative (i.e., when $I_b > I_o + MI_s$), I_p is essentially zero because there is no physical

[4] W. M. Webster, "On the Variation of Junction-Transistor Current-Amplification Factor with Emitter Current," *Proc. I.R.E.*, Vol. 42, p. 914, June, 1954.

[5] J. M. Early, "Effects of Space-Charge Layer Widening in Junction Transistors," *Proc. I.R.E.*, Vol. 40, p. 1401, November, 1952.

mechanism by means of which holes can flow in large numbers from the base to the emitter. In other words, because a negative I_p denotes a reverse bias on the emitter junction, the value of I_p is essentially zero.

When the base current is zero, the collector current for small values of collector voltage is approximately α_{cb} times the current which would flow between collector and base with the emitter current equal to zero, I_{co}. As the collector voltage increases, the value of the factor M increases, the denominator of Equation (5) diminishes to zero, and I_p increases to infinity. The collector current is equal to $M(\beta I_p + I_s) + I_o$, as shown in Figure 2. However, because I_o and I_s are small compared to I_p in the region of breakdown, and M and β are approximately equal to unity, the collector current is approximately equal to I_p. The increase of collector current with M is slightly enhanced by the fact that M appears in the numerator. However, this increase is unimportant compared to the effect of the term M—1 in the denominator. The curve for $I_b = 0$ shown in Figure 1, therefore, is in qualitative agreement with Equations (5) and (2). In addition, the curves in region II may be explained by means of the same equations.

Equation (5) may be used to predict the "onset" voltage at which the collector current suddenly increases. A voltage V_0 represents the point at which $M-1 = 1/\alpha_{cb}$ (the point at which I_c becomes infinite for $I_b = 0$, as shown in Figure 1). For voltages greater than V_0, the denominator of Equation (5) is, in a sense, already zero. If the numerator is zero or negative, I_c is finite and has the value $MI_s + I_o$. This value is equal to I_{co}, the current which flows between collector and base when the emitter current is zero. When the numerator becomes positive, however, I_c increases abruptly. This change occurs when

$$I_b \leqq MI_s + I_o = I_{co}\,. \tag{6}$$

The shape of the collector voltage–current characteristic in the negative resistance region may be calculated by solution of Equation (5) for M and combination of the result with Equation (2). Although Equation (2) applies to collector–base voltage, the collector–emitter voltage, V_c is only negligibly different. Thus,

$$\frac{V_c}{V_b} = \left[\frac{I_p/\alpha_{cb} - I_o - I_s + I_b}{I_p + I_p/\alpha_{cb} - I_o + I_b}\right]^{1/n}. \tag{7}$$

For the high values of I_p found in the breakdown region, I_p dominates the denominator of Equation (7) and may be set equal to I_c as indicated above. This simplifies the relation to yield:

$$\frac{V_c}{V_b} \approx \left[\frac{1}{\alpha_{cb}} - \frac{I_o + I_s - I_b}{I_c} \right]^{1/n} \tag{8}$$

Equation (8) indicates that V_c decreases as I_c increases, and that a negative resistance appears at voltages in excess of V_0, provided other terms in the equation are held constant. In practice, α_{cb} decreases as I_c increases, and I_o increases with V_c. Both effects tend to decrease the magnitude of the negative resistance, and the variation of α_{cb} may make the slope positive if it is sufficiently strong. The analysis of the DCC effect described in this section has been verified by experiments with good agreement.

The factors which determine breakdown time are not at all clear. Although emitter–collector spacing undoubtedly should be small for good performance, the usual simplifying assumptions of space-charge neutrality and pure diffusion flow in the base are invalid. It is probable that a substantial field exists in the base during breakdown as a result of the majority carriers produced in the collector junction. Such a field encourages flow of minority carriers injected at the emitter and reduces their transit time below the "steady-state" or diffusion value. This effect accounts for the fact that the rise time of multivibrators operating in the DCC mode is much shorter than that predicted from small-signal parameters. Current decay, however, is probably limited by diffusion of minority carriers stored in the base. This effect results in a recovery time which is proportional to the square of the emitter–collector spacing and longer than the breakdown time.

It can be seen, therefore, that DCC operation is similar to thyratron behavior. The input impedance is high, output impedance low or negative, and firing time shorter than recovery time. The important distinction is that the control electrode (base) of the transistor retains at least some control after breakdown, while the thyratron grid loses control completely.

Circuit Applications of DCC Mode

The DCC mode can be used in many circuit applications. The variation of I_{co} and α_{cb} with V_c and I_c produces a variety of collector characteristics in different transistors, as shown in Figure 3. The circuits described in this section of the paper include an amplifier, an impedance switch, a voltage regulator, and circuits requiring a negative resistance characteristic, such as multivibrators and pulse amplifiers. The negative resistance may also be used in sine-wave oscillators.

1. Amplification in the DCC Mode

If a load line is drawn in the DCC region of Figure 3(a), an amplifier having low output impedance can be obtained. A number of transistors having positive impedance characteristics were used as amplifiers; input impedance was high, output impedance was low, and power gain was about the same as that obtained with the same transistors in normal operation. When a transistor exhibiting a controllable negative resistance was used in a wide-band amplifier, considerable improvement in gain was obtained. Typical curves of voltage gain as a function of frequency for such amplifiers are shown in Fig-

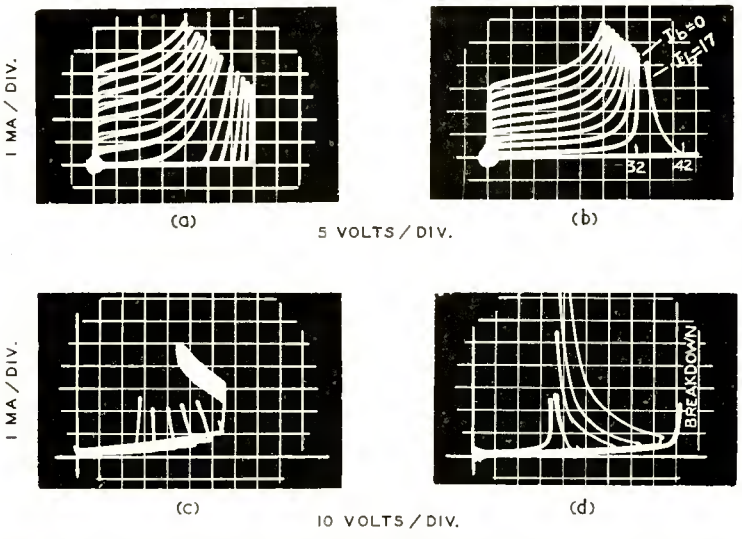


Fig. 3—Some typical collector characteristics which show different impedances in the DCC mode.

ures 4 and 5. In these figures, curve A is for conventional operation, and curve B is for operation in the DCC mode. The increased gain is not a simple regenerative effect because the gain–bandwidth product is also increased. It is apparent that the reduced output impedance of the DCC amplifier improves the frequency response considerably. The output signal voltage under DCC conditions, however, is restricted to smaller values than those obtainable in conventional operation.

2. Voltage Regulator

An extremely simple and efficient controlable voltage regulator can be designed by use of the characteristics of the constant-voltage pre-breakdown region. A photograph of the constant-voltage characteristics for supply voltages of 45, 55, and 65 volts is shown in

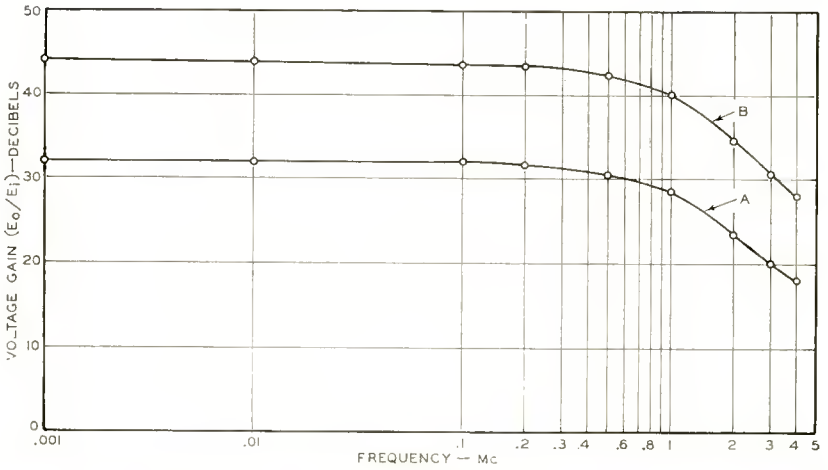


Fig. 4—Frequency response of experimental high-frequency transistor. Curve A—conventional operation, $V_c = 22.5$ volts; Curve B—DCC mode operation, $V_c = 50$ volts. Both curves taken with load resistor of 4,700 ohms.

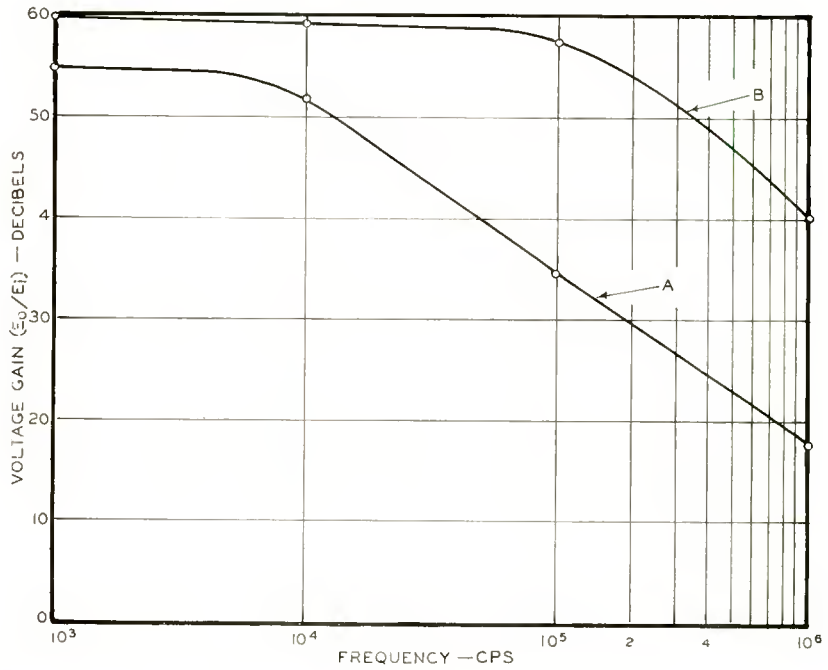


Fig. 5—Frequency response of audio-frequency transistor. Curve A—conventional operation, $V_c = 22.5$ volts; Curve B—DCC mode operation $V_c = 110$ volts. Both curves taken with load resistor of 56,000 ohms.

Figure 6(a). These curves were obtained in the voltage regulator circuit shown in Figure 6(b). It is evident that excellent regulation is obtained. This circuit is unlike other voltage regulators (e.g., gas tubes and silicon diodes) in that the output voltage may be adjusted by variation of the base current. In this way, the collector output voltage has been controlled over a range of more than 20 volts simply by variation of the external base resistance. The voltage regulator will pass currents corresponding to the maximum power rating of the transistor. Voltages of either polarity may be regulated because either p-n-p or n-p-n transistors may be used.

3. Impedance Switch

The conditions necessary for use of the transistor as a transient-free switch are shown in Figure 7(a). The horizontal curves represent grounded-base operation, and the vertical curve the DCC mode. An

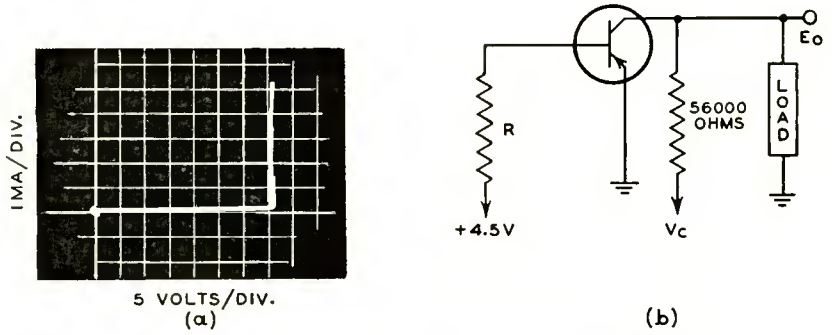


Fig. 6—Voltage-regulator characteristic and circuit.

impedance change greater than 1000 to 1 can be accomplished with no change in d-c operating point by an effective change from DCC to grounded-base operation. Many arrangements can be designed using this phenomenon; Figure 7(b) shows one such circuit. The switching transistor, T_1, operates in the conventional mode with constant base voltage or in the DCC mode with constant reverse base current, depending on whether or not the diode, D_1, is conducting. In the absence of a positive switching pulse, the switching transistor has a high output impedance in shunt with the load R_L. Transistor T_2 then operates at full gain. When a positive pulse is applied to the base of T_1, diode D_1 begins to conduct and the reverse-bias circuit in the base of T_1 assumes control. As a result, T_1 operates in the DCC mode and its collector impedance becomes very low. Because this impedance is in shunt with R_L, the output of T_2 is greatly attenuated. The circuit

can be designed to provide attenuations proportional to the ratio of the collector impedances of T_1 in DCC and conventional operations. The value of R_1 is selected so that the collector current and voltage of T_1 are the same in both modes. Because the switching circuit has no capacitors, it responds to d-c signals, and no transient should be observed provided the bias conditions of the two modes are adjusted correctly.

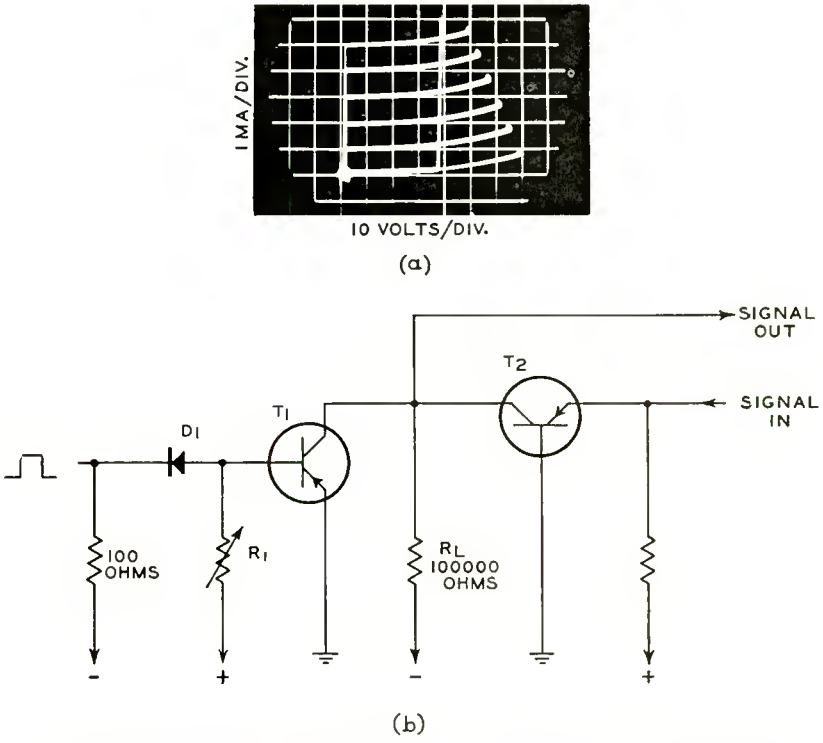


Fig. 7—Collector characteristics and circuit required for DCC and conventional operation for impedance switch.

4. Multivibrators and Similar Circuits

Circuits which work in the negative-resistance region can be analyzed in a manner similar to that used for point-contact transistors[6] because the emitter–collector current gain, α_{ce}, of the junction transistor exceeds unity when the collector impedance becomes negative. In some circuits, the two types of transistor will operate interchangeably except for adjustment of bias.

[6] A. W. Lo, "Transistor Trigger Circuits," *Proc. I.R.E.*, Vol. 40, pp. 1531-1541, November, 1952.

The advantages of junction transistors over point-contact transistors for DCC operation are several. Because the reverse base current of junction units is much smaller than that of point-contact units for the same applied collector voltage, higher voltage operation for a given dissipation is allowed. Also, higher gains are obtainable with junction transistors in pulse amplifiers. In addition, a greater range of circuit operations is possible with junction transistors because they have essentially all the characteristics of both types.

The operation of pulse circuits generally includes the transition between two ore more regions, particularly including region V of Figure 1, the negative-resistance region. The various multivibrator

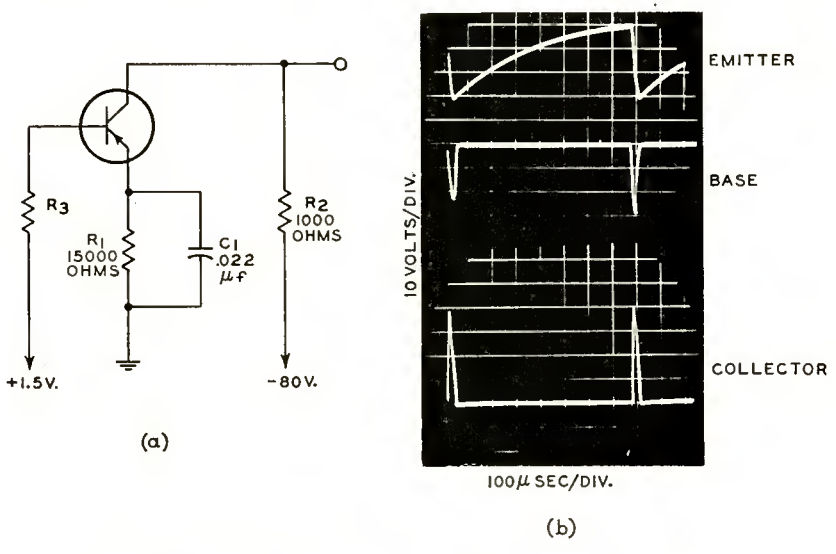


Fig. 8—Circuit and waveforms for relaxation oscillator.

and flip-flop circuits are designed for operation in regions III, IV, or V, depending on the type of operation required. Because region IV, the constant-voltage region, has no exact counterpart in point-contact-transistor operation, new circuits can be designed utilizing this characteristic.

(a) *Astable Circuits*—A relaxation oscillator of the "astable" or "free-running" type is shown in Figure 8(a). If the operating point of the collector is in the unstable region V, it will move to either region III or IV, depending on the previous condition of the circuit. If C_1 is initially charged negatively and the transistor is cut off (region III), only leakage current flows. As C_1 discharges through R_1 and the high reverse emitter impedance, the voltage between the

emitter and base finally reaches a point at which high conduction occurs in the negative-resistance region V. The collector voltage immediately drops (in less than 0.1 microsecond) to region IV, the constant-voltage region (positive resistance). At this point, the emitter starts to conduct, and current flows through the emitter–collector circuit. C_1 charges as the emitter voltage increases negatively. The charging time shown in Figure 8(b) is in the order of 20 microseconds. When C_1 is charged, the emitter becomes more negative than the base, and the transistor is triggered off. The collector voltage then rises to the supply voltage, and the cycle repeats. The waveforms of this relaxation oscillator are shown in Figure 8(b).

Basically, this circuit provides a means for the charging and dis-

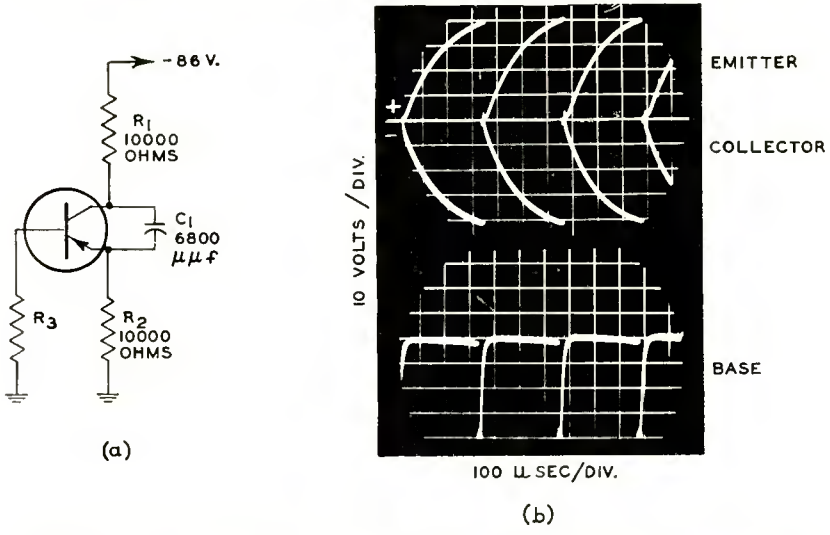


Fig. 9—Circuit and waveforms for push-pull sawtooth generator.

charging of a capacitor through different resistances which utilizes the negative resistance of the transistor. Many other circuit arrangements are possible. Sine-wave stabilization can be added to this relaxation oscillator to increase the frequency stability. Stabilization can be obtained by the use of suitable frequency-determining circuits in the collector, emitter, or base lead.

The circuit shown in Figure 9 provides a 40-volt, push-pull, sawtooth output from a single transistor. Because essentially all the current can be made to flow in the emitter and collector resistors, the waveforms are equivalent if the resistors are equal. This circuit can be operated "free-running," or it can be triggered at the base electrode if desired. If different ratios of positive and negative pulses are

desired, R_1 and R_2 can be varied accordingly.

(b) *Monostable Circuits*—Astable circuits of the type described above can be biased in the "off" or "nonconducting" state and become monostable circuits. Under these conditions, very little collector current flows until the base or emitter voltages are changed to the forward direction, or the collector voltage is increased by means of an applied pulse. When any of these changes occur, the circuit becomes violently

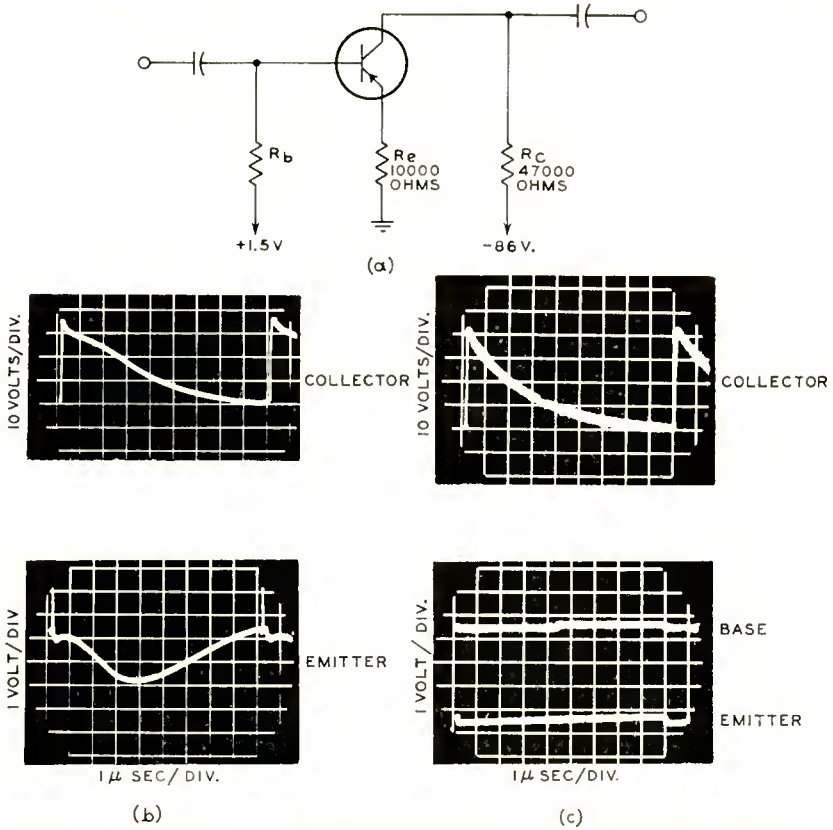


Fig. 10—(a) Circuit for regenerative amplifier, (b) Waveforms with sine-wave trigger, (c) Wave forms with square-wave trigger.

regenerative and experiences the same action that occurs in the relaxation oscillator. The only difference between the two circuits is that the cycle is not repetitive in the monostable circuit because the conditions for regeneration are present only during the pulse. A useful monostable circuit in the form of a regenerative amplifier is shown in Figure 10(a). The circuit can utilize fixed bias or self bias, and can be triggered in the emitter, base or collector circuits. The largest gain or highest sensitivity is obtained by the use of a negative pulse

for triggering the base (for p-n-p types). Figures 10(b) and 10(c) are oscillograms of the regenerative amplifier waveforms obtained with sinusoidal and square-wave inputs, respectively. The switching time is very short, the pulse duration being of the order of 0.1 microsecond. When the pulse was applied directly to the deflection plates of an oscilloscope, a rise time of 0.02 microsecond was observed. The regenerative amplifier provides voltage gains up to about 600 with base triggering. These values are believed to be higher than those obtained with point-contact regenerative amplifiers.

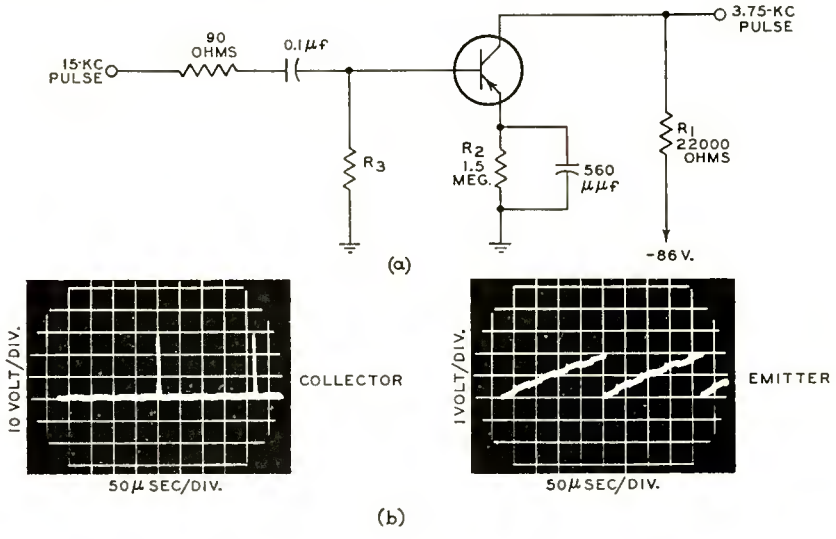


Fig. 11—Circuit and waveforms for a frequency divider.

The regenerative amplifier may also be used as a frequency divider, as shown in Figure 11. The waveforms in Figure 11 show the collector and emitter characteristics for divider operation. A 30-volt pulse is shown at the collector. The divider in this application is counting down from 15 to 3.75 kilocycles.

(c) *Single-Shot Multivibrator*—A single-shot multivibrator can be designed by biasing the transistor in the stable cutoff region III. This device is similar to the regenerative amplifier in that it becomes momentarily regenerative when the operating point shifts across the negative resistance. Figures 12(a) and 12(b) show the circuit and the waveforms observed. The rise time of the collector voltage is extremely short (less than 0.1 microsecond), but the fall time is considerably longer (in the order of 1 or 2 microseconds) for the circuit shown in Figure 12(a). The delay time, T_d, is approximately equal to

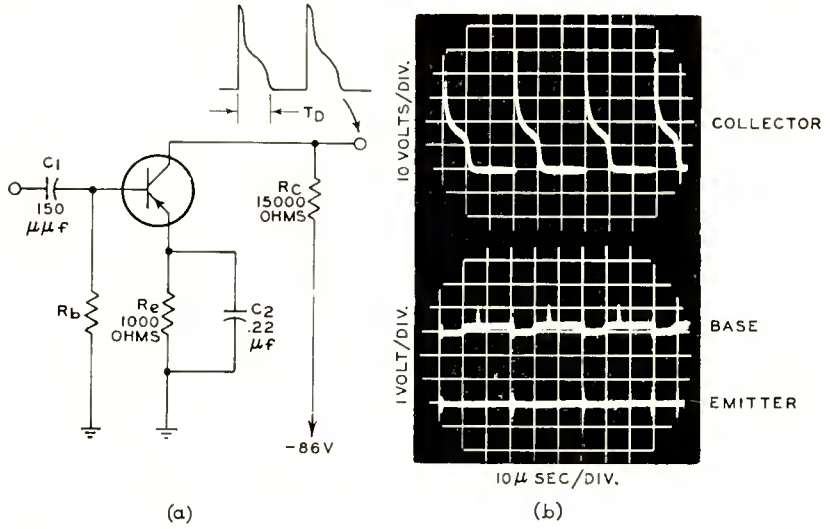


Fig. 12—Circuit and waveforms for a single-shot multivibrator.

$\frac{C_1 R_b}{R_e R_c}$. Capacitor C_2 is used to bypass R_e and does not affect the delay time appreciably.

(d) *Bistable Circuits*—In a bistable circuit such as that shown in Figure 13(a), the load line must intersect the collector-characteristic

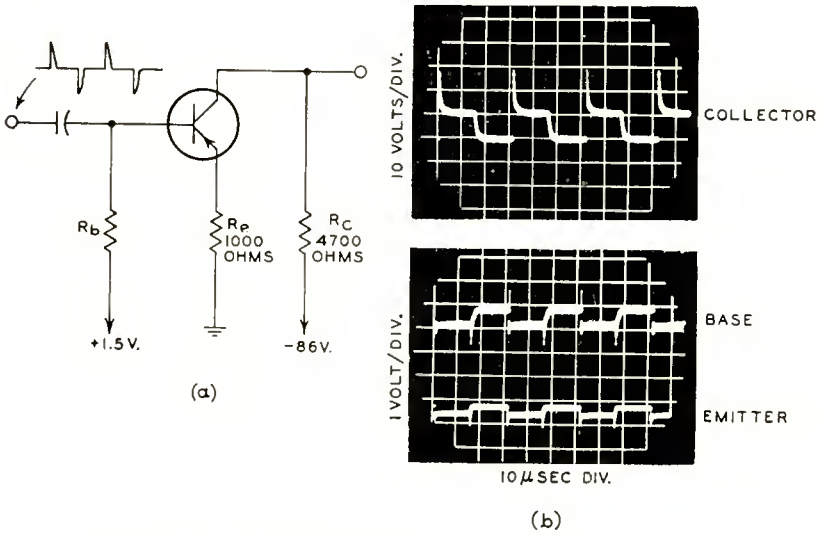


Fig. 13—Circuit and waveforms for a bistable circuit.

curve in three regions. The high-current and low-current conditions for this circuit are stable operating points in region III and IV, and the intermediate intersection is an unstable point in region V. As the transistor is pulsed from one stable state into the negative-resistance region, the regeneration of the circuit causes it to "flip" or "flop" to the other condition. Waveforms of a single junction-transistor flip-flop circuit are shown in Figure 13(b). A positive pulse on the base triggers the transistor off, and a negative pulse triggers it on. When flip-flop operation has been established, a binary counter can be made by the use of "pulse steering" (i.e., circuitry in which pulses of the same polarity cause the transistor to go from one state to the other). Pulse steering can be accomplished with diodes in conventional circuits.

ACKNOWLEDGMENT

The authors wish to acknowledge the interest, encouragement, and assistance to C. M. Sinnett, G. L. Grundmann, F. B. Stone, and E. W. Herold.

COMPARATIVE HIGH-FREQUENCY OPERATION OF JUNCTION TRANSISTORS MADE OF DIFFERENT SEMICONDUCTOR MATERIALS*

By

L. J. Giacoletto

Research Laboratory, RCA Laboratories,
Princeton, N. J.

Summary—*The high-frequency performance of junction transistors is determined in part by the time required for the injected minority carriers to traverse the base region. Large minority carrier mobility contributes to reducing this transit time. However, the high-frequency performance is also affected by factors such as base-lead resistance and collector-to-base junction capacitance, and it is shown that large majority carrier mobility will reduce the detrimental effects of these factors.*

An important result of the analysis of high-frequency operation is the discovery that both minority and majority carrier mobilities are of about equal importance. It follows that the high-frequency performance of an n-p-n transistor should be about the same as that of a geometrically identical p-n-p transistor. Furthermore, for evaluation of new semiconductors, a knowledge of the values of both mobilities is required, and a figure of merit is proposed which is formed by the product of the two drift mobilities divided by the square root of the dielectric constant.

Introduction

CHARGE carriers in semiconductors are of two types—minority carriers and majority carriers. In an n-type semiconductor, the electrons are the majority carriers and are normally present; the holes are the minority carriers, often injected from one of the boundaries. The reverse is true for p-type semiconductors.

Because, in the conventional junction transistor, it is the minority carriers which flow out of the emitter, through the base, and into the collector, the role of the *majority* carriers is often neglected. This leads to an incorrect concept of the factors which enter into high-frequency behavior. It is true that the high-frequency performance of a junction transistor is affected by the time required for the minority carriers to traverse the base region, and that large minority carrier mobility contributes to the reduction of this transit time, but this is only part of the story. Since charge neutrality must exist, minority carriers flowing into, and out of, the base region are accom-

* Decimal Classification: R282.12.

panied by flow of *majority* carriers also into, and out of, the base region. This produces a high-frequency voltage drop in the base-lead resistance which is determined in part by the *majority* carrier mobility. In addition, *majority* carrier mobility affects the very important collector-to-base junction capacitance.

An important result of this analysis of high-frequency operation is discovery of the fact that the majority and minority carrier mobilities are of nearly equal importance. It follows that, in general, p-n-p type transistors should not be very different from n-p-n type transistors in high-frequency performance. Furthermore, for evaluation of new semiconductors, a knowledge of the values of both majority and minority carrier mobilities is required. A figure of merit for semiconductors is proposed which is formed by the product of the two *drift* mobilities, divided by the square root of the dielectric constant. With the aid of this figure of merit, semiconductor materials that have been considered good because of their large electron Hall mobility are shown to be poor for use in transistor devices.

In this article, after a discussion of the general equation for high-frequency performance, a comparison is drawn, first between transistors of *different polarity* (n-p-n and p-n-p) but made with the same semiconductor material, and second between transistors of the same polarity but made with *different semiconductor materials*. This report covers only the analysis, but the conclusions drawn are in accord with extensive small-signal experimental data on germanium and silicon junction transistors.

High-Frequency Operation

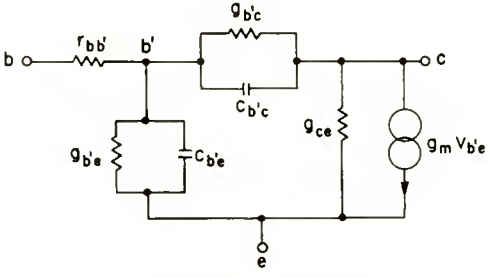
The hybrid-π common-emitter equivalent circuit as shown in Figure 1a will be employed as the basis for the study of high-frequency operation.[1]

At higher operating frequencies the conductances shown in Figure 1a are negligible in comparison with the capacitive susceptances, and the equivalent circuit which is applicable is that shown in Figure 1b. An approximate analysis of this circuit indicates that the maximum power amplification is given by

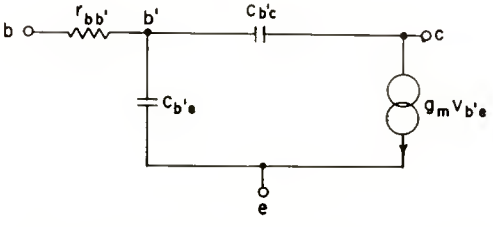
$$\text{P.A.} \approx \frac{g_m}{4\omega^2 r_{bb'} C_{b'c} C_{b'c}}. \qquad (1)$$

[1] L. J. Giacoletto, "The Study and Design of Alloyed-Junction Transistors," 1954 National Convention Record of the I.R.E., Part 3—Electron Devices and Component Parts, pp. 99-103, and id., "Study of P-N-P Alloy Junction Transistor from D-C through Medium Frequencies," *RCA Review*, Vol. 15, pp. 506-562, December, 1954.

With the aid of equations for the transistor parameters and some approximations, this equation can be written in terms of the semiconductor constants and the operating point. Expressions will be written of general validity, but, for clarity, the subscript notation will be that for an n-p-n transistor. In order to obtain the expressions for a p-n-p transistor it is only necessary to interchange n and p subscripts and d and a (donor and acceptor) subscripts wherever they occur. The development throughout this article is for abrupt impurity transitions as found in alloy-type junction transistors. In grown-type



(a) Complete circuit.



(b) Circuit applicable at high frequencies.

Fig. 1—Hybrid-π common-emitter equivalent circuit.

junction transistors the impurity transition is generally gradual. Although the detailed results differ, the basic conclusions are the same for both alloy- and grown-type junction transistors.

For a reasonably good transistor, emitter and collector currents are approximately equal, so that $g_m \approx \Lambda I_E$ where Λ is 39 volts^{-1} at room temperature* and I_E is the emitter current. The base-lead resistance, $r_{bb'}$ is a function of transistor geometry. From measurements, $r_{bb'}$ is known to be inversely proportional to the conductivity of the base-region semiconductor, σ_b, so that $r_{bb'} = 1/(G\sigma_b)$ where G is the

* For other temperatures, $\Lambda = \dfrac{e}{kT}$ where e is the electronic charge and k is Boltzmann's constant. Hence $\Lambda = 11600/T$ where T is in degrees K.

constant of proportionality involving the transistor geometry. The expression for the diffusion capacitance, $C_{b'e}$ is

$$C_{b'e} = \Lambda^2 I_E \frac{W_b^2}{2\mu_n}$$

where μ_n is the electron mobility in the p-type base semiconductor and W_b is the base thickness between points of zero electrostatic voltage gradient. Throughout this report, W_b will be considered constant.[2] The collector-to-base capacitance, $C_{b'c}$, generally consists of two factors—a barrier capacitance and a transit time capacitance. In high-frequency transistors where W_b is small and σ_b is large, the collector junction barrier capacitance will predominate, so that

$$C_{b'c} \approx A_c \left[\frac{K_e \epsilon_o \sigma_b}{2\mu_p V_{CE}} \right]^{1/2}$$

where A_c is the collector area, K_e is the dielectric constant of the base semiconductor, ϵ_o is the permittivity of free space, μ_p is the hole mobility in the p-type base semiconductor, and V_{CE} is the collector voltage. With the expression $\sigma_b \approx eN_a\mu_p$ (where e is the electron charge and N_a is the density of the base impurity acceptor atoms), Equation (1) can be written

$$\text{P.A.} \approx \frac{G\mu_p\mu_n}{\omega^2 A_c \Lambda \left[\frac{2K_e\epsilon_o}{eN_a V_{CE}} \right]^{1/2} W_b^2}. \qquad (2)$$

As mentioned above, this expression is applicable to alloy-type junction transistors where the collector junction barrier occurs in the base semiconductor.

Examination of Equation (2), at first glance, makes it seem advisable to choose a high collector bias, V_{CE}, and a high base impurity density, N_a. However, the "breakdown" phenomenon of the collector, considered in the next section, is such that, when N_a rises, the maximum V_{CE} falls. Furthermore, when N_a rises sufficiently, the mobilities μ_p and μ_n begin to decrease noticeably. The next two sections discuss these interrelationships.

[2] When W_b is very small, or the collector voltage very large, the dependence of W_b upon the collector voltage may have to be considered.

Effect of Collector Voltage

In so far as the operating point is concerned, Equation (2) indicates that the high-frequency power amplification is independent of emitter current but directly proportional to the square root of the collector voltage. For improved high-frequency performance V_{CE} should be made as large as possible. There are two possible limits to the magnitude of V_{CE}. First, the collector voltage may be so large that the collector junction barrier thickness is comparable to W_b. This is not likely to occur in high-frequency transistors where σ_b is relatively large. As mentioned before,[2] the variation of W_b with V_{CE} must be considered in this region of operation; other factors which are considered in the last section may also be significant. Second, the collector voltage may be so large that collector breakdown occurs. This is the more likely limitation on the magnitude of V_{CE}. The practical maximum collector voltage will generally be less than the collector breakdown voltage. Two mechanisms, Zener breakdown[3] and avalanche breakdown[4], have generally been used to account for collector breakdown. Zener breakdown is due to internal field emission and should in theory occur when the collector-to-emitter voltage is

$$V_Z = \frac{K_e \epsilon_o}{2eN_a} E_Z^2. \tag{3}$$

The critical Zener field, E_Z, is proportional to the "band gap" and the lattice spacing of the material. Avalanche breakdown is due to electron-hole pair formation by ionization similar to Townsend discharges in gases. It is now generally believed that the avalanche mechanism plays the predominant role in collector breakdown. Preliminary measurements of avalanche breakdown indicate that the breakdown voltage is proportional to $N_a^{-0.7}$. The semiconductor surface also contributes to the breakdown characteristics, but in a manner which tends to obscure the actual breakdown. For purposes of this article, the differences between Zener and avalanche breakdown dependence on N_a are minor so, for simplicity, it will be assumed that collector breakdown voltage is proportional to $1/N_a$.

The variation in high-frequency performance with operating point (collector voltage) has been considered. The improvement in high-frequency performance that can be obtained by altering the transistor design will now be considered.

---

[3] C. Zener, "A Theory of the Electrical Breakdown of Solid Dielectrics," *Proc. Roy. Soc.* (London), Vol. 145A, pp. 523-529, July, 1934.

[4] K. G. McKay, "Avalanche Breakdown in Silicon," *Phys. Rev.*, Vol. 94, pp. 877-884, May 15, 1954.

Optimum Transistor Design

In accordance with Equation (2), improved performance can be obtained by designing for a large geometrical factor (small $r_{bb'}$), G, small collector area (small $C_{b'c}$), A_e, and small base thickness (small $C_{b'e}$), W_b. In addition, the semiconductor material can be "doped" more heavily to increase N_a. However, this will cause μ_p and μ_n to decrease and will necessitate a reduction in V_{CE}. If V_{CE} is always adjusted to some fixed fraction of the breakdown voltage, then $N_a V_{CE}$ will be constant. In this event, relatively pure (e.g. $\sigma_b < 1$ mho per centimeter) semiconductor material should be used since $\mu_p \mu_n$ will then be maximum. This may be better understood by considering that the change in $r_{bb'}$, for different semiconductor conductivities, is exactly counterbalanced by a change in $C_{b'c}$ when V_{CE} is readjusted in proportion to the breakdown voltage.

The situation may be considerably different if, in the operation of the transistor, the collector voltage is to be held constant independent of the semiconductor conductivity. In this event the power amplification as given by Equation (2) can be optimized by choosing σ_b so that $(eN_a)^{1/2} \mu_p \mu_n \equiv (\sigma_b \mu_p)^{1/2} \mu_n$ is maximum, keeping in mind that the μ's are functions of N_a and σ_b. Accordingly, the optimum value of σ_b can be determined by plotting $(\sigma_b \mu_p)^{1/2} \mu_n$ or by satisfying the relation

$$\frac{1}{\sigma_b} + \frac{1}{\mu_p} \frac{\partial \mu_p}{\partial \sigma_b} + \frac{2}{\mu_n} \frac{\partial \mu_n}{\partial \sigma_b} = 0. \tag{4}$$

Similar expressions with n and p subscripts interchanged are applicable for a p-n-p transistor. Equation (4) states that the optimum value of σ_b is such that the sum of the fractional change in base conductivity plus the fractional change in majority carrier mobility plus twice the fractional change in minority carrier mobility is zero. The optimum value of σ_b can be determined if the functional relationship between μ_n and μ_p and σ_b is known or if measured data is available. Thus, in the case of germanium,[5] the optimum σ_b is approximately 8 mhos per centimeter for an n-p-n transistor and 22 mhos per centimeter for a p-n-p transistor.

The design implications of optimum σ_b, when the collector voltage is fixed, can be summarized as follows. The σ_b to be employed is the smallest of the two values determined from (a) the optimum σ_b as given by Equation (4) and (b) the σ_b corresponding to a collector breakdown voltage somewhat larger than the operating collector voltage.

---

[5] M. B. Prince, "Drift Mobilities in Semiconductors I. Ge," *Phys. Rev.*, Vol. 93, pp. 681-687, November 1, 1953.

Comparison of N-P-N and P-N-P Transistors Made from the Same Semiconductor Material

If an n-p-n and a p-n-p transistor are made from the same semiconductor material containing equal densities of impurity atoms, with identical geometries including the same W_b, and the two are operated at the same collector voltage and frequency, then from Equation (2)

$$\frac{\text{P.A.}_{\text{n-p-n}}}{\text{P.A.}_{\text{p-n-p}}} = \frac{\mu_p\mu_n]_{\text{p-material}}}{\mu_n\mu_p]_{\text{n-material}}} \approx 1. \tag{5}$$

That is, the high-frequency power amplification of the two transistors would be approximately equal. The approximation comes about only because of the relationship between mobility and impurity density, but is quite valid for germanium even with relatively large impurity densities (with conductivities of around 1 mho per centimeter, the accuracy is within 2 per cent).

If, instead of equal impurity densities, they have equal conductivities of the base material, then

$$\frac{\text{P.A.}_{\text{n-p-n}}}{\text{P.A.}_{\text{p-n-p}}} = \frac{(\mu_p)^{1/2}\mu_n]_{\text{p-material}}}{(\mu_n)^{1/2}\mu_p]_{\text{n-material}}} \approx b^{1/2} \tag{6}$$

where b is the ratio of electron to hole mobility in intrinsic material. The $b^{1/2}$ result ($2.05^{1/2} = 1.43$ for germanium and $2.4^{1/2} = 1.55$ for silicon), although exact only for small conductivities approaching that of intrinsic material, is still approximately correct for relatively large conductivities. Thus, for 1 mho per centimeter germanium, the more exact evaluation is 1.34.

The optimum value of σ_b forms a somewhat better basis for comparing n-p-n and p-n-p transistors. This comparison can be carried out when values for minority and majority carrier mobilities as a function of σ_b are available. For germanium when σ_b is optimum, the $\text{P.A.}_{\text{n-p-n}}/\text{P.A.}_{\text{p-n-p}}$ ratio is 0.95.

The comparisons above are for equal collector voltages. If the collector voltages are adjusted in proportion to the breakdown voltages, then in all cases the ratio is that given by Equation (5) so that the two transistors have approximately equal high-frequency power amplifications.

Comparison of Transistors Made from Different Semiconductor Materials

According to Equation (2), the power amplification of an n-p-n transistor is proportional to $[N_a V_{CE}/K_c]^{1/2}\ \mu_p\mu_n$. Consider two n-p-n

transistors made of different semiconductor materials with the same impurity densities, identical geometries including the same W_b, and operating at the same d-c collector voltage and frequency. In order for the first semiconductor to be superior to the second, it is necessary that the $\mu_p\mu_n/K_e^{1/2}$ factor of the former be larger than that of the latter. Similarly, if instead of the same impurity densities, the same base conductivities are used, the appropriate factor is $[\mu_p/K_e]^{1/2}\ \mu_n$.

If the collector voltages are adjusted to the same proportion of their respective collector breakdown voltages, the appropriate comparison factor is $\mu_p\mu_n M$ independent of the base conductivities. M is a factor, dependent upon the material, which would be equal to E_Z if the collector breakdown were due to a Zener mechanism.

In order to compare p-n-p transistors it is only necessary to interchange the p and n subscripts in the above factors. It is important to remember that although interchanging subscripts does not alter the form of some of the factors, the numbers substituted will generally be different when considering "doped" semiconductors. The numbers should be those of the majority and minority mobilities in the base material used.

It is noted that the comparison factor changes, depending upon the basis of comparison. Generally, detailed theoretical or measured data concerning minority and majority carrier mobilities in a semiconductor is needed before transistors of different semiconductor materials can be compared analytically. As a convenient basis of comparison, a factor of $\mu_p\mu_n/K_e^{1/2}$ for the intrinsic material can be used as a figure of merit. A few values of this factor are tabulated in Table I. It is important to note here that the mobilities used are drift mobilities. When drift mobility data is not available, rough comparisons can be made using mobility data obtained by other methods. However, it should be understood that these comparisons may not be highly significant in so far as transistor operation is concerned.

Table I—Semiconductor Figure of Merit.

Material	μ_p	μ_n	K_e	$\frac{\mu_p\mu_n}{K_e^{\frac{1}{2}}}$
Germanium[5]	1900	3900	16	1.85×10^6
Silicon[6]	500	1200	12	0.17×10^6
90% Ge-10% Si Alloy[7]	900 est.	1300	15.6 est.	0.30×10^6

[6] M. B. Prince, "Drift Mobilities in Semiconductors II Si," *Phys. Rev.*, Vol. 93, pp. 1204-1206, March 15, 1954.

[7] S. M. Christian and E. R. Johnson, "Energy Gap and Mobility Measurements on Ge-Si Alloys," Session IV-A, Conference on Semiconductor Device Research, Minneapolis, Minn., June 28-30, 1954.

Further Discussion

It was mentioned that W_b is a function of the collector voltage and that this must be considered when W_b is small. W_b decreases with larger collector voltages so that the improvement in power amplification should be somewhat greater than that predicted by Equation (2). In so far as this equation is concerned, the power amplification should vary inversely as W_b^2. In the derivation of Equation (1), the output self-conductance term, g_{ce}, (see Figure 1a) was neglected. This term is inversely proportional to W_b and may no longer be negligible when W_b is very small. The modified expression for the power amplification including g_{ce} is

$$\text{P.A.} \approx \frac{g_m}{4\omega^2 r_{bb'} C^2_{b'e}\left[\frac{C_{b'c}}{C_{b'e}} + \frac{g_{ce}}{g_m}\right]}. \tag{6}$$

Therefore, when g_{ce} dominates the effect of $C_{b'c}$, the power amplification should vary at an even faster rate than formerly, namely inversely proportional to W_b^3.

If constructional techniques should be developed such that $r_{bb'} \rightarrow 0$, then both $C_{b'e}$ and $C_{b'c}$ could be "tuned out" with suitable circuits. In this event, the maximum high-frequency performance of the transistor would be limited by the transit time of the minority carriers through the base, and the drift mobility of the minority carriers would be of primary importance.

Modified methods of junction transistor construction[8,9] will alter some of the results, but generally drift mobilities of both carriers will still be important. In all cases it is desirable to make the base width as small as possible, to use a most favorable geometry including small junction areas, to use a base material whose conductivity is optimum, and to operate at a large collector-to-emitter voltage.

---

[8] H. Krömer, "Der Drift Transistor," *Die Naturwissenschaften*, Vol. 40, No. 22, pp. 578-579, 1953. See also id., "Zur Theorie des Diffusions und des Drifttransistors," *Archiv der Electrishen Übertragung*, Vol. 8, Part I, pp. 203-228, May, 1954: Part II, pp. 363-369, August, 1954; Part III, pp. 499-504, November, 1954.

[9] J. M. Early, "P-N-I-P and N-P-I-N Junction Transistor Triodes," *Bell Sys. Tech. Jour.*, Vol. 33, pp. 517-533, May, 1954.

A PHASE ROTATION SINGLE-SIDEBAND GENERATING SYSTEM*

By

J. R. Hall

Engineering Products Division, Radio Corporation of America, Camden, N. J.

Summary—*This paper describes a new system of phase rotation single-sideband generation which eliminates the critically adjusted 90° radio-frequency phase-shift network. Undesired sideband suppression is adjusted through a simple audio level control in one phase branch of the audio modulating signal, thus reducing to a single control the balancing out of the undesired sideband for normal changes in operating frequency.*

SINGLE sideband generators of the phase-rotation type rely, in general, on the accuracy of two ninety-degree phase-shift networks to maintain suppression of the undesired sideband. One of these networks must supply two outputs of equal amplitude and ninety degrees phase difference over the desired audio passband,[1-4] and the other network in the radio frequency circuit usually supplies two outputs equal in amplitude, and ninety degrees phase difference at only one frequency. This network must be readjusted for a change in operating frequency to maintain a constant amount of sideband suppression.

The generation system described in this paper similarly uses two such networks, but with one exception. The radio-frequency network again supplies two outputs equal in amplitude and with a ninety-degree phase difference; however, the ninety-degree phase difference is necessary only for maximum efficiency, while a deviation from this value will not result in a loss of ability to suppress the undesired sideband.

---

* Decimal Classification: R423.52.

[1] R. B. Dome, "Wideband Phase Shift Networks," *Electronics*, Vol. 19, pp. 112-115, December, 1946.

[2] S. Darlington, "Realization of a Constant Phase Difference," *Bell. Sys. Tech. Jour.*, Vol. 29, p. 94, January, 1950.

[3] H. J. Orchard, "Synthesis of Wideband Two-Phase Networks," *Wireless Engineer*, Vol. 27, p. 72, March, 1950.

[4] O. G. Villard, "Cascade Connections of 90-degree Phase-Shift Networks," *Proc. I.R.E.*, Vol. 40, p. 334, March, 1952.

The conventional system[5,6] is shown in block diagram form in Figure 1. It consists of two carrier-suppressing balanced modulators, which are essentially two pairs of amplitude modulators. Each pair is so connected that their carrier components are out of phase in the load and cancel while the sideband components are in phase and add. The output from each balanced modulator consists of an upper- and a lower-sideband component plus any distortion products.

The radio-frequency drives to the two balanced modulators as mentioned before have a phase difference of ninety degrees. Likewise, the audio modulating signals to the two modulators have a ninety-degree phase difference over the desired audio passband. A series of vector diagrams is shown in Figure 2 to illustrate the carrier and

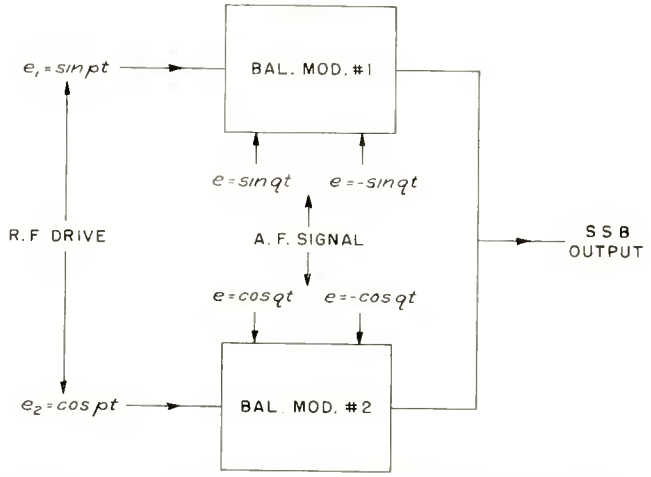


Fig. 1—Block diagram of conventional phase rotation system using two balanced modulators.

undesired sideband suppression operation. Carrier terms are represented by vectors labeled "C" and upper and lower sideband terms are represented by vectors "U" and "L" respectively.

The system, as shown in block diagram in Figure 3, consists of four linear amplitude modulators whose inputs and outputs are connected in pairs. Modulators one and four are supplied with one radio-frequency drive, and modulators two and three with the other. The two driving sources are equal in amplitude and have a phase difference between zero and one hundred eighty degrees, preferably ninety degrees for maximum efficiency.

[5] K. Henny, *Radio Engineering Handbook*, 3rd edition, McGraw-Hill Book Co., New York, N. Y., 1941, pp. 552-553.

[6] M. A. Honnell, "Single Sideband Generator," *Electronics*, Vol. 18, pp. 166-168, November, 1945.

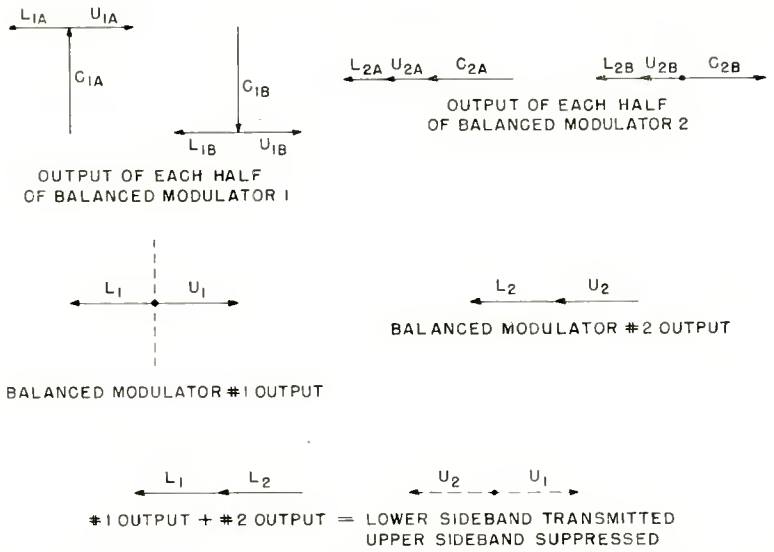


Fig. 2—Vector diagrams of balanced modulator output in conventional system.

The audio modulating signal is supplied to modulators one and two, equal in amplitude and opposite in phase. This signal is also supplied to modulators three and four in parallel, but with an amplification factor K and a phase difference of ninety degrees over the desired passband. The outputs of modulators one and two are added together as are those of three and four. The two resulting signals,

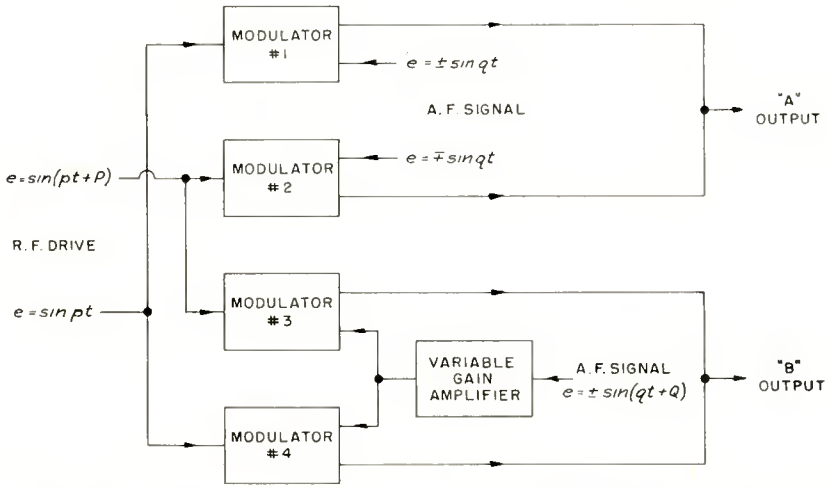


Fig. 3—Block diagram of new phase rotation system using four linear modulators.

when combined subtractively as in a push-pull tank circuit, produce a single-sideband suppressed-carrier signal.

If we let $\sin(pt+P)$ represent the ratio-frequency drive, and $\sin(qt+Q)$ the audio-frequency modulating signal, then the outputs of the modulators may be written as follows:

For modulator #1 $(P=0,\ Q=0)$

$$\begin{aligned} e_1 &= A_0\,[1+m_1 \sin qt]\sin pt \\ &= A_0\,[\sin pt + m_1 \sin pt \sin qt] \\ &= A_0 \sin pt + A_0 \frac{m_1}{2}\{\cos(p-q)t - \cos(p+q)t\}. \end{aligned}$$

For modulator #2 $(P=P,\ Q=180°)$

$$\begin{aligned} e_2 &= A_0\,[1+m_2 \sin(qt+180°)]\sin(pt+P) \\ &= A_0\,[\sin(pt+P) + m_2 \sin(pt+P)\sin(-qt)] \\ &= A_0 \sin(pt+P) + A_0 \frac{m_2}{2}[\cos\{(p+q)t+P\} \\ &\qquad - \cos\{(p-q)t+P\}]. \end{aligned}$$

For modulator #3 $(P=P,\ Q=+90°)$

$$\begin{aligned} e_3 &= A_0\,[1+m_3 \sin(qt+90)]\sin(pt+P) \\ &= A_0\,[\sin(pt+P) + m_3 \sin(pt+P)\cos qt] \\ &= A_0 \sin(pt+P) + A_0 \frac{m_3}{2}[\sin\{(p+q)t+P\} \\ &\qquad + \sin\{(p-q)t+P\}]. \end{aligned}$$

For modulator #4 $(P=0,\ Q=90°)$

$$\begin{aligned} e_4 &= A_0\,[1+m_4 \sin(qt+90)]\sin pt \\ &= A_0\,[\sin pt + m_4 \sin pt \cos qt] \\ &= A_0 \sin pt + A_0 \frac{m_4}{2}[\sin(p+q)t + \sin(p-q)t]. \end{aligned}$$

Adding the outputs of modulators #1 and #2, where $m_1 = m_2 = m_A$,

$$\begin{aligned} e_1 + e_2 = E_A &= A_0 \sin pt + A_0 \frac{m_A}{2}[\cos(p-q)t - \cos(p+q)t] \\ &\quad + A_0 \sin(pt+P) + A_0 \frac{m_A}{2}[\cos\{(p+q)t+P\} \\ &\quad - \cos\{(p-q)t+P\}]. \end{aligned}$$

The carrier terms are

$$E_{A(C)} = A_0 \sin pt + A_0 \sin (pt+P)$$
$$= A_0 \sin (pt+P/2) \sqrt{(1+\cos P)^2 + \sin^2 P}$$
$$= A_0 \sin (pt+P/2)\ 2\sqrt{\frac{1+\cos P}{2}}$$
$$= A_0 \sin (pt+P/2)\ (2\cos (P/2)).$$

Similarly, the sideband terms are

$$E_{A(SB)} = A_0 \frac{m_A}{2} [\cos (p-q)t - \cos (p+q)t]$$
$$+ A_0 \frac{m_A}{2} [\cos \{(p+q)t + P\} - \cos \{(p-q)t + P)]$$
$$= A_0 \frac{m_A}{2} [\cos \{(p-q)t - 90 + P/2\}]\ 2\sin (P/2)$$
$$+ A_0 \frac{m_A}{2} [\cos \{(p+q)t + 90 + P/2\}]\ 2\sin (P/2)$$
$$= A_0 \frac{m_A}{2} [2\sin (P/2)]\ [\cos \{(p-q)t - 90 + P/2\}$$
$$+ \cos \{(p+q)t + 90 + P/2\}].$$

The resultant output at "A" is

$$E_A = A_0 \sin (pt+P/2)\ (2\cos (P/2)) + A_0 \frac{m_A}{2} (2\sin (P/2))$$
$$[\cos \{(p-q)t - 90 + P/2\} + \cos \{(p+q)t + 90 + P/2\}].$$

Similarly, adding the outputs of modulators #3 and #4 where $m_3 = m_4 = m_B$,

$$e_3 + e_4 = E_B = A_0 \sin (pt+P/2)\ (2\cos (P/2))$$
$$+ A_0 \frac{m_B}{2} (2\cos (P/2))$$
$$[\sin \{(p+q)t + P/2\} + \sin \{(p-q)t + P/2\}].$$

The result of subtracting output B from output A is

(1) for carrier terms:

$$E_{0(C)} = A_0 \sin (pt+P/2)\ (2\cos (P/2))$$
$$- A_0 \sin (pt+P/2)\ (2\cos (P/2)) = 0$$

(2) for upper sideband terms:

$$E_{0(UP)} = A_0 \frac{m_A}{2} (2 \sin (P/2)) \cos [(p+q)t + 90 + P/2]$$

$$- A_0 \frac{m_B}{2} (2 \cos (P/2)) \sin [(p+q)t + P/2]$$

$$= 2A_0 \left(\frac{m_A}{2} \sin (P/2) \right) [- \sin \{(p+q)t + P/2\}]$$

$$+ 2A_0 \left(\frac{m_B}{2} \cos (P/2) \right) [- \sin \{(p+q)t + P/2\}]$$

$$E_{0(UP)} = A_0 [- \sin \{(p+q)t + P/2\}]$$

$$[m_A \sin (P/2) + m_B \cos (P/2)]$$

(3) for lower sideband terms:

$$E_{0(LOW)} = A_0 [- \sin \{(p-q)t + P/2\}]$$

$$[- m_A \sin (P/2) + m_B \cos (P/2)].$$

It is seen that both the upper and lower sideband amplitudes are modified by similar terms of

$$m_A \sin (P/2) + m_B \cos (P/2) \text{ and}$$

$$- m_A \sin (P/2) + m_B \cos (P/2) \text{ respectively.}$$

Taking the usual case where the phase difference between the two radio frequency driving signals is ninety degrees, the amplitude of the lower sideband transmitted is

$$E_{0(LOW)} = -m_A \sin (90/2) + m_B \cos (90/2)$$

$$= - m_A (1/\sqrt{2}) + m_B (1/\sqrt{2})$$

or, if $m_A = m_B$

$$E_{0(LOW)} = - m (1/\sqrt{2}) + m (1\sqrt{2}) = 0.$$

By equating the amplitude of the lower sideband to zero, it is seen that for any value of P, m_A and m_B, the respective modulation percentages of channels A and B may be adjusted to obtain full suppression

$$m_B \cos (P/2) - m_A \sin (P/2) = 0$$

$$m_B \cos (P/2) = m_A \sin (P/2)$$

$$\frac{m_B}{m_A} = \frac{\sin (P/2)}{\cos (P/2)} = \tan (P/2).$$

The amplitude of the transmitted sideband in the usual phase rotation system is equal to the sum of the two in-phase outputs of the two balanced modulator channels. In the system described here, the transmitted sideband amplitude is $A = m_B \cos (P/2) + m_A \sin (P/2)$ or, for $P = 90°$ and m_A and m_B adjusted for maximum suppression of the undesired sideband, the transmitted signal amplitude is $\cos (90/2) + \sin (90/2) = 1.414$. This is a reduction of 3 decibels in the maximum output compared to the usual system. If the undesired output of both systems is zero, the suppression ratio in both cases would still be infinite. However, assuming the balance obtained results in some minimum undesired output voltage, as is often the case, then the maximum degree of suppression of the new system is reduced by 3 decibels, all other factors being optimum.

It is interesting to note that for a $\pm 30°$ variation of P from 90°, the maximum transmitted signal amplitude can be computed as follows:

$$P = 60°, \quad P/2 = 30°,$$

$$m_B/m_A = \tan 30° = .5774.$$

If the maximum value of m_A is unity, at which time $m_B = .5774$, then $A_{max} = .5774 \times .8660 + 1 \times .5 = 1$.

This is a reduction in maximum output of 3 decibels and correspondingly, a 3-decibels reduction in maximum possible sideband suppression ratio. A similar variation of 30° in the radio-frequency phase difference of the usual phase rotation sideband generator would result in a maximum obtainable sideband suppression of approximately 12 decibels which is far below a useable value.

Since the amplitude of the transmitted sideband is $A = 2 \sin (P/2)$ for values of P between 0 and 90°, and $A = 2 \cos (P/2)$ for values of P between 90° and 180°, it is desirable to operate at a nominal value of $P = 90°$. In practice, this would mean the radio-frequency phase network would be adjusted to provide a difference of ninety degrees in the center of the desired operating range, and the maximum range would be limited to the point where the difference varies $\pm 30°$ from the nominal 90° value. Since maximum suppression can be controlled by means of audio gain in one channel, a feedback arrangement might be utilized to make the system suppression automatic.

Figure 4 is a series of vector diagrams illustrating the preceeding mathematical development. As before, carrier vectors are labeled "C" and upper and lower sideband vectors are "U" and "L" respectively. Two examples are shown; one where the radio-frequency phase difference is 90° as in the conventional system, and the other for a radio-frequency phase difference of 60°. It is obvious in the later case that

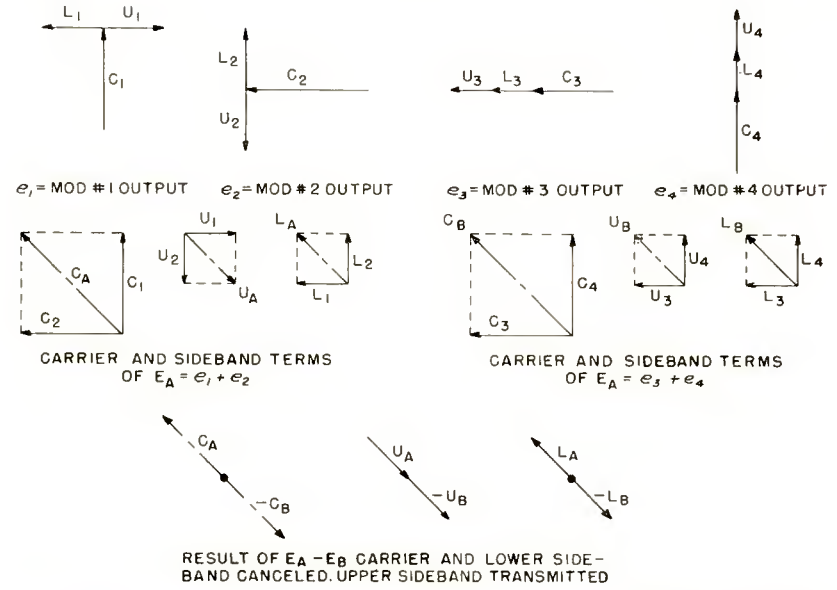


Fig. 4—Vector diagrams of modulator outputs of new system for radio- and audio-frequency and phase differences, both at 90°. $m_1 = m_2$, $m_3 = m_4$.

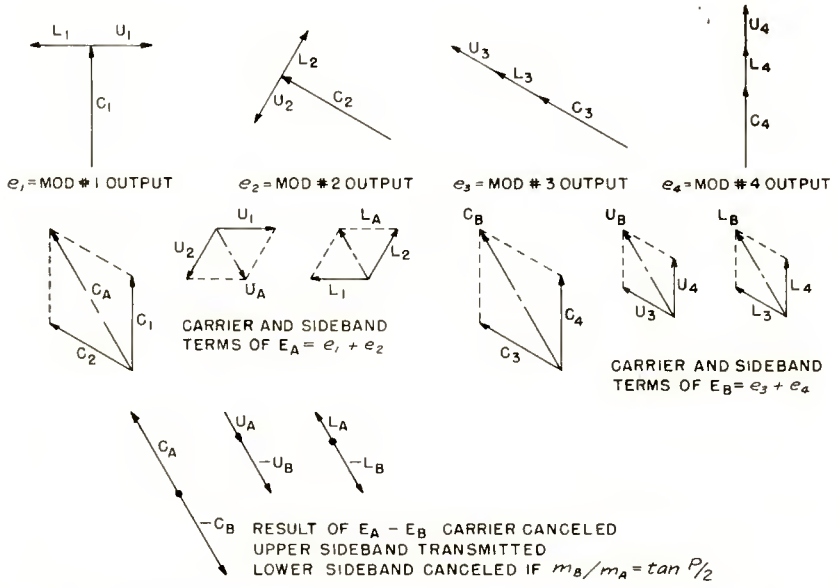


Fig. 5—Vector diagrams of modulator outputs of new system for radio-frequency phase difference of 60°. $m_1 = m_2$, $m_3 = m_4$.

to obtain complete suppression of the undesired sideband, one merely has to change the modulation percentage of one pair of modulators.

In both systems, selection of the sideband to be transmitted may be accomplished by reversing the output of audio-frequency phase shift network so that the opposite pair of modulators receives the signal delayed by 90°.

The adjustment technique for a large frequency and corresponding phase change would be to monitor the suppressed sideband with signal level indicating device and vary the levels of the audio channels A and B differentially for a minimum output signal.

This system of phase rotation single-sideband generation should prove especially useful for compact portable or mobile equipment where ease of adjustment is desired, or when operated by personnel with limited technical ability.

ON THE MEASUREMENT OF THE AVERAGE TIME DELAY IN SECONDARY EMISSION*

By

M. H. Greenblatt

Research Laboratory, RCA Laboratories,
Princeton, N. J.

Summary—*An effort was made to measure the time delay in secondary emission in as direct a manner as possible. The method employed consists of a comparison between the duration of short bursts of primary electrons and the duration of the consequent bursts of secondary electrons. Short bursts of primary electrons are produced by deflecting an electron beam at a 400-megacycle rate across a circular aperture. The bursts of secondaries are analyzed by synchronously deflecting them across a fluorescent screen, or presentation electrode. Since the instantaneous deflection varies while the burst of secondary electrons is travelling between the deflection plates, the shape of the pulse of secondaries corresponds to a variation in light intensity across the fluorescent screen. Measurements show that the duration of the pulses of secondaries is larger than that of the primaries by an amount which depends on the accelerating voltage. Pulses of secondaries having a width of* 1.3×10^{-10} *second have been produced by pulses of primaries having a width of* 6×10^{-11} *second. The difference is mainly due to a dispersion in transit times. The dispersion in time delay in secondary emission must, therefore, be smaller than* 7×10^{-11} *second, and is probably much smaller.*

Introduction

THE magnitude of the time delay in the secondary emission process is a quantity of increasing practical importance. There are a few reports of measurements of this time delay in the literature,[1-4] but all of these measurements are of the indirect type. R. R. Law[5] has recently reported results of an experiment with power multipliers. One of the theories proposed by Law to explain his results requires the assumption that secondary emission has an expo-

---

* Decimal Classification: 530.

[1] C. C. Wang, "Reflex Oscillators Utilizing Secondary Emission Current," *Phys. Rev.*, Vol. 68, p. 284, 1945.

[2] H. M. Wagner and W. R. Ferris, "The Orbital-Beam Secondary-Electron Multiplier for Ultra-High-Frequency Amplification," *Proc. I.R.E.*, Vol. 29, pp. 598-602, November, 1941.

[3] M. H. Greenblatt and P. H. Miller, "A Microwave Secondary Electron Multiplier," *Phys. Rev.*, Vol. 72, p. 160, 1947.

[4] G. Diemer and J. H. Jonker, "On the Time Delay of Secondary Emission," *Phillips Research Reports*, Vol. 5, June, 1950.

[5] R. R. Law, "Emission Time of Secondary Electrons," *Nucleonics*, Vol. 10, p. 38, 1952.

nential decay with a decay constant of 3×10^{-10} second. This assumption, if true, would be very important both in future uses of multiplier tubes when speed of response is important, and in the theoretical aspects of secondary emission. Existing commercial photomultipliers cannot be used to investigate a time delay in secondary emission such as Law assumes because the interstage transit-time spread is too great —about 5×10^{-10} second. Law took this fact into account in his work by building exceedingly small multipliers in which the transit-time dispersion is correspondingly smaller.

The present work was undertaken in order to measure, in as direct a manner as possible, the time delay in secondary emission. It should be mentioned that only the *dispersion* in time delays limits the frequency response of secondary emission multipliers or the resolving time of photomultiplier coincidence circuits. The quantity measured here is actually the dispersion in time delays, but it will be referred to simply as time delay, in deference to common usage.

Description of Experiment

In considering various ways of measuring the time delay in secondary emission, it soon becomes apparent that practical considerations of capacity, resistance, the required resolving time, and limitations on existing amplifier circuits make it difficult to use standard measuring techniques. The method used here is relatively straightforward. Basically, it consists of directing very short pulses of primary electrons at a secondary-emission target, and then examining the resultant pulses of secondary electrons to see if these pulses are lengthened by the secondary-emission process. Short primary pulses are obtained by sweeping an electron beam across an aperture at a radio-frequency (r-f) rate and directing the electrons which pass through the aperture onto a secondary-emitting target. The resultant bursts of secondaries are analyzed by focusing the electrons into a beam which is then deflected at the same frequency as the beam of primaries. If the deflection frequency is high enough, the deflection voltage varies appreciably as the pulse travels between the deflection plates and different portions of the pulse are deflected through different angles. The distribution of the number of secondaries as a function of time is thus displayed as a variation in current density in a direction at right angles to the beam. If the frequency and amplitude of the sweep are constant, the subsequent pulses of secondaries will coincide. If a phosphor screen is placed on the end face of the tube, the shape of the pulse of secondaries as a function of time corresponds to a variation in light intensity as a function of distance across the face of the tube.

The experiment consists of a comparison between the duration of a pulse of secondary electrons and the duration of the originating pulse of primaries. Primary pulses shorter than 10^{-10} second are produced by deflecting the primary beam past an aperture at a 400-megacycle rate. A calculation of the time between half-height points of the primary pulses requires a knowledge of the frequency, the peak-to-peak r-f deflection voltage (ΔV) and of the d-c deflection voltage difference (δV) between half-current points as the beam is moved across the aperture. ΔV is measured by a slide-back method in which the d-c equivalent of the peak-to-peak r-f voltage is measured. δV can be measured from the static characteristics of the tube. An expression for the time interval between half-height points of a primary pulse can be obtained by considering the time required for a radius vector of length $\Delta V/2$ which is rotating at an angular frequency ω to go past an aperture δV. This time is $2\delta V/\omega\Delta V$ where δV, of course, varies with primary voltage.

The analysis of the time corresponding to the width at half-height of the secondary pulse is similar to the above claculation. The expression for the duration of the secondary pulses is $2d/L\omega$ where d is the distance between half-intensity points along the face of the tube and L is the length of the trace that would be caused by the second r-f deflection if the current from the target were constant. It is important, of course, to have the pulse of secondaries fall in the middle portion of the second r-f scan in order to use the above expression. Otherwise, corrections for the nonlinearity of the scan must be made.

Thus, the half-width of the primary pulses can be calculated and the half-width of the resulting secondary pulses can be measured. It would appear that these measured and calculated half-widths would agree if a low radio-frequency is used, and that the better time resolution obtained with a higher radio-frequency would show the existence of a difference between the measured and calculated values. The difference would probably take the form of a simple broadening of the secondary pulse. A broadening could be caused by a dispersion of the transit times of the secondary electrons, by a dispersion in the time delay in secondary emission, or by some anomalous electron-optical effect in the measurement or production of short pulses. The possibility of an anomalous electron-optical effect was investigated by performing this experiment at a low radio-frequency (260 kilocycles), in addition to the experiment at 400 megacycles. The use of a low radio-frequency changes the time scale so that transit-time dispersions, etc. are not noticeable. It does not, however, change the electron optics from that of the high-frequency experiment.

The effect of transit-time dispersion was investigated by varying

the accelerating voltages. This has the effect of varying the transit-time dispersion. By performing the experiment at several different over-all voltages, data is obtained from which it is possible to extrapolate to find the behavior at infinite accelerating voltages, in which case there would be no transit-time dispersion.

It is worthy of note that transit-time dispersions and other error effects tend to increase the measured pulse width. Because of this, the results of this experiment represent upper limits on the time delay in secondary emission.

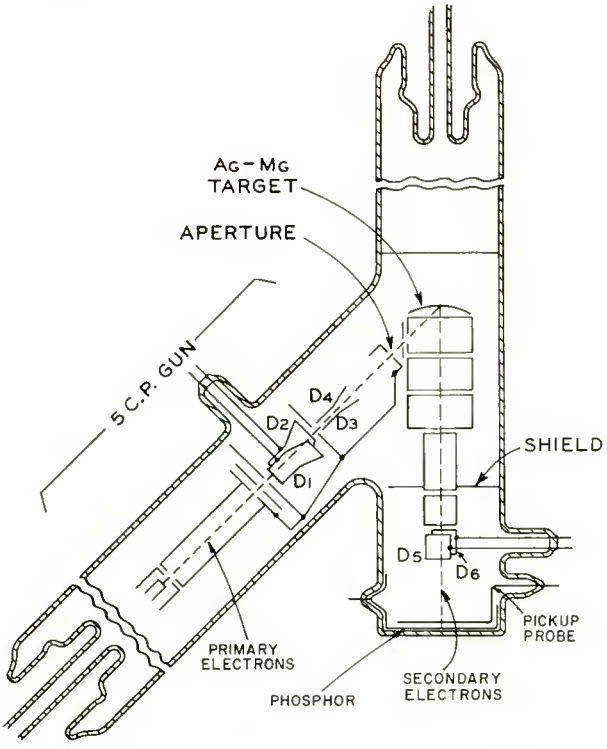


Fig. 1—Diagram of experimental tube No. 5.

Experimental Tube

A tube which operates in the previously mentioned manner is shown in Figure 1. The beam of primary electrons is aimed through the aperture at the silver–magnesium secondary-emissive target. The target has a very small amount of willemite on it in order to facilitate the initial adjustment of the beam. Secondaries from the target are made to form a beam and this beam is aimed at the end of the tube which

has a phosphor, as is shown in Figure 1. An r-f voltage is applied to the deflection plates, $D_{5,6}$, and the beam of secondary electrons is deflected back and forth across the phosphor. If the same r-f frequency is now applied to $D_{1,2}$, secondary electrons will not be present at all times during the r-f cycle, but rather there will be two pulses of primaries getting through the aperture, 180° apart. The two consequent pulses of secondaries, still 180° apart, will, in general, strike the phosphor at two different places. Proper phasing of the r-f on $D_{5,6}$ relative to that on $D_{1,2}$ can cause both pulses to strike in essentially the same spot. This spot is the middle of the r-f deflection sweep, where the scanning velocity is the highest. In order to simplify the resultant pattern on the phosphor, some of the r-f from $D_{1,2}$ can be applied in quadrature to $D_{3,4}$. The scan of primary electrons is then somewhat elliptical because $D_{1,2}$ and $D_{3,4}$ are at right angles to each other and a small d-c deflection voltage on $D_{3,4}$ can move the primary scan so that only one pulse per r-f cycle gets through the aperture. The deflection voltage can be varied so that the other pulse in each cycle can also be seen, alone, on the phosphor.

With the secondary pulses striking the phosphor near the middle of the r-f deflection sweep, a plot of the number of electrons landing on the phosphor as a function of position along the phosphor will be a curve of the shape of the secondarily emitted pulse as a function of time. The error introduced by considering the middle 2/3 of a sinusoidal scan to be linear is only 9 per cent.

The density of electrons as a function of position along the phosphor can be measured in several ways. A photograph of the phosphor screen can be taken and a densitometer will then give the desired curve. Another way, the one actually used, presents the curve directly on an ordinary oscilloscope. This is accomplished by means of a pickup probe in the form of a wire placed parallel to and near the phosphor, and in the plane of $D_{5,6}$. Audio-frequency voltage is superimposed on the r-f applied to $D_{5,6}$ and the same frequency is applied to the horizontal amplifier of the oscilloscope. The electrical signal picked up by the probe is amplified and applied to the vertical deflection terminals of the oscilloscope. Since the effect of superimposing audio voltage on the r-f is equivalent to mechanically moving the wire at the audio frequency, the current-density distribution in the pulse of secondaries is sampled at an audio rate and is presented to the oscilloscope as an amplitude-modulated, low-frequency signal.

The magnitude of the current picked up by the signal probe determines the magnitude of the signal available. It varies directly with the total amount of current in the secondary beam and inversely with

the size of the focused spot at the phosphor. This latter effect is due to the fact that a given size wire intercepts a larger fraction of a well focused spot than it does of a poorly focused spot, if the spot is larger than the wire diameter. The total amount of current in the secondary beam depends on the efficiency with which the secondaries are collected. The size of the focused spot of electrons, however, in general varies with the collection efficiency if one requires that the beam of collected secondaries be capable of deflection by reasonable electrostatic means. As a result, it is difficult to combine a well-focused spot with a high collection efficiency.

Two versions of the tube were built—the one in Figure 1 (designated tube No. 5), which collects approximately 50 per cent of the

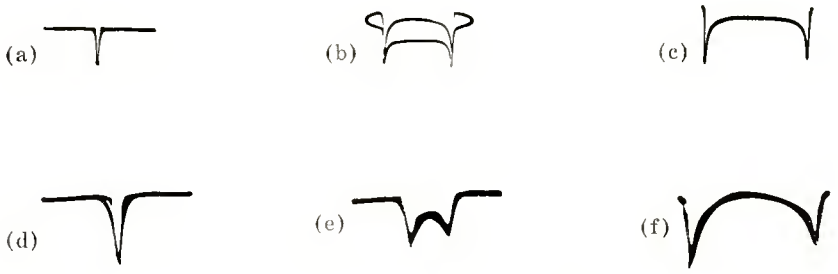


Fig. 2—Photographs of typical traces: (a) tube No. 2—audio on $D_{5,6}$, no r-f on $D_{1,2}$; (b) tube No. 2—audio plus some r-f on $D_{5,6}$, no r-f on $D_{1,2}$; (c) tube No. 2—audio plus r-f on $D_{5,6}$, no r-f on $D_{1,2}$; (d) tube No. 5—audio on $D_{5,6}$, no r-f on $D_{1,2}$; (e) tube No. 5—audio plus some r-f on $D_{5,6}$, no r-f on $D_{1,2}$; (f) tube No. 5—audio plus r-f on $D_{5,6}$, no r-f on $D_{1,2}$.

secondary emission current but has a focused spot about ¼ inch in diameter, and another one (designated tube No. 2) which collects only about 5 per cent of the current but has a focused spot less than .070 inch in diameter.

Figure 2 shows photographs of the typical behavior of tube No. 2 when the probe wire is used. Figure 2a shows the oscilloscope trace caused by a d-c beam of primaries and only audio deflection on $D_{5,6}$. The width at half-height of this pulse is considerably narrower than the widths to be measured.

With the audio voltage on $D_{5,6}$, r-f is applied to $D_{5,6}$. Application of enough r-f to cause somewhat less r-f deflection than audio deflection produces a trace similar to that of Figure 2b. The trace is peaked at the ends since the r-f is sweeping the beam faster in the

middle than at the ends. Hence, the average current collected by the probe is smaller in the middle. In all the subsequent photographs, the r-f deflection on $D_{5,6}$ is equal to the audio deflection. In the trace of Figure 2c there is no r-f on $D_{1,2}$. This establishes the length of the trace due to r-f only on $D_{5,6}$ as being equivalent to the horizontal deflection of the oscilloscope regardless of whether or not there is r-f on $D_{1,2}$. This procedure establishes a time scale for measurement purposes. If the angular frequency of the r-f is ω and the length of the r-f, or audio, trace is D, then the sweep speed near the center of the trace is $\omega D/2$ centimeters per second.

If r-f is applied to $D_{1,2}$, the resultant pattern is shown in Figure 3a. The top and bottom halves of the elliptical trace in Figure 3a are identical, and only one need be considered. The ellipticity in Figures

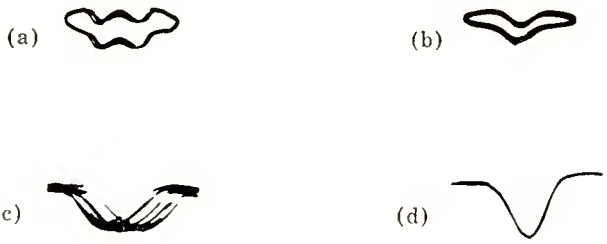


Fig. 3—Photographs of typical traces: (a) tube No. 2—audio plus r-f on $D_{5,6}$, r-f on $D_{1,2}$ not phased correctly; (b) tube No. 2—audio plus r-f on $D_{5,6}$, r-f on $D_{1,2}$ phased correctly; (c) tube No. 5—audio plus r-f on $D_{5,6}$, r-f on $D_{1,2}$ phase varying; (d) tube No. 5—audio plus r-f on $D_{5,6}$, r-f on $D_{1,2}$—retouched photograph of a single trace while phase was varying.

2b and 3a has been achieved by coupling some of the horizontal audio into the vertical amplifier of the oscilloscope and by using a sinusoidal rather than a sawtooth scan. The amount of ellipticity is, of course, variable. The two pulses in the lower half of Figure 3a, for example, are 180° apart as measured on the r-f time scale. They are symmetrical about the center of the trace. Such a pattern can be produced only when the r-f on $D_{3,4}$ is small enough to allow both pulses in an r-f cycle to get through the aperture. The phase of the r-f on $D_{5,6}$ relative to that on $D_{1,2}$ can now be adjusted so that these two pulses are essentially superimposed and Figure 3b results. If the r-f on $D_{3,4}$ is now increased, the centering voltage on $D_{3,4}$ can be varied so that one or the other pulse of each r-f cycle appears.

Figures 2a, 2b, 2c, 3a and 3b are photographs of traces derived

from tube No. 2 at a frequency of 200 megacycles. These photographs are reproduced without retouching. The final experiment was performed on tube No. 5 operating at 400 megacycles. Figures 2d, 2e, and 2f are photographs of traces from tube No. 5 at 400 megacycles, and it is seen that they are less clear than those in Figures 2a, 2b, and 2c. In fact, when an attempt was made to duplicate Figure 3b with tube No. 5 at 400 megacycles, Figure 3c resulted. In this photograph, the audio trace is much less elliptical, but the pulses are not in registration with each other. This is believed to be due to spurious noise modulation of the phase of the r-f on $D_{1,2}$. In order to avoid this multiplicity of traces, an exposure short enough to record only a few of these traces is taken. Such a photograph is completely satisfactory for measurement purposes, but does not have enough contrast to be suitable for reproduction unless it is retouched. Figure 3d is such a retouched photograph.

Associated Equipment

The r-f power required by $D_{5,6}$ for satisfactory operation is of the order of 5 watts. A small portion of this r-f power is used to drive a tuned amplifier which uses a 6252 twin-tetrode amplifier tube, the output of which is sent through a line stretcher to $D_{1,2}$. The output of the tuned amplifier is varied by varying the supply voltage. Both sets of deflection plates are at the ends of separate, tuned Lecher-wire systems. The r-f power for each set of deflection plates is sent from the amplifier or the oscillator over 50-ohm cable past a double-stub tuner to a resonant loop which is coupled to the end of the Lecher wire. Double-stub tuners are used to match the deflection plates and Lecher wires to the line in order to be able to use the line stretcher to change phase without changing amplitude. A diagram of the arrangement of the 400-megacycle driver and amplifier is shown in Figure 4a. The equipment for 200 megacycles is very similar, an 829B tube being used for the tuned amplifier circuit.

For the measurements at 260 kilocycles it was necessary to utilize an LRC method of phase shifting, due to the much longer wavelengths involved. The diagram for the 260-kilocycle oscillator, phase shifter, and amplifier is shown in Figure 4b. The Lecher wires, double stub tuners, and line stretcher cannot be used at this low frequency.

Supplementary Measurements

The method used to vary the effective relative phases of the r-f on $D_{1,2}$ and $D_{5,6}$ deserves mention. When the line stretcher is used, the length must be variable from 0 to $\lambda/2$, which for 400 megacycles is

37 centimeters. The same effect can be achieved by varying the transit time of electrons in the path from $D_{1,2}$ to $D_{5,6}$. At 400 megacycles, this can be done by varying the over-all voltage. This method was used when only small changes in phase were needed.[6]

The measurement of the peak-to-peak r-f deflection at the aperture is the only r-f measurement required. By varying the d-c bias on $D_{1,2}$, the beam may be shifted so that it passes through the aperture only at the extreme positive or extreme negative excursion of the r-f

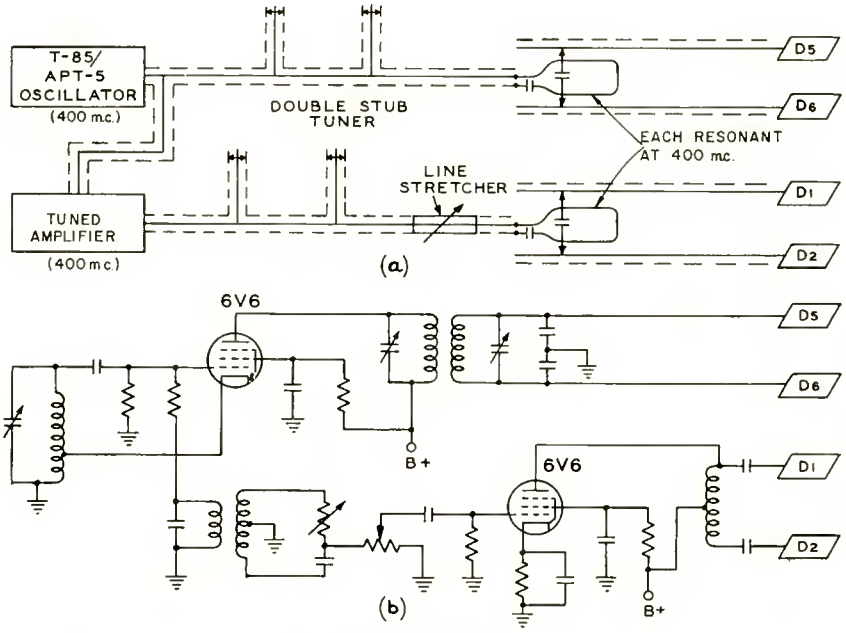


Fig. 4—Schematic diagram of oscillator and tuned amplifier: (a) 400 megacycles; (b) 200 kilocycles.

deflection sweep. The difference in bias between these two positions represents the d-c equivalent of the r-f deflection. This d-c equivalent is the ΔV referred to previously.

When the above measurement was made, it was ascertained that the r-f on $D_{5,6}$ was still equal in effect to the audio voltage on these same

[6] A convenient method of measuring the transit time from one set of deflection plates to the other is as follows: The over-all voltage can be adjusted so that the pulses fall at the ends of the r-f or audio sweep. An arbitrary change in the over-all voltage now causes the pulses to appear somewhere in the middle of the sweep, and a large enough change causes the pulses again to appear at the ends of the r-f sweep. Since the difference in transit time from $D_{1,2}$ to $D_{5,6}$ is then ½ of an r-f cycle, and since transit times are proportional to $V^{-\frac{1}{2}}$, two equations can be written which can be solved for the actual transit time. Such a measurement was made, and the transit time was found to be of the order of 10^{-8} second.

plates. This was the case when the data of Figure 2(f) was obtained. There is still a possibility, however, that the magnitude of the r-f on $D_{5,6}$ might have changed when the r-f was applied to $D_{1,2}$. This possibility can be eliminated while the peak-to-peak voltage is being measured. If the d-c centering of the beam at $D_{1,2}$ is varied, the position of the pulse on the oscilloscope relative to the audio scan will correspond to the position of the aperture relative to the primary r-f scan. When a d-c deflection sufficient to allow only the peak r-f voltage to bring the beam back to the aperture has been reached, the pulse on the oscilloscope will have moved over to the end of the audio trace, if the r-f level on $D_{5,6}$ has not changed.

Results

The measurments of time delay in secondary emission are shown in Figures 5a, 5b, and 5c. Each of these is a plot of the width at half-height of the secondary-emission pulse as a function of the reciprocal of the peak-to-peak r-f voltage on $D_{1,2}$, at different primary voltages, V_1. The width at half-height of these pulses corresponds to a definite fraction of an r-f cycle. The fraction of an r-f cycle that the pulse occupies has a time scale associated with it and so does the peak-to-peak r-f deflection on $D_{1,2}$. The plots, therefore, are essentially curves of measured half-widths of secondary pulses against calculated half-widths of primary pulses. The solid line in Figures 5a, 5b, and 5c is the theoretical curve which is drawn from the origin through the point calculated for a peak-to-peak r-f voltage of 100 volts. This point is calculated from the expression for the half-width, $\frac{2\delta V}{\Delta V}$. This straight line, from the origin through the calculated point, is meaningless for very large values of ΔV since the measured pulse width can not be smaller than the audio-only beam width (Figure 2a). The dotted line shows an appropriate correction applied to the theoretical curve.

The large solid dots in Figure 5 are the results of measurements taken at a radio frequency of 260 kilocycles. At this frequency, it is expected that these points will lie on the theoretical curve, since the sweep speed is too slow to show any time dispersion. Experiment shows that this is the case.

The small dots are measurements taken at 400 megacycles. At this frequency, the unit distance along the ordinate scale is 10^{-10} second. The points are seen to lie along a line which is roughly parallel to the theoretical line and displaced from it by an amount τ. τ is the excess of the measured duration of the pulse over the calculated duration. This excess is a time dispersion and could be due to transit time dispersion or to a dispersion in time delays in secondary emission.

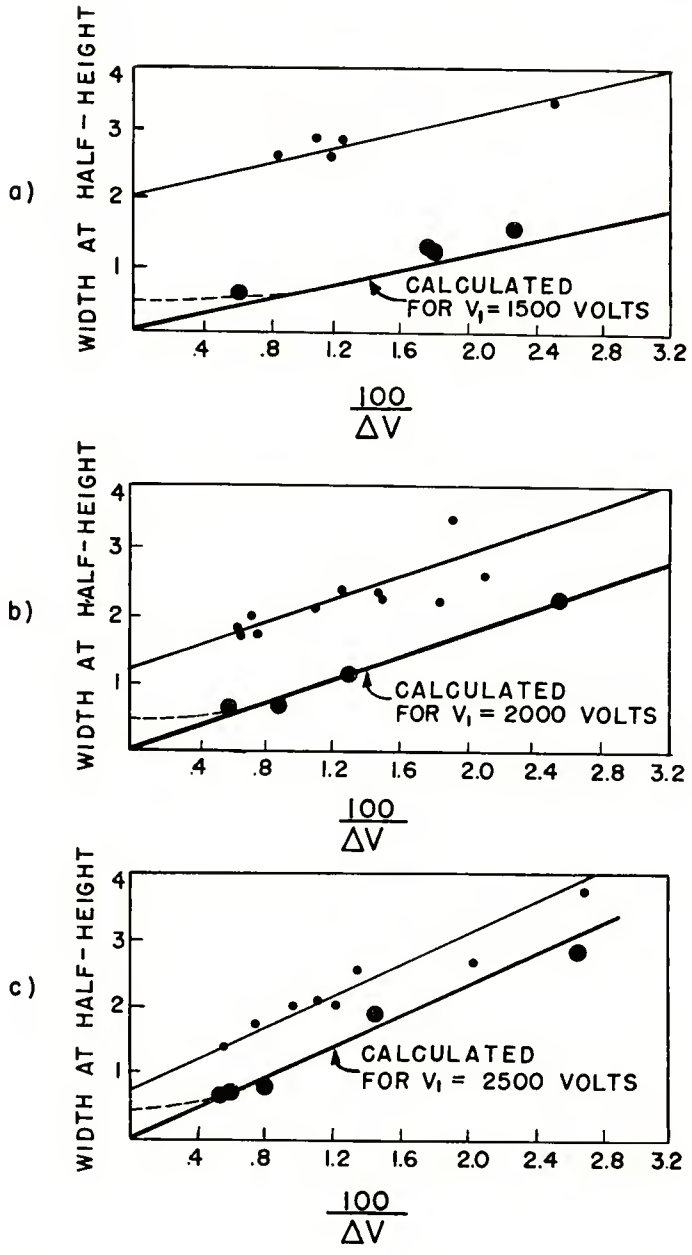


Fig. 5—Observed half-width of secondary emission pulses versus reciprocal of peak-to-peak voltage on $D_{1,2}$. These curves are for different values of accelerating voltage, V_1. The large solid dots are measurements at 260 kilocycles, the small dots at 400 megacycles. Unit distance along the ordinate is 10^{-10} second for the 400-megacycle measurements.

If the time dispersion is due partly to transit-time dispersion and partly to time delay in secondary emission, then a plot of the observed τ's as a function of $1/V_1$ should result in a straight line, the zero intercept of which would indicate the dispersion in time delay in secondary emission alone. The present experiment has not yielded sufficiently accurate data to give such a zero intercept. The values of τ are seen to decrease monotonically as V_1 is increased. The fact that τ does decrease as the accelerating voltage is increased implies very strongly that the observed excess is caused at least partly by transit-time dispersion, and is surely less than the smallest measured τ which is 7×10^{-11} second.

An analysis was made of the expected transit-time dispersion. This was similar to an analysis made by Malter.[7] It showed that a transit-time dispersion of the order of 10^{-10} second at an accelerating voltage of the order of 2000 volts should be expected due to the initial velocities of the secondaries. The smallest τ might, therefore, be largely transit-time dispersion.

A stronger, but more specialized, conclusion can be drawn from the above data. If the distribution in time delays is not symmetrical but is one-sided as, for example, a simple exponential delay would be, then the width at half-height of this distribution could not be larger than the *half*-width at half-height of the measured pulses. This corresponds to an upper limit of about 3×10^{-11} second if the delay is exponential, or similarly asymmetrical.

Discussion

The results given are valid only if the assumptions concerning the behaviour of the experimental tube and r-f equipment were correct. An effort was, therefore, made to verify the basic description of events. For example, it seemed possible that the r-f applied to $D_{1,2}$ or to $D_{5,6}$ might have an excessive harmonic content. Measurements with a wavemeter showed that the second- and third-harmonic contents were negligibly small.

An auxiliary experiment was performed to ascertain just how short the primary pulses could actually be made. A tube was built and tested in which the beam of primaries could go straight through the aperture, straight through a secondary "gun" and deflection plates and strike the phosphor. As expected, pulses shorter than 10^{-10} second were produced.

Another point which was checked was the possibility that a small

[7] L. Malter, "The Behavior of Electrostatic Electron Multipliers as a Function of Frequency," *Proc. I.R.E.*, Vol. 29, pp. 587-598, November, 1941.

fraction of the secondary-emission current in the secondary beam emerges with very little, if any, time delay, and that a substantial fraction of that current is then emitted over a much longer time interval. The area under one of the oscilloscope traces represents a certain current. This current is proportional to the area under the oscilloscope trace, and inversely proportional to the product of the length of the audio trace and the gain of the amplifier. It should be equal to the current picked up by the probe wire, which can be read on a microammeter. It was found that the calculated current (approximately 4×10^{-8} ampere) was at least 80 per cent of the measured current. The fact that the measured and calculated currents do not agree more closely is attributed to difficulties in measuring the small d-c currents.

Still another adverse possibility is that only a very small fraction of all the secondaries is collected and used in the secondary beam. This brings up the possibility that there is a small fast component and a much larger slow component in secondary emission and that only the fast component, for example, has been collected and examined in this experiment. Measurements were, therefore, made of the d-c target current and of the current collected by the phosphor. (The target current is of the order of 100 microamperes.) A measurement shows that approximately 50 per cent of the total secondary emission current is being investigated. The fact that this fraction is not higher is due to a compromise between the fraction of secondaries collected and the size of the undeflected spot of electrons at the phosphor. It may seem surprising that a much larger fraction than 50 per cent of the secondaries is not formed into the beam, but the necessity of deflecting the secondaries with practical r-f voltages imposes serious restrictions. Another limitation on the electron optics is the fact that the field at the surface of the secondary emitter must be strong in order to minimize the transit time dispersion. Because of these considerations, a collection efficiency of 50 per cent is not considered to be unduly low.

Conclusion

Pulses of secondary electrons as short as 1.3×10^{-10} second duration have been measured. These pulses were produced by pulses of primary electrons 6×10^{-11} second in duration. The broadening of 7×10^{-11} second is shown to be at least partly due to transit-time dispersion. The time dispersion due to the secondary-emission process itself must, therefore, be less than 7×10^{-11} second.

OPTIMUM DESIGN OF PERIODIC MAGNET STRUCTURES FOR ELECTRON BEAM FOCUSING*

By

Kern K. N. Chang

Research Laboratory, RCA Laboratories,
Princeton, N. J.

Summary—*The peak axial magnetic field produced by a periodic magnet constructed of axially magnetized ring-shaped magnets depends on the geometry of the ring magnet as well as on the characteristics of the magnetic material used. The soft iron shims which are inserted between the rings to help reduce asymmetrical transverse fields have a great influence upon the peak field. Design curves for periodic magnets of various dimensions, of minimum weight, and of maximum field are given. The curves check very well with experimental measurements.*

Introduction

THE development of the theory[1-4] of periodic-field beam focusing has led to problems involving the actual design of the magnet, namely, how to best shape a periodic magnet as regards weight and size, in order to focus an electron beam of given current, voltage, and diameter.

The most convenient way of constructing a periodic magnet is by assembling a series of identical axially magnetized ring magnets end to end, with adjacent sections oppositely polarized. The magnetic field along the axis of these rings is primarily determined by the ratios of their inner and outer radii to the length of one section and also by the operating point on the demagnetization curve of the material from which the magnets are formed. A periodic magnet for beam focusing usually requires nearly perfect axial symmetry in order to minimize effects of the defocusing asymmetrical transverse fields. For this reason, symmetrical pole pieces, or shims, are used between each pair

---

* Decimal Classification: R138.312.

[1] J. R. Pierce, "Spatially Alternating Magnetic Fields for Focusing Low Voltage Electron Beams," *Jour. Appl. Phys.*, Vol. 24, p. 1247, 1953.

[2] A. M. Clogston and H. Heffner, "Focusing of an Electron Beam by Periodic Fields," *Jour. Appl. Phys.*, Vol. 25, pp. 436-447, April, 1954.

[3] J. J. Mendel, C. F. Quate, and W. H. Yocom, "Electron Beam Focusing with Periodic Permanent Magnet Fields," *Proc. I.R.E.*, Vol. 42, pp. 800-810, May, 1954.

[4] K. K. N. Chang, "Beam Focusing by Periodic and Complementary Fields," *Proc. I. R. E.*, Vol. 43, pp. 62-71, January, 1955.

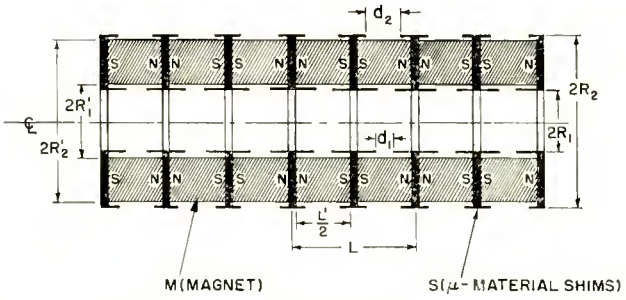


Fig. 1—Periodic permanent magnet.

of magnets to serve as equipotential boundaries and to ensure the field symmetry. These shims will also have considerable effect on the peak value of the field along the axis.

In actual practice, it may be desirable that the magnetic field along the axis be adjustable for different focusing conditions. To achieve this, an adjustable magnetic shim can be placed outside the magnet, thus enabling one to change the peak field.

Magnetic Field along the Axis

The physical picture of the periodic magnet is shown in Figure 1. M's are permanent magnets of ring shape which, after having been individually magnetized in the axial direction, are assembled between shims with polarities as illustrated. The magnetic field between the axis and the inner surface of the rings (i.e., $R < R_1$) will be computed by analysis of the model illustrated in Figure 2. The upper part of Figure 2 indicates the assumed periodic equipotential surfaces formed by the shims, which are assumed to have negligible thickness. The resulting magnetic scalar potential on this cylindrical surface $R = R_1$, is shown in the lower part of the figure. A linearly varying potential

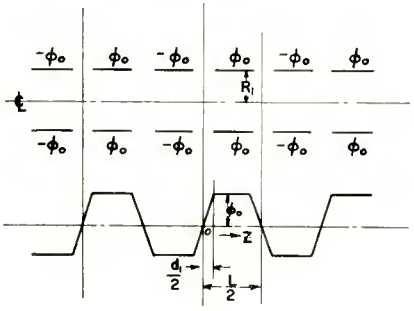


Fig. 2—Potential distribution on the boundary of the periodic magnet shown by Figure 1.

is assumed between two adjacent shims. With this axially symmetrical cylindrical geometry, the potential $\phi_1(z,r)$ inside can be obtained by solving Laplace's equation.[5] It is

$$\phi_1(z,r) = \sum_{n=1}^{\infty} A_{n1} \sin\left(\frac{2n\pi}{L} z\right) J_0\left(\frac{i2n\pi}{L} R\right) \tag{1}$$

With the boundary condition given by Figure 2, the coefficient A_{n1} can be determined:

$$A_{n1} = \frac{1}{J_0\left(\frac{i2n\pi}{L} R_1\right)} \frac{2}{L} \left[\int_0^{d_1/2} \frac{2\phi_0}{d} z \sin\left(\frac{2n\pi}{L} z\right) dz + \right.$$

$$\int_{d_1/2}^{1/2\,(L-d_1)} \phi_0 \sin\left(\frac{2n\pi}{L} z\right) dz + \int_{1/2\,(L-d_1)}^{1/2\,(L+d_1)} \frac{2\phi_0}{-d}\left(z - \frac{L}{2}\right) \sin\left(\frac{2n\pi}{L} z\right) dz$$

$$+ \int_{1/2\,(L+d_1)}^{L-d_1/2} -\phi_0 \sin\left(\frac{2n\pi}{L} z\right) dz + \int_{L-d_1/2}^{L} \frac{2\phi_0}{d} (z-L)$$

$$\left. \sin\left(\frac{2n\pi}{L} z\right) dz \right] \tag{2}$$

where ϕ_0 is the potential on the shims.

Let
$$\frac{d_1}{L} = \sigma_1 \tag{3}$$

where d_1 is the shim separation, and L is the period.

Then
$$A_{n1} = \frac{2\phi_0}{J_0\left(\frac{i2n\pi}{L} R_1\right)} \frac{1}{\sigma_1 (n\pi)^2} \left[(1 - \cos n\pi) \sin (\sigma_1 n\pi) \right] \tag{4}$$

The axial component of magnetic field, B_z, is obtained by differentiating the magnetic scalar potential of Equation (1) with respect to axial distance z:

---

[5] V. K. Zworykin and G. A. Morton, *Electron Optics and the Electron Microscope*, John Wiley & Sons, New York, N. Y.

$$B_z = \mu_0 \frac{\partial \phi_1}{\partial z}$$

$$= \mu_0 \sum_{n=1}^{\infty} \frac{4\phi_0 (1 - \cos n\pi)}{L J_0 \left(\frac{i2n\pi}{L} R_1 \right)} \frac{\sin (\sigma_1 n\pi)}{\sigma_1 n\pi} \cos \left(\frac{2n\pi}{L} z \right) J_0 \left(\frac{i2n\pi}{L} R \right). \quad (5)$$

$$0 \leq R < R_1 .$$

The magnetic field along the axis B_{z0} $(R = 0)$ is then

$$B_{z0} = \mu_0 \sum_{n=1}^{\infty} \frac{4\phi_0 (1 - \cos n\pi)}{L J_0 \left(\frac{i2n\pi}{L} R_1 \right)} \frac{\sin (\sigma_1 n\pi)}{\sigma_1 n\pi} \cos \frac{2n\pi}{L} z. \quad (6)$$

The total flux ψ_1 across the plane $z = 0$ inside the ring is:

$$\psi_1 = \int_0^{R_1} B_z \, 2\pi R dR$$

$$= \mu_0 \sum_{n=1}^{\infty} \frac{-iJ_1 \left(\frac{i2n\pi}{L} R_1 \right)}{J_0 \left(\frac{i2n\pi}{L} R_1 \right)} \frac{\sin \sigma_1 n\pi}{\sigma_1 n\pi} \frac{4\phi_0 (1 - \cos n\pi)}{n} R_1 . \quad (7)$$

Stray Magnetic Field

The stray or non-useful magnetic flux will have great influence upon the available peak field along the axis. The outside radius of the ring magnet will be determined by the total magnetic flux required, usable and non-usable. The stray magnetic field dealt with here is that due to the flux outside the ring magnets (i.e. $R > R_2$) shown in Figure 1. The potential $\phi_2 (z, r)$ outside the ring should vary radially as the modified Bessel function of the second kind of zero order: $iH_0^{(1)} (2\pi niR/L)$ and is equal to

$$\phi_2 (z, r) = \sum_{n=1}^{\infty} A_{n2} \sin \left(\frac{2n\pi}{L} z \right) \left[iH_0^{(1)} \left(\frac{i2n\pi}{L} R \right) \right]. \quad (8)$$

Here A_{n2} as determined in the same way as A_{n1} is equal to

$$A_{n2} = \frac{2\phi_0}{iH_0^{(1)}\left(\frac{i2n\pi}{L} R_2\right)} \frac{\sin(\sigma_2 n\pi)}{\sigma_2 (n\pi)^2} (1 - \cos n\pi) \tag{9}$$

where $\sigma_2 = d_2/L$. (10)

The external axial magnetic field B_z is similarly obtained by differentiating Equation (7) with respect to the axial distance z:

$$B_z = \mu_0 \frac{\partial \phi_2}{\partial z}$$

$$= \sum_{n=1}^{\infty} \frac{4\phi_0 (1 - \cos n\pi)\ \mu_0 \sin(\sigma_2 n\pi) \cos(2\pi nz/L)\ iH_0^{(1)}(2\pi niR/L)}{LiH_0^{(1)}(2\pi niR_2/L)\ \sigma_2 n\pi} \tag{11}$$

$R_2 < R < \infty$.

The total flux ψ_2 across the plane $z = 0$ outside the ring magnet is

$$\psi_2 = \int_{R_2}^{\infty} B_z 2\pi R dR$$

$$= \mu_0 \sum \frac{-H_1^{(1)}\left(\frac{i2n\pi}{L} R_2\right)}{iH_0^{(1)}\left(\frac{i2n\pi}{L} R_2\right)} \frac{\sin \sigma_2 n\pi}{\sigma_2 n\pi} \frac{4\phi_0 (1 - \cos n\pi)}{n} R_2. \tag{12}$$

Permanent Magnet Materials

The design of a periodic magnetic field is not complete until we know the operating characteristics of the permanent magnet which produces the fields computed above. The length of magnet (equal to $L'/2$ shown in Figure 1), capable of delivering the required magnetomotive potential ϕ_0, is[6]

$$\frac{L'}{4} = \frac{\phi_0}{H_m} \tag{13}$$

[6] H. F. Olson, *Elements of Acoustical Engineering*, D. Van Nostrand Co., Inc., New York, N. Y., pp. 183.

where H_m is equal to the operating demagnetizing force on the demagnetization characteristic of the permanent-magnet materials. Assuming that flux flow within each ring is everywhere axial, the required cross-sectional area of the permanent magnet is the total flux $\psi_1 + \psi_2$ (given by Equations (7) and (12)) divided by the magnetic induction B_m corresponding to H_m on the demagnetization characteristic; i.e.

$$q\pi\ (R_2^2 - R_1^2) = \frac{\psi_1 + \psi_2}{B_m}, \quad q = \frac{R_2'^2 - R_1'^2}{R_2^2 - R_1^2} \tag{14}$$

where q is a correction factor arising from the fact that the cross section of the magnet with radii R_1' and R_2' is different from that of geometrical ring with radii R_1 and R_2. It is seen from Figure 1 that q is always less than 1.

The demagnetization curves of various permanent magnet materials are shown in Figure 3. The materials are Magnadur,[7] Ferrimag,[7] a special RCA ceramic,[8] and Bismanol.[9] The B-H characteristics of these materials shown by Figure 3(a) are fairly linear and also have relative permeabilities in the neighborhood of two. Alnico V is not shown here because it is not very good for producing a periodic field. Its low coercive force (690 oersteds) makes it very sensitive to self-demagnetization during assembly of the periodic structure.

Design Curves for Periodic Magnets

To construct design curves for periodic magnets, let us substitute ϕ_0 expressed by Equation (13) into Equations (6), (7), (12), and (14), making the following normalizations:

$$\overline{R}_1 = \frac{R_1}{L} \tag{15}$$

$$\overline{R}_2 = \frac{R_2}{L} \tag{16}$$

$$\overline{\hat{B}}_{z0} = \frac{\hat{B}_{z0}}{\mu_0 H_m} \tag{17}$$

[7] J. J. Went, G. W. Rathenau, E. W. Gorter, and G. W. Van Oosterhaut, "Ferroxdure, a Class of New Permanent Magnet Materials," *Philips Technical Review*, Vol. 13, No. 7, January, 1952.

[8] Unpublished report.

[9] Edmond Adams, "A New Permanent Magnet from Powdered Manganese Bismuthide," *Rev. Mod. Phys.*, Vol. 25, pp. 306-309, January, 1953.

$$\overline{B}_m = \frac{B_m}{\mu_0 H_m} \tag{18}$$

$$\bar{\psi}_1 = \frac{\psi_1}{4\mu_0 H_m p L^2} \tag{19}$$

$$\bar{\psi}_2 = \frac{\psi_2}{4\mu_0 H_m p L^2}, \tag{20}$$

where $\bar{B}_{z0}$ is the peak field along the axis and

$$p = \frac{L'}{2L}. \tag{21}$$

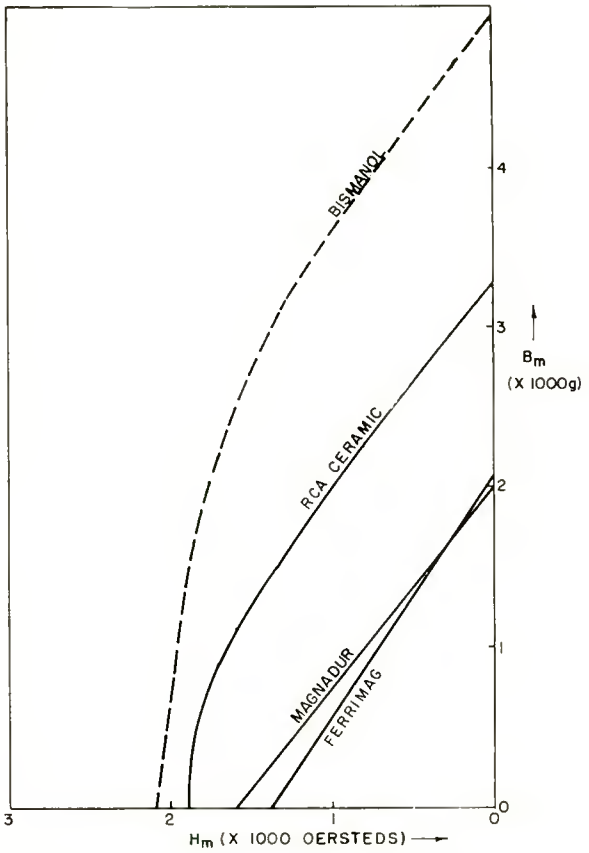


Fig. 3(a)—*B–H* demagnetization characteristics of permanent magnets.

Then, one obtains

$$\bar{B}_{z0} = 2p \sum_{n=1}^{\infty} \frac{(1 - \cos n\pi)}{J_0(i2n\pi\bar{R}_1)} \frac{\sin(\sigma_1 n\pi)}{\sigma_1 n\pi} \tag{22}$$

$$\bar{\psi}_1 = \frac{1}{2} \sum_{n=1}^{\infty} \frac{iJ_1(i2n\pi\bar{R}_1)}{J_0(i2n\pi R_1)} \frac{\sin(\sigma_1 n\pi)}{\sigma_1 n\pi} \frac{(1 - \cos n\pi)}{n} \bar{R}_1 \tag{23}$$

$$\bar{\psi}_2 = \frac{1}{2} \sum_{n=1}^{\infty} \frac{-H_1^{(1)}(i2n\pi R_2)}{iH_0^{(1)}(i2n\pi\bar{R}_2)} \frac{\sin(\sigma_2 n\pi)}{\sigma_2 n\pi} \frac{(1 - \cos n\pi)}{n} \bar{R}_2 \tag{24}$$

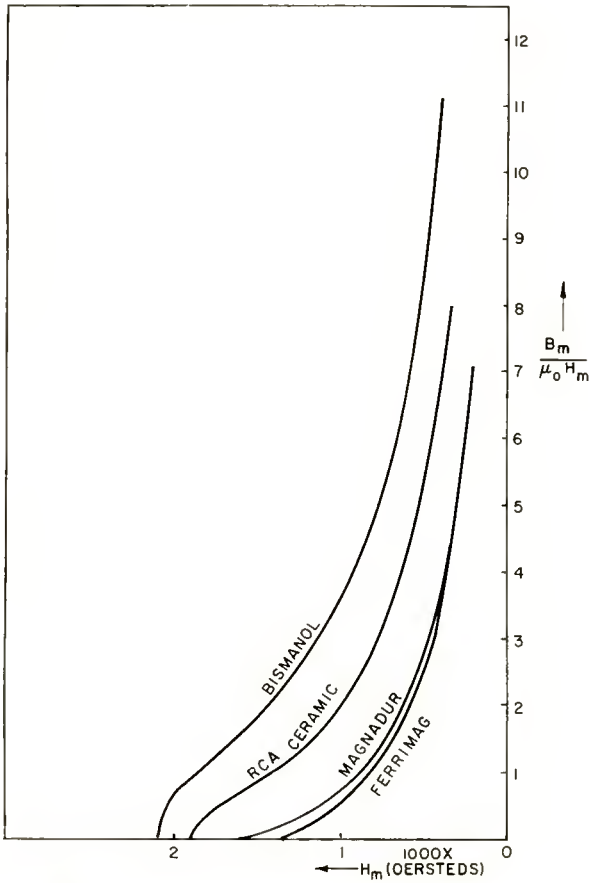


Fig. 3(b)— $\frac{B_m}{\mu_0 H_m}$ versus H_m curves of different permanent magnets.

$$\bar{\psi}_1 + \bar{\psi}_2 = \frac{\overline{B}_m \pi}{4} \frac{q}{p} (\overline{R}_2^2 - \overline{R}_1^2). \tag{25}$$

Equation (22) gives $\frac{\overline{B}_{z0}}{p}$ as a function of $\overline{R}_1$. This is plotted in Figure 4 for different values of σ_1. The relationship between $\overline{R}_1$ and $\overline{R}_2$ can be found graphically by solving Equations (23), (24), and (25)

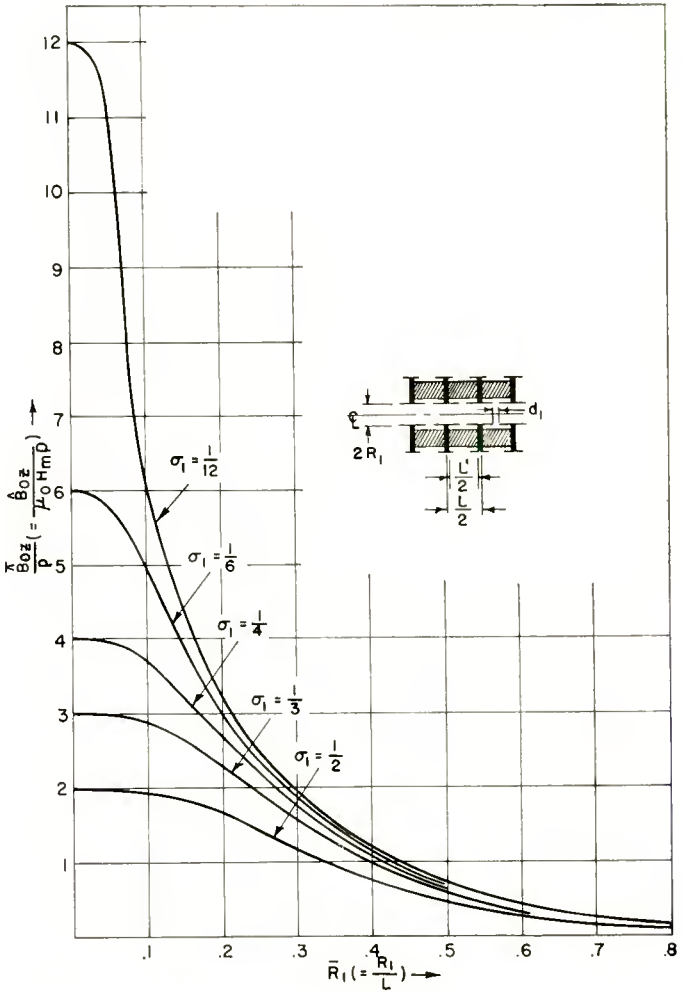


Fig. 4—The peak magnetic field along the axis $\hat{B}_{z0}$ as a function of the demagnetizing force H_m of the magnet, the normalized radius $\overline{R}_1$, the shimming parameter σ_1, and the period correction factor p.

simultaneously. To do so, first plot $\bar{\psi}_2$ as a function of $\overline{R}_2$, according to Equation (24), then substitute sets of the obtained values of $\bar{\psi}_2$ and $\overline{R}_2$ in Equation (25), and, lastly, plot $\bar{\psi}_1$ as a function of $\overline{R}_1$ from Equations (23) and (25). The intersections of these two ψ_1 curves yield values of $\overline{R}_2$ corresponding to those of $\overline{R}_1$. The $\overline{R}_2$-versus-$\overline{R}_1$ curves are plotted in Figure 5, with different values of $\overline{B}_m$ and σ_1. The plot of Figure 5 shows only the case $\sigma_2 = 1/2$, which is its optimum value. That it is optimum is shown in Figure 6, in which $\overline{R}_1$ is plotted

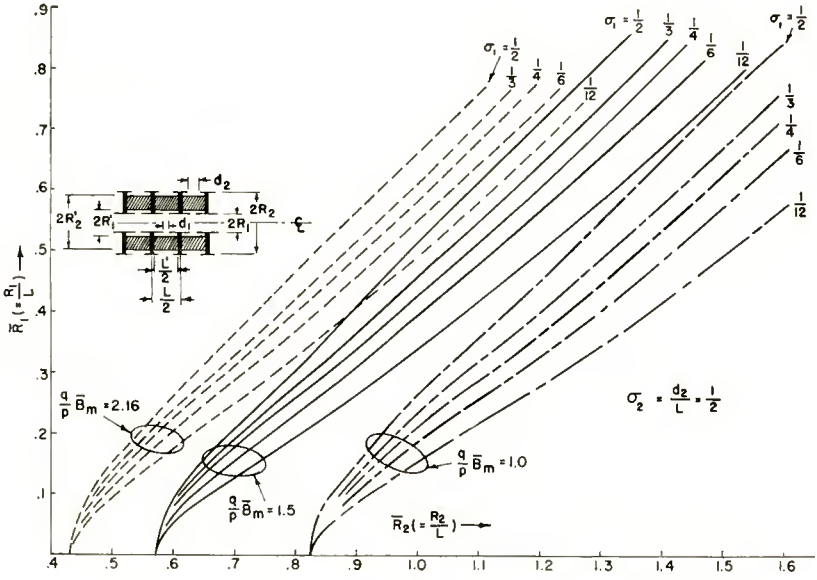


Fig. 5—The inner radius of the magnet R_1 as a function of the outer radius R_2, the period L, the shimming parameter σ_1, the normalized magnet field $\overline{B}_m$, the period correction factor p and the cross section correction q.

versus $\overline{R}_2$ *at* $(q/p)\ \overline{B}_m = 1.5$ for different values of σ_1 and σ_2. Figure 6 indicates that a larger value of $\overline{R}_2$ must be chosen with a value of σ_2 less than $1/2$.

In Figure 7, the value of $\overline{R}_2^2 - \overline{R}_1^2$ is plotted against $\overline{R}_1$ for different values of $\overline{B}_m$ and σ_1. The figure reveals how the cross-sectional area (or the weight) of the magnet varies as a function of $\overline{R}_1$.

In computing the summation

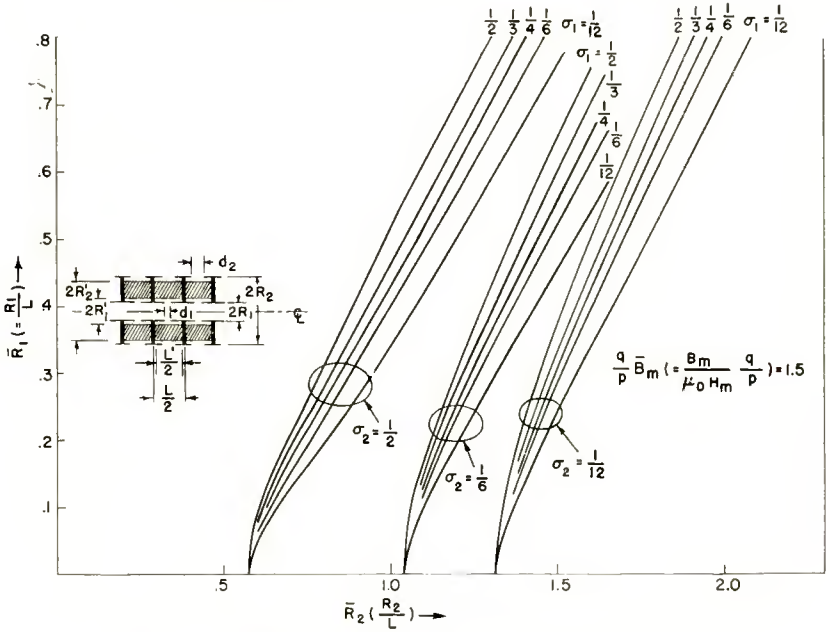


Fig. 6—$\bar{R}_1$ versus $\bar{R}_2$ curves with different values of σ_1 and σ_2.

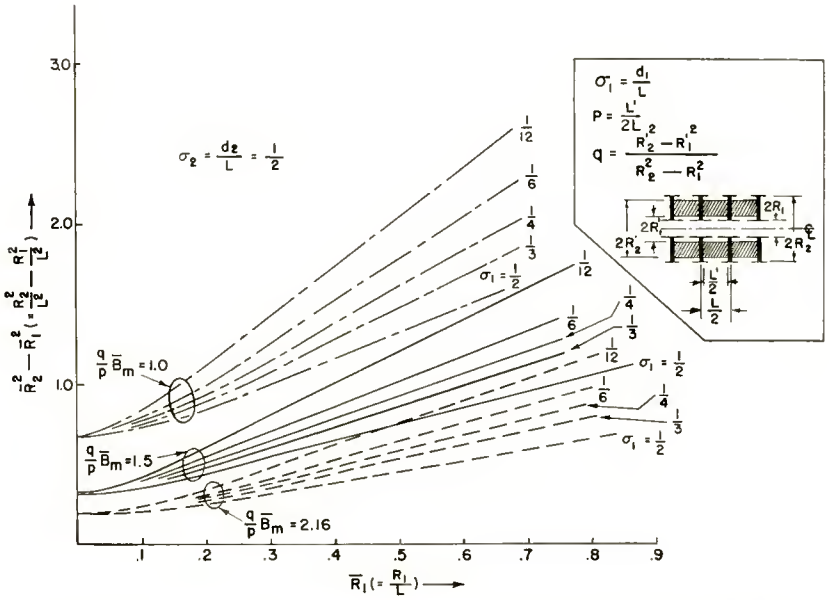


Fig. 7—The weight of the magnet as a function of $\bar{R}_1$, σ_1 and B_m.

$$\sum \frac{1}{J_0(i2n\pi\bar{R}_1)} \frac{\sin \sigma_1 n\pi}{\sigma_1 n\pi},$$

which enters into Equation (22) for $\bar{B}_{z0}$, more terms should be taken at small values of $\bar{R}_1$ and σ_1 to insure the accuracy of the summation.

To demonstrate how one applies the above design curves, a few examples are shown in the following table. For comparison, measured results are also shown in the last row.

Table I

Magnet Material	"Magnadur"*	"Ferrimag"** I	"Ferrimag" II	Special Ceramics
$2R_1$ (in.)	1.0	.400	.400	.456
$2R_2$ (in.)	3.125	1.50	1.50	1.0
L (in.)	1.25	.75	.625	.50
σ_1	.454	.5	.5	.5
σ_2	.454	.5	.5	.5
$\bar{R}_1 = R_1/L$	.40	.266	.32	.456
$\bar{R}_2 = R_2/L$	1.25	1.00	1.2	1.0
p	.45	.5	.5	.5
q	.8	1.0	1.0	1.0
$\frac{q}{p}\bar{B}_m = \frac{q}{p}\frac{B_m}{\mu_0 H_m}$ (from Figure 5)	.90	1.05	.7	1.40
$\frac{\hat{B}_{z0}}{p\mu_0 H_m}$ (from Figure 4)	.75	1.35	1.1	.6
H_m (oesteds) (from Figure 3b)	1100	1040	1140	1580
$\hat{B}_{z0}$ (gausses computed)	372	710	630	474
$\hat{B}_{z0}$ (gausses measured)	340	740	640	490

The table reveals that the measured values of $\hat{B}_{z0}$ are in good agreement with those computed.

In actual practice, for minimum weight and for best efficiency, the geometry of the magnet should be chosen such that the magnet operates at the point on the demagnetization curve where the magnet delivers the maximum energy. Figure 8 shows design curves for

* Registered Trade Mark of the Ferrocube Corp.
** Registered Trade Mark of the Crucible Steel Co.

(q/p) B_m equal to 2.85. This particular value is chosen simply because it is a reasonable value for a practical geometry of most of the commercial ferrite materials operating at the maximum energy point. A numerical example below illustrates the method of using the design curves.

Design Procedure

Step 1. Given: the peak field of amplitude $\bar{B}_{z0}$ and of period L (as obtained from References (3) and (4)), the magnet material employed

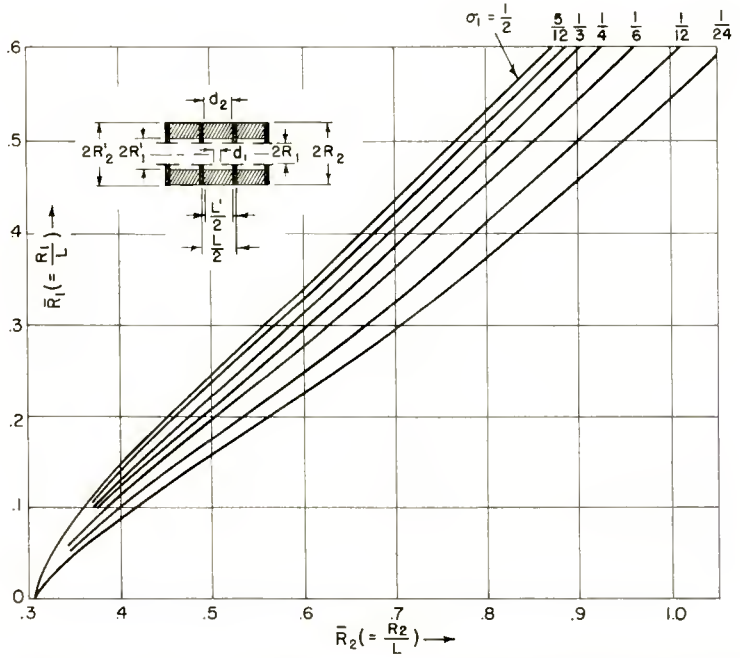


Fig. 8—The inner radius of the magnet R_1, as a function of the outer radius R_2, the period L, and the shimming parameter σ_1. The value of (q/p) B_m has been fixed at 2.85; that of σ_2, 5/12, where $\bar{B}_m = B_m/\mu_0 H_m$, $P = L'/2L$, $q = (R_2'^2 - R_1'^2)/(R_2^2 - R_1^2)$, $d_2 = L'/2$, and $R_2 = R_2'$.

and the length of the magnet $(L'/2)$. It is recommended that for good design the part of the shim which separates each pair of magnets should be made as thin as possible, such that L' is almost equal to L. First determine the value of $\bar{B}_{z0}/\mu_0 H_m p$. Suppose

$$\bar{B}_{z0} = 600 \text{ gausses},$$

$$p = \frac{L'}{2L} = \frac{d_2}{L} = .417,$$

and the RCA ceramic magnet is used. The maximum energy point is (from Figure 3)

$$\overline{B}_m = 1.5$$

and

$$H_m = 1190 \text{ oersteds}$$

then

$$\frac{\hat{B}_{z0}}{\mu_0 H_m p} = \frac{600}{1190 \times .417} = 1.2.$$

Step 2. Choose the value of the shim separation d_1, keeping in mind that the smaller the shim separation, the less the amount of possible asymmetrical transverse field. From Figure 4 find the inner radius of the magnet (R_1) for the given values of σ_1 ($= d_1/L$) and $\hat{B}_{z0}/\mu_0 H_m p$.

Suppose

$$\sigma_1 = \frac{d_1}{L} = \frac{1}{12}$$

then

$$\overline{R}_1 = \frac{R_1}{L} = .40.^*$$

Step 3. Estimate the correction factor q which arises from the fact that the cross section of the magnet is smaller than that of the assumed ring with radii R_1 and R_2. With known values of R_1, p, q, σ_1, σ_2, and $\overline{B}_m$, a corresponding value of the outer radius of the magnet is obtained by using Figure 8.

Suppose

$$q = \frac{R_2'^2 - R_1'^2}{R_2^2 - R_1^2} = .80$$

then

$$\frac{q}{p} B_m = \frac{.80}{.417} \times 1.5 = 2.85$$

* Sometimes the calculated value of R_1 may be too small to allow the insertion of the glass tube envelope inside the magnet. In that case, it is not possible to design the magnet at the maximum energy point. Instead of using the value of H_m at the maximum energy point, a larger value must be chosen. Then find the outer radius by using curves in Figures (5) and (6).

and from Figure 8,

$$\bar{R}_2 = .79.$$

In some instances, however, the designer may be interested only in realizing a large peak field without considering the efficiency of the magnet. By assuming a linear demagnetization curve (which is a very good approximation for most of the commercial magnet material such as Magnadur, Ferrimag, etc.), one is able to express the peak field ($\bar{B}_{z0}$) as a function of the shim separation (d_1) with given dimensions of the magnet of given material.

Let
$$H_m = -\frac{B_m}{\mu} + H_c \tag{26}$$

where H_c is the coercive force of the magnet and μ is the slope of the demagnetization curve. By combining Equations (25), (6), (13), and (14), one obtains

$$\bar{B}_{z0} = \frac{2\mu_0 p H_c \sum_{n=1}^{\infty} \frac{(1-\cos n\pi)}{J_0(i2n\pi\bar{R}_1)} \frac{\sin \sigma_1 n\pi}{\sigma_1 n\pi}}{1 + \frac{2\mu_0 p}{q\mu\pi\,(\bar{R}_2^2 - \bar{R}_1^2)} \sum_{n=1}^{\infty} \left(\frac{1-\cos n\pi}{n} \right) \cdot \Psi} \tag{27}$$

where $\Psi = \left[\frac{-iJ_1(i2n\pi\bar{R}_1)}{J_0(i2n\pi\bar{R}_1)} \frac{\sin \sigma_1 n\pi}{\sigma_1 n\pi} \bar{R}_1 + \frac{-H_1^{(1)}(i2n\pi\bar{R}_2)}{iH_0^{(1)}(i2n\pi\bar{R}_2)} \frac{\sin \sigma_2 n\pi}{\sigma_2 n\pi} \bar{R}_2 \right]$

As a specific example, assume

$$\frac{4\mu_0 p}{\mu q \pi} = .55$$

and
$$\sigma_2 \pi = 1.31$$

which are reasonable values for periodic magnets of practical geometry and of commercially available material. $\bar{B}_{z0}$ can then be plotted as a function of $\bar{R}_1$ and $\bar{R}_2$ as shown by Figure 9. It is interesting to note from Figure 9 that for every set of $\bar{R}_1$ and $\bar{R}_2$, there is an optimum value of the ratio (σ_1) of the shim separation d_1 to the period L which yields a maximum peak field. This optimum value of d_1/L increases

with decreasing $\overline{R}_2$ and with increasing $\overline{R}_1$. That σ_1 has an optimum value has been experimentally verified in a periodic magnet assembly which has been designed and tested.† This magnet, constructed of "Indox*," has the dimensions $\overline{R}_2 = 1.69$, $\overline{R}_1 = .368$, $\sigma_2\pi = 1.31$, $L = .720$ inch, $L' = .610$ inch. The measured optimum peak field is 620 gausses found at $\sigma_1 = .080$. The computed value, using Equation (27) and Figure 9 is 660 gausses, at $\sigma_1 = .081$.

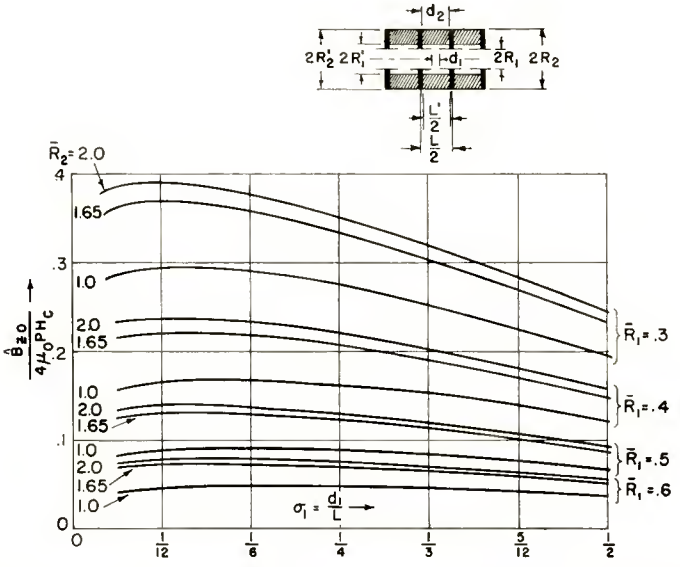


Fig. 9—The normalized peak magnetic field $\hat{B}_{z0}/4\mu_0 pH_c$ as a function of the shim separation d_1, the period L, the inner radius R_1, and the outer radius R_2. The value of $4\mu_0 p/\mu q\pi$ has been fixed at 0.55; that of σ_2, 5/12.

Summary and Conclusions

The design curves of a periodic magnet discussed in the preceding sections are calculated under the assumptions of a linearly magnetic scalar potential on the boundary $R = R_1$ and of no leakage flux internal to an individual magnet. Also, it is assumed that the shims which form the boundary have negligible thickness. Nevertheless, the calculated values check well with the measured ones.

As shown by the curves, the magnetic field along the axis is a

† This work was carried out by the Traveling-Wave Tube Advanced Development Group of the RCA Tube Division, Harrison, N. J.

* Trade Mark of the Indiana Steel Co.

function of the inner and the outer radii of the magnet, the length of the magnet, and the B-H characteristic of the magnet material. For a given peak magnetic field along the axis, and with given magnetic scalar potential on the shims, smaller shims (i.e., large σ_1) result in smaller radii of the magnet and thus in less weight (see Figures 4 and 5). On the other hand, for constant weight and with given operating point on the demagnetization curve of the magnet, it is seen from Figure 7 that the smaller ratio of the inner radius to the period and the smaller the shim separation, σ_1, the stronger is the peak field which will be developed along the axis. In practical cases, however, the inner radius of the magnet, fixed by the glass envelope of the traveling wave tube, has a minimum limit. In addition, the period of the magnet has an upper limit, determined by the stability criteria of a given electron flow.[3, 4] Thus, the ratio of the inner radius of the magnet to the period of the magnet has a minimum value. If, then, the designer prefers heavy shims in order to minimize the asymmetrical transverse magnetic field for focusing purposes, he has to sacrifice either the low weight or the large peak field. In any case, for good design it is always recommended that the geometry of the magnet should be chosen such that the magnet operates at the maximum energy point.

As is also indicated by the curves, for the design of a small magnet capable of producing a large peak field along the axis, the choice of a magnet material of a high B-H energy product is always desirable. Magnets with larger H_m at a given ratio of B_m to $\mu_0 H_m$ would generate more flux density (see Figures 4 and 5). This fact reveals that Bismanol (see Figures 3(a) and 3(b)) is the best material of those tested, for constructing a small periodic magnet with a strong field.

For periodic magnets possessing nearly linear demagnetization characteristic, the shim separation (d_1) has an optimum value for a maximum peak field as shown in Figure 9. These curves allow the designer to construct a magnet delivering the largest peak field with given magnetic material.

STUDIES OF EXTERNALLY HEATED HOT CATHODE ARCS*

Part III — Plasma Density Distributions in the Anode-Glow Mode

By

E. O. Johnson† and W. M. Webster‡

Summary—*In Part I of this series the qualitative aspects of the various discharge modes were treated. In Part II the ion generation and loss relations in the anode-glow mode were analyzed and compared with experiment. Experimental evidence was presented on the form of the plasma density distributions. In this paper these distributions are treated analytically and criteria presented for the conditions which delimit the existence of the anode-glow mode of discharge. In particular, the density distributions are shown to be expressible in terms of lowest mode solutions of Laplace's equation. Distributions corresponding to different values of geometric and gas parameters are illustrated. In addition, an electrostatic analogue solution is presented to show how errors can arise in plasma density determinations when probes are introduced into a tube.*

From the analysis of the density distributions it is shown that the anode-glow mode is capable of existence only when (1) the gas pressure lies between certain limits, (2) the tube geometry is appropriate, and (3) the emission capability of the cathode is at least twice as large as the anode current. Experimental evidence to support this analysis is presented along with observations on the transition out of the anode-glow mode.

I. Introduction

PART I[1] of this series gave a qualitative description of the various possible modes of operation assumed by a hot-cathode arc. Such aspects as the potential distribution, volt–ampere characteristics, visible appearance of the discharge, and electron temperatures were presented. Part II[2] was concerned with a detailed study of one of these modes, the anode-glow (AG) mode. This mode turned out to be particularly convenient to analyze and treat experimentally,

---

* Decimal Classification: R337.1.

‡ Formerly, Research Laboratory, RCA Laboratories, Princeton, N. J.; now Tube Division, Radio Corporation of America, Harrison, N. J.

† Research Laboratory, RCA Laboratories, Princeton, N.J.; presently on leave of absence at the Swiss Federal Institute of Technology, Zurich, Switzerland.

[1] L. Malter, E. O. Johnson, and W. M. Webster, "Studies of Externally Heated Hot Cathode Arcs, Part I—Modes of the Discharge," *RCA Review*, Vol. XII, pp. 415-435, September, 1951.

[2] W. M. Webster, E. O. Johnson and L. Malter, "Studies of Externally Heated Hot Cathode Arcs, Part II—The Anode-Glow Mode," *RCA Review*, Vol. XIII, pp. 163-182, June, 1952.

mainly because the sources and sinks for the ions and electrons are geometrically separable. The processes of ion generation and loss in a cylindrical diode were treated analytically and then equated to obtain an expression for the volt–ampere characteristic. Experimental data which supports this analysis and illustrates the form of the plasma density distribution was also presented in Part II.

This paper is devoted to an analytical treatment of the plasma density distributions in the AG mode. The pertinent physical aspects

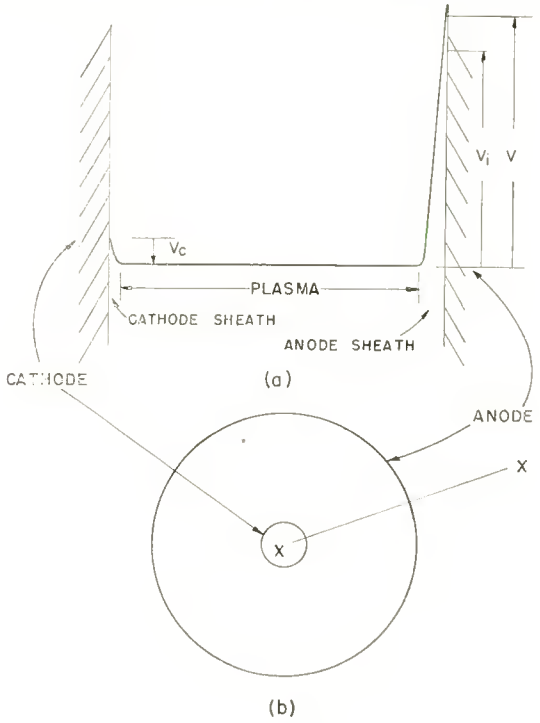


Fig. 1—Anode-glow discharge
a. Potential distribution.
b. Cross section of tube geometry.

can be discussed with reference to the potential distribution of the AG mode which is as shown in Figure 1a. This distribution corresponds to that in the cross section x-x of the cylindrical tube geometry shown in Figure 1b. The slight potential depression V_c at the cathode is a necessary consequence of the cathode having a total available emission many times greater than the circuit current. The only electrons that get into the plasma enter through the cathode sheath. Since they enter the plasma against a retarding field, it is to be expected that the temperature of the electrons in the plasma would be

the same as their temperature when they left the cathode. Probe studies have shown this to be the case. The electrons from the cathode drift through the plasma to its anode edge because of plasma density gradients and small electric fields. From the anode edge of the plasma the electrons are accelerated in the anode electron sheath, and when their energies exceed the ionization potential of the gas atoms, V_i, some of these electrons generate ions. Others generate excited atoms which give rise to the "anode glow" from which the name of this discharge mode was taken. Because the anode sheath thickness is small compared to the electron mean free paths for excitation and ionization, most of the electrons give all of their energy to the anode where it is dissipated as heat. For example, in helium about one electron in three hundred produces an ion.

The ions generated in the sheath fall back into the plasma and sustain its density against diffusion losses at the ends of the structure and, as pointed out in Part II, to the cathode surface. The losses to the end plates and cathode must be exactly equal to the generated ion current in the anode sheath. In Part II the analysis of this ion balance was shown to be in good agreement with experiment if the fact that ion losses do occur at the cathode is taken into account.

In the structures studied, the volume recombination losses within the plasma are insignificant compared to the diffusion losses of the particles to the ends. Because of the low electron energies within the plasma (mean energy of about 0.1 electron volt) there is no appreciable ion generation in the plasma. Both of these conditions help make the AG plasma an ideal one to analyze. This analysis of cylindrical structures, carried out in Section II of this paper, consists of setting up and solving the particle flow and continuity equations subject to the simple boundary conditions suggested by the above discussion. The usual simplifying assumption of equal ion and electron densities within the plasma is used. Also in Section II an electrostatic analogous solution of a plasma density distribution is presented and used to show what can happen when plasma density measurements are attempted with probes.

In Section III, the existence criteria of the AG mode are developed from the density distributions, and it is shown that these criteria involve conditions at the cathode as well as at the anode. The cathode conditions have to do with electron emission capability, cathode diameter with respect to anode diameter, and gas pressure. Anode conditions have to do with gas pressure and the aspect ratio of the tube, i.e., the ratio of anode diameter to the axial length of the tube. Experimental evidence for the analytical picture is introduced.

In Section IV, transitions out of the AG mode into the Langmuir

and ball-of-fire modes are briefly discussed. This discussion includes the sequence of events which leads to the transition and the new boundary conditions which exist after the transition. Several experimental observations to support the discharge model are presented.

II — Plasma Density Distribution

The relations that express the behavior of the ions and electrons in the AG plasma are the particle flow equations

$$\Gamma_e = -n\mu_e E - D_e \operatorname{grad} n, \tag{1}$$

$$\Gamma_p = n\mu_p E - D_p \operatorname{grad} n, \tag{2}$$

and the equations of continuity for the particles

$$\operatorname{div} \Gamma_e = 0 \tag{3}$$

$$\operatorname{div} \Gamma_p = 0 \tag{4}$$

where Γ is the particle flow rate density, μ is the mobility, E is the electric field, and D is the diffusion coefficient. The subscripts e and p refer to electrons and ions, respectively. The usual assumption that the electron density n_e is everywhere equal to the ion density n_p, i.e., $n_e = n_p = n$, has been made. This assumption greatly simplifies the analysis and can be expected to introduce little error for the plasma densities under discussion except, possibly, near the boundary regions where a plasma-sheath transition occurs. In the steady-state AG plasma there is no ion generation and insignificant volume recombination so that the right-hand sides of Equations (3) and (4) are justifiably zero.

If the electric field is eliminated from Equations (1) and (2), one obtains

$$\mu_p \Gamma_e + \mu_e \Gamma_p = -(\mu_p D_e + \mu_e D_p) \operatorname{grad} n.$$

Since the ambipolar diffusion coefficient D_a is defined by

$$D_a = \frac{\mu_p D_e + \mu_e D_p}{\mu_p + \mu_e},$$

or (since $\mu_e >> \mu_p$) by

$$D_a \cong \frac{\mu_p D_e + \mu_e D_p}{\mu_e},$$

ne obtains

$$\Gamma_p + \frac{\mu_p}{\mu_e}\Gamma_c = -D_a \operatorname{grad} n. \tag{5}$$

By taking the divergence of both sides of this Equation, one finally obtains Laplace's equation

$$\triangle^2 n = 0. \tag{6}$$

For the cylindrical structure with coordinates and dimensions as shown in Figure 2, the general solution to Equation (6) is

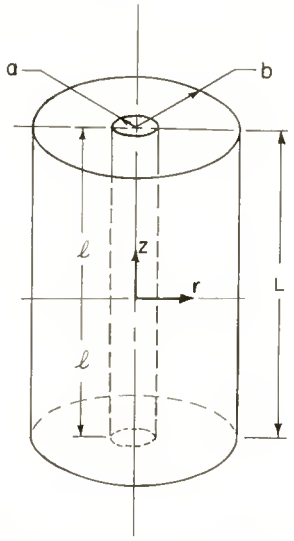


Fig. 2—Geometry and coordinate system of cylindrical diode used for analysis of plasma density distributions.

$$n = \sum_{m=o}^{\infty} [A_m J_o(jmKr) + B_m j H_o^1(jmKr)]$$
$$[c_m \cos(mKz) + D_m \sin(mKz)] \tag{7}$$

Here A_m, B_m, C_m, and D_m are constants of integration, J_o is a zero-order Bessel function and H_o^1 is a zero-order Hankel function. Both J_o and H_o^1 have imaginary arguments. If the symmetry of the structure in Figure 2 is considered, and it is assumed that the plasma density is essentially zero at each end plate, then $D_m = 0$ and $K = \pi/L$. Further, if the amplitudes of the higher density modes are negligible, as experiment[2] indicates, Equation (7) reduces to:

$$n = \left[A_o J_o\left(j\frac{\pi}{L} r \right) + B_o j H_o^1\left(j\frac{\pi}{L} r \right) \right] \cos\frac{\pi}{L} z. \tag{8}$$

If the end-plate spacing L is large compared to the anode diameter b, then the axial solution can be dropped and the solution of Equation (6) becomes simply

$$n = A' \ln r + B', \tag{9}$$

where A' and B' are again constants which depend on boundary conditions.

Equation (9) is especially simple to evaluate since the assumption of a large value of L, compared to that of b, is equivalent to making the diffusion losses to the end plates negligible. Then, Γ_p in Equation (5) is negligible at any radius r. Consequently, for this boundary condition Equation (9) yields

$$A' = -\frac{b\,\Gamma_{eb}}{D_a}\frac{\mu_p}{\mu_e}, \tag{10}$$

where Γ_{eb} is the electron flow density at the anode surface and is introduced by the relation

$$\Gamma_e = \frac{b\,\Gamma_{eb}}{r}. \tag{10a}$$

The second boundary condition is introduced by letting the plasma density at the edge of the anode sheath[3] be n_b, where n_b is defined by the kinetic theory relation

$$\frac{n_b \bar{c}_e}{4} = \Gamma_{eb}. \tag{11}$$

Here, $\bar{c}_e$ is the mean thermal velocity of the plasma electrons. Thus, Equation (9) finally becomes

$$n = \Gamma_{eb}\left[\frac{4}{\bar{c}_e} + \frac{b}{D_a}\frac{\mu_p}{\mu_e}\ln\frac{b}{r}\right] \tag{12}$$

This shows the plasma density is everywhere directly proportional to the anode current. It also shows that the plasma density increases

[3] It can be shown that this definition of the anode edge of the plasma is nearly the same as defining the edge by $E = -\frac{kT}{e\lambda_e}$, where k is Boltzmann's constant, T is the electron temperature, e is the electronic charge, and λ_e is the electron mean free path.

as the cathode is approached, the increase being directly proportional to the gas pressure because of the inverse dependence of D_a on gas pressure. This increase in density is greater with the heavy noble gases since the quantity $\frac{1}{D_a}\frac{\mu_p}{\mu_e}$ increases with atomic mass. These conclusions are physically reasonable because the cathode is the particle (electron) source and the anode is the sink. And, just as in any analogous flow system, the gradient becomes greater as the flow impedance, in this case the factor $\frac{1}{D_a}\frac{\mu_p}{\mu_e}$, becomes greater.

For the more general case, where the ion losses are appreciable, Γ_p is not negligible and varies in value throughout the plasma. For this case the first boundary condition has to be developed in terms of the equality existing between the rate at which ions are fed into the plasma at the anode edge and the rate at which they are lost by ambipolar diffusion at the ends of the structure. The total ion flow rate into the plasma volume S, bounded by the end plates and the radius r, is

$$2\pi r\Gamma_{pro}\int_{-l}^{l}\cos\frac{\pi}{L}z\,dz = 8rl\Gamma_{pro}, \tag{13}$$

where $L = 2l$ and Γ_{pro} is the ion flow rate density in the plane $z = 0$ at the radius r. The ions are lost from this volume S at a rate given by

$$\frac{1}{\tau}\int_S n\,dv, \tag{14}$$

where τ is the mean life time of an ion in this volume. The quantity τ, applicable here because only the lowest diffusion mode is present, has the value

$$\tau = \frac{\Lambda^2}{D_a} = \frac{\left(\frac{L}{\pi}\right)^2}{D_a}, \tag{14a}$$

with $\Lambda = L/\pi$ being the characteristic diffusion length of this plasma. If the axial plasma density distribution is expressed by

$$n = n_{or} \cos \frac{\pi}{L} z,$$

with n_{or} being the density in the plane $z = 0$ at the radius r, then Equation (14) becomes

$$\frac{1}{\tau} \int_a^r \int_{-l}^{l} n_{or} \left(\cos \frac{\pi}{L} z \right) 2\pi r dr dz = \frac{8l}{\tau} \int_a^r r n_{or} dr.$$

Equating this to Equation (13) it is found that

$$\Gamma_{pro} = \frac{1}{r\tau} \int_a^r r n_{or} \, dr;$$

and from this, and Equation (10a), one obtains

$$\frac{\Gamma_{pro}}{\Gamma_{ero}} = \frac{1}{b\tau \Gamma_{ebo}} \int_a^r r n_{or} \, dr. \tag{15}$$

This relation can now be substituted, along with Equation (8), into Equation (5) to get the first relation linking the constants A and B. A second relation between A and B can be obtained, as before, by using Equation (11). Simultaneous solution of these two relations gives the values of A and B in terms of known constants. Using these values, Equation (8) for the plasma density distribution becomes

$$\frac{n}{n_b} = \frac{D_o}{\Delta} J_o \left(j \frac{r}{\Lambda} \right) + \left(\frac{F_o}{\Delta} \right) j H_o^1 \left(j \frac{r}{\Lambda} \right), \tag{16}$$

where

$$\Delta = J_o \left(j \frac{b}{\Lambda} \right) H_1^1 \left(j \frac{a}{\Lambda} \right) + j J_1 \left(j \frac{a}{\Lambda} \right) j H_o^1 \left(j \frac{b}{\Lambda} \right)$$

$$D_o = H_1^1 \left(j \frac{a}{\Lambda} \right) + \frac{b}{a} Q j H_o^1 \left(j \frac{b}{\Lambda} \right)$$

$$F_o = j J_1 \left(j \frac{a}{\Lambda} \right) - \frac{b}{a} Q J_o \left(j \frac{b}{\Lambda} \right)$$

$$Q = \Lambda \frac{\bar{c}_e}{4D_a} \frac{\mu_p}{\mu_e}$$

$$= \frac{3}{4} \Lambda \frac{1}{\lambda_e \left(1 + \frac{T_p}{T_e}\right)} . \qquad \text{(See Footnote (4))}$$

J_1 and H_1^1 are first-order Bessel and Hankel functions, respectively. The quantity Λ was previously defined by relation (14a); λ_e is the plasma electron mean free path, and T_p and T_e are the ion and electron temperatures, respectively.

The density distributions as computed from Equation (16) for

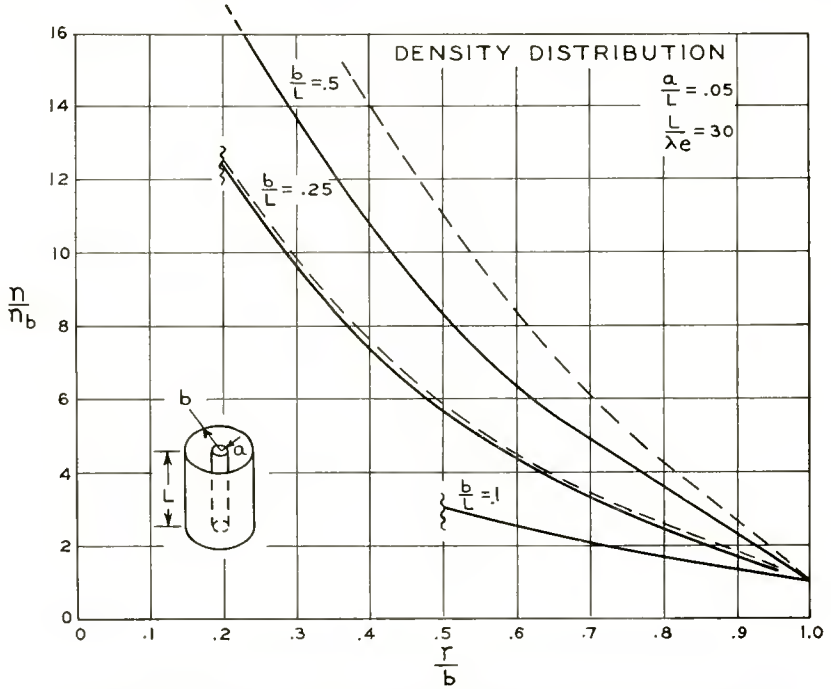


Fig. 3—Plasma density distributions in cylindrical diode with tube aspect ratio as parameter.

several different values of the geometrical constants are shown in Figure 3. The ratio of plasma density at the fractional radius r/b to the density at the anode edge of the plasma is plotted for different values of the aspect ratio b/L. For these curves $L = 2.0$ centimeters, $a = 0.10$ centimeter, and $\lambda_e = 0.07$ centimeter, corresponding to a

[4] Derived from classical kinetic theory relations.

helium pressure of about 1 millimeter. The dashed curves in this figure represent the approximate density distributions as computed from Equation (12), wherein ion losses to the end plates are not taken into account. In this approximate treatment the ions are considered as being present but as not having any net drift velocity. It is seen that this approximation is very good when $b/L < 0.25$ and gives, as it should, densities which are too large when $b/L > 0.25$.

Figure 4 shows solutions of Equation (16) for different values of λ_e, i.e., for different gas pressures. As described in connection with the approximate solution, Equation (12), the density gradients are seen to increase with pressure. It is to be noted that the densities in

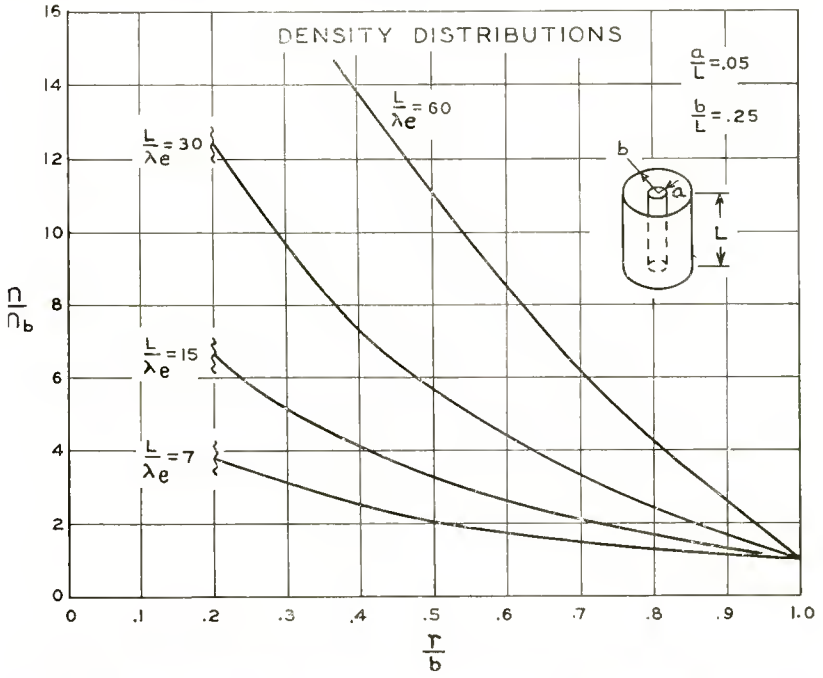


Fig. 4—Plasma density distributions in cylindrical diode with electron mean free paths as parameter.

the vicinity of the cathode are ten times or more those near the anode. Data of the sort previously presented[2] has confirmed the existence of the behavior shown in both Figures 3 and 4.

The fact that the density distributions in the AG plasma are described for the general case by Laplace's equation suggests that solutions for structures with complicated geometries can be solved by analogues, such as the electrolytic tank, resistance paper, etc. To illustrate this possibility and at the same time show the effect of

probes on plasma density distributions, a cylindrical structure of the geometry shown in Figure 5 was studied. The dimensions of the simulated structure are indicated in the figure; the actual analogue was built to scale with an anode diameter of about 4 inches. The simulated structure is a coaxial one of infinite length with the anode surrounding the cathode. Two probes, parallel to the cathode, are located diametrically opposite each other at a radius halfway between the cathode and anode surfaces. One probe has a diameter of 30 mils and the other 5 mils. The analogue itself consisted of metal electrodes mounted on resistance paper. With an electric current flowing through the model, the voltage on the surface of the resistance paper, measured with a high impedance voltmeter, simulated plasma density. The

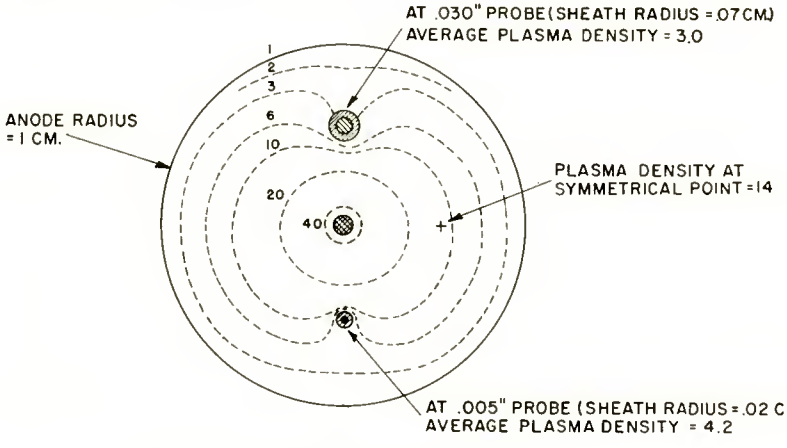


Fig. 5—Electrostatic analogue simulation of plasma density distributions in cylindrical diode containing probes.

scaling relations between the analogue and the structure it simulates are developed as follows. If i refers to the type of particle leaving and j to the particle not leaving, then Equation (5) reduces to

$$\text{grad } n = -\frac{\frac{\mu_j}{\mu_e}\Gamma_i}{D_a}.$$

Now, from relation (11),

$$\Gamma_i = \frac{n\bar{c}_i}{4},$$

so

$$\text{grad } n = -\left[\frac{\bar{c}_i}{4}\frac{\frac{\mu_j}{\mu_e}}{D_a}\right] n.$$

Using classical kinetic relations for the various quantities on the right, it is found that

$$\text{grad } n = \left[\frac{3}{4}\frac{K_i}{\lambda_i\left(1+\frac{T_j}{T_i}\right)}\right] n,$$

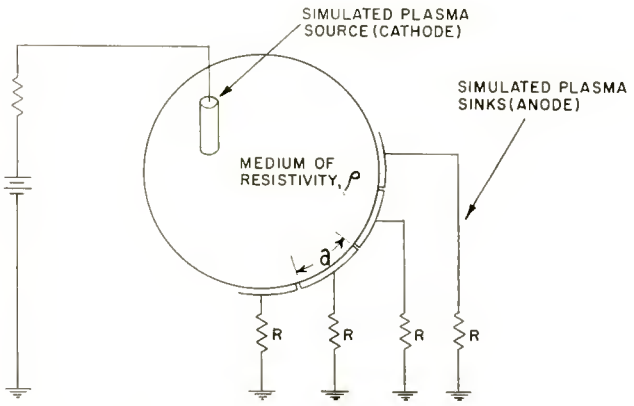


Fig. 6—Basic circuit for electrostatic analogue.

where $K_i = 1$ when electrons leave, and $K_i = 0.85$ when ions leave.

Consider the sketch of an analogue and accompanying circuit shown in Figure 6. If the potential of a boundary strip of area "a" is ϕ, then the current density J to it will be ϕ/aR. Also, if the resistivity of the analogue medium is ρ, then

$$\text{grad } \phi = -J\rho$$

everywhere. At the boundary,

$$\text{grad } \phi = \frac{\rho}{aR}\phi$$

which is a relation of the same form as the boundary condition on plasma density. To maintain the correct scale between the analogue and the simulated structure it is only necessary that

$$\frac{\rho}{a\,R} = \frac{3}{4}\,\frac{K_i}{\lambda_i\left(1+\dfrac{T_j}{T_i}\right)}.$$

From this it is seen that an increase in pressure, which reduces the mean free path λ_i, can be simulated by an increase in the quantity ρ/R. In the particular analogue in question there were 26 strips of length $a = 1$ centimeter evenly distributed around the periphery of the analogue. If no probes were present, so that geometry was perfectly symmetrical, then the strips could be replaced by a continuous ring, and the resistance to ground reduced accordingly. The value of the resistance R was 3600 ohms; the value of the resistances connected with the simulated probes back to ground were 1100 ohms for the large probe and 3750 ohms for the small probe. This simulated probes that were supplied with a negative voltage sufficient to saturate the ion current. Sheath radii as noted in Figure 5 were assumed. The above resistances combined with a resistance paper having a resistivity of about 4×10^4 ohms per square, represents a helium pressure of about 1.5 millimeters.

It is particularly to be noted that the effect of the probes is to depress the plasma density in their vicinity. This occurs because each probe constitutes a plasma "sink." The following qualitative conclusions bearing on the effect of probes are fairly obvious from the preceding analysis and Figure 5:

1) Because it is a plasma sink, whether collecting ions, electrons, or a combination of both, a probe will always tend to depress the plasma density in its vicinity. This means that the density, as measured by the probe, will always tend to be less than that which would be present in the absence of the probe.
2) The larger the probe with respect to the plasma being probed, the larger will be this discrepancy.
3) The larger the plasma flow impedance, i.e., the higher the gas pressure or the lower the mobility of the plasma particles, the greater will be this discrepancy.

Whereas it might be difficult to compute the effect that a probe has on a plasma density distribution in any particular geometry, tech-

niques such as briefly outlined above can be of great utility in obtaining estimates of this effect when conditions are such that Equation (6) is satisfied.

III. Anode-Glow Mode Existence Criteria

Thus far in the analysis it has been assumed that the discharge operates in the AG mode, yielding the boundary conditions and, consequently, the plasma solutions previously discussed. It will now be shown that there are a number of conditions, having to do with the gas parameters, geometry, and cathode emission, that must be satisfied if the discharge is to exist in the AG mode. All of these conditions relate to the ability of the plasma to carry electron current. Specifically, a section of plasma of cross-section area A and density n can carry an electron current which is no greater than the space current

$$\frac{n\,e\,\bar{c}_e A}{4}. \tag{17}$$

When the electron current is equal to this quantity, the current-carrying ability of the plasma is saturated and there will be a strong tendency for an electron sheath to form in this region. Probe theory specifically depends upon this fact. The question then arises as to where the electron sheath should be in the AG plasma. Why, for example, should it be at the anode as shown by experiment and has been assumed in this analysis? Why doesn't it appear at the cathode, or at some intermediate radius from the cathode? In arriving at an answer to these questions, consider the curves plotted in Figure 7. These represent the behavior of the quantity rn/bn_b, which is the ratio of the electron current carrying ability of the plasma at the fractional radius r/b to its value at the anode. These curves are plotted for several of the cases presented in Figure 3. It is seen that all of the curves in Figure 7, with the sole exception of the one for $b/L = 2$, exceed unity at all radii smaller than b. This means that the current-carrying ability of the plasma, except in the last case, is least at the anode edge of the plasma and, thus, it is here that the electron sheath might be expected to form. On the other hand, the case where $b/L = 2$ is not possible for the AG mode because the sheath would tend to form close to where $r/b = 0.75$, instead of at $r/b = 1$. As will be explained later, this situation forces the discharge to operate in the ball-of-fire mode, wherein a ball of luminosity projects from the anode surface.

From experiment one notices that the AG mode is present at low

anode currents even when the ratio b/L, as computed from tube geometry, is excessive. This, however, is to be expected since at low anode currents the electron sheath at the anode is very thick so that the true value of b/L, which must be computed from the plasma dimensions, is not excessive. As the anode current is increased, this sheath becomes thinner according to the 3/2-power space-charge law (modified slightly by the positive ions generated in the sheath). Because of this, the true b/L ratio increases with anode current. Along with this, as is apparent from the behavior of the curves in Figure 7, the slope of rn/bn_b at the anode edge of the plasma must also decrease in magnitude. Finally, this slope becomes equal to zero if the tube

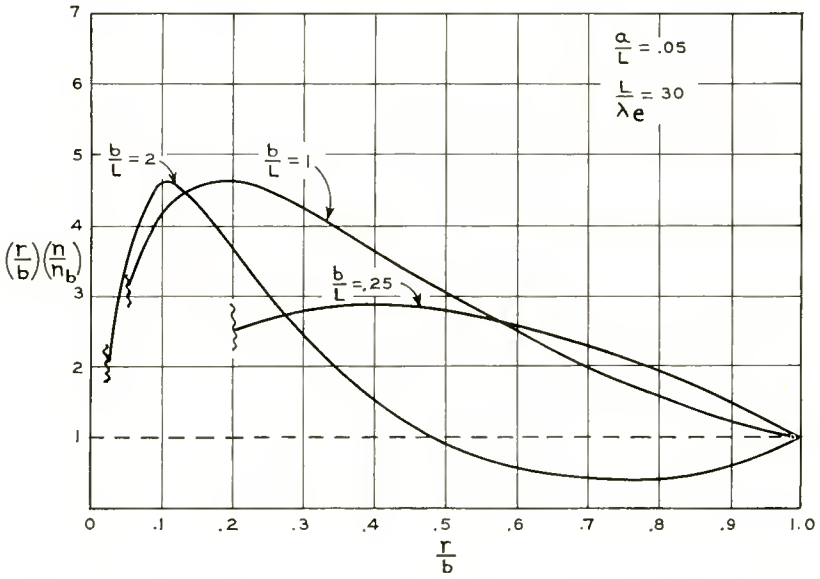


Fig. 7—Distribution of current-carrying ability of plasma with tube aspect ratio as parameter.

geometry is such that the b/L value computed from tube dimensions exceeds a critical value. Any further increase in the b/L of the plasma will cause the minimum in the rn/bn_b curve to move away from the anode surface, thereby invalidating the previous description of the AG mode. A description of what is thought to ensue beyond this point deserves a detailed treatment and is reserved for Section IV.

Physically, the preceding behavior is clear. As the true b/L ratio increases, the plasma is getting larger and requires larger ion currents to sustain its density. This means that the anode edge of the plasma, where ions pour in from the electron sheath, must become more impor-

tant as a source of plasma particles (ions). The density at the anode edge of the plasma is thus not so low as for the case where the anode is purely a sink. For an extreme case the density at this plasma edge becomes so large that the quantity given by expression (17) is larger at b than at some smaller radius. Therefore, it is at this smaller radius that the sheath might be expected to form. Another factor, which can be of equal importance to the action of the anode sheath, is the increase in ion losses at the cathode as the anode current is increased.[2] As the anode current is increased, the retarding potential V_o (see Figure 1) must necessarily decrease to allow a larger current to flow into the plasma from the cathode. The decrease in V_o exposes to the plasma more sections or "patches" of the cathode that can act as ion sinks. Thus, as the anode current increases, the tube structure becomes more "lossy" and this, in effect, is equivalent to an anomolous increase in the value of b/L.

Consideration of the curves in Figure 7 makes it apparent that the criterion for the maximum tolerable b/L ratio is the set of circumstances which lead to

$$\frac{d(rn)}{dr_{r=b}} = 0. \tag{18}$$

The critical ratio of ion to electron current flow at the anode edge follows directly from Equations (18), (5), and (11), and is

$$\left(\frac{\Gamma_p}{\Gamma_e}\right)_b = \frac{\mu_p}{\mu_e}\left[1 - \frac{4D_a}{b\bar{c}_e}\frac{\mu_e}{\mu_p}\right].$$

If appropriate values are substituted for D_a, $\bar{c}_e$, and μ_p, one obtains

$$\left(\frac{\Gamma_p}{\Gamma_e}\right)_b = \frac{\mu_p}{\mu_e}\left[1 \pm \frac{4}{3}\frac{\lambda_e}{b}\left(1 + \frac{T_p}{T_e}\right)\right], \tag{19}$$

where the negative sign refers to the case at hand and the positive sign to the inverted case where the cathode and anode have been interchanged. Carrying out the same argument with the plane parallel case one obtains

$$\left(\frac{\Gamma_p}{\Gamma_e}\right)_b = \frac{\mu_p}{\mu_e}. \tag{20}$$

Equations (19) and (20) show how the critical ratio of $\left(\frac{\Gamma_p}{\Gamma_e}\right)_b$ com-

pares to the mobility ratio with respect to geometry and the gas constants; they say nothing about the actual ratio of $\left(\frac{\Gamma_p}{\Gamma_e}\right)_b$. Nevertheless, these equations are of interest because they show that the inverted case is more tolerant of a large flow ratio than the other geometries. For example, the introduction of spurious plasma losses, such as at the cathode or at probes, would have a smaller tendency to disrupt the AG mode of discharge in an inverted structure than in either of the other cases. This expectation is well substantiated by experiment. For example, a structure of the form shown in Figure 8, which is an approach to the inverted case, was found to operate in the AG mode with far greater tenacity than a comparable cylindrical

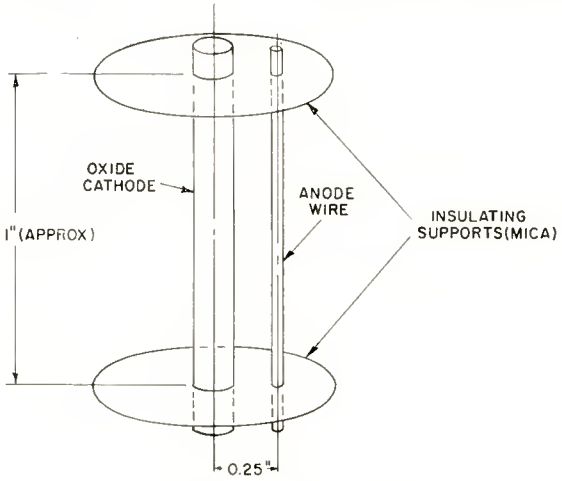


Fig. 8—Fine-wire diode.

structure of the form shown in Figure 2. This inverted structure, with the same end-plate spacing, cathode, cathode–anode spacing, and gas parameters, was far less affected by ion losses at the cathode as evidenced by an ability to pass a current 10 to 100 times greater than was possible in the more conventional structure before the discharge made a transition to the ball-of-fire mode.

Although relation (18) can be directly evaluated to obtain the critical geometry and gas conditions, it is more instructive to compare the actual value of $\left(\frac{\Gamma_p}{\Gamma_e}\right)_b$ to the critical value of this ratio as given by Equation (19). The actual value of $\left(\frac{\Gamma_p}{\Gamma_e}\right)_b$ for the normal cylindrical case can be obtained directly from Equation (15) into

which is substituted Equation (16). When this is done it is found that,

$$\left(\frac{\Gamma_p}{\Gamma_c}\right)_b = \frac{1}{b\tau\Gamma_{cb}}\int_a^b rn\,dr$$

$$= \frac{\mu_p}{\mu_e}\frac{1}{Q}\left[\frac{B}{\Delta}\left[H_1^1\left(j\frac{b}{\Lambda}\right) - \frac{a}{b}H_1^1\left(j\frac{a}{\Lambda}\right)\right]\right.$$

$$\left. - \frac{A}{\Delta}\left[jJ_1\left(j\frac{b}{\Lambda}\right) - \frac{a}{b}jJ_1\left(j\frac{a}{\Lambda}\right)\right]\right] \qquad (21)$$

In Figure 9 both the critical value (solid lines) of $\left(\frac{\Gamma_p}{\Gamma_c}\right)_b$, as given by Equation (19), and the actual value (dashed line) as given by the above equation are plotted. These quantities are plotted as a function of b/L for different electron mean free paths (hence gas pressures). The value of the cathode radius "a" has little effect on the results unless it is comparable to the value of the anode radius b. It does not appear at all in Equation (19) and has but slight influence on Equation (21). The maximum tolerable ratio of b/L for a given electron mean free path (hence gas pressure) occurs at the value of abscissa where the curve for the actual value of $\left(\frac{\Gamma_p}{\Gamma_c}\right)_b$ intersects the curve for the critical value, both curves corresponding to the same electron mean free path. From these intersections (marked by circles) it is seen that the AG mode can tolerate larger b/L ratios as the pressure is increased. The increase in the actual current ratio as pressure is reduced is easily observed in a structure of the form shown in Figure 2 which is supplied with metal end plates instead of micas. With these plates biased several volts negative with respect to the cathode, the entire ion current loss, with the exception of ions lost at the cathode,[2] can be measured. It is then found that this current falls several-fold as the helium pressure is increased from, say, 100 to 400 microns.

An estimate of the AG low-pressure limit,[5] as deduced from the curves in Figure 9, will now be compared with experiment. These curves indicate that the maximum tolerable electron mean free path of about 0.30 centimeter is reached when $b/L = 0.25$. If this mean free path is converted to gas pressure by using kinetic theory relations

[5] The low-pressure limit seems to be the reason why the AG mode was not seen by early observers who almost always worked with low gas pressures.

and data,[6] the values listed under "Theory" in Table I are obtained.

Table I

Gas	Minimum Pressure (microns) Theory	Experiment
Helium	280	250 to 400
Argon	220	80 to 150
Xenon	100	50 to 80

The experimental values listed in the third column were obtained from

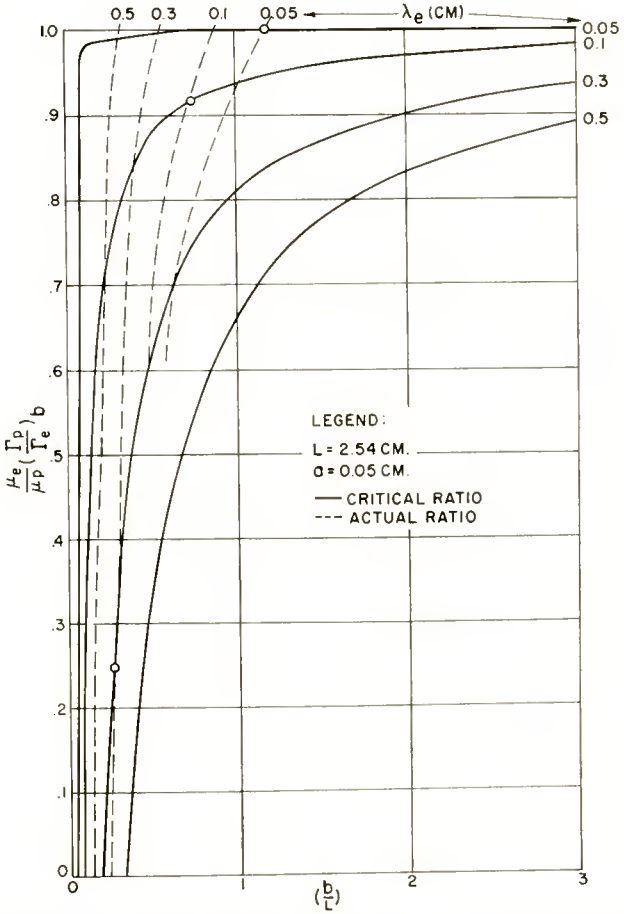


Fig. 9—Critical and actual ratios of ion to electron current densities at anode edge of plasma with electron mean free path as parameter.

[6] The paucity of direct data on electron mean free paths for the energies present in these experiments (~ 0.1 volt) necessitates this expedient.

a tube with $L = 2.5$ centimeters and the effective value of $b = 0.64$ centimeter. An oxide cathode of 0.1 centimeter diameter was used in the tube. With the anode current held at a constant value (approximately 50 milliamperes) the pressure was reduced until the discharge ceased operating in the anode-glow mode. This transition at the low-pressure limit is not the clearcut one that can be observed at higher pressures but, instead, is marked by the gradual onset of a fuzzy-appearing anode glow. Because of this, and the deviation from ideal conditions caused by ion losses at the cathode,[7] the experimental values cannot be ascertained with any greater certainty than that indicated in the table. In light of this, and the possible errors that can arise from the conversion of electron mean free paths to pressure, the indicated agreement between theory and experiment can only be considered as a rough check on the validity of the model.

In addition to the low-pressure limit, experiment has also disclosed the presence of a high-pressure limit. In the tube just described, for example, the transition out of the AG discharge occurs when the gas pressures for the various gases are as follows: helium, 25 millimeters; argon, 3 millimeters; and xenon, 1 millimeter. The reason for this is not clear. Whereas no detailed studies have been made of this effect, it is believed to arise either because of sheath instability or because of excess plasma density near the cathode surface. The sheath instability may come about because of the transition from free fall to mobility conditions within the sheath. On the other hand, the excess plasma density near the cathode is sufficient, for the case of helium noted above, to make the plasma space currents near the cathode comparable in magnitude to the emission current density of the cathode.[8] This would result in the elimination of the retarding potential V_c (see Figure 1) at the cathode and cause a transition directly into the Langmuir mode. The visual appearance of the discharge seems to justify this explanation. For the case of argon and xenon, however, the computed densities are not nearly large enough for such an effect. Visual observation shows that the anode glow with these gases merely undergoes a several-fold increase in thickness after the transition out of the AG mode.

The curves in Figure 5 imply that there might be a set of circumstances (i.e., for small cathodes and low gas pressures) where rn/bn_b could fall below unity at the cathode. If such were the case then it is to be expected that an AG to Langmuir mode transition should take

[7] The fact that the electron mean free path approaches plasma dimensions also serves to invalidate the assumed conditions.

[8] About 1 ampere per cm^2 for the cathodes used in these experiments.

place. Visual observation in tubes which should be prone to such an effect because of a small cathode diameter, have indeed shown the presence of such a phenomenon. The observed gas pressures for these transitions, however, were 5 to 10 times smaller than those predicted by an analysis which states that the transition should occur when the electron mean free path is from one to a few times the cathode radius, depending on the geometry of the rest of the tube. Perhaps this is because of the dubious validity of the discharge model at these low pressures where the particle mean free paths are comparable with plasma dimensions.

Existence criteria for the AG discharge are related to conditions in the cathode region in yet another way. This has to do with the ratio of the anode current to the field-free emission of the cathode. In short, the AG mode of discharge is not possible if the anode current exceeds roughly one-half of the field-free emission of the cathode. At the cathode edge of the anode-glow plasma, the net electron current can be broken into two components. The first of these is a current entering the plasma from the cathode whose magnitude is

$$I_o^{-\frac{eV_c}{KT}},$$

where I_o is the field-free emission of the cathode, T is the temperature of the cathode, and V_c is the potential between the cathode surface and the plasma as shown in Figure 1. The second current is the plasma space current, I_s, of density $J_s = \frac{ne\bar{G}_e}{4}$, which flows from the plasma into the cathode. The difference between these two current densities constitutes the circuit current density I_c, i.e.,

$$I_c = I_o^{-\frac{eV_c}{KT}} - I_s.$$

With the discharge in the AG mode operating at a fixed anode current, a decrease in cathode emission necessitates a decrease in V_c. If the emission is progressively decreased there will be a point at which V_c finally becomes zero. Then

$$I_c = I_o - I_s. \tag{22}$$

Also, since there is no longer any retarding effect on ions approaching the cathode, the latter will become a strong ion sink causing a depression in the surrounding plasma density. This reduces n near

the cathode and, consequently, I_s so that I_c may increase further. However, the circuit current must be carried through the plasma by the space current and the inequality $I_c \leqslant I_s$ may not be violated without having the discharge go out of the AG mode. When this does happen a sheath will form at the cathode and the discharge will go into the Langmuir mode. The cathode emission, however, is not yet saturated because $I_c < I_o$. An electron-space-charge well of a diameter slightly greater than that of the cathode will be observed to appear between the cathode surface and the edge of the Langmuir plasma. The ion current I_p flowing out of this plasma into the cathode must then be related to the electron current I_e flowing into this plasma from the cathode by the Langmuir double-sheath relation[9]

$$\frac{I_p}{I_e} = \sqrt{\frac{m}{M}}, \tag{23}$$

where M is the ion mass and m the electron mass. This specification insures that the electric field at each end of the sheath is essentially zero.

When the transition takes place, $I_c = I_s$ and Equation 22 reduces to

$$I_c \approx \frac{1}{2} I_o \tag{25}$$

so that the circuit current is about half of the saturated cathode emission. This equality is approached as a limit, as V_c approaches zero and the plasma density in the vicinity of the cathode becomes progressively smaller. The above arguments also hold if the cathode emission capability is held constant and the anode current is progressively increased. The criterion also applies to the transition out of the ball-of-fire mode into the Langmuir mode.

A typical example of the foregoing is shown in the oscilloscope photograph of Figure 10. This shows the volt–ampere characteristic of a cylindrical diode such as shown in Figure 2. This particular diode contained argon at a pressure of 0.250 millimeter and had the following dimensions: $L = 2.5$ centimeters, $a = .002$ centimeter, $b = 1.0$ centimeter. The cathode was a tungsten wire filament. The origin of the current–voltage axis is marked by the bright dot at the left of the photograph. Positive anode potential is plotted to the right with the maximum excursions of the trace corresponding to about 16 volts.

[9] I. Langmuir, "The Interaction of Electron and Positive Ion Space Charges in Cathode Sheaths," *Phys. Rev.*, Vol. 33, p. 954, June, 1929.

Increasing anode current is represented upwards with the maximum value on the characteristic corresponding to 150 milliamperes. In accordance with previous descriptions[1, 2] the current first becomes appreciable as the voltage approaches the ionization potential of the gas. In this first part of the characteristic the tube is operating in the AG mode. The first discontinuity, where the voltage abruptly falls, marks the transition into the ball-of-fire mode. The second discontinuity, where the voltage abruptly changes again, marks the transition into the Langmuir mode. The knee at the very top of the characteristic marks the point at which the cathode has become temperature limited. It is to be noted that the second discontinuity occurs at a current which is just about one-half of the temperature limited cur-

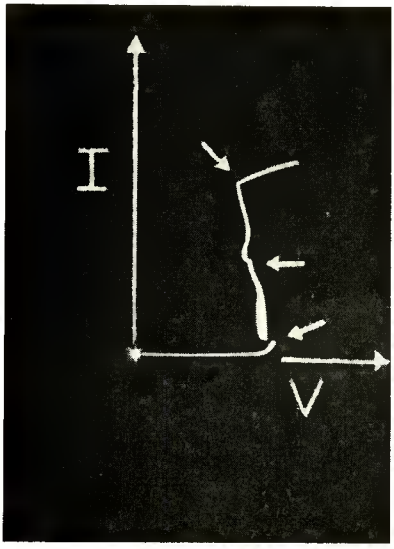


Fig. 10—Oscilloscope photograph of volt–ampere characteristic of hot-cathode diode showing discharge transition points.

rent. In the great many cases observed with both pure metal and coated cathodes, this discontinuity was found to occur at a current which ranged between about 40 and 60 per cent of the temperature limited current. Whereas pure metal cathodes usually give clean-cut results, the temperature-limited emission of oxide coated cathodes is often hard to ascertain because of poorly defined "knees." In their studies of hot cathode gas tubes the authors have found the half-current point to be a particularly useful indicator of the emission capabilities of a cathode.

IV. Transitions Out of the Anode-Glow Mode

As long as the portion of the AG plasma least able to carry current is adjacent the anode surface there will be an electron sheath at this surface and the discharge will have the potential distribution of Figure 1. Now, as the anode electron current is increased the ratio $\left(\frac{\Gamma_p}{\Gamma_e}\right)_b$ will increase because of the two reasons mentioned previously, i.e., increasing ion losses to the cathode surface and increasing size of the plasma due to the diminution of sheath thickness. As this ratio increases, the transition to the ball-of-fire (BF) mode will suddenly occur. The chain of events which lead to this transition can be considered in two different ways, quantitatively so alike that it is difficult to make a choice between them either by experiment or analysis. The first has to do with events at the anode edge of the plasma; the second, with events in the anode sheath.

When the ratio $\left(\frac{\Gamma_p}{\Gamma_e}\right)_b$ increases to the point where relation (18) becomes applicable, any further increase will cause the minimum in the plasma current carrying ability to move away from the anode. The sheath edge will necessarily move with it thereby causing an increase in the sheath thickness. This does not violate space-charge relations within the sheath providing that there is an appropriate increase in ion generation within the sheath. This increase in ion generation must terminate, however, when the limiting value[9]

$$\left(\frac{\Gamma_p}{\Gamma_e}\right) = \sqrt{\frac{m}{M}}$$

is reached. This value specifies that the electric field at each end of the sheath is zero, a condition consistent with the existence of a plasma at both ends of the sheath.

The second way this same end result may be reached starts out with a consideration of sheath conditions. Here, as the ratio increases, the potential distribution within the sheath must change progressively from form (1) to some form such as (2) shown in Figure 11(a). It can be shown that there is a critical value, where

$$\left(\frac{\Gamma_p}{\Gamma_e}\right)_b \approx 0.7\sqrt{\frac{m}{M}}$$

beyond which the ion generation in the sheath increases regeneratively

until the preceding limiting value is reached. This regeneration follows from an unstable relationship between the ratio $\left(\frac{\Gamma_p}{\Gamma_e}\right)_b$, sheath shape, and the ensuing rate of ion generation within the sheath. Whether the first chain of events triggers the second, or vice versa, or whether both occur cooperatively, is not known. However, in any event, with a plasma present at both ends of the sheath the discharge

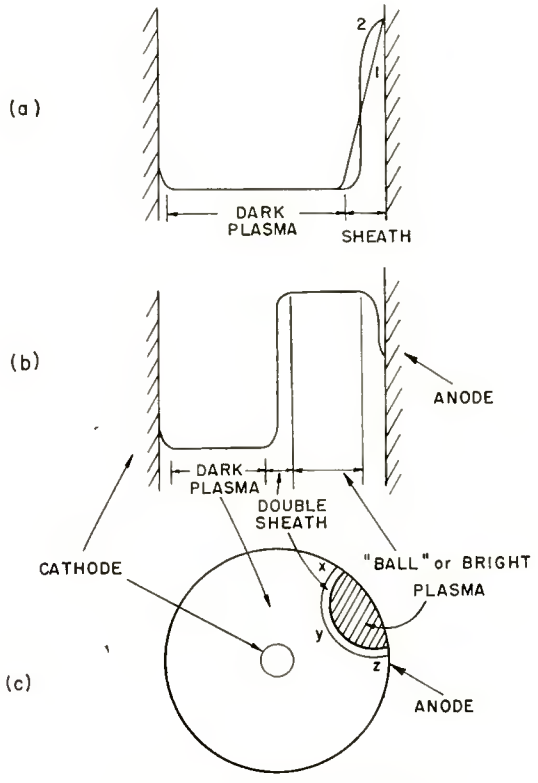


Fig. 11—Ball-of-fire discharge.
a. Potential distribution prior to transition.
b. Potential distribution after transition.
c. Cross section of tube showing discharge anatomy.

is confronted with a new set of boundary conditions which are presumed to be

1) The edge XYZ of the lower, or "dark" plasma, will be located (see Figures 11(b) and (c)) such that the plasma current-carrying ability is least at this edge.
2) The size and conditions in the ball or "bright" plasma are such

that the rate of ion generation in this plasma is adequate to supply the losses in the dark plasma. This condition is apparently often achieved by the bright plasma taking on a shape close to the hemispherical[10] as shown in Figure 11(c).

3) Since the double sheath ties two plasmas, the bright and the dark, together the field at each end of this sheath must be essentially zero, at least compared to the field inside the sheath itself. This requires that the ratio of ion to electron current densities be as specified above.

Using a toothed wheel for observation of the discharge during the AG–BF transition, it was clearly seen that the ball started as a small spot of light on the anode surface and expanded out into the dark plasma. For these observations the anode of the tube was subjected to a 60-cycle voltage wave so that the discharge increased in current in the AG mode, broke over into the BF mode, increased in current up to a maximum and then decreased, going back through the same steps.

To illustrate that the edge of the ball tends to form adjacent to a region where the dark plasma current carrying capacity is minimum, a cylindrical structure with a movable "plasma sink" was built. The structure had the following dimensions: $L = 2.6$ centimeters, $b = 1.5$ centimeters, and $a = 0.05$ centimeter. The anode was made from a section of drawn tubing of very uniform diameter; the oxide cathode was accurately centered along the axis of this anode. The movable plasma sink was a ceramic rod 2 centimeters long and 0.3 centimeter in diameter. This rod was free to roll around on the inside surface of the anode, its particular position being determined by gravity and the physical orientation of the tube. When the cathode of this tube had been well aged to insure a maximum uniformity of its surface conditions, it was found that the ball always hugged the ceramic rod as the latter was allowed to roll around on the inner surface of the anode. In this experiment the rod serves to depress the plasma density in its vicinity because of the ions and electrons which flow to it and recombine. This simple experiment is backed up by a great many observations to the effect that the ball always forms close to any plasma sink (such as a probe, insulator, etc.) in an otherwise symmetrical structure. If the ball locates itself between two identical negative probes, for example, then it will move closer to the probe upon which is placed the larger negative potential, since this probe

[10] A ball with its brightest portion of spherical shape is often observed. There is evidence to indicate that the BF discharge can exist in different modes. For example, one mode exists with an anode potential slightly below ionization potential; another, with only a few volts of anode potential.

will become a larger plasma sink than the other, because of its larger ion sheath.

V. Conclusions

The plasma flow and continuity relations, along with the appropriate boundary conditions, have been used to show that the plasma density distribution in an AG plasma is described by Laplace's equation. As in any flow system, the density is shown to be highest nearest the source (cathode) and lowest near the sink (anode), with the largest gradients occurring in the presence of the largest flow impedances.

An electrostatic analogue solution of a diode structure containing probes is used to illustrate the depression in plasma density that can be caused by probes. This shows that the density, as measured by a probe, will tend to be less than that which would be present in the absence of the probe. The larger the plasma flow impedance (i.e., the higher the gas pressure or the lower the mobility of the plasma particles) the greater will be this discrepancy.

The plasma density distributions are combined with a consideration of electron space currents to show that the existence of the AG mode requires that 1) the gas pressure be above a critical value, 2) the tube aspect ratio be below a critical value, and 3) the cathode emission be at least twice the circuit current. In addition to these there is a further, less well understood, requirement that the pressure be below a critical value.

In a brief treatment of the phenomena incident upon the ball-of-fire transition it is illustrated that the edge of the ball tends to locate itself in a region where the current-carrying ability of the dark, or cathode, plasma is a minimum.

VARIATION OF THE CONDUCTIVITY OF THE SEMITRANSPARENT CESIUM-ANTIMONY PHOTOCATHODE*

By

W. Widmaier and R. W. Engstrom

Tube Division, Radio Corporation of America, Lancaster, Pa.

Summary—*At low temperatures, phototubes having semitransparent cesium–antimony cathodes appear to yield subnormal photoresponse. It has been found that this effect is due to the resistance of the thin cathode, and not to a reduction in photoemission. Even at room temperature and normal operating voltages, the photocurrent may be resistance limited. When the tube is placed in operation after a period of idleness, the cathode conductivity is low initially but increases with operating time to a normal value. The rate of increase in conductivity at different temperatures satisfies an exponential equation. Similar conductivity variations but having a much smaller time constant have been observed with silver-oxide–bismuth–cesium semitransparent cathodes.*

Introduction

THE lower limit of usefulness of a phototube in detecting and measuring weak light signals is determined by its dark current. A major component of the dark current is thermionic emission from the cathode, which is amplified by secondary emission. This thermionic emission may be reduced by refrigeration, thus extending the useful range of the tube to lower light levels, provided sensitivity is not appreciably reduced by the change in temperature. Measurements on a 1P21 multiplier phototube,[1] which employs a cesium–antimony cathode backed by solid nickel, indicate only a small reduction in sensitivity at temperatures as low as —175°C. However, preliminary measurements on a 5819 multiplier phototube at low temperatures indicated a considerable loss in sensitivity. The 5819 employs a semitransparent cesium–antimony cathode mounted on the glass faceplate of the tube.

Conductivity Variations

A more thorough investigation of the semitransparent cesium–antimony cathode has shown that the apparent sensitivity loss is the

* Decimal Classification: 535.38.

[1] R. W. Engstrom, "Multiplier Photo-Tube Characteristics: Application to Low Light Levels," *Jour. Opt. Soc. Amer.*, Vol. 32, p. 420, June, 1947.

result of an increase in the resistivity of the cathode layer. The resistance of this layer between the area of photoemission and the metallic cathode contact causes a drop in potential and a resultant loss in the efficiency of collection of photoelectrons. This decrease in conductivity of the photocathode layer is observed not only at low temperatures but also after a period of storage in darkness. Normal conductivity is regained following a period of operation. Recovery time after dark storage increases with decreasing temperature. Similar effects are observed in unilluminated cesium–antimony layers when electric currents are passed through the layer by means of specially connected electrodes.

The apparent loss of sensitivity is most noticeable at high levels of cathode illumination. It may be minimized by a short period of operation or by the application of a high voltage between the cathode and the first dynode.

Temperature Effects

In order to obtain consistent and reproducible temperature characteristics for the 5819 photocathode, it was found necessary to store the tube at room temperature in darkness for about 16 hours. Just prior to taking data, the temperature was lowered to the desired point. Thermocouples placed close to the cathode measured the tube temperature. The tube was operated as a diode with 250 volts applied to the first dynode, which served as an anode for the cathode. After temperature equilibrium was established, a flux of 0.060 lumen (2870°K color temperature) was exposed on a ⅜-inch circle at the center of the 2-inch cathode. Measurements of photocurrent were made as a function of time after light exposure at temperatures from —83°C to +25C°. Temperature characteristics for one tube are shown in Figure 1.

In the positive-slope regions of the curves of Figure 1, an increase in the first-dynode voltage produces an increase in photocurrent. At the peak values of current and beyond, an increase in first-dynode voltage produces little change in measured photocurrent. It appears therefore that immediately after storage in darkness the current is limited by cathode resistance.

Light-Level Effects

Measurements at lower light levels at a temperature of —60°C are shown in Figure 2. These were obtained under the same conditions as those shown in Figure 1. The lower-light-level readings were obtained by momentary insertion of neutral filters in the 0.06-lumen

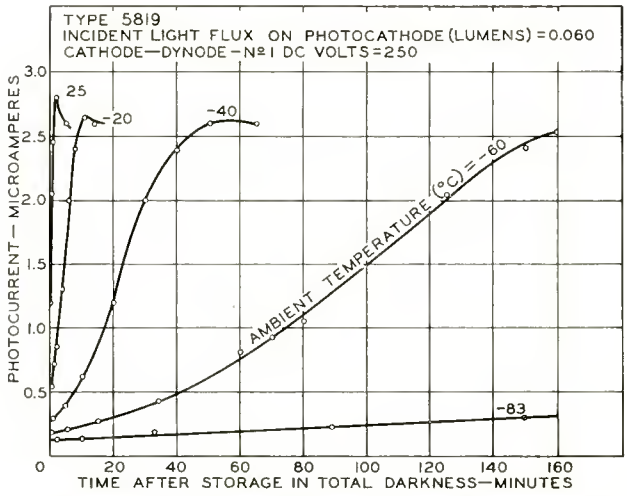


Fig. 1—Rise in photocurrent after storage in total darkness as a function of time at several different temperatures.

light beam. It can be seen that maximum photocurrent occurs earlier at the lower light levels.

The effect of the exposure level on the rise of photoresponse after dark storage was determined by means of two tests made at room temperature. In one test, the center of the cathode was continuously exposed to a luminous flux of 0.060 lumen, and photocurrent readings were taken at frequent time intervals. In the second test, a neutral

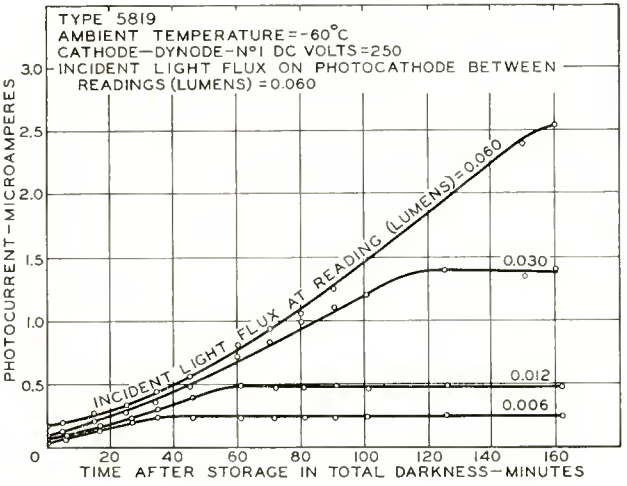


Fig. 2—Rise in photocurrent after storage in total darkness as a function of time.

filter was placed in the light beam to reduce the light flux to 0.012 lumen. At frequent intervals the filter was momentarily removed to obtain a photoresponse at 0.060 lumen. As shown in Figure 3, the rate of rise of photoresponse is reduced when the exposure light level is reduced.

Resistance Measurements

Several special photodiodes were made having cesium–antimony cathodes on glass and metal probes in the center of the cathodes. The use of these tubes provided a means of direct measurement of the cathode resistance between the center and the outside edge of the cathode when a d-c voltage was applied between the probe and the cathode con-

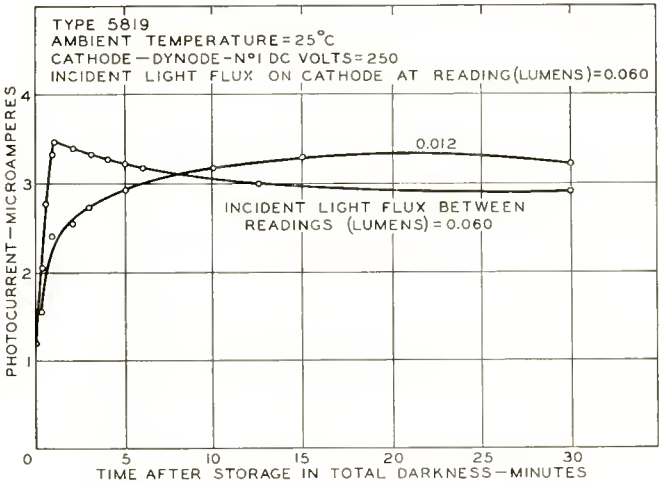


Fig. 3—Rise in photocurrent at two different light levels after storage in total darkness.

nection. These measurements were made with the tube in total darkness. It was found that when a d-c potential was applied the current was initially low and increased with time in the same manner that the photocurrent increases after storage in darkness. The results of this test are shown graphically in Figure 4. Following this test, the tube was again stored in total darkness and photoresponse measurements were made, as shown in Figure 4. The two curves are very similar, indicating that the mechanism of conductivity increase is related to the passage of current through the semiconducting cesium-antimony layer.

Silver-Oxide–Bismuth–Cesium Cathode

Tests were also made on the semitransparent silver-oxide–bismuth–cesium cathode used in the 6217 multiplier phototube to determine

whether behavior similar to that of the cesium–antimony cathode was observed. At room temperature no effect was observed. However, after a 6217 was stored in darkness for several days, its photoresponse was measured at a temperature of —95°C. A resistive effect was observed, as shown in Figure 5. The time required to reach maximum response is much shorter for the silver-oxide–bismuth–cesium cathode at —95° than for the semitransparent cesium–antimony photocathode at —83°C. From a practical point of view, it is possible to increase the response at low temperatures by pre-exposure of the cathode at

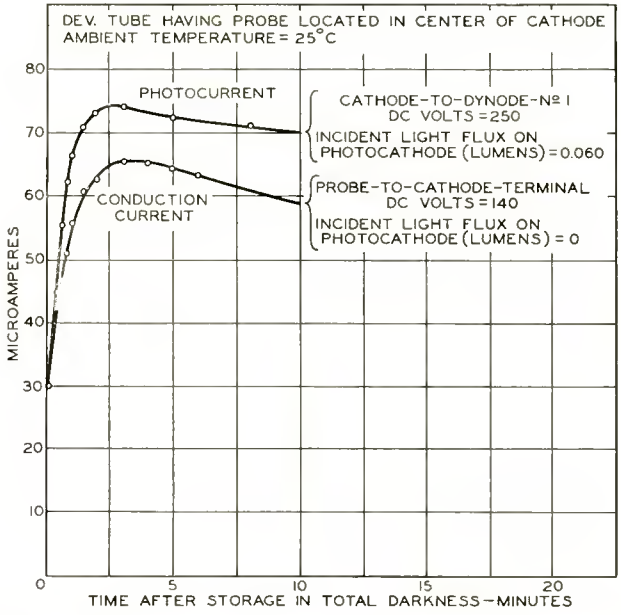


Fig. 4—Rise in conduction current and photocurrent after storage in total darkness.

the low temperature at a high light level and with voltage on the first dynode only. Voltage should be removed from all other dynodes to prevent excessive dynode currents. Exposure at a higher temperature until maximum conductivity is reached does not yield maximum conductivity at the lower temperature.

Conclusions

In all of the tests mentioned above, it is necessary for the cathode current to traverse a large unilluminated area of the cathode. It is well known that the cesium–antimony photocathode behaves as a semiconductor. However, it appears that the conductivity of this semi-

conductor depends upon its previous history. When the tube is stored in complete darkness, with no applied electric field, current carriers (electrons, holes, or ions) tend to disappear and the conductivity diminishes to a low value. When current is passed through the cathode, the carriers reappear. This effect is observed not only in the cesium–antimony cathode, but also in the silver-oxide–bismuth–cesium cathode. A comparison of the curves shown in Figure 3 indicates that the rate of reappearance of the current carriers is dependent upon the current through the cathode. The curve of conduction current shown in Figure 4 indicates that light is not necessary in increasing the conductivity of the cathode. In fact, the similarity between the two curves shown

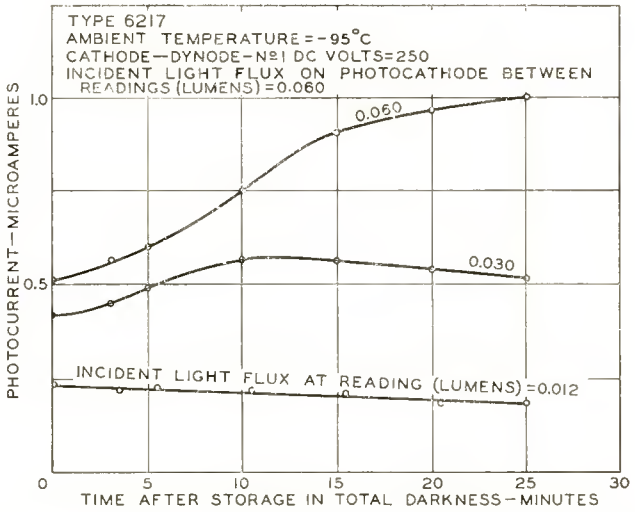


Fig. 5—Rise in photocurrent after storage in total darkness for a semi-transparent silver-oxide–bismuth–cesium cathode.

in Figure 4 indicates that the mere passage of an electrical current plays the dominant role in producing carriers in the photocathode.

When the temperature of the photocathode is reduced, the process of creating carriers is greatly slowed down and closer examination of the data is possible. It was found that, until the maximum photoresponse is obtained at the collector voltage used, the current–voltage characteristic is not saturated. The current obtained is limited by the cathode resistance, and changes in voltage produce corresponding changes in current. If the collector voltage is increased sufficiently, the limiting effect of the cathode resistance can be eliminated. At room temperature and a light level of 0.060 lumen, a collector voltage of 800 volts yields maximum photoresponse immediately.

Inspection of the rate of rise of photoresponse, as shown in Figure 1, suggests an exponential relationship between the rate of increase and temperature. In Figure 6, the logarithm of the rate of increase in collected current is plotted as a function of the reciprocal of the

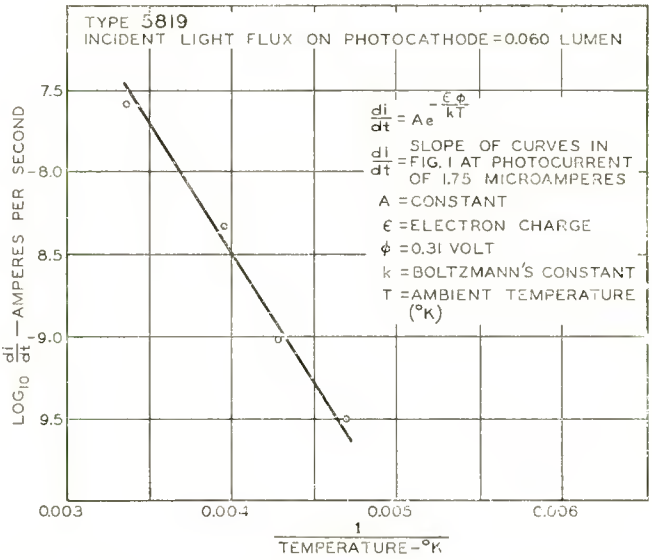


Fig. 6—Rate of rise of photocurrent versus reciprocal of temperature.

absolute temperature. The curve fits an exponential equation of the type

$$R = Ae^{-\frac{e\phi}{kT}}$$

where R is the rate of rise in amperes per second, A is a constant, e is the electronic charge, ϕ is an activation energy level, k is Boltzmann's constant, and T is temperature in degrees Kelvin. The slope of the curve shown in Figure 6 yields a value of 0.3 volt for ϕ

CALCULATION OF RADIANT PHOTOELECTRIC SENSITIVITY FROM LUMINOUS SENSITIVITY*

By

Ralph W. Engstrom

Tube Division, Radio Corporation of America,
Lancaster, Pa.

Summary—*This paper describes a simple method by which the luminous sensitivities usually specified for commercial photoelectric devices may be converted to radiant sensitivities. Mathematical expressions are presented to illustrate the derivation of the relationship, and tables are given to facilitate practical conversion. A typical analysis is included for a phototube having an S-1 spectral-response characteristic.*

MOST sensitivity standards and specifications for commercial photoelectric devices are given in terms of lumens. The tungsten lamp provides an excellent and reproducible standard operated at the color temperature agreed upon nationally, 2870 degress Kelvin. However, in certain applications involving other radiant sources, it is necessary to know the sensitivity of the photoelectric device in terms of radiant energy at a specific wavelength. Although it is possible to provide absolute calibration for the measurement of the spectral-response characteristic of the device by means of a nonselective comparator such as a radiation thermocouple, this procedure is not always convenient. Furthermore, it is desirable that all sensitivities be referred to the same standard. This paper describes a simple method for the conversion of luminous sensitivities to radiant sensitivities.

Mathematical expressions are presented for converting luminous to radiant sensitivities as a function of the relative spectral response of the device, the emission characteristic of the 2870°K source, and the spectral-response characteristic of the average human eye. Tables are given to facilitate practical conversion, and a typical analysis is given for the S-1 spectral-response characteristic.

In the following discussion, $R\ (\lambda)$ represents the relative spectral response of the photosensitive device normalized to unity at the wavelength of maximum sensitivity. The specific device is considered to be a phototube. The absolute sensitivity at the peak wavelength in amperes per watt is designated by σ. The complete absolute spectral response of the phototube, therefore, is $\sigma R\ (\lambda)$.

---

* Decimal Classification: 535.3.

The relative sensitivity of the average human eye is represented by $\overline{y}(\lambda)$. Since the sensitivity of the eye at the wavelength of maximum sensitivity is 680 lumens per watt, the complete absolute expression for the luminous equivalent of each wavelength is $680\overline{y}(\lambda)$.

The distribution of energy in watts originating from the tungsten lamp and irradiating the phototube during test may be represented by $W(\lambda)$. This function need not be stated in absolute terms provided the same relative expression is used throughout the analysis. The luminous sensitivity of the phototube tested with the tungsten source in amperes per lumen is indicated by S.

When the total radiation of the test flux is allowed to fall on the phototube, the response in amperes is

$$\int \sigma R(\lambda)\ W(\lambda)\ d\lambda. \tag{1}$$

This light flux evaluated in terms of lumens is equal to

$$\int 680\,\overline{y}(\lambda)\ W(\lambda)\ d\lambda. \tag{2}$$

The luminous sensitivity in amperes per lumen may be expressed as

$$S = \frac{\sigma \int R(\lambda)\ W(\lambda)\ d\lambda}{\int 680\,\overline{y}(\lambda)\ W(\lambda)\ d\lambda}. \tag{3}$$

From expression (3) the maximum radiant sensitivity σ may be obtained in amperes per watt:

$$\sigma = \frac{680\,S \int \overline{y}(\lambda)\ W(\lambda)\ d\lambda}{\int R(\lambda)\ W(\lambda)\ d\lambda}.$$

Tabulations of $\overline{y}(\lambda)$, $W(\lambda)$ (in relative terms), and the product $\overline{y}(\lambda)\ W(\lambda)$ are given in Table I to facilitate specific caluculations of σ. These integrals may be evaluated by numerical or graphical methods.

The evaluation of the peak radiant sensitivity of a phototube having an S-1 response (silver-oxygen-cesium photosurface) is given

as an illustration of this conversion method. A tabulation of the S-1 relative response is included in Table I. This tabulation is based on a tentative standard curve* which is part of an industry-wide standards accumulation of the Joint Electron Tube Engineering Council (JTC4 Committee) functioning under RETMA.

When the micron is used as a unit of wavelength, the following results may be obtained by numerical integration:

$$\int \bar{y}(\lambda)\ W(\lambda)\ d\lambda = 0.037836$$

$$\int R(\lambda)\ W(\lambda)\ d\lambda = 0.28870.$$

Thus, $\sigma = 89.1\ S$ amperes per watt for the S-1 characteristic when S is expressed in amperes per lumen.

Similar calculations on the S-4 response characteristic give a value of σ equal to 1010 S amperes per watt when S is given in amperes per lumen.

When appropriate units are used, these formulas may be applied to photovoltaic and photoconductive devices, as well as to photoemissive devises, provided response is proportional to the radiant flux.

Table I — Spectral Sensitivity Characteristics — Relative Units

Wavelength Microns (λ)	*Tungsten Lamp at 2870°K* $W(\lambda)$†	*Human Eye* $\bar{y}(\lambda)$	*S-1 Response* $R(\lambda)$	$\bar{y}(\lambda)\,W(\lambda)$	$R(\lambda)\,W(\lambda)$
0.30	0.0034	0.0000	0.090	0.0000	0.000
0.31	0.0049		0.270		0.001
0.32	0.0070		0.520		0.004

* It should be noted that specific response characteristics of particular phototubes having S-1 response may differ from the tentative S-1 standard curve presented in Table I just as characteristics of particular human eyes differ from the internationally agreed upon standard tabulated here.

† For lack of better data, $W(\lambda)$ is represented by Planck's black-body radiation function:

$$J(\lambda) = \frac{C_1}{\lambda^5} \frac{1}{\exp\,(C_2/\lambda T) - 1}$$

where the value of C_1 is such that $J(\lambda)$ has a maximum value of 1 at 1.0097 microns and C_2 equals 14388.4 micron degrees.

Wavelength Microns (λ)	*Tungsten Lamp at 2870°K* $W(\lambda)$	*Human Eye* $\overline{y}(\lambda)$	*S-1 Response* $R(\lambda)$	$\overline{y}(\lambda)W(\lambda)$	$R(\lambda)W(\lambda)$
0.33	0.0096		0.810		0.008
0.34	0.0130		1.105		0.014
0.35	0.0171	0.0000	1.350	0.0000	0.023
0.36	0.0221	0.0000	1.480		0.033
0.37	0.0281	0.0000	1.350		0.038
0.38	0.0351	0.0000	1.100		0.039
0.39	0.0433	0.0001	0.820		0.036
0.40	0.0526	0.0004	0.561	0.0000	0.030
0.41	0.0631	0.0012	0.360	0.0001	0.023
0.42	0.0748	0.0040	0.235	0.0003	0.018
0.43	0.0878	0.0116	0.190	0.0010	0.017
0.44	0.1020	0.0230	0.177	0.0023	0.018
0.45	0.1174	0.0380	0.171	0.0045	0.020
0.46	0.1341	0.0600	0.170	0.0080	0.023
0.47	0.1518	0.0910	0.174	0.0138	0.026
0.48	0.1706	0.1390	0.182	0.0237	0.031
0.49	0.1905	0.2080	0.194	0.0396	0.037
0.50	0.2113	0.3230	0.209	0.0682	0.044
0.51	0.2330	0.5030	0.227	0.1172	0.053
0.52	0.2554	0.7100	0.248	0.1813	0.063
0.53	0.2785	0.8620	0.271	0.2401	0.075
0.54	0.3022	0.9540	0.295	0.2883	0.089
0.55	0.3265	0.9950	0.321	0.3249	0.105
0.56	0.3511	0.9950	0.349	0.3493	0.123
0.57	0.3760	0.9520	0.378	0.3580	0.142
0.58	0.4011	0.8700	0.408	0.3490	0.164
0.59	0.4264	0.7570	0.439	0.3228	0.187
0.60	0.4517	0.6310	0.471	0.2850	0.213
0.61	0.4769	0.5030	0.504	0.2399	0.240
0.62	0.5020	0.3810	0.539	0.1913	0.271
0.63	0.5269	0.2650	0.575	0.1396	0.303
0.64	0.5515	0.1750	0.611	0.0965	0.337
0.65	0.5758	0.1070	0.647	0.0616	0.373
0.66	0.5996	0.0610	0.683	0.0366	0.410
0.67	0.6230	0.0320	0.718	0.0199	0.447
0.68	0.6458	0.0170	0.753	0.0110	0.486
0.69	0.6681	0.0082	0.787	0.0055	0.526
0.70	0.6898	0.0041	0.820	0.0028	0.566
0.71	0.7108	0.0021	0.851	0.0015	0.605
0.72	0.7312	0.0010	0.881	0.0007	0.644
0.73	0.7508	0.0005	0.909	0.0004	0.682
0.74	0.7698	0.0003	0.933	0.0002	0.718
0.75	0.7879	0.0001	0.954	0.0001	0.752
0.76	0.8053	0.0001	0.971	0.0001	0.782
0.77	0.8219	0.0000	0.984	0.0000	0.809
0.78	0.8378	0.0000	0.993	0.0000	0.832

Wavelength Microns (λ)	*Tungsten Lamp at 2870°K* $W(\lambda)$	*Human Eye* $\overline{y}(\lambda)$	*S-1 Response* $R(\lambda)$	$\overline{y}(\lambda)W(\lambda)$	$R(\lambda)W(\lambda)$
0.79	0.8528	0.0000	0.998	0.0000	0.827
0.80	0.8671	0.0000	1.000	0.0000	0.867
0.81	0.8806		0.998		0.879
0.82	0.8932		0.993		0.887
0.83	0.9051		0.983		0.890
0.84	0.9162		0.969		0.888
0.85	0.9266	0.0000	0.951	0.0000	0.881
0.86	0.9362		0.929		0.870
0.87	0.9450		0.903		0.853
0.88	0.9531		0.874		0.833
0.89	0.9605		0.841		0.808
0.90	0.9672	0.0000	0.803	0.0000	0.777
0.91	0.9732		0.761		0.741
0.92	0.9786		0.716		0.701
0.93	0.9833		0.670		0.659
0.94	0.9874		0.622		0.614
0.95	0.9909	0.0000	0.573	0.0000	0.568
0.96	0.9938		0.523		0.520
0.97	0.9961		0.473		0.471
0.98	0.9978		0.423		0.422
0.99	0.9991		0.375		0.375
1.00	0.9998	0.0000	0.329	0.0000	0.329
1.01	1.0000		0.285		0.285
1.02	0.9998		0.245		0.245
1.03	0.9990		0.208		0.208
1.04	0.9980		0.175		0.175
1.05	0.9964	0.0000	0.148	0.0000	0.147
1.06	0.9944		0.124		0.123
1.07	0.9921		0.104		0.103
1.08	0.9894		0.088		0.087
1.09	0.9864		0.074		0.073
1.10	0.9830	0.0000	0.063	0.0000	0.062
1.11	0.9793		0.053		0.052
1.12	0.9754		0.044		0.043
1.13	0.9711		0.036		0.035
1.14	0.9666		0.029		0.028
1.15	0.9618	0.0000	0.023	0.0000	0.022
1.16	0.9568		0.017		0.016
1.17	0.9516		0.012		0.011
1.18	0.9462		0.008		0.008
1.19	0.9406		0.005		0.005
1.20	0.9347	0.0000	0.002	0.0000	0.002
1.21	0.9288		0.000		0.000
1.22	0.9226		0.000		0.000
1.24	0.9099	0.0000	0.000	0.0000	0.000
1.26	0.8967				
1.28	0.8830				

Wavelength Microns (λ)	*Tungsten Lamp at 2870°K* $W(\lambda)$	*Human Eye* $\overline{y}(\lambda)$	*S-1 Response* $R(\lambda)$	$\overline{y}(\lambda)W(\lambda)$	$R(\lambda)W(\lambda)$
1.30	0.8690	0.0000	0.000	0.0000	0.000
1.32	0.8547				
1.34	0.8401				
1.36	0.8524				
1.38	0.8106				
1.40	0.7956	0.0000	0.000	0.0000	0.000
1.50	0.7210				
1.60	0.6490				
1.70	0.5817				
1.80	0.5200				
1.90	0.4643				
2.00	0.4144	0.0000	0.000	0.0000	0.000
2.10	0.3700				
2.20	0.3307				
2.30	0.2959				
2.40	0.2651				
2.50	0.2379	0.0000	0.000	0.0000	0.000
2.60	0.2139				
2.70	0.1927				
2.80	0.1738				
2.90	0.1572				
3.00	0.1424	0.0000	0.000	0.0000	0.000
3.10	0.1292				
3.20	0.1174				
3.30	0.1070				
3.40	0.0976				
3.50	0.0892				

DEVELOPMENT OF A 21-INCH METAL-ENVELOPE COLOR KINESCOPE*

By

H. R. Seelen, H. C. Moodey, D. D. VanOrmer, and A. M. Morrell

RCA Tube Division,
Lancaster, Pa.

Summary—*The major features of a 21-inch metal-envelope color kinescope are described, and considerations involved in the choice of the size and construction used are discussed. Basic operational principles of the tube are explained, and geometrical relations governing its design dimensions are given. A discussion of a number of factors affecting register is included.*

Introduction

THIS paper describes some of the main technical features of the recently announced 21AXP22 formed shadow-mask color kinescope. A photograph of the tube showing its round metal envelope appears in Figure 1. The picture size of 260 square inches is obtained with an over-all tube length of less than 25½ inches. This was made possible by the use of a short gun and a deflection angle of 70°. The spherical faceplate, upon which the phosphor screen is deposited, is of 77 per cent transmission gray glass for high contrast under ambient light operation.

A comparison of the physical dimensions of this tube with those of a conventional 21-inch black-and-white tube may be of interest. Table I gives the picture and over-all tube dimensions of the 21ZP4-B and the 21AXP22. The color tube has a somewhat greater picture size. The height of the round tube is, of course, greater than that of the rectangular 21ZP4-B; however, its diameter falls between the horizontal and diagonal dimensions of the black-and-white tube. The two-inch difference in over-all length is taken up entirely in the neck portion. The longer neck of the color tube is necessitated not by gun length, which is about equal in the two tubes, but by the space requirements of circuit components used on the neck of the tube.

A cross section of the tube is shown diagrammatically in Figure 2. The chrome-iron shell is made in two pieces having matching flanges which are joined by inert-gas welding in the final bulb-sealing operation. The "top cap" consists of the faceplate sealed to the rim portion, within which is mounted a light frame supporting the formed shadow

* Decimal Classification: R138.31.

mask. The lower, conical section of the envelope is sealed at its small end to a glass funnel-neck assembly.

The round tube shape was chosen in preference to a rectangular shape for greater stability of both envelope and internal parts, and for the lower cost of the simpler structure. The decision as to tube size was based on the popularity of 21-inch tubes for black-and-white television, which combine a picture size adequate for most homes with moderate deflection angle and over-all tube length. The use of metal rather than glass for the envelope contributes to ease of handling the tube, which is at least 10 pounds lighter than a comparable glass tube. In addition, the metal envelope provided more flexibility for design and development.

Fig. 1—Photograph of 21AXP22 color kinescope.

Table I

	BLACK-AND WHITE KINESCOPE 21ZP4-B	COLOR KINESCOPE 21AXP22
Picture Dimensions (Min.)		
Area	245 sq. in.	260 sq. in.
Width	19-1/8 in.	19-5/16 in.
Height	14-3/16 in.	15-1/4 in.
Tube Dimensions		
Width	20-3/8 in.	——
Height	15-11/16 in.	——
Diagonal or Diameter	21-11/32 in.	20-11/16 in.
Over-all Length	23-13/32 in.	25-3/8 in.
Deflection Angle (Diagonal)	70°	70°
Weight	24 lbs.	28 lbs.

Placement of the phosphor screen on the inner surface of the tube face puts the color picture on the same basis as black-and-white so far as the ratio of picture size to bulb size is concerned. The same freedom from raster distortion shown by white-screen tubes is obtained.

Operating Principles

As a basis for explaining some of the advances made in the art of color-tube design and fabrication, a brief review of the principles of shadow-mask[1] operation may be in order. In the three-beam type of tube, exemplified by the 21AXP22, each electron beam is used to excite one of three interlaced phosphor-dot arrays forming the screen. Because each of the dot arrays emits one of the three primary colors,

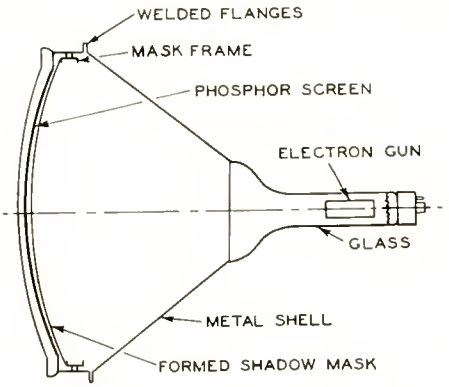


Fig. 2—Schematic diagram of 21-inch color kinescope.

red, green, or blue, control of the relative beam intensities permits mixing the primary colors as desired to produce any of a wide range of hues and chromas. The mechanism by which each beam is caused to strike the dots of only one array is based on the fact that electrons travelling in field-free space move in straight lines. Hence, a perforated shield or mask interposed in the path of a beam shadows portions of the screen from that beam, yet permits it to strike other portions. By the use of three beams so adjusted as to approach the mask and screen from different directions, i.e., from different apparent sources in the deflection plane, each beam may be caused to strike different, non-overlapping portions of the screen. Attainment of this condition requires proper dimensions and spacings of the tube parts. Furthermore, correct operation requires proper placement of the phosphor dots relative to the mask apertures.

[1] For further details see H. B. Law, "A Three-Gun Shadow-Mask Color Kinescope," *Proc. I.R.E.*, Vol. 39, pp. 1186-1194, October, 1951.

Phosphor-Dot Screen Fabrication

Accurate deposition of the dots is accomplished by a photographic exposure process carried out in an apparatus called a "lighthouse." The lighthouse contains a light source and top cap holder so arranged that the source may be moved successively to three positions. These positions or "color centers" correspond to the apparent electron sources in the deflection plane of the finished tube.

The procedure for fabrication of the phosphor-dot screen consists essentially of the following steps:

1. Coating the inner surface of the tube face with a uniform layer of one of the phosphors and a photoresist;
2. Exposing the photoresist through the mask apertures and developing the array of dots thus exposed;
3. Repeating these steps with each of the other two phosphors in turn; and
4. Applying an aluminum backing to the screen surface.

The phosphors used are:

for blue—a silver-activated zinc sulfide;
for green—a manganese-activated zinc orthosilicate or willemite;
for red—an improved manganese-activated zinc phosphate.

Red emitting phosphors, in general, exhibit relatively lower efficiencies than the other two primaries, and thus limit the light output of any tricolor tube. An improvement in efficiency of the red phosphor, therefore, shows up directly in the brightness capabilities of the tube. In the 21AXP22, a substantial increase in efficiency of the zinc phosphate has been obtained by improved techniques and controls in its manufacture. In addition, a loss of efficiency formerly suffered by the phosphate during application has been minimized. In contact with water, it had tended to hydrolyze, a decomposition which is now being prevented by use of an excess of zinc oxide in phosphor manufacture. Together these improvements raised the red light output by about 30 per cent over that obtained in earlier color tubes.

The successive phosphor coatings are deposited on the faceplate by a new method which gives improved quality at a considerable saving in time. The conventional process consisted of allowing the phosphor crystals to settle out of an aqueous suspension. After the phosphor was dried, a layer of photosensitive resist was spun over the surface and dried. The coating was then ready for exposure. As each of these operations is time-consuming, and as each of the three primaries must be applied in turn, the total time required for the fabrication of one complete screen amounted to about 4¾ hours. The new process con-

sists of flowing a thin paste or slurry containing phosphor and photoresist over the faceplate surface in such a manner as to provide a uniform coating. For this operation, a machine (Figure 3) is used which gives the top cap a rotating and tilting motion at suitable speeds and angles. The proper amount of slurry is poured onto the center of the face and motion of the machine spreads the mixture outward. Excess slurry is thrown to the edge, from which it is drained off. The total time required for the application of the entire screen by the slurry method is about 2¼ hours, a saving of 2½ hours per screen over the settling method.

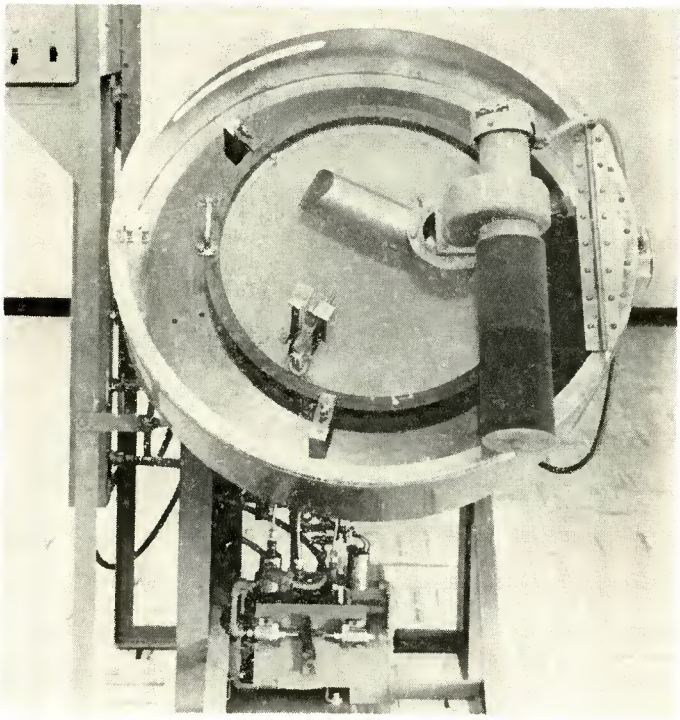


Fig. 3—Photograph of slurry equipment used for application of phosphors.

To aid in the exposure process, several recent improvements have been built into the lighthouse. They include alteration of mechanical features to increase accuracy and ease of use, and a complete redesign of the light assembly. The new assembly not only is more efficient, but also provides approximately uniform light distribution over the entire screen surface. The resulting printed dots are uniform in size. In tube operation this dot uniformity contributes importantly to equalization of the brightness of each color field and particularly to white uniformity in black-and-white reception.

Mask Design and Fabrication

The mask material is a cupro-nickel alloy, 95 per cent copper, 0.008 inch thick, which is highly satisfactory for etching. Inasmuch as the aperture array is fabricated by a photoengraving process, the latter characteristic is very important. Approximately 357,000 apertures 0.010 inch in diameter are etched within the useful area of each mask. The mask is formed from the flat etched sheet to the proper curvature edged by a short cylindrical section. This border fits closely around the light, C-section frame and is welded to it. Thus, the mask is supported by a light, rigid platform which is sufficiently rugged to hold the mask in shape during its repeated insertions into the top cap.

Mask Mounting

Satisfactory mounting of the frame within the top cap has been the subject of considerable developmental effort. To appreciate the logic behind the present mounting design, one should bear in mind certain facts concerning tube construction and processing. In the first place, the mask, frame, and envelope are of different metals having different coefficients of expansion, and these assembled parts must be heated to about 400°C during the exhaust process. In the second place, any net displacement of mask apertures which results from this processing causes a shift of the electron beams from their proper landing positions on the phosphor screen, a condition of misregister which produces color dilution. From the beginning of color tube work, it was clearly understood that some means was required to permit differential expansion of parts during tube processing and simultaneously to prevent any permanent change in the shape and position of the parts.

In the shadow-mask tubes having flat masks and screens, this problem was met by a system of three pins whose heads fitted into radial grooves in the phosphor plate. The pins, mounted from the heavy mask frame, could slide in the grooves as required during heating and cooling. This system with its rigid frame automatically brought the mask back to its initial position upon return to room temperature.

In the 21AXP22, however, it was desired to avoid loss of any picture area which would be required by a mounting device on the faceplate surface. The frame supports were therefore attached to the metal rim of the top cap. A set of posts or studs, projection welded to the metal, proved to be a sturdy, satisfactory form of support.

The next consideration was how best to mount the mask frame from these studs. Previous experience with a three-point radial sliding mechanism suggested a modification of this method for the new construction. It was found, however, that friction between the metal

parts, particularly after exhaust treatment, resulted in binding, or at best erratic slipping, of one part against another. A spring mounting method completely free of the vagaries of friction coefficients was then tried. Figure 4 shows the highly satisfactory spring mounting used in the 21AXP22. Because of the stability of the mask frame, three leaf springs have proved adequate for its support. These springs are fastened to the inside of the frame, each with a hole near one end which seats on the tapered head of the corresponding stud. The combination of the taper on the stud with the flexing of the spring provides the necessary self-aligning or indexing feature of this device. As will be covered more fully in the following discussion of misregister, the

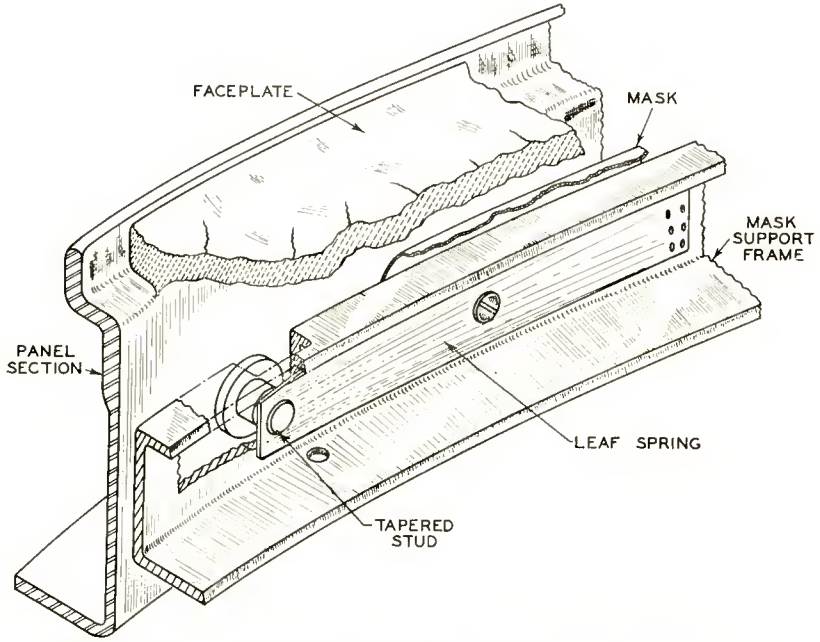


Fig. 4—Detail of spring mounting for shadow-mask support frame.

frictionless frame mounting consistently permits the mask assembly to return after exhaust bakeout to its initial position relative to the phosphor screen.

Electron-Gun Assembly

The short triple-beam gun assembly used in the 21AXP22 (Figure 5) is an improved version of the internal pole piece gun developed in 1952-1953 and demonstrated in January 1954. The gun assembly incorporates electrostatic focus and combined mechanical and magnetic convergence. The three electron guns, each including a heater,

cathode, control grid, accelerating grid, focusing grid, and high-voltage grid, are inclined at a small angle to the axis of the assembly to converge the beams at the center of the mask. Cathodes, control grids, and accelerating grids are all electrically isolated to permit cathode and/or grid modulation and adjustment of drive characteristics to fit receiver requirements. The new gun assembly is so designed as to economize on neck length and at the same time produce a narrow, high-current-density beam. The focusing grid structure has been designed in relation to the accelerator grid to form a small cross-over at high voltage for sharp focus. The "prefocusing" action between these electrodes, together with the shortness of the focusing

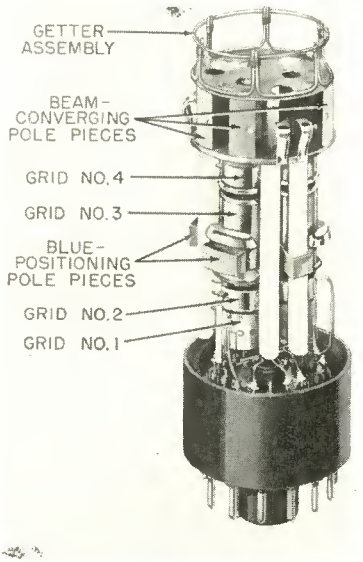


Fig. 5—Photograph of electron-gun assembly.

grid, provides a narrow beam in the focusing and deflection regions, a characteristic which has triple usefulness. A narrow beam leads to high gun efficiency, reduced aberration in the focusing lens, and improved screen tolerance in respect to color purity. This increased tolerance results from the fact that the diameter of the fluorescent dot in a shadow-mask tube is a function of the beam diameter in the deflection field. Thus, a narrow beam provides greater tolerance in landing position of the electrons with respect to the phosphor dot. The short gun has a further advantage in that alignment of its parts is less critical.

In addition to the mechanical convergence built into the gun, which superimposes the beams only at the center of the mask, dynamic convergence is needed in synchronism with the scanning to maintain

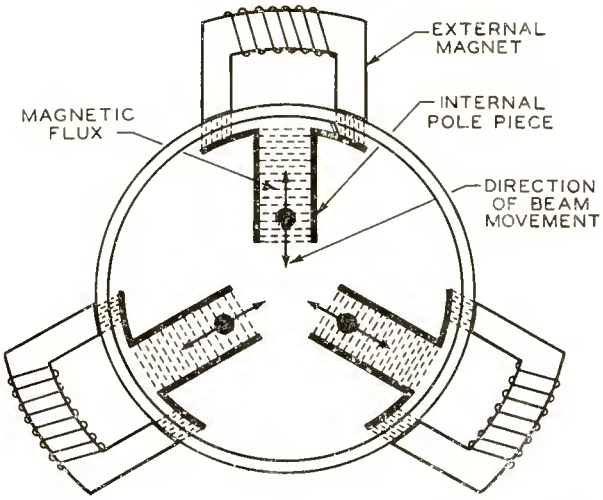


Fig. 6—Schematic of beam-converging pole pieces.

convergence over the entire mask. The beam angle variations needed for this purpose are produced by a separate magnetic field acting on each beam. To form these fields three pairs of radial pole pieces, located at the top of the gun (Figure 6), are excited by three external electromagnets. Thus an appropriate field variation may be applied to each beam individually as required by the geometry and electron optics of the system. Any inaccuracy of mechanical convergence may be corrected by application of direct current fields, as needed, to the three radial pairs of pole pieces and an additional tangential pair located further back on the "blue" gun (Figure 7). These blue posi-

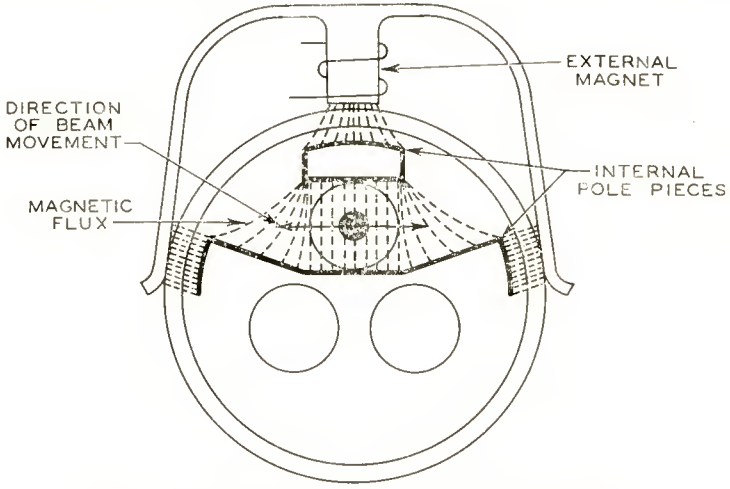


Fig. 7—Schematic of blue-positioning pole pieces.

tioning pole pieces are shown in Figure 5. The combined mechanical and magnetic convergence incorporated in this gun design offers a number of distinct advantages over earlier convergence methods. The maximum voltage brought through the base is that of the focusing electrode, nominally 4500 volts. This low voltage, together with a new base design, eliminates serious leakage troubles. As the pole-piece structure and neck coating operate at the same potential, a number of getters may be placed at the top of the gun and flashed to produce a large active area. Mechanical convergence of the beams provides good stability of convergence regardless of voltage fluctuations. Finally, the application of a magnetic field to each of the three beams separately provides the flexibility needed to obtain accurate dynamic convergence over the entire picture area, even at the relatively wide deflection angle used. All of these factors add up to the production of a high fidelity, high definition, stable picture devoid of color fringing and color shading.

Geometrical Considerations

The basic design dimensions of a shadow-mask tube are determined by straightforward geometry. When the mask and screen are curved, the geometry becomes a little more complex than in the case of a flat screen assembly because of the additional variables introduced, namely, the curvatures of face and mask. The geometry of a tube having a curved structure is worked out by first assuming a suitable radius for the faceplate and then determining the coordinate mask radius needed to provide constant magnification.

The term "constant magnification" may best be illustrated by consideration of a planar assembly[1] (Figure 8). The flat mask and phosphor plate are shown parallel to each other, spaced at distances p and L, respectively, from the source, O, which is a distance S from the axis of the assembly. The center-to-center spacing of the apertures in the mask is a and the center-to-center spacing of the phosphor dots is D.

In the phosphor-screen application process and later during tube operation, the mask pattern is projected onto the phosphor plate, forming an image having dimensions D/a times the corresponding mask dimensions. This ratio D/a, which is equal to L/p, is termed the magnification.

The magnification in this planar structure is constant, regardless of the angle at which the rays strike the mask. Consequently, if the mask-hole spacing is uniform, as indicated by the dimension a, the phosphor-dot spacing in each array, D, is likewise uniform. By the same reasoning, the dot size and the dot-to-dot spacing within each

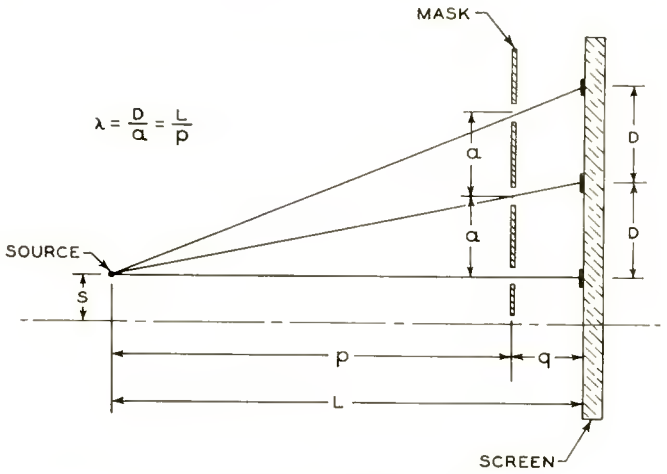


Fig. 8—Geometry of planar system.

triad are also uniform and each triad of dots is a replica of every other triad. The desirability of constant magnification stems from this uniformity of picture element spacing, and, in this planar case, the resultant optimum utilization of the screen area.

Although the spherically curved system does not permit attainment of precisely constant magnification, the slight deviations required turn out to be negligible. In Figure 9 the source, O, again lies a distance S from the axis and a distance L along the axis from the phosphor

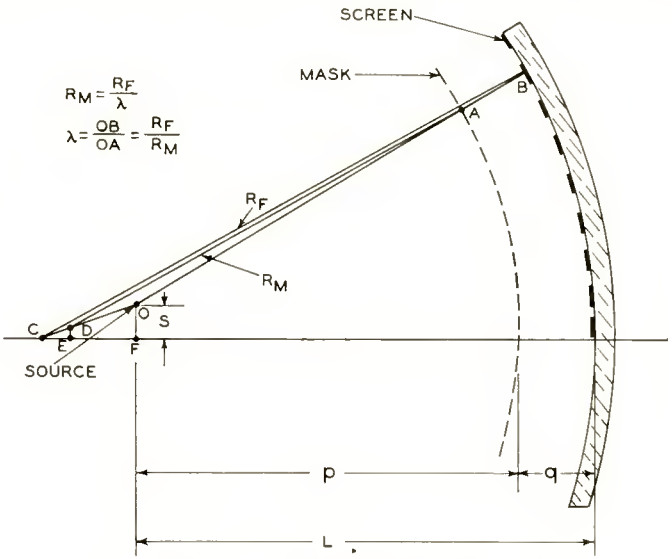


Fig. 9—Geometry of spherical system.

screen. B is any point on the screen, and A is the point where the line OB intersects the mask. If a constant magnification, λ, is required throughout the screen, then $\frac{OB}{OA}$ must be equal to λ. To meet this condition, first make the mask radius, R_m, equal to $\frac{1}{\lambda}$ times the face radius, R_f. The center of curvature of the face is at point C on the axis. Lay R_m parallel to R_f, terminating at point A. The center of curvature of the mask then falls at point D on the line connecting C and O. Since $\frac{OB}{OA}$ and $\frac{R_f}{R_m}$ both equal λ, triangles OAD and OBC are similar and $\frac{CO}{DO}$ also equals λ. It follows that $\frac{DE}{OF}$ is equal to $\left(1-\frac{1}{\lambda}\right)$, from which is obtained $DE = S\left(1-\frac{1}{\lambda}\right)$.

Thus far only one of the three sources has been considered. Because these sources are symmetrically disposed about the axis, it is desirable to shift the center of curvature of the mask from the point D, as shown, to point E on the axis, a distance $S\left(1-\frac{1}{\lambda}\right)$. In the 21AXP22, $S = 0.274$ inch and $\lambda = 1.04$, so that $DE = 0.011$ inch. It may be shown that such a compromise results in a deviation from constant magnification of a negligible 0.1 per cent.

An additional design dimension, the axial distance between mask and screen, is given by $q = R_f - R_m - CE$. Since $\frac{CE}{CF} = \left(1-\frac{1}{\lambda}\right)$ and $CF = R_f - L$, the expression for q becomes $L\left(1-\frac{1}{\lambda}\right)$.

The geometry of the spherical system, in contrast to that of the planar system, is such as to produce asymmetrical instead of equilateral triads toward the edge of the screen. Because of this asymmetry, the basic design criterion of constant magnification may be modified slightly to provide optimum phosphor dot coverage of the screen area.

Factors Affecting Register

The preceding geometrical criteria are basic in the design of the tube. However, consideration must also be given to probable deviations from the assumed mechanical dimensions and operational characteristics. From the printing industry the term *register* has been borrowed to convey the idea of the accurate landing required of the

electron beams on the appropriate phosphor dots. Perfect register, then, indicates that the center of each fluorescent spot coincides with the center of the corresponding phosphor dot. Naturally, the chief concern is with *mis*register which generally results in impure or even incorrect colors appearing in parts of the picture.

After the tube has been designed as indicated in the discussion of its geometry, and the screen photographically deposited according to this geometry, it is well to consider what factors can cause misregister in the picture produced by the completed tube. The causes are either mechanical or electrical. Mechanical effects will be considered first.

Processing Deformations

The tube and its parts are subjected to various mechanical stresses between the time of screen application and the completion of the tube. The top cap must be baked in air, and then welded to the lower cone. The tube must then be exhausted and baked while under vacuum. Thus, the tube parts may suffer a deformation due to thermal effects and to mechanical stresses on the envelope, including atmospheric pressure. Extensive tests on the 21AXP22 have indicated that resulting deformations occur, but are rather small. They are composed almost entirely of a radial component of misregister with respect to the tube axis, due principally to a minute flattening of the faceplate from atmospheric pressure. In effect, the mask is mechanically isolated from its environment by the frictionless, spring-type mounting, which permits one or more of the springs to absorb distortion of the envelope without transmitting this motion to the frame and mask.

Inaccuracies of Parts Assembly

Another potential source of mechanical misregister is misalignment between gun and screen. As already indicated, it is necessary for proper operation that the electron beams reaching the screen appear to originate at the color centers previously set by the light source during screen fabrication. This alignment is achieved by using a set of reference points repeatedly during tube manufacture. The reference points consist of three dimples approximately 120° apart stamped in the top cap flange. The dimples mate with three V-grooves in the lighthouse and thus position the top cap uniquely with respect to the light source. The flange of the lower cone is likewise stamped with a set of V-grooves, with reference to which the neck and gun seals may be made. Finally, after screen fabrication and mask insertion, the dimples and V-grooves in the flanges are mated to position the top cap uniquely with respect to the lower cone, and the final weld is made. This procedure of referencing parts positions always from the same set of three dimple-and-groove locations reduces misalignment between

gun and screen to a small value. As will be indicated later, compensation may readily be applied for this small residual error.

Mask Expansion

Motion of the mask apertures results from heating under electron bombardment. The mask's shadowing function of course requires that it intercept a large fraction of the total electron current — about 85 per cent. If heating effects are such as to move an aperture a short distance *along* the path of the beam which passes through it, the change cannot be noticed. However, if the aperture moves *transversely* to the beam path, the fluorescent spot moves about the same distance and misregister ensues.

Therefore, since mask heating and resultant expansion are unavoidable, the desirable condition is that the whole mask move a small distance toward the screen as it expands. If it were restrained at the edge, it would bow toward the screen and the apertures in a wide zone between center and edge would have an undesirable component of lateral motion. On the other hand, if it were unrestrained at the edge but fixed as to distance from the screen, the lateral aperture displacement would increase from zero at the center to a high value at the edge.

To prevent restraint at the edge and provide the desired motion toward the screen, a series of holes is punched around the border of the array (Figure 10) leaving a set of tabs or hinges connecting the array and the frame. These hinges are formed at about 45° to the side of the frame so that the tangential thrust of the mask expanding with respect to the frame causes the hinges to rise away from the frame and lessen the distance from mask to screen.

The electrical factors in misregister may be of several sorts.

Yoke Effects

As has been known for many years, the effective center of deflection of a yoke moves toward the screen as the deflection angle is increased. This motion appears whether a modern deflecting yoke having a flared "front end" and extensive fringe fields, or a hypothetical simplified yoke having "flat" fields and no fringes is considered. Its magnitude is a function of yoke length and beam deflection angle. Figure 11 illustrates this point. Starting with a very small deflection angle, for which the apparent center of deflection is at the midplane of the yoke, O, the deflection is increased to A'. The apparent deflection center for this ray is point A, while the center from which ray B' appears to come is point B. Now, because of this motion and the fact that the phosphor dots were exposed simultaneously from a stationary source of light, the electron paths are dissimilar from those of the light rays. The resulting error is radial with respect to the tube axis. Figure 12

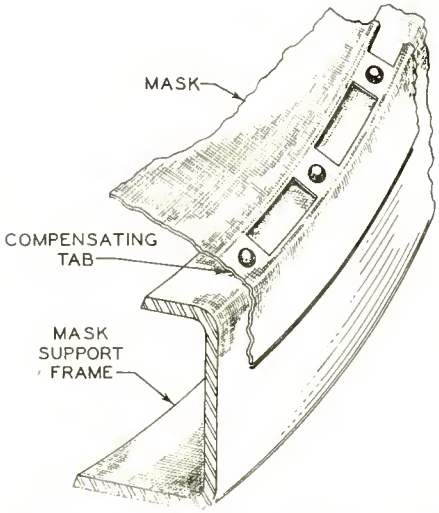


Fig. 10—Detail of shadow mask showing thermal compensating tabs.

indicates measured values of this radial component of misregister as a function of distance from the tube axis. The ordinate R is the magnitude of the radial component, with R positive when the beam is displaced toward the tube axis from its proper landing position. Adjustment of yoke position shifts the error as indicated by curves 1, 2, and 3. Thus the error may be compromised between different zones in the screen (curve 2), with results as shown in Figure 13. In this figure, the light rays which exposed the screen are indicated by dashed lines, emanating from a small source at E, while D, E, and F represent

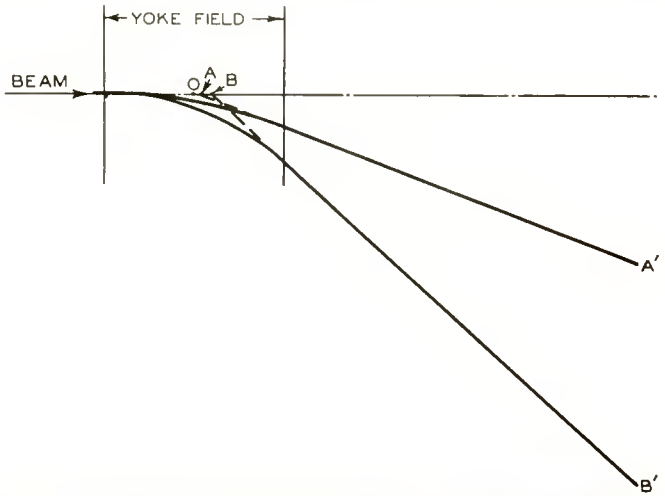


Fig. 11—Sketch illustrating motion of deflection centers.

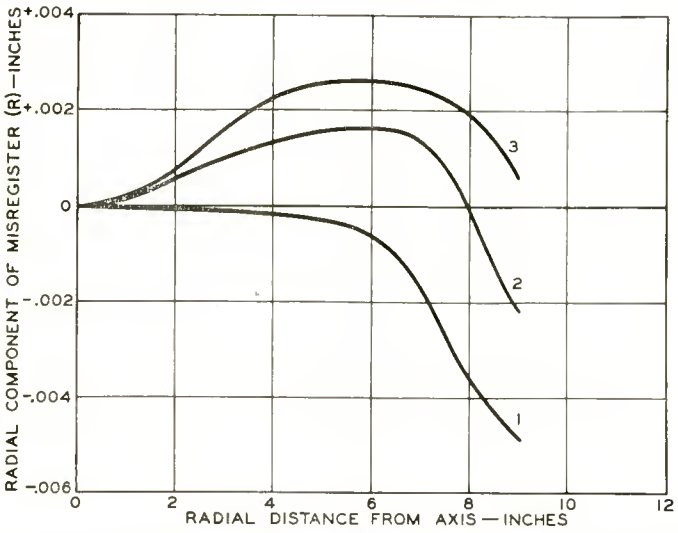


Fig. 12—Radial component of misregister as a function of distance from tube axis.

successive deflection centers as the electron beam is deflected to wider and wider angles during tube operation. The axis of the electron beam leaving deflection center E (in this case the original light source position) lands precisely on the center of phosphor dot B, while the wide-angle electron ray from F falls radially outside the center of dot A and the narrow-angle ray from D falls inside the center of dot C. To remove this error, precompensation may be built into the tube by suitable correction during tube fabrication. Thus the radial error due not only to the yoke, but also to the mechanical deformation mentioned previously, is eliminated.

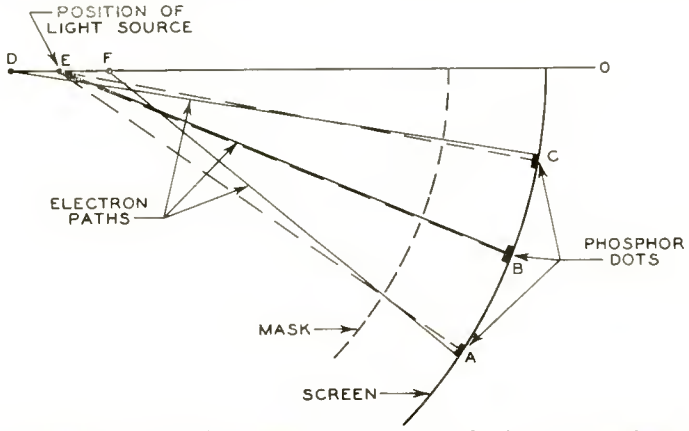


Fig. 13—Sketch illustrating radial error with yoke in compromise position.

Dynamic Convergence Effects

As mentioned previously, dynamic convergence involves variation of the relative angles of the three beams as they scan the screen in order to maintain convergence throughout the raster. Because the beam angles must be altered by a convergence means located in the gun some distance back of the yoke, the separation of the beams as they pass through the deflection plane changes simultaneously with the dynamic field. Figure 14 illustrates the process by showing for simplicity two beams in the plane of the axis. As the screen is scanned out to its edge, the convergence angle of the beams must be decreased, thus moving their deflection centers from A to A' and from B to B'. Now, the triad of fluorescent spots produced by the three beams passing through an aperture is, in effect, a demagnified image of the three

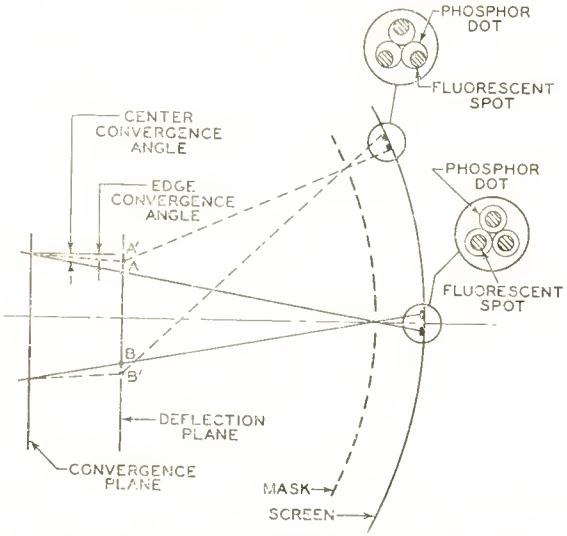


Fig. 14—Sketch illustrating effect of dynamic convergence.

beams as cut by the deflection plane. Thus, when dynamic convergence fields are applied, the object in the deflection plane AB is enlarged to $A'B'$ and its image, the fluorescent triad, is correspondingly increased in size. The phosphor screen was exposed from light sources at A and B, and the undeflected beams pass through these color centers. When deflection and dynamic convergence fields are applied, the fluorescent spots are "degrouped" with respect to the phosphor dot triads. Several promising approaches for elimination of this error are being investigated. In the interim, of course, the error may be minimized by splitting it between center and edge of the screen.

Effect of the Earth's Field

Although the earth's magnetic field is of small magnitude and the

electrons travel at high velocity, the influence of the field on the beam direction and landing may be appreciable in terms of misregister. If the force on the beams were uniform throughout the tube, the effect could be compensated by a fairly simple shift. In the neck region, where it is nearly uniform, a transverse field is generally applied over the gun by the "purifying magnet," which serves simultaneously to correct for any slight misalignment between gun and screen. However, in the cone region the effect of the earth's field on the beam varies with the angle of deflection, so that the resulting misregister is far from uniform over the screen area. Furthermore, this misregister pattern changes with the orientation of the tube in the earth's field

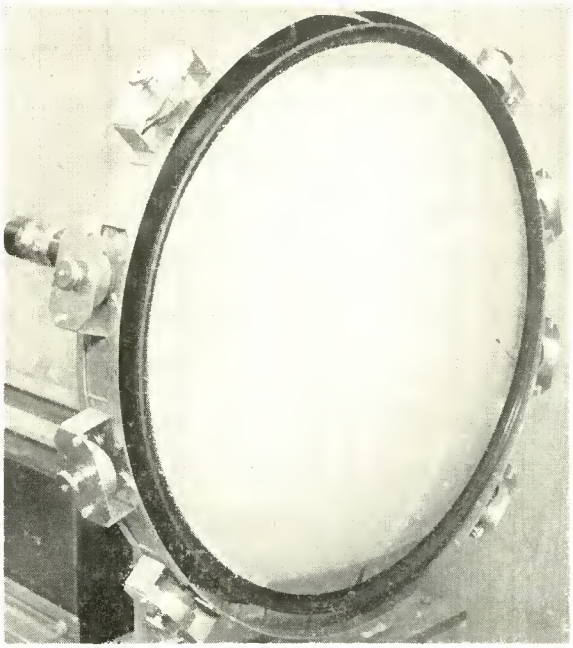


Fig. 15—Photograph of magnetic-field equalizer in position on the 21AXP22.

The metal envelope of the tube provides partial shielding and tends to redistribute the field. The resulting misregister is such as to be readily corrected by a magnetic field system placed around the front end of the tube. This "magnetic-field equalizer" (Figure 15) consists of eight small magnets, with associated pole pieces, which are adjustable as to strength and direction.

ACKNOWLEDGMENT

The authors wish to give credit to their many colleagues and associates who helped to evolve this successful design.

DEFLECTION AND CONVERGENCE OF THE 21-INCH COLOR KINESCOPE*

By

M. J. Obert

Tube Division, Radio Corporation of America,
Camden, N. J.

Summary—*This paper describes new deflection and convergence components and circuitry used to achieve successful operation of the 21AXP22 21-inch color kinescope. The effects of deflecting-yoke characteristics on the performance of horizontal- and vertical-deflection circuits, and circuitry required to energize the dynamic-convergence assembly, are discussed.*

Various methods of presenting information on flux distribution of deflecting yokes are reviewed, including a new method for obtaining three-dimensional flux plots. Considerations involved in the design and development of the yoke, the horizontal- and vertical-output transformers, and the converging-magnet assembly are also reviewed. Operating circuits are shown, and typical performance data given.

OPERATION of the new 21AXP22 21-inch color kinescope resulted in development of new deflection and convergence components and circuitry which are different from those included in the original 15-inch color receiver. Figure 1 shows the 21-inch kinescope with the required components in their correct relative positions. These components include the deflecting yoke, converging-magnet assembly, purifying magnet, blue-positioning magnet, and magnetic field equalizer. The new horizontal-output and high-voltage transformer and the vertical-deflection-output transformer are shown in Figure 2.

Figure 3 illustrates the basic operating principles of the kinescope which the components must satisfy to produce a good color picture. The shadow mask and faceplate are constructed to provide color changes conditional upon the direction of arrival of the impinging electron beams. After scanning, each individual beam should have maintained or regained its original relative position within the composite beam array, and the three beams should converge at the shadow mask. Since pure color fields are attained when each electron beam lands on its proper phosphor dot, corrections are made in phosphor-dot placement which effectively maintain the deflection center of the yoke and the kinescope color centers in register throughout the entire scanning cycle.

Decimal Classification: R583.13.

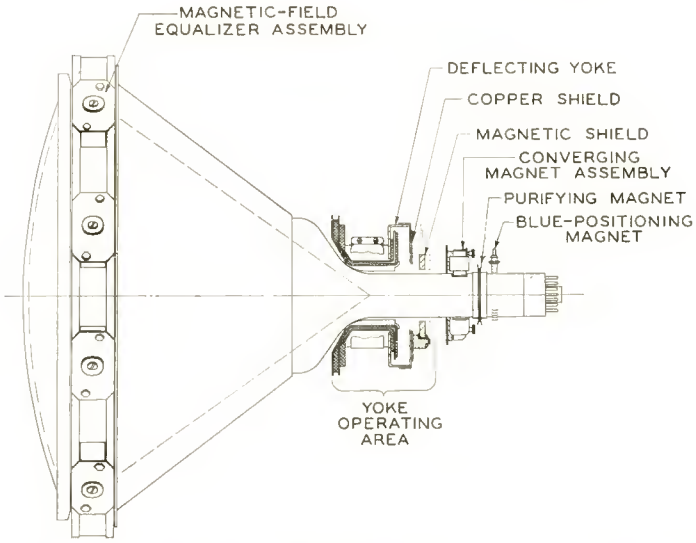


Fig. 1—Sketch showing relative placement of components on 21-inch color kinescope.

The development of 70-degree-deflection color television components has been expedited by the use of a simplified test set for which a block diagram is shown in Figure 4. An adequate signal source is the TG-2A sync and bar generator. If color pictures are required for demonstration, a flying-spot scanner may be added. Most performance tests are

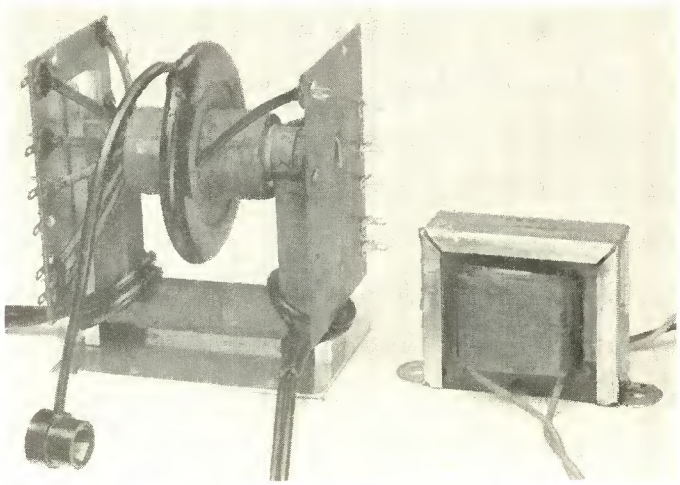


Fig. 2 — Developmental horizontal-output and high-voltage transformer (left) and vertical output transformer 247FT1.

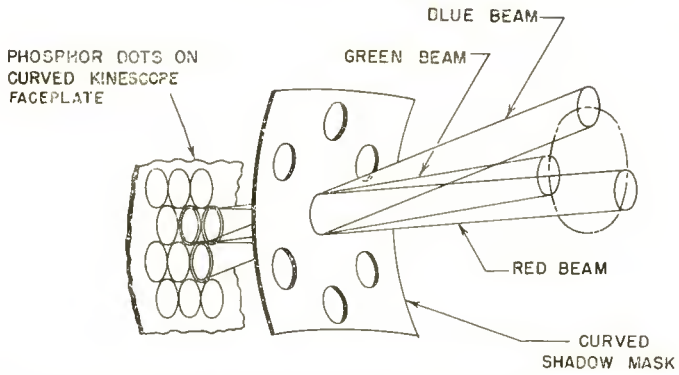


Fig. 3—Sketch illustrating basic operating principles of 21-inch shadow-mask color kinescope.

made using a white-bar test pattern, experience having proved this pattern to be superior to a dot pattern for the measurement and recording of spacing between beam centers. In addition, "ringing" effects (slight irregular vertical displacement of horizontal lines) are clearly visible when bars are used.

Deflecting Yoke

The 230FD1 deflecting yoke, shown in Figure 5, was designed to provide the required color purity and convergence characteristics for operating the 21-inch color kinescope, when used with associated com-

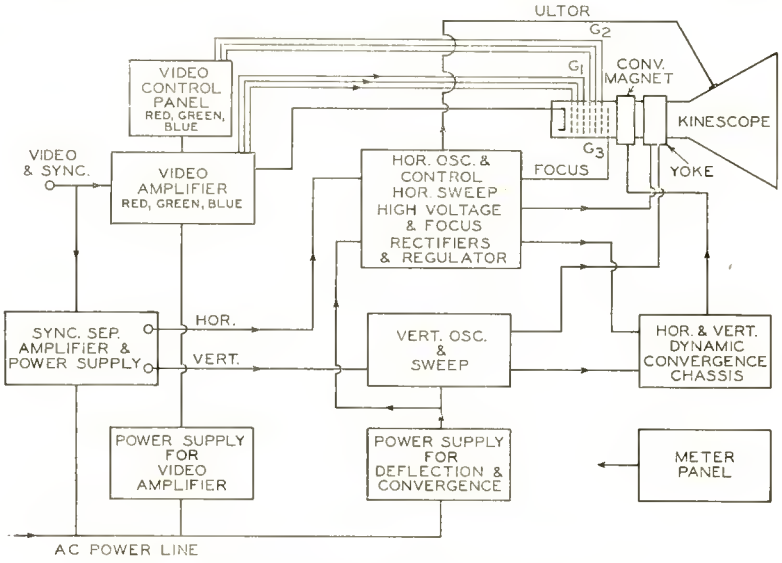


Fig. 4—Block diagram of test chassis.

ponents and circuitry. Expanded-coil and expanded-front-core construction provides shaped flux fields for full 70-degree deflection, and maintains within close tolerances red-, green-, and blue-beam convergence over the entire raster. Particular attention was given to the maintenance of red and green convergence on horizontal lines in the corners, so that the need for additional costly dynamic-convergence corner-correction circuitry was eliminated. Inductance and insulation characteristics were selected for use in auto-transformer deflection circuits of the energy-recovery type, with ultor voltages up to 25 kilovolts.

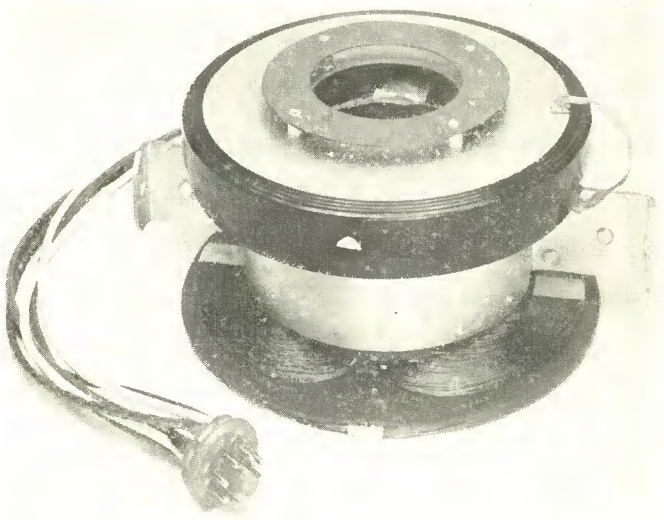


Fig. 5—Deflecting yoke 230FD1 for 21-inch color kinescope.

Yoke-Flux Plotting

Although various methods may be used for the measurement of the direction and magnitude of the three-dimensional magnetic field of a yoke at any point within or about the yoke, it is difficult to depict graphically all of the data needed to define the yoke field. A rapid method of obtaining iron-filings flux plots showing equipotential magnetic lines of force uses Ozalid*-sensitized linen (activated by light and ammonia) cut to fit the yoke section being studied. The linen is reinforced with cardboard and uniformly covered with iron particles of 300-mesh size. The yoke is then energized with d-c and the assembly is vibrated by means of low-audio-frequency sound waves. After the

*Trade mark of the General Aniline and Film Corp.

flux pattern forms, it is exposed to ultraviolet light until the color of the linen changes from yellow to white. The field and light are then cut off, the powdered iron is removed, and the paper is run through the developing section of the Ozalid machine. Sample patterns obtained are shown in Figures 6, 7, and 8; the approximate area traversed by the beam trio is marked on each pattern.

Three-dimensional flux plots have been obtained by the use of a suspension of finely divided flakes of iron in a clear casting resin. A rubber mold made to fit the inside yoke contour is used as a container for the mixture. The viscosity of the resin changes quickly after the promoter and the hardener are added. The mixture is poured into the mold about six minutes after the promoter is added. Current is then applied through the coils, and the iron flakes in the resin align themselves with the flux field.

Application of the current to the coils intermittently keeps the pattern well defined until the plastic gels, without pulling the iron particles out of position. Immediately after the mixture gels, the casting is placed in an oven at a temperature of about 85° C for ¾ hour to prevent cracking during the exothermic expansion.

After the casting has hardened, it is machined and polished so that the pattern becomes visible. Coating of the finished casting with a lacquer provides a hard, smooth finish and considerably reduces the amount of polishing needed. Ordinarily, the casting is cut axially or transversely in order to display the flux pattern distribution to best advantage. Figure 9 shows results obtained when this procedure was used.

Coil-Turns Distribution and Electron-Beam Paths

The specific turns distribution and the coil shape of the 230FD1 yoke and some of the various designs tried during development are shown in Figure 10. The relation $d = T \cos \theta$ defines the popular cosine distribution associated with conventional monochrome yoke-coil design. When a constant wire diameter is used throughout the coil, the turns distribution varies as the thickness d at the angle θ. As may be seen from the chart, neither $\cos \theta$ nor $\cos^2 \theta$ turns distribution has been used in the design of successful color yokes.

In addition to performance tests using d-c-driven multi-tapped yoke coils, flux plots of various configurations were studied and related to the actual path of an electron beam as deflected by an expanded-coil yoke. Figure 11 shows the path traced by an electron beam deflected with the 230FD1 color yoke. This data was obtained with a kinescope having a fitted, phosphor-coated, mica piece in the funnel area. The

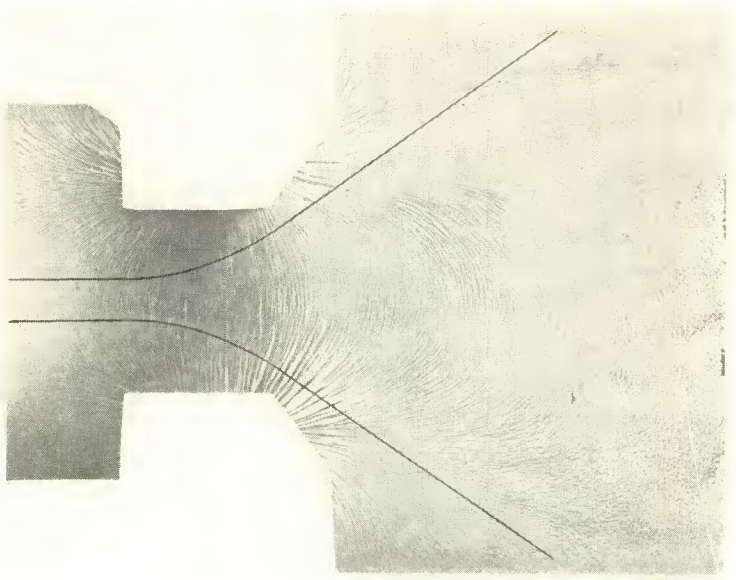


Fig. 6—Flux pattern of vertical coils of deflecting yoke 230FD1. Solid lines approximate border of area swept by beam trio.

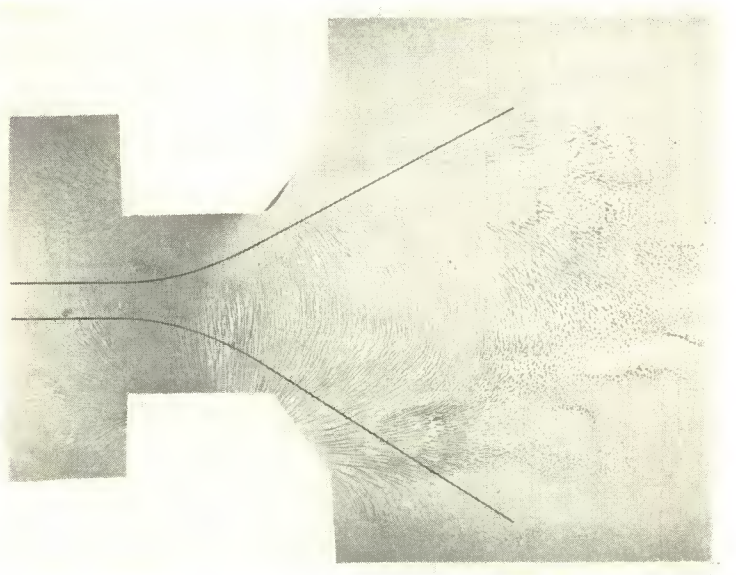


Fig. 7—Flux pattern of horizontal coils of deflecting yoke 230FD1. Solid lines approximate border of area swept by beam trio.

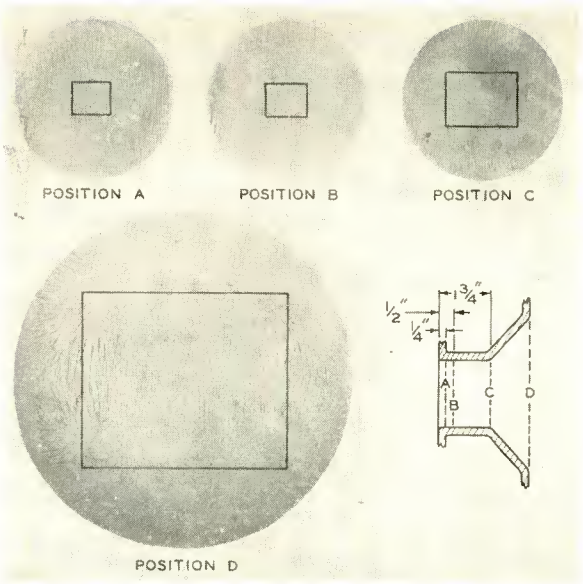


Fig. 8—Flux pattern of vertical coils of deflecting yoke 230FD1 taken in positions indicated. Solid lines approximate border of area swept by beam trio.

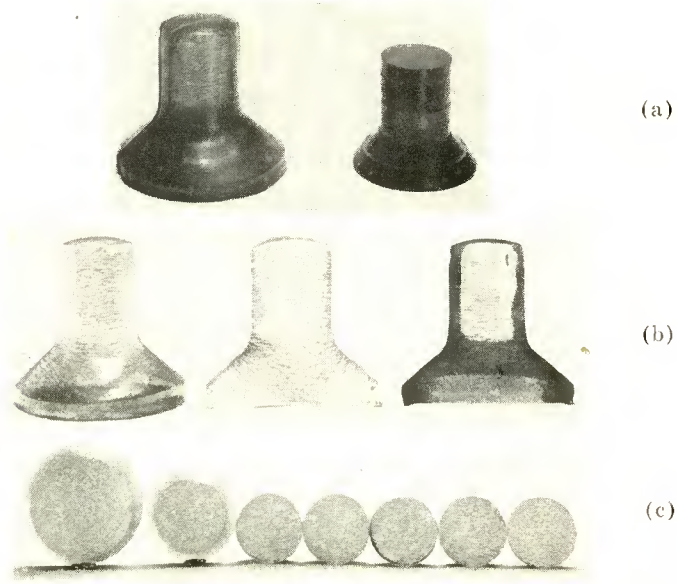
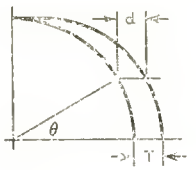


Fig. 9—Three-dimensional flux patterns of an expanded-front deflecting yoke: (a) complete casting made with iron flakes suspended in a clear resin; (b) sections of the casting cut parallel to yoke axis; (c) sections cut normal to yoke axis.



$d = T\,(\cos\theta)^n$

$\frac{d}{T} = (\cos\theta)^n$

d = DIMENSION OF COIL AT ANGLE θ

				DEV. 45° COLOR YOKE WITH STRAIGHT SIDES			DEV. 62° COLOR YOKE WITH STRAIGHT SIDES			EXPANDED-FRONT 70° COLOR YOKE 230FD1		
θ	$\cos\theta$	$\cos^2\theta$	$d\ (\cos\theta)$	d	d/t	n	d	d/t	n	d	d/t	n
0	1.000	1.000	0.125	0.125	1.000	0	0.125	1.000	0	0.125	1.000	0
15	0.966	0.933	0.121	0.127	1.016	-0.457	0.125	1.000	0	0.125	1.000	0
30	0.866	0.750	0.108	0.131	1.048	-0.326	0.125	1.000	0	0.125	1.000	0
45	0.707	0.500	0.083	0.136	1.088	-0.243	0.125	1.000	0	0.125	1.000	0
60	0.500	0.250	0.062	0.142	1.136	-0.184	0.125	1.000	0	0.136	1.088	-0.122
75	0.259	0.067	0.032	0.149	1.192	-0.130	0.125	1.000	0			
90	0											

Fig. 10—Specifications for deflecting yoke 230FD1 and for 45-degree and 62-degree developmental models.

coated mica was marked with a cross-hatched grid having known dimensions, and the beam path was observed visually.

Insulation and Core Considerations

The voltage developed across the horizontal coils during flyback is approximately 3,000 volts peak-to-peak at 15,750 cycles per second; the coils are series-connected to provide minimum voltage between turns and to permit connection of neutralizing capacitors across each coil. The vertical coils are also series-connected, with an operating voltage across the coils of approximately 200 volts peak-to-peak at 60

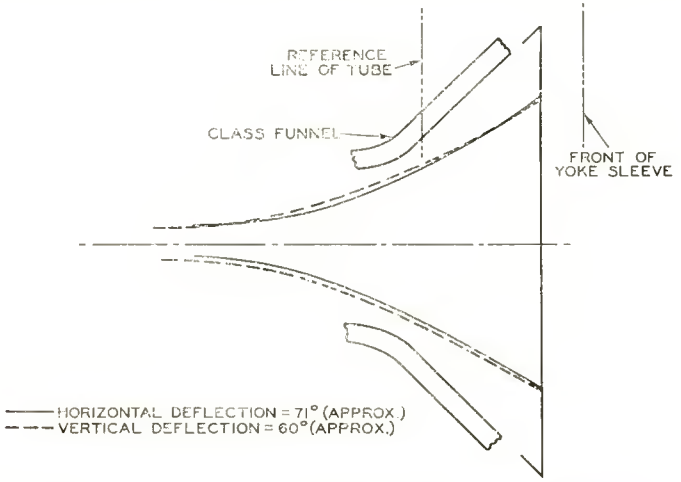


Fig. 11—Electron-beam paths in field of deflecting yoke.

cycles per second. Voltage induced by the horizontal retrace pulse of 1,250 volts peak appears between the vertical coils and ground.

Insulation between coils is provided by specially compounded molded synthetic rubber shaped to fit the coils. This insulation must withstand operating temperatures up to 85°C, at an absolute maximum voltage rating of 4,500 volts for a pulse of approximately 10 microseconds duration, and at a repetition rate of 15,750 cycles per second. Continuous life tests made at 6,000 volts peak at rated 85°C operating temperature proved the effectiveness of the insulation and the coil construction.

The ferrite cores used for the flux return path provide high permeability, high resistivity, and low losses at the operating frequencies. Permeability alone may be used in the evaluation of core materials for use at the vertical scanning frequency because losses in the ferrite at 60 cycles per second are calculated to be 5.9 milliwatts. The calculation is based on the core used in the 230FD1 yoke, which weighs 1.73 pounds and operates at an average flux density of 254 gausses. The copper loss in the vertical-deflection-coil windings, which have a total d-c resistance of 54 ohms and require an r.m.s. current of 0.145 milliampere, is 1.12 watts. This value is more than 190 times greater than the core loss, indicating the relative unimportance of core losses at the vertical-deflection frequency.

At the horizontal-scanning frequency, the core loss of the 230FD1 yoke is approximately 0.285 watt, based upon an average flux density of 230 gausses. The copper loss in the horizontal-deflection windings, which have a total d-c resistance of 7.3 ohms and require an r.m.s. current of 520 milliamperes for full scan, is 2.0 watts. The core loss is, therefore, approximately 14 per cent of the copper loss in the horizontal deflection windings.

At the low operating flux density, ferrite permeability and losses are relatively stable, and the magnetic circuit is considerably diluted by the large air cap. Careful observation of the performance of a yoke during a heat run indicated little change in performance at temperatures from 25°C to 80°C. Figure 12 shows the variation of permeability with temperature for this type of ferrite.

The ferrite core has an expanded front for flux shaping, and is designed in four 90-degree segments for better manufacturability. The volume of core material used provides optimum deflection sensitivity and good mechanical strength consistent with low cost. Because concentricity of the core and coils is required for good convergence, purity, and pattern rectangularity, the cores are ground to close tolerances.

Compressible gaskets between the core segments and the coils help maintain concentricity during assembly. A heavy cold-rolled steel band

locks the core securely in position and provides means for mounting the yoke in the receiver. Figure 13 shows the disassembled yoke.

Shielding and Casing

The cap-and-shield assembly includes a thin copper disc to minimize the coupling of horizontal-frequency yoke fields into the pole-piece assembly, with resulting complication of the required convergence correction. A magnetic shield provides similar action with respect to the 60-cycle-per-second vertical-deflection-frequency flux at the back

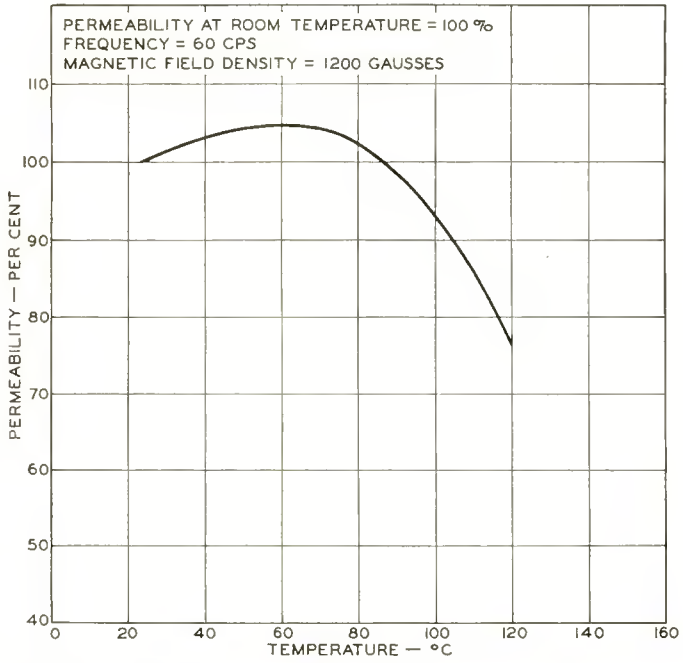


Fig. 12—Variation of permeability with temperature for ferrite mix.

of the yoke. The magnetic shield consists of a ¼-inch thick ferrite disc, or several 0.004-inch thick oriented silicon-steel discs mounted with grain structures in quadrature for symmetry.

The molded case-and-terminal assembly is made of thermo-setting general-purpose phenolic, and the terminals are assembled directly to the case. The characteristics of the phenolic meet the mechanical and electrical requirements of this application. This material was chosen in preference to lower-cost injection-molded thermoplastics because of its mechanical stability at elevated temperature, freedom from distortion, and mechanical strength.

Coil Winding, Assembly, and Test

The coils are wound to approximately final shape with thermoplastic-coated insulated wire by means of special tension devices and arbors. After being wound, the turns are temporarily fixed in place by current passed through the coils sufficient for heating and softening the thermo-plastic coating. With the wire sizes used, #24 and #25 wound bifilar, current required for the horizontal coils is 13 amperes for 9 seconds. The wound coils are impregnated in vinyl compound, dried thoroughly, then heated, shaped, and cooled in the mold to lock the turns into their final position.

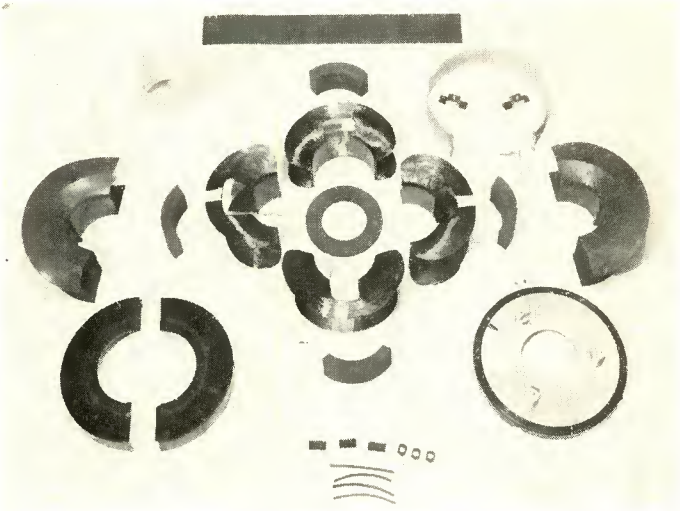


Fig. 13—Disassembled parts of deflecting yoke 230FD1.

The horizontal coil is made with 145 turns each of #24 and #25 wire bifilar wound. The vertical coil consists of 465 turns of #26 wire.

The individual coils are carefully gauged for physical size; then tested for shorted turns, number of turns, and continuity; induced-voltage-tested at 4,000 volts peak for corona; and finally calibrated for inductance and resistance. Matched pairs are assembled into complete yokes, which are then tested for interconnection of the coils, cross talk, "hipot," and induced voltage. The horizontal-deflection coils are induced in series at a peak voltage of 5,000 volts, and the vertical coils at a peak voltage of 900 volts. The performance of the completed yokes is then tested in a deflection circuit. The yokes are held to close tolerances for deflection sensitivity, ultor power output, pattern rectangularity, and convergence with color purity.

Convergence Specification

When recommended components and dynamic-convergence circuitry are used on an average 21AXP22 kinescope, the maximum distance between line centers of a rectangular-grid test pattern measured in the corners is not allowed to be greater than 1/16 inch with the center converged. In the center of the top, bottom, right, and left edges, the

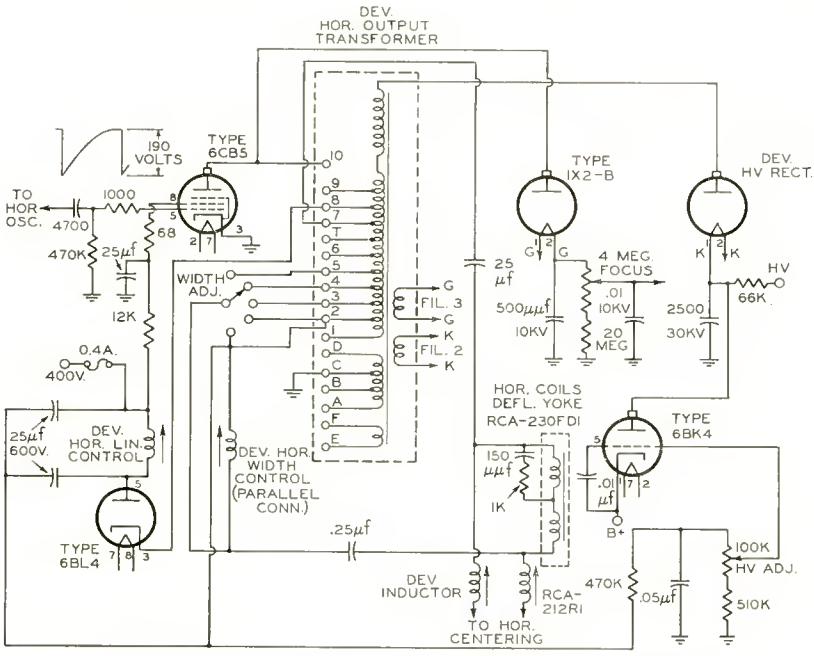


Fig. 14—Horizontal-deflection and high-voltage circuit.

maximum distance between lines of the cross-hatch pattern will not exceed 1/32 inch. These measurements are made approximately 1 inch from the kinescope edge.

Over-all performance of the 230FD1 yoke on the 21-inch color kinescope has been considered equal or superior to that previously obtained with narrow-angle color kinescopes having smaller picture area.

Horizontal-Deflection and High-Voltage Transformer

The advantages of auto-transformer flyback operation dictated the use of the circuit shown in Figure 14 for conversion of the practical value of yoke inductance of 11.8 millihenries to the required value of plate-load inductance for the driver tube. In addition to providing

the required matching ratio, the transformer supplies the high-voltage ultor power and the horizontal-deflection-frequency pulse voltages needed for color-receiver operation. The new developmental transformer represents an improvement over that used in earlier circuits because the ultor and focus voltages require two rectifiers instead of three. The 1X2-B provides 5.5 kilovolts at currents up to 75 microamperes with good regulation for focusing, plus a current of approximately 150 microamperes for the bleeder. The developmental high-voltage diode rectifies the full pulse voltage of the transformer, providing the 25-kilovolt ultor voltage at a current of 800 microamperes.

Electrical Requirements

A conventional oscillator–discharge circuit is used to provide the driving voltage for correct operation of the circuit. Other sawtooth-generating circuits such as the stabilized multivibrator may be used to provide the required voltage amplitude and waveform. Sufficient reserve driving-voltage amplitude should be available to permit adjustment for the excess-drive condition characterized by the appearance of a bright vertical line near the center of the raster. The correct driving-voltage amplitude is slightly less than the overdrive condition.

Tube type 6BK4 is used in a shunt-regulator circuit capable of maintaining the 25-kilovolt ultor voltage within ± 5 per cent from zero beam current to full 800-microampere load current. Ultor-voltage regulation is used to prevent changes in picture brightness, size, convergence, color purity, and horizontal linearity which tend to accompany large variations in beam current. The unloaded transformer is designed to deliver about 32 kilovolts. The regulator load in shunt with the kinescope draws a total current of about 800 microamperes from the rectifier at the 25-kilovolt setting, with zero kinescope current drain. Dark picture scenes, which draw little current from the kinescope, and lighter scenes requiring heavier current do not change the total load in the deflection circuit because the current through the regulator tube drops as the kinescope current increases. The regulator tube is capable of dissipating all the energy provided by the system, and does so when the kinescope is dark.

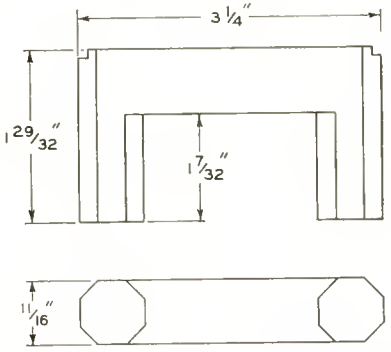
Transformer-Design Considerations

Available literature[1] includes several excellent detailed procedures on deflection-transformer design which may be somewhat simplified in

[1] O. H. Schade, "Electro-Optical Characteristics of Television Systems," *RCA Review*, Vol. IX, March, June, September, and December, 1948.

practice. When the yoke sensitivity and inductance are known, the tube complement is defined, and the supply voltage and boosted B power requirements are indicated, the following design method may be used:

1. *Core Size*—Determine the ferrite-core size, using previous designs as a guide. The ultor voltage specified will require a coil-winding structure and spacing from the grounded core which necessitate the use of an adequate window area. The core used for the developmental transformer is made of manganese ferrite and has a cross-sectional area of 2.55 square centimeters. Dimensions are given in Figure 15.



ORDER OF WINDING	TURNS	COIL DATA TAPS	WIRE SIZE	WINDING WIDTH (CAM)	CROSSES PER TURN (X/T)
Secondary #1	68	23, 32	0.0100 SNHF	2"	1/4
Secondary #2	15	-	0.0100 SNHF	2"	1/4
Primary	810	22, 33, 50, 63, 108, 243, 477, 518, 788	2X 0.0100 SNHF	2"	1/4
Tertiary	1700	-	0.0045 DNSF	3/16"	4

Fig. 15—Core and winding data for developmental horizontal-output and high-voltage transformer.

2. *Secondary Turns*—Calculate the number of turns required for the transformer secondary. Although an inductance of 5 to 10 times the yoke inductance is usually satisfactory, the use of a yoke requiring high deflection current assures the storage of adequate energy for good regulation and permits the use of higher ratios. The core is gapped to minimize acoustic radiation and to dilute the magnetic circuit, thus reducing the variation in inductance between cores of different permeability.

The following expressions are used to determine the number of turns required for the yoke portion of the transformer secondary winding.

$$L_S = KN_y^2,$$

where N_y = number of transformer secondary turns for yoke portion of the winding,

K = proportionality factor for the core used

$= 6.8 \times 10^{-7}$,

L_S = required inductance of yoke section of transformer secondary

$= 11 \times 11.8 = 130$ millihenries

$$N_y = \sqrt{\frac{130 \times 10^{-3}}{6.8 \times 10^{-7}}} = 435 \text{ turns required.}$$ (The final optimized design included 455 turns.)

3. *Flux Density*—Calculate the flux density, B_{max}, using the formula

$$B_{max} = \frac{\hat{e} \times 10^8}{\omega N_y A}$$

where $\hat{e}$ = peak voltage across the yoke,

A = cross-sectional area of the core = 2.55 square centimeters

$\omega = 2\pi f$

f = retrace frequency

$$B_{max} = \frac{2970 \times 10^8}{2\pi \times 50{,}000 \times 435 \times 2.55}$$

$= 855$ gausses

$$f = \frac{1}{2T}$$

T = retrace time (assume 10 microseconds)

$$f = \frac{1}{2 \times 10 \times 10^{-6}} = 50 \text{ kilocycles.}$$

This flux density, which is lower than that normally used in monochrome television transformers, minimizes core losses and resulting ultor power slump over the range from room temperature to the maximum operating temperature of 100°C.

4. *Transformer Turns Ratios*—Practical ratios between driver and damper tube are determined by losses in the system, and are usually about 1.5 or 2.0 to 1. The optimum value is best determined empirically by measurement of performance on a sample

transformer having extra taps. The sawtooth of current through the yoke results in a voltage drop, E_y, across the yoke section of the secondary which reduces the plate voltage of the driver tube.

$$E_y = L\frac{di}{dt} = \frac{11.8 \times 10^{-3} \times 0.9}{27 \times 10^{-6}} = 393 \text{ volts.}$$

The plate voltage of the driver must not become less than the voltage needed to supply the required peak plate current or low enough to cause excessive screen dissipation.

$E_{bb\,(\text{assumed})}$ = Boosted B voltage of $2(+B)$ = 800 volts

E_b = Minimum voltage on plate of driver = 100 volts

$$\frac{N_p}{N_y} = \frac{(E_{bb} - E_b)}{E_y} = \frac{(800 - 100)}{(393)}$$

where N_p = turns from a-c ground to driver plate,

$N_y = 435$,

$$N_p = 435 \times \frac{700}{393} = 775 \text{ turns (788 turns used on optimized design)}$$

Improvement in efficiency resulted when the damper cathode was connected to a transformer tap having a voltage approximately 10 per cent higher than that of the yoke.

5. *High-Voltage Tertiary*—The exact number of turns required in the tertiary cannot readily be calculated because of resonance effects which add to or subtract from the step-up voltage based upon the yoke retrace pulse. The approximate turns, however, may be calculated as follows:

The theoretical retrace voltage developed across the yoke, $\hat{e}_y$, is given by

$$\hat{e}_y = \hat{\imath}\omega L = 0.9\ (2\pi \times 50{,}000)\ 11.8 \times 10^{-3} = 3300 \text{ volts,}$$

$$\hat{e}_t = \hat{e}_y \left(\frac{N_t}{N_y}\right)$$

where $\hat{e}_t$ = peak tertiary voltage $= \hat{e}_{\text{ultor}} - \hat{e}_y$

$\hat{e}_y$ = peak pulse voltage across yoke

N_t = number of total turns

$N_y = 435$.

Therefore, 25,000 (less boosted B voltage of 800 volts) -3300

$$= 3300 \left(\frac{N_t}{N_y} \right)$$

$N_t = 2755$ turns.

A sample transformer was checked for performance as turns were removed; 2510 total turns were used on the final transformer design.

6. *Coil Configuration*—The primary and secondary coils, listed in Figure 15, are universal-wound wide coils for tight coupling and good deflection sensitivity; the high-voltage tertiary is made in a narrow coil for reduced distributed capacity to improve retrace time. The wire diameters used, which are listed in the tabulated data, are large enough to keep the winding-temperature rise within safe limits. Insulation used is adequate for the operating voltages.

7. *Power-Conversion Efficiency*—The 230FD1 yoke requires relatively high peak-to-peak current — 1.8 amperes for full 70-degree horizontal deflection. This condition is favorable for the design of a transformer which must operate at a relatively high conversion efficiency.

Power input for horizontal scanning and high voltage $= E_B(I_p) = 400\ (0.191) = 76.5$ watts.

$$\text{Conversion efficiency} = \left[\frac{\text{Focus} + \text{Ultor} + \text{Boosted B+ Rectifier Filament Power Output}}{\text{Power Input}} \right]$$

$$= \left[\frac{5{,}500\,(225 \times 10^{-6}) + 25{,}000\ (800 \times 10^{-6}) + 800\,(0.002) + 3.15\,(0.2)}{76.5} \right]$$

$$= .307 = 30.7\%.$$

This value is appreciably higher than the 10- to 15-per cent conversion efficiency provided by deflection systems presently used for 21-inch 90-degree monochrome kinescopes.

Plate Dissipation on 6CB5

A dynamic test was made to obtain temperature readings on the 6CB5 bulb at normal operating level. A static run was made, during

Table I — Typical Performance Data on Circuit Using Developmental Horizontal-Output and High-Voltage Transformer and Deflecting Yoke 230FD1

B voltage	400 volts
Ultor Power	20 watts
Boosted B Voltage	800 volts
Horizontal-Output Tube 6CB5	
Peak Positive-Pulse Plate Voltage	6200 volts
Grid-No. 2 Voltage	180 volts
Grid-No. 1 Voltage (Sawtooth)	190 volts p-p
Plate Current	193 milliamperes
Grid-No. 2 Current	18 milliamperes
Cathode Current	211 milliamperes
Grid-No. 2 Input	3.24 watts
Damper Tube 6BL4	
Peak Inverse Plate Voltage	3860 volts
Plate Current	193 milliamperes
Peak Heater–Cathode Voltage	4260 volts
Developmental Horizontal-Output and High-Voltage Transformer	
Pulse Voltage at Terminals (Measured to Ground, Terminal C)	
Terminal D	+250 volts
Terminal B	— 60 volts
Terminal A	—220 volts
Terminal E to F	100 volts
Vertical-Output Tube 6BL7-GT	
Plate Supply Voltage	400 volts
Plate Voltage	354 volts
Peak Positive-Pulse Plate Voltage	1440 volts
Grid Voltage (Referenced to cathode)	— 23 volts
Plate Current	28 milliamperes

Heat run—The final core temperature was 68°C, the transformer primary winding was 65°C, and the width-control winding was 50.5°C in a 24°C ambient temperature. Tests were made in the open, uncovered, with 10 per cent increase in $+B$ voltage and the width-control switch on terminal #4.

which screen power input was maintained unchanged from the dynamic test condition, and d-c plate power input was plotted as a function of bulb temperature. Plate dissipation was 23 watts, which is equal to the design-center value for this tube.

Table II — Deflection-Yoke 230FD1 Sensitivity

Horizontal-Deflection Coils:	
Current required for full 70-degree deflection (peak-to-peak)	1.8 amperes
Vertical-Deflection Coils:	
Current required for full 55-degree deflection (peak-to-peak)	0.5 ampere

Vertical-Deflection Transformer

The 230FD1 yoke requires 500 milliamperes (approximately) peak-to-peak current to scan the 21-inch color kinescope vertically. The circuit shown in Figure 16 includes a vertical-blocking-oscillator transformer (208T2) used with one triode section of a 6SN7-GT to provide the driving signal for the 6BL7-GT vertical-deflection amplifier. The two triode sections of the 6BL7-GT, are connected in parallel.

Good performance was achieved with a low-cost transformer (247FT1) having a ratio of 8:1, made with a square stack of the ⅞-inch center leg 0.014 inch XXX grade lamination. The primary was 3,400 turns of 0.0063 inch diameter enamel wire; the secondary

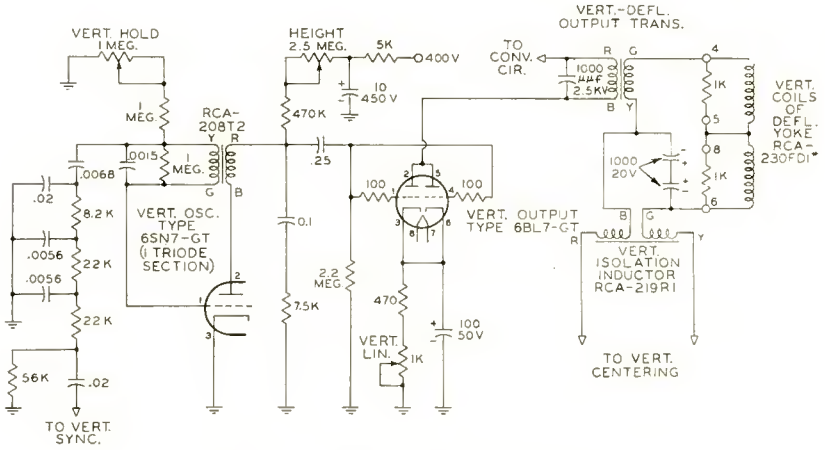


Fig. 16—Vertical-deflection-output circuit.

was 425 turns of 0.0113 inch diameter enamel wire. Nominal primary impedance with a 4-mil gap spacer, measured at 30 volts and 60 cycles per second with 0.025 ampere d-c is 10,000 ohms.

Convergence Components

The deflection of a three-beam triangular array originally converged in the center of the kinescope has been shown to result in misconvergence proportional to the separation between the beams in the deflection plan and the deflection angle scanned. The 21AXP22 kinescope is designed to allow correction for this misconvergence, and has an electron gun equipped with pole pieces for static and dynamic control of each individual beam. Each pole piece supplies a magnetic field with a d-c flux component to obtain center convergence, plus dynamic components of flux to maintain convergence at full screen deflection.

The converging magnet assembly shown in Figure 17 consists of three sets of cores and coils, one to control each of the three beams. Each core assembly consists of two L-shaped pieces butted together to form a U core; the butting surfaces are slotted to accept a circular magnet. The magnet is polarized normal to its axis, and may be rotated to provide the d-c flux for center-screen convergence. Two separate bifilar wound coils on each core permit flexibility in the design of the circuit for providing the necessary horizontal- and vertical-dynamic convergence voltages.

Convergence of the three beams in the center of the screen is obtained with four separate controls. The three radial position adjustments are obtained with the converging magnet assembly. The blue-

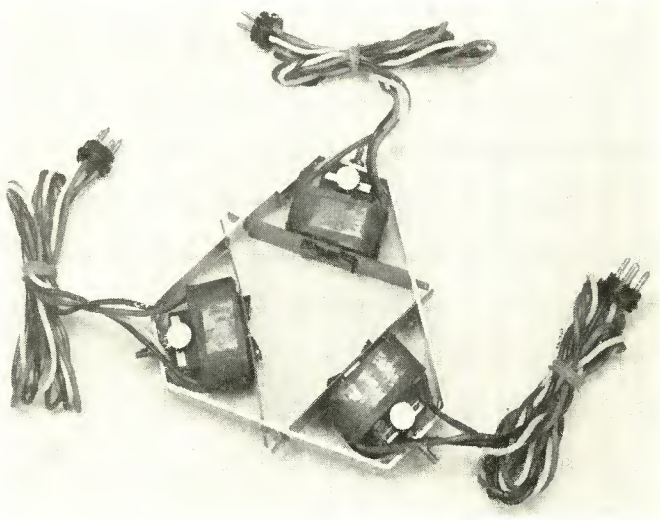


Fig. 17—Converging-magnet assembly.

positioning magnet shown in Figure 18 permits lateral movement of the blue beam so that all three beams may be converged in the center of the screen.

The Purifying Magnet shown in Figure 18 permits adjustment of the angle of approach of the three-beam array so that it strikes the shadow mask at the correct angle to excite the phosphor dots symmetrically, and produce uniform color rendition. Rotation of the separate magnets together, or with respect to each other, provides a magnetic field adjustable in direction and intensity from 0.1 to 9 gausses.

A magnetic-field equalizer assembly positioned around the faceplate end of the 21AXP22 color kinescope provides sectionalized magnetic

fields to permit compensation in localized areas for the effects of stray magnetic fields and the earth's magnetic field on color purity

The equalizer consists of eight separate magnet-and-shunt assemblies, each mounted in position over four flux-conducting pole pieces. The magnets may be adjusted to change direction and magnitude of the flux in the peripheral area of the screen adjacent to each magnet to provide beam-landing position adjustment.

The assembly is mounted between the flange and faceplate. Adequate insulation must be provided between the equalizer assembly and the metal shell of the kinescope to permit operation of the equalizer assembly at ground potential.

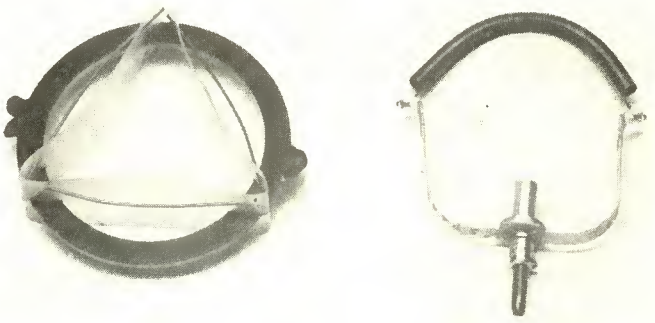


Fig. 18—Purifying magnet (left) and blue-positioning magnet.

Dynamic-Convergence Considerations

Early attempts to scan wide-angle color kinescopes were made using complex convergence circuits which had four tubes and included two special corner-correction amplifiers. Improvements made in the deflecting yoke minimized red–green separation in the corners and permitted elimination of this special type of correction, with resulting circuitry simplification and cost reduction.

Convergence Circuits

The departure of a blue horizontal line from a straight line was checked, using a cross-hair telescope rigidly mounted on a calibrated rack and pinion. The curve obtained, shown in Figure 19, approximates a parabola for which the equation is $y = 2\ (0.070)\ x^2$.

Several convergence circuits have been used successfully to obtain properly shaped current waveforms at the convergence-magnet assembly. One of the circuits, shown in Figure 20, included the use of tube amplifiers to provide the basically parabolic waveforms. Pro-

vision was also made for adding positive- or negative-sloped sawtooth shaping as required for symmetry correction.

Dynamic-convergence correction in the top and bottom picture areas may be obtained by use of the shaping network in the plate circuit of the vertical-output tube. Shaping for symmetry correction is provided as required with a current sawtooth waveform developed across the 5,000-ohm center-tapped pots.

Dynamic-convergence correction at the raster sides is obtained with similarly shaped current waveforms at the horizontal-deflection rate. One triode section of the 6BL7-GT is used as a sawtooth amplifier referenced to the horizontal-output transformer. The other section of the 6BL7-GT and the two sections of a 12BH7 are used as blue-,

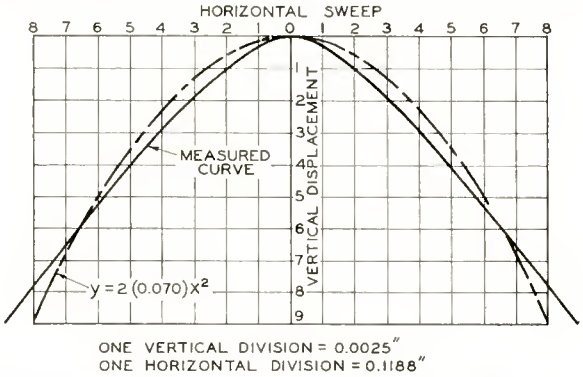


Fig. 19—Departure of a blue horizontal scanning line from a straight line.

red- and green-convergence amplifiers. A positive or negative pulse from the deflection transformer added to the sawtooth provides the required waveform.

Figure 21 shows the bar pattern completely converged.

Figure 22 shows the typical pattern obtained with an average 230FD1 yoke and 21AXP22 kinescope, with the center converged, and with no dynamic correction applied.

Figure 23 shows the horizontal stripes only, with no dynamic convergence correction. Blue is low at the sides, requiring appreciable correction, but red and green are converged within less than 0.045 inch throughout the raster. The yoke-flux pattern is designed to provide this characteristic, which makes practicable the use of low-cost horizontal dynamic-convergence correction circuitry.

Figure 24 shows the converged lines after the adjustment of horizontal and vertical dynamic-convergence amplitude.

Currently used vertical-convergence circuitry provides the required parabolic waveforms of sufficient amplitude. Figure 25 shows a vertical

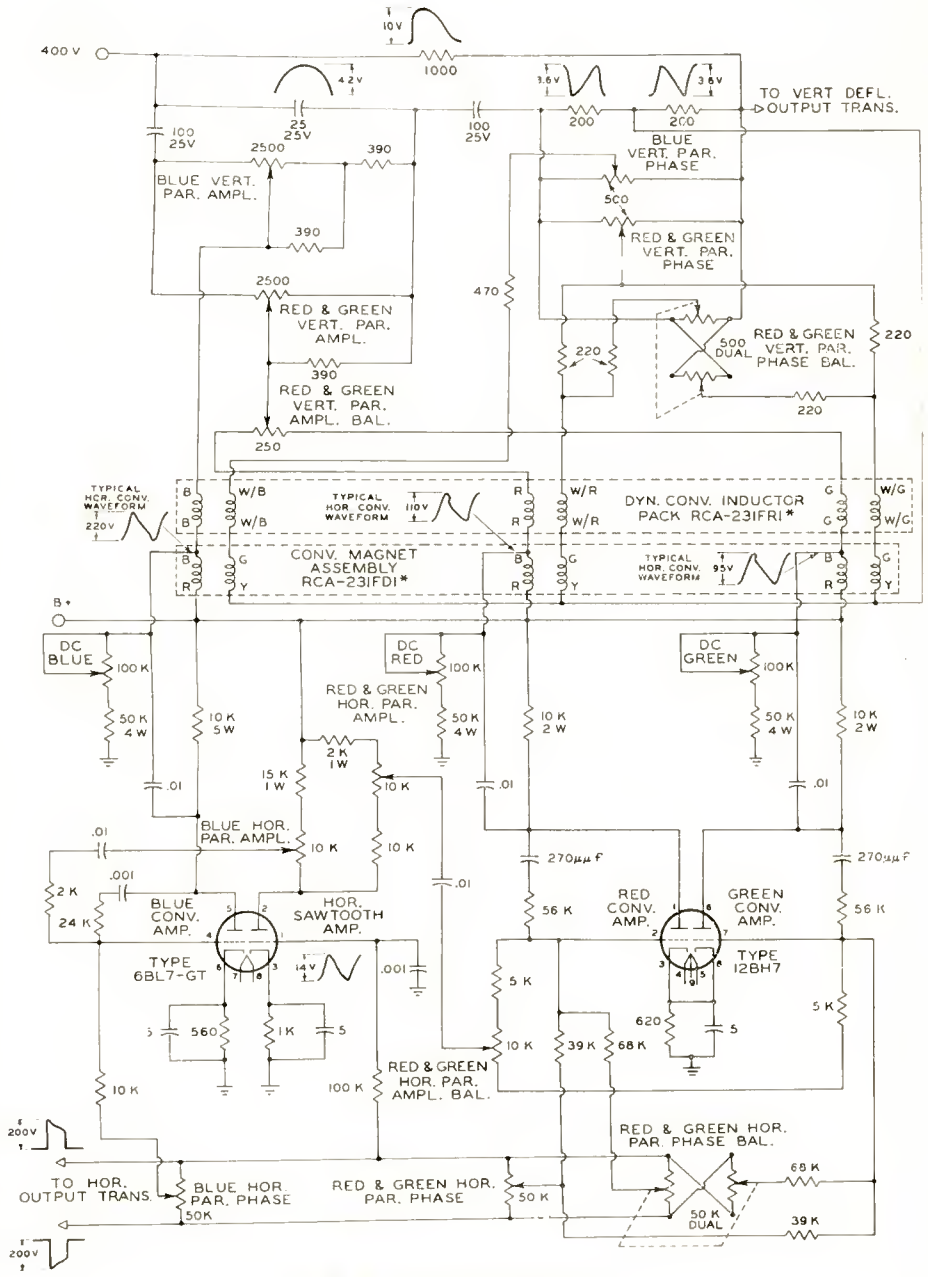


Fig. 20—Convergence circuit.

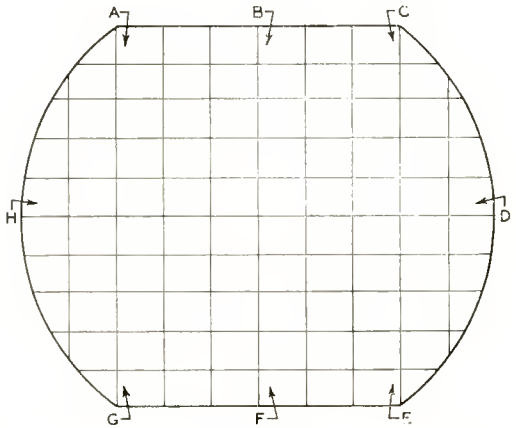


Fig. 21—Bar pattern illustrating proper convergence. The red, green, and blue horizontal and vertical bars are converged over the entire raster within 1/32 inch at B, D, F, and H, and within 1/16 inch at A, C, E, and G.

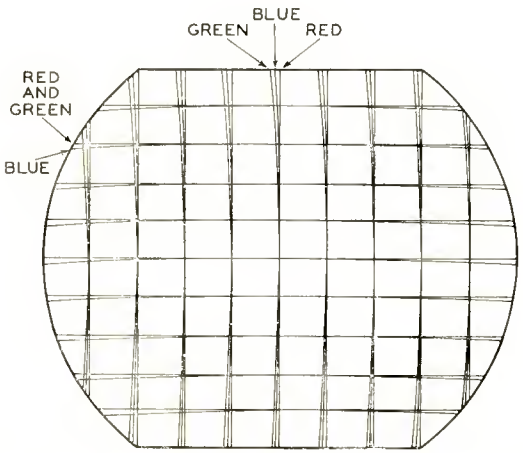


Fig. 22—Typical test pattern obtained with average yoke and kinescope. No dynamic-convergence correction. Red and green horizontal bars are converged on horizontal lines within 0.045 inch. Blue horizontal bars are low at sides of pattern, indicating need for dynamic-convergence correction. The center of the blue vertical bar at the edges of the pattern is within 1/16 inch of the center line between the red and green vertical bars.

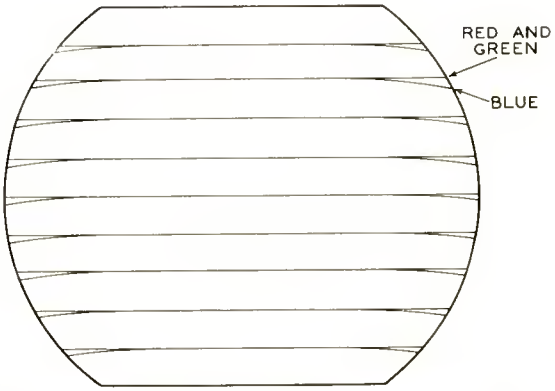


Fig. 23—Test pattern showing horizontal bars only. No dynamic-convergence correction applied. Red and green horizontal bars are converged on horizontal lines within 0.045 inch. Blue horizontal bar is low at edges of pattern, indicating need for dynamic-convergence correction.

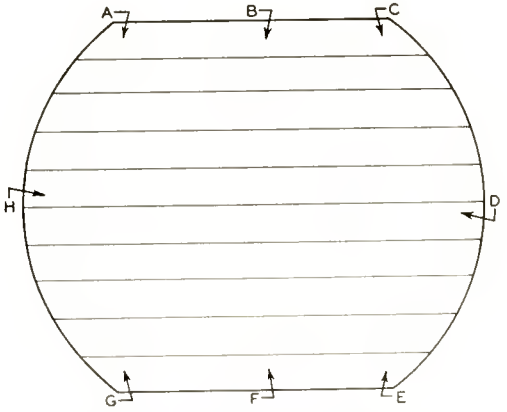


Fig. 24—Test pattern showing horizontal bars only after dynamic-convergence correction is applied by adjustment of horizontal and vertical dynamic-convergence controls. The red, green, and blue horizontal bars are converged over the entire raster within 1/32 inch at B, D, F, and H, and within 1/16 inch at A, C, E, and G.

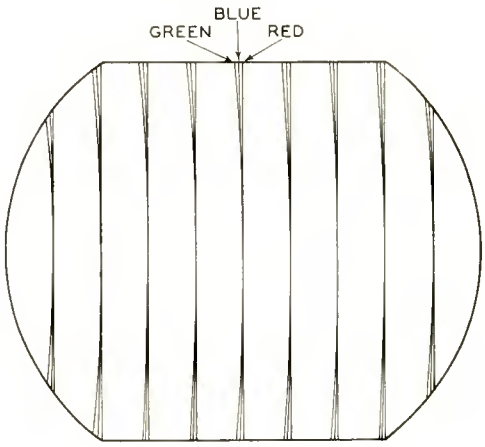


Fig. 25—Test pattern showing vertical bars only. No dynamic-convergence correction applied. The center of the blue vertical bar at the edges of the pattern is within 1/16 inch of the center line between the red and green vertical bars.

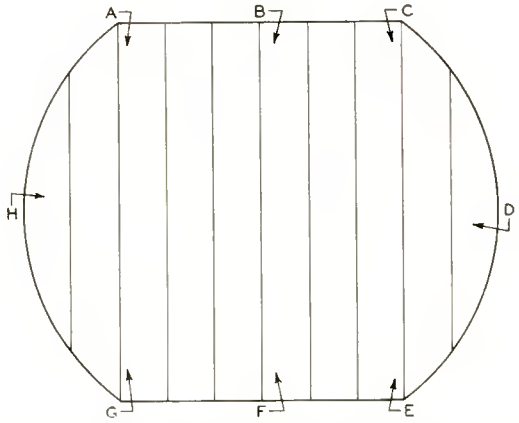


Fig. 26—Test pattern showing vertical bars only after dynamic-convergence correction is applied by adjustment of vertical dynamic-convergence controls. The red, green, and blue vertical bars are converged over the entire raster within 1/32 inch at B, D, F, and H, and within 1/16 inch at A, C, E, and G.

line pattern with the center converged, and with no dynamic correction. Red and green are bowed toward each other, almost symmetrically about the horizontal axis.

Figure 26 shows the red and green vertical lines straightened and made parallel with the vertical dynamic-convergence amplitude control. The vertical-deflection-frequency dynamic-tilt control is used to cause the maximum beam movement to occur above or below the center.

Acknowledgment

The horizontal-deflection transformer was developed by B. V. Vonderschmitt. The yoke development and product design was by W. H. Barkow, J. K. Kratz, and C. C. Matthews. L. E. Annus and P. E. Wiseman helped with the component testing. M. Bayer built the deflection test set, and flux patterns were made by D. J. Coyle and B. McHugh.

Appendix — Measurement of Losses in Ferrite Core of Deflecting Yoke 230FD1

The core loss was measured in two different ways. In the first method, a thermal meter was used to measure the voltage induced in test coils wound around the core segment, in the positions shown in Figure 27. This voltage

$$e = N \frac{d\phi}{dt} \times 10^{-8}$$

would have the waveform shown in Figure 27(a), if the flux corresponds to Figure 27(b).

With the flux variation

$$\phi_1 = K_1 t_1,$$

then $$e_1 = K_1$$

from Figure 27(b) $$e_1 = \frac{\Phi \times 10^{-8}}{26.55 \times 10^{-6}} = K_1$$

$$e_2 = \frac{\Phi \times 10^{-8}}{5.15 \times 10^{-6}} = K_2.$$

The induced voltage E_{rms} (for a single turn at horizontal frequency) $$= \sqrt{\frac{K_1^2 t_1 + K_2^2 t_2}{T}}.$$

$$= \sqrt{\frac{\Phi^2 \times 10^{-4}}{(26.55)^2}(53.1 \times 10^{-6}) + \frac{\Phi^2 \times 10^{-4}}{(5.15)^2}(10.3 \times 10^{-6})}{(63.4 \times 10^{-6})}}$$

$$= 8.55 \times 10^{-4}\,\Phi$$

$$\Phi = \frac{E_{rms} \times 10^{-4}}{8.55} = 1170\,E_{rms} \text{ maxwells.}$$

With the search coil at position 1, $E_1 = 1.34$ volts for 2 turns, and 0.67 for 1 turn. Cross sectional area core $= 5.56$ cm^2, $B = \frac{\Phi}{A}$

$$B_1 = \frac{1170\ (0.67)}{5.56} = 141 \text{ gausses.}$$

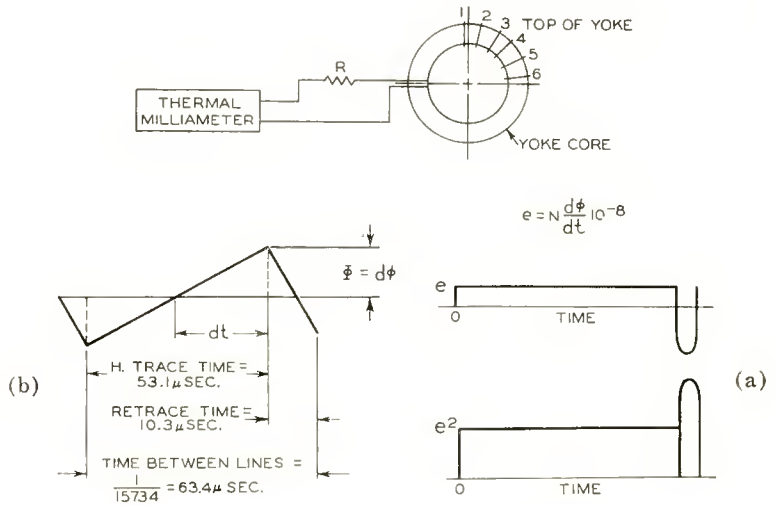


Fig. 27—Theoretical waveforms used in yoke-core flux density determination at the horizontal-deflection frequency.

At vertical-deflection frequency, trace time $= 1/60$ second minus return time of about 4 per cent, or .016 second. (See Figure 28.) The voltage induced in the test coils, with just the vertical coils energized is

$$e = N\frac{d\phi}{dt_1} \times 10^{-8}$$

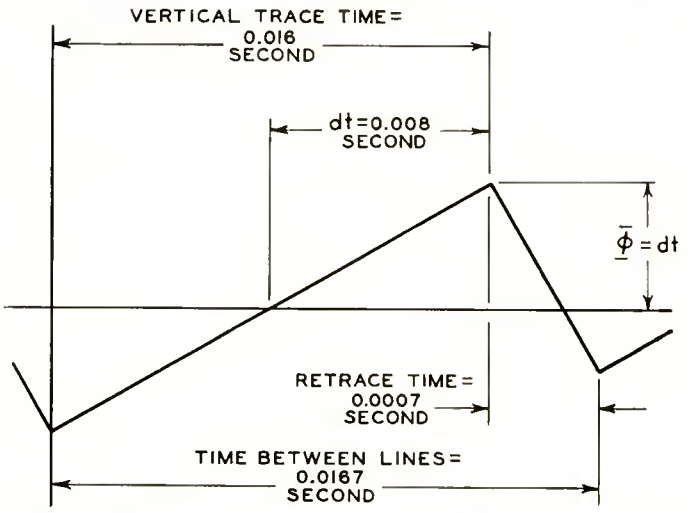


Fig. 28—Theoretical vertical-deflection frequency waveform.

Table III — Dynamic Test Method for Determining Flux Density

	Horizontal Frequency						
	Position	Measured Current (amperes)	Measured Resistance (ohms)	Volts	Volts per turn	Φ (max-wells)	B (gausses)
Seg. 1	1	0.042	31.97	1.34	0.67	782.	141.
Seg. 2	2	0.055	31.96	1.757	0.878	1025.	185.
Seg. 2	3	0.052	31.96	1.66	0.83	972.	175.
Seg. 2	4	0.064	31.97	2.042	1.021	1200.	216.
Seg. 2	5	0.082	31.97	2.62	1.31	1535.	276.
Seg. 2	6.	0.086	31.99	2.75	1.375	1610.	290.
					Avg. (Seg. 2) =		228. gausses

	Vertical Frequency				Total Flux Density (gausses)
	Position	Volts (20-turn coil)	Φ (max-wells)	B (gausses)	
Seg. 1	1	0.19	1570.	282.	423.
Seg. 2	2	0.23	1900.	342.	527.
Seg. 2	3	0.21	1737.	312.	487.
Seg. 2	4	0.18	1490.	268.	484.
Seg. 2	5	0.14	1160.	209.	485.
Seg. 2	6	0.094	777.	140.	430.
			Avg. (Seg. 2) =	254. gausses	

$$E_{rms} = \sqrt{\frac{\frac{\Phi^2 \times 10^{-10}}{8^2}(16 \times 10^{-3}) + \frac{\Phi^2 \times 10^{-10}}{.35^2}(0.7 \times 10^{-3})}{16.7 \times 10^{-3}}}$$

$$= .600\Phi \times 10^{-5} \text{ volts per turn and, for a 20-turn coil}$$

$$\Phi = \frac{E_{rms} \times 10^5}{20 \times .600} = 8340\, E_{rms} \text{ maxwells,}$$

$$B_{\max} \text{ (vertical coil)} = \frac{\Phi}{A} = \frac{8340 \times 0.19}{5.56} = 285 \text{ gausses. (position 1)}$$

This method of determining flux density was checked with a static test, using a ballistic galvanometer, energizing the coil with d-c as required for full deflection, and then opening the circuit. Values obtained were reasonably in agreement with those determined by the first method.

BIBLIOGRAPHY

O. H. Schade, "Characteristics of High Efficiency Deflection and High Voltage Supply Systems for Kinescopes," *RCA Review*, Vol. XI, March, 1950.

H. C. Moodey and D. D. Van Ormer, "Three-Beam Guns for Color Kinescopes," *RCA Review*, Vol. XII, September, 1951.

A. W. Friend, "Deflection and Convergence in Color Kinescopes," *RCA Review*, Vol. XII, September, 1951.

RCA TECHNICAL PAPERS†

Fourth Quarter, 1954

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

"A Bridge for the Measurement of Cathode Impedance," R. M. Matheson and L. S. Nergaard, *RCA Review* (December) 1954
"Cable Equalization for Television Studio Circuits," R. C. Kennedy, *RCA Review* (December) 1954
"Characteristics of a Transmission Control Viewing Storage Tube with Halftone Display," M. Knoll, H. O. Hook, and R. P. Stone, *Proc. I.R.E.* (October) 1954
"Citizens Radio," H. B. Scott and L. G. Sands, *Radio-Electronic Engineering* (October) 1954
"Color Test Techniques for TV Broadcasters," J. W. Wentworth, *Electronics* (November) 1954
"Comment on 'Ages of Creativeness of Electronic Engineers'," E. A. Laport, *Proc. I.R.E.* (December) (Letter to the Editor) 1954
"Determination of Surface Zinc Oxide on Zinc Sulfide Phosphors," S. Larach and S. M. Thomsen, *Analytical Chemistry* (October) 1954
"Determination of Typical Operating Conditions for RCA Tubes Used as Linear RF Power Amplifiers," A. P. Sweet, *RCA Ham Tips* (December) 1954
"'Electrofax' — Direct Electrophotographic Printing on Paper," C. J. Young and H. G. Greig, *RCA Review* (December) 1954
"Elements of Boolean Algebra for the Study of Information-Handling Systems," R. Serrell, *Proc. I.R.E.* (October) 1954
"Experimental Closed-Circuit Television," R. G. Neuhauser, *Electronic Equipment* (December) 1954
"Improvement of Base Adherence on Electron Tubes," A. M. Seybold, *Ceramic Age* (December) 1954
"Investigations of Noise in Audio Frequency Amplifiers Using Junction Transistors," M. B. Herscher and P. L. Bargellini, *RCA Industry Service Laboratory Bulletin LB-964* (October 1) 1954
"Magnetic Anisotropy Constants of Ferromagnetic Spinels," R. S. Weisz, *Phys. Rev.* (November 1) (Letter to the Editor) 1954
"Microwave Applications of Gas Discharges," F. R. Arams, *Electronics* (November) 1954
"Observed Diurnal Variations in Frequencies and Signal Qualities between New York and Central Europe," J. H. Nelson, *RCA Review* (December) 1954
"Photoconductivity," A. Rose, *L'Onde Electrique* (October) 1954
"Photoconductivity and Crystal Imperfections in Cadmium Sulfide Crystals," R. H. Bube and S. M. Thomsen, *Jour. Chem. Phys.* (December) 1954

† Report all corrections or additions to RCA Review, RCA Laboratories, Princeton, N. J.
* *RCA Industry Service Laboratory Bulletins* are not published and are issued only as a service to licensees of the Radio Corporation of America.

"Practical Considerations in the Design of Low-Microphonic Tubes," T. M. Cunningham, *RCA Review* (December) 1954

"Semiconducting Cadmium Telluride," D. A. Jenny and R. H. Bube, *Phys. Rev.* (December 1) 1954

"Shielding and Potting," S. J. Burruano, E. F. Bailey, and S. Cramer, *Electronics* (October) 1954

"Sine-Squared Pulses Test Color-TV Systems," R. C. Kennedy, *Electronics* (December) 1954

"Study of P-N-P Alloy Junction Transistor from D-C through Medium Frequencies," L. J. Giacoletto, *RCA Review* (December) 1954

"TG-2A Studio Sync Generator," A. C. Luther, Jr., *Broadcast News* (December) 1954

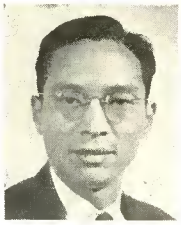
"Transmitter Space Diversity Applied to Shipboard Reception," G. E. Hansell, *Transactions of the I.R.E. Professional Group on Communication Systems* (November) 1954

"The Traveling-Wave Tube," S. Bloom, *Yale Scientific Magazine* (December) 1954

"U.H.F. and V.H.F. Antenna," R. F. Kolar, *Radio-Electronics* (December) 1954

AUTHORS

WARREN H. BLISS studied electrical engineering at Michigan State College, receiving a B.S. degree in 1928 and an M.S. degree in 1931. During the period, 1931-1943 he taught electrical engineering and communications courses at the University of Maine. After spending several summers with the research group at the RCA Communications terminal in New York City he joined RCA Laboratories as a full time research engineer in 1943. His work has included radio facsimile, high-speed-pulse radiotelegraphy, secrecy communication, and radar range circuit calibration. At present he is concerned with electronic counters and special cathode-ray tube display devices. Mr. Bliss is a Member of the Institute of Radio Engineers and the American Institute of Electrical Engineers.



K. K. N. CHANG received the B.S. degree from the National Central University, Nanking, China, in 1940; the M.S. degree in Electrical Engineering from the University of Michigan in 1948, and the D.E.E. degree in 1954 from the Polytechnic Institute of Brooklyn. From 1940 to 1945, he was associated with the Central Radio Manufacturing Works, Kunming, China, working on radio receivers and from 1945 to 1947, he was a radio instructor in the Office of Strategic Service, U. S. Army, China Theatre. Since 1948, he has been at RCA Laboratories, Princeton, N. J., where he is presently engaged in research on microwave tubes. Dr. Chang is a member of Sigma Xi.

RALPH W. ENGSTROM received the B.S. degree from St. Olaf College in Northfield, Minnesota in 1935. He subsequently attended Northwestern University in Evanston, Illinois, and received the M.S. degree in 1937 and the Ph.D. in 1939. He joined the RCA Tube Division in Harrison, N. J. in 1941, and was engaged for several years in the development of multiplier phototubes and related devices. He is now in charge of the applications engineering activity on pickup, photo, and special tubes in the Cathode-Ray and Power-Tube Organization at Lancaster, Pennsylvania. Dr. Engstrom is a member of Sigma Pi Sigma and of the honorary physical society Sigma Psi, and is a Fellow of the American Physical Society.

L. J. GIACOLETTO received the B.S. degree in Electrical Engineering from Rose Polytechnic Institute, Terre Haute, Indiana, in 1938, and the M.S. degree in Physics from the State University of Iowa in 1939. From 1939 to 1941, while a Teaching Fellow in the department of Electrical Engineering at the University of Michigan, he engaged in frequency-modulation research. He received the Ph.D. degree in Electrical Engineering from the University of Michigan in 1952. He was associated with the Collins Radio Company during the summers of 1937 and 1938, and with the Bell Telephone Laboratories in 1940. From 1941 to 1945 he served with the Signal Corps and returned to inactive status as a major in the Signal Corps Reserve in May, 1946. Since June, 1946, he has been serving as a research engineer at RCA Laboratories, Princeton, N. J. Dr. Giacoletto is a member of the American Association for the Advancement of Science, Gamma Alpha, Iota Alpha, Phi Kappa Phi, Tau Beta Pi, and Sigma Xi.

MANUEL H. GREENBLATT received the B.A. degree from the University of Pennsylvania in 1943. In 1949, he received the Ph.D. degree in Physics from the same University. While studying there he was an Assistant Instructor in Physics and he also engaged in fundamental research in the field of solid-state physics and electronics. In 1948, he joined the RCA Laboratories, Princeton, N. J., where he is at present engaged in research in the field of photomultipliers and scintillation counting. Dr. Greenblatt is a member of the American Physical Society and of Sigma Xi.

JAMES R. HALL received the B.S. degree in Electrical Engineering from Seattle University in 1951. He joined RCA in 1951 where he was a specialized trainee. In 1952 he became a member of the Optics, Sound and Special Engineering Section. With this group he participated in development of tape mechanisms for use as information storage units in computers. Since 1953 he has been engaged in magnetic recording of high-frequency signals and development of associated electronic circuitry. The investigation of single-sideband systems has been a part of Amateur radio activities, where he operates with the call K2AER. Mr. Hall is an Associate Member of the Institute of Radio Engineers.

WERNER HASENBERG received the B.S. Degree in Electrical Engineering in 1921 at Mittweida Engineering School, Germany. He joined the Research Laboratories of the Osram Company, Berlin, Germany, where he was engaged in studies of gas discharge phenomena and tungsten single-crystal development. In 1928 he joined the German Radio Corporation Telefunken, Berlin, Germany to organize and manage the Tube Application Engineering Division. During 1934 he emigrated to France, where he became associated with the Fotos Tube Factory until, in 1935, he was sent to Brazil by the Philips Company as a technical assistant to the management of their Rio de Janeiro branch. In 1940 he became associated with the Televox Company, Rio de Janeiro, Brazil, where he developed and manufactured switchless Intercoms of his own invention. He joined RCA Victor, Rio de Janeiro, Brazil in 1946 and became engaged in operating and servicing electron microscopes. He came to the United States in May 1948 and joined RCA Laboratories, Princeton, N. J., where he was engaged in patent research work. In 1949 he joined the Tube Application and Semiconductor Coordination Group at RCA's Television Engineering Division, Camden, N. J. Mr. Hasenberg is a Senior Member of the I.R.E. and Vice Chairman of the Professional Group on Electron Devices of the Philadelphia Section of the I.R.E.

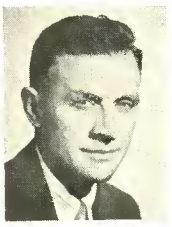
EDWARD O. JOHNSON received the B.S. degree in Electrical Engineering from Pratt Institute of Brooklyn in 1948. From 1941-1945 he served as an electronic technician in the U. S. Navy. In 1948 he joined the RCA Laboratories, Princeton, N. J., where he was engaged in research on gaseous electronics. Last fall he was given a leave of absence by RCA for a year of academic study abroad. He is at present attending the Swiss Federal Institute of Technology in Zurich, Switzerland. Mr. Johnson is a 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.

HANNAH C. MOODEY received the A.B. degree from Smith College in 1927, the M.Sc. degree in Physics from Rutgers University in 1933, and the B.S. degree in Electrical Engineering from Massachusetts Institute of Technology in 1936. For several years she was employed in the Long Lines Engineering Department of the American Telephone and Telegraph Company. Since 1936, she has been with the Tube Division of the Radio Corporation of America in Harrison, New Jersey, and in Lancaster, Pennsylvania, specializing in cathode-ray-tube design. For the past five years she has had engineering responsibility for the design and development of color kinescopes. Miss Moodey is a Senior Member of the Institute of Radio Engineers.

ALBERT M. MORRELL received the B.S. degree in Electrical Engineering from Iowa State College in 1950, and is currently doing graduate work at Franklin and Marshall College. From 1943 to 1946 he served with the U. S. Army Air Force. Since 1950 he has been with the Tube Division of RCA at Lancaster, Pa. as a design engineer on color kinescopes. Mr Morrell is a member of Phi Kappa Phi, Eta Kappa Nu, and the Institute of Radio Engineers.

MAXIMILIAN J. OBERT attended Drexel Institute of Technology in Philadelphia, Penna., from 1929 to 1932. He joined the Radio Corporation of America, Camden, N. J., in 1933 as a process engineer on transformers. He later attended evening courses at Drexel Institute, and received a diploma in Electrical Engineering in 1939. He was a manufacturing development engineer from 1940 to 1945, and then became a design engineer on transformer and deflection components. Since 1950, he has been Manager of Transformer Development for the Tube Division. Mr. Obert is a Senior Member of the Institute of Radio Engineers.



JOHN E. RUEDY received the A.B. degree from Western Reserve University in 1924 and the Ph.D. degree in Physics from Cornell University in 1933. From 1925 to 1927 he was an engineer with Bell Telephone Laboratories, from 1927 to 1934 held teaching and research assistantships at Cornell University, from 1934 to 1936 was with the Evaporated Metal Films Company, and from 1936 to the present time he has been with RCA, first with the RCA Manufacturing Company and later with RCA Laboratories. Dr. Ruedy is a member of Sigma Xi, the American Physical Society, and is a Fellow of American Association for the Advancement of Science.

HARRY R. SEELEN received the B.S. degree from Providence College in Rhode Island in 1929. Shortly thereafter he was employed as an engineer in the Radio Corporation of America's Tube Plant at Harrison, New Jersey, and was later made Manager of the Tube Development Shop. When part of the engineering activities was moved to Lancaster, Pennsylvania in 1942, he was transferred there as Manager of Engineering Services. He became Manager of Engineering at the Lancaster plant in 1949, and was recently appointed Operations Manager for Color Kinescopes for the RCA Tube Division. During his 25 years with RCA, he has been associated with such major tube developments as acorn tubes, miniature and metal receiving tubes, VHF and UHF transmitting tubes, and various camera tubes and kinescopes. Mr. Seelen is a member of the Institute of Radio Engineers, the American Association for the Advancement of Science, and Sigma Pi Sigma.

DAVID B. VANORMER received the B.S. degree in Physics from the University of Delaware in 1947, and has since done graduate work at Indiana University and Franklin and Marshall College. Since March, 1948 he has been employed at the Tube Division plant of RCA at Lancaster, Pa. as a cathode-ray-tube design engineer. At the present time, he is the engineering leader of the color-kinescope product design group of the engineering section. Mr. VanOrmer is a member of Phi Kappa Phi, Pi Mu Epsilon, and Sigma Pi Sigma.

WILLIAM M. WEBSTER studied physics at Rensselaer Polytechnic Institute and at Union College as a Navy V-12 student. He received a B.S. degree in Physics in 1945. He was released from active duty in 1946 and joined the RCA Laboratories Division, Princeton, N. J. in October of that year. He received the Ph.D. degree from Princeton University, where he was enrolled on a part-time basis. In December, 1954 he transferred to the Tube Division in Harrison, N. J. Dr. Webster is a member of Sigma Xi and a Senior Member of the Institute of Radio Engineers.

WILLIAM WIDMAIER received the B.S. Degree in 1934 and the M.S. Degree in 1935 from Pennsylvania State College. He joined the RCA Tube Division at Lancaster, Pennsylvania in 1949, and was engaged in developmental engineering on gas tubes for two years. He is now a member of the phototube design activity of the Cathode-Ray and Power-Tube Organization at Lancaster. Mr. Widmaier is a member of the physical society Sigma Pi Sigma.