RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH • ENGINEERING

VOLUME XVI DECEMBER 1955 NO. 4

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 | DECEMBER, 1955 | NUMBER 4

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

TRANSISTORS — I

A new book, *Transistors — I*, containing 33 original research papers by RCA authors, will be ready for distribution in the first quarter of 1956. These papers are all previously unpublished material and will not, in fact, appear in any other public medium. *Transistors — I*, therefore, will differ from prior books by RCA REVIEW which for the most part contained material published elsewhere.

The purpose of publishing *Transistors — I* is to provide a convenient reference source which should prove extremely useful to persons working with transistors, both those concerned with the fabrication of transistors and those engaged in circuit design. For the sake of completeness, the book will contain 10 additional papers which have been, or will be, published elsewhere. This makes a total of 43 papers. The composition of the book is approximately one half on semiconductor materials and devices and the other half on circuits and applications.

An announcement concerning the price and availability of *Transistors — I* will be sent in the near future to all *RCA Review* subscribers.

The Manager, *RCA Review*

IMPROVEMENT IN COLOR KINESCOPES THROUGH OPTICAL ANALOGY

By

D. W. Epstein,* P. Kaus,* and D. D. VanOrmer†

Summary—*In color kinescopes wherein the phosphor dots are deposited by the conventional optical exposure, the movement of the deflection center with deflection angle causes a radial misregister between the phosphor dots and electron spots. This misregister has been eliminated by interposing a thin aspheric lens between the light source and the aperture mask during exposure of the phosphor screen in the manufacture of the tube.*

Introduction

In the elementary theory of the shadow mask tube[1] it is assumed that the electrons in the three beams move in straight lines after deflection, appearing to originate at three fixed points lying in a plane perpendicular to the axis. This plane is known as the plane of the color centers. It is possible to determine the relationship between phosphor dots on the screen and the holes in the shadow mask by using light beams emanating from small light sources located at the red, green and blue color centers. When a complete tube is operated, electron beams take the place of these light beams. To obtain color purity the yoke is so placed that its plane of deflection is located in the plane of the color centers and the electron guns are so located that the electron beams intersect the plane of deflection just where the light source was placed during exposure. This procedure is accurate for relatively small deflection angles, but the larger the deflection angle the less exact becomes the register between phosphor dots and electron dots.[2] This is due to the fact that the electrons do not appear to originate from a fixed point as is assumed in the elementary theory. The plane of deflection of the yoke moves closer to the shadow mask as the deflection angle is increased.

The object of this paper is to discuss the causes of such misregister

* RCA Laboratories, Princeton, N. J.
† RCA Tube Division, Lancaster, Pa.

[1] H. B. Law, "A Three-Gun Shadow-Mask Color Kinescope," *RCA Review*, Vol. XII, pp. 466-486, September, 1951.

[2] H. R. Seelen, H. C. Moodey, D. D. VanOrmer and A. M. Morrell, "Development of a 21-inch Metal Envelope Color Kinescope," *RCA Review*, Vol. XVI, pp. 122-137, March, 1955.

in detail, to indicate some of the means that have been used in the past to minimize the effects, and to describe a new optical approach which eliminates the misregister almost completely.

Movement of the Plane of Deflection

It is well known that in present deflection yokes, the deflection plane travels forward with the angle of deflection. To be exact, the plane of intersection of the original and the final beams moves away from the source of the electrons as the deflection is increased. The precise amount of this movement depends on the magnetic field distribution of the yoke involved. However, a first-order expression may be obtained by assuming a uniform deflection field of length l. If the deflection angle is given by $2u$ and the forward movement by Δp, the relation is[3]

$$\Delta p = (l/2) \tan^2 (u/2). \tag{1}$$

If the yoke has an effective length of 5 inches, then the forward movement will be approximately .25 inch for a 70° deflection angle.

Misregister Caused by Movement

The phosphor on the faceplate is deposited at those spots where a light beam emanating from a fixed light source strikes the plate after having passed through the shadow mask. Measured along the axis (see Figure 1), the light source is at a distance p from the shadow mask, and the shadow mask is a distance q from the faceplate along the axis. The faceplate and shadow mask have an approximately equal radius of curvature R. The inclination of the faceplate with the vertical as a distance r from the axis is given by the angle u_R. The tangent of the angle u_R is

$$\tan u_R = r/\sqrt{R^2 - r^2}. \tag{2}$$

The effect of the forward movement of the deflection plane on misregister is illustrated in Figure 1, where the effect is greatly exaggerated for explanatory purposes. Assume that the light source S is located to agree with the apparent source of the electron beam for small deflection angles. The location of the phosphor dots is determined by

[3] A. W. Friend, "Deflection and Convergence in Color Kinescopes," *RCA Review*, Vol. XII, pp. 612-644, September, 1951.

the light rays originating from S. Thus at the larger deflection angle shown in Figure 1, the phosphor dot is deposited on the faceplate at a distance r from the tube axis. The electron beam passing through the same hole in the shadow mask and, thus, corresponding to this light ray, appears from a source at a distance $p - \Delta p$ from the mask and strikes the faceplate at a distance $r + \Delta r$ from the tube axis. The distance Δp (Equation 1) is the amount of forward movement of the deflection centers and the distance Δr is the amount of radial misregister.

The misregister Δr is determined from the displacement Δp through the geometry of the tube. A quantitative relationship derivable from

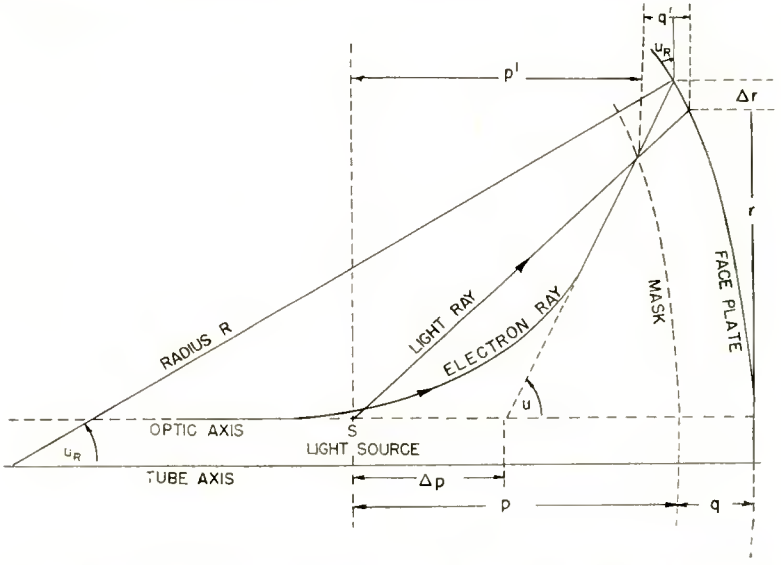


Fig. 1—Schematic diagram of electron ray and light ray geometry.

Figure 1 which holds to a good approximation is given by

$$\Delta r = \Delta p \frac{q}{p} \frac{\tan u}{1 + \tan u_R \tan u} \tag{3}$$

wherein:

Δr = radial misregister,

Δp = displacement of deflection plane (Equation 1),

p = distance from mask to light source,

q = distance from mask to faceplate,

u_R = inclination of faceplate (Equation 2),

u = half deflection angle.

This relationship takes the obliquity due to the curvature of the faceplate into consideration, but makes the approximation that $q/p = q'/p'$ (see Figure 1).

A relatively simple explicit expression for the misregister Δr is obtained by making the reasonable approximation that

$$\tan u_R = [(p+q)/R] \tan u \tag{4}$$

and substituting it and Equation (1) into Equation (3). This results in

$$\Delta r = \frac{l}{2}\frac{q}{p}\tan^2\frac{u}{2}\frac{\tan u}{1+[(q+p)/R]\tan^2 u} \tag{5}$$

which gives Δr as a function of variable deflection angle $2u$ and the constant parameters of the tube R, q and p, and the length of the deflection field l.

Figure 2 shows the misregister Δr as a function of the deflection angle for various idealized yoke lengths. The radius R was assumed to be 26.4 inches, q and p were taken as .535 and 15.225 inches respectively. The broken line is an observed misregister curve with a practical yoke. The fact that this curve does not agree perfectly with any of the theoretical curves, is not due to the approximations made in the evaluation of Equation (5), but to the assumptions under which Equation (5) was derived. First, the field of the practical yoke is not uniform as was assumed in Equation (1). Second, the mask and faceplate are not exactly spherical nor do they have exactly the same radius of curvature. Third, the forward movement of the deflection plane is not the only cause of radial misregister. There are mechanical causes which can probably account for as much as one mil of misregister. However, the observed misregister curve, which reaches about 5 mils is always within .5 mil of the theoretical $l = 5$ curve, which corresponds to the effective length of actual deflection yokes. Forward movement of the deflection plane, therefore, is the main factor involved in radial misregister. Figure 2 shows that for a yoke with $l = 5$ inches, a radial misregister of 3 mils will occur at the edge of a 55° tube and a misregister of 5 mils at the edge of a 70° tube.

Methods of Correction

There are several methods of correction for radial misregister that have been used in the past. The first is a reduction of the diameters of the apertures of the mask. This increases the tolerance with respect to misregister at the expense of some loss in brightness. Another method of correction is to move the yoke away from the mask. In that case the deflection plane starts in back of the color centers and in traveling forward as a function of deflection angle passes the color centers and ends up ahead of them, but not by as great an amount as without this corrected position. The misregister is thus split into a negative part for small deflection angles and a positive part for large deflection angles. The maximum amount can be reduced to 3 mils.* When the yoke position is thus compromised to improve misregister,

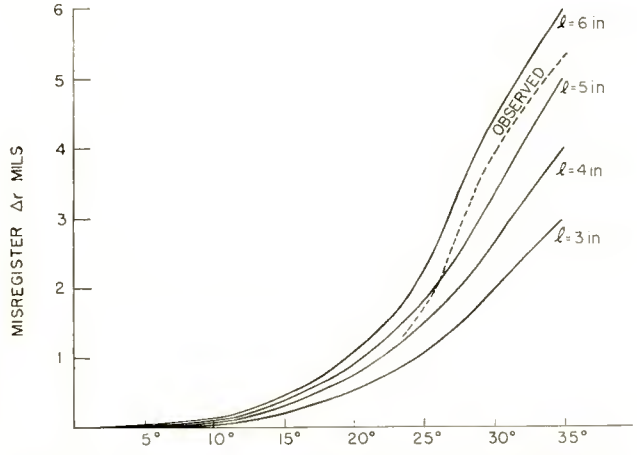


Fig. 2—Theoretical and observed radial misregister.

one is no longer free to adjust yoke position to correct for other deviations from the norm such as statistical tube variations.

It will now be shown that all purely radial misregister can be entirely eliminated by interposing a lens between the light source and the shadow mask, during the exposure of the photo-resist. This lens has the property that the virtual source of light as seen from the shadow mask, moves forward in exactly the same way as the plane of deflection of the yoke does as a function of deflection angle. The beam of light after passage through the lens will therefore coincide exactly with an electron beam, and a phosphor dot will be placed on the faceplate exactly where an electron beam will strike.

* For details, see Reference (2).

A first-order correction using the spherical aberration of a simple spherical lens provides a forward movement of the virtual deflection plane. However, the over-all fit to the type of observed misregister curve shown in Figure 2 is then only approximate. Since in the mass production of kinescopes the cost of a lens is a minor item, it is generally preferable to use a more exact correction lens.

To design an exact correction lens, the optical path is split into three portions forming the three angles shown as u_1, u_2, and u_3 in Figure 3: u_1 is the inclination of the initial ray with the axis; u_2 is the inclination of the ray in the lens; and u_3 is the inclination of the final ray. The final ray inclination is determined by the fact that it must coincide with an electron beam and therefore must appear to

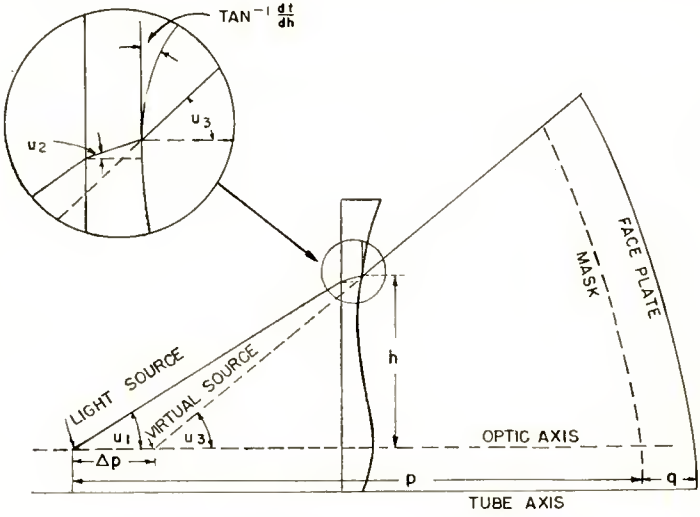


Fig. 3—Schematic diagram of correction lens.

originate from a distance $p - \Delta p$ on the optic axis. Therefore, starting with an arbitrary u_1, one can calculate the height at which this ray will emerge. An observed misregister curve, such at the dotted curve of Figure 2, is then used with Equation (3) to obtain the Δp appropriate for this angle. The angle u_3 is then determined. The slope of the lens at this point is then determined from the formula

$$\frac{dt}{dh} = \frac{\sin u_3 - \sin u_1}{N \cos u_2 - \cos u_3} \tag{6}$$

where N is the index of refraction of the glass.

Such a lens has the great advantage of being easy to produce and easy to use. Once this lens is placed in the lighthouse and properly aligned, the exposure of the photo-resist is no more complicated than it was before, and in the finished tube all radial misregister is eliminated. It is interesting to note that the lens design is determined primarily by the effective length of the yoke and not by the tube dimensions.

The shape of the lens is approximately as shown in Figure 3. It is aspheric and has a diameter of about 10 inches; the thickness first decreases, has a minimum at a radius of approximately 3 inches and then increases again quite sharply towards the end. This lens has been produced and tested. It completely eliminated the systematic radial misregister. It is in use in the production of the 21AXP22 color kinescope.

STUDIES OF EXTERNALLY HEATED HOT CATHODE ARCS*

Part IV—The Low-Voltage Form of the Ball-of-Fire Mode (The Low-Voltage Arc)

By

Edward O. Johnson

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

Summary—*In Part I of this series the qualitative aspects of the various discharge modes were treated. In Parts II and III the first of these modes, the anode glow, was treated in considerable detail. In this paper an analytical and experimental treatment is made of the second mode, the ball of fire. In particular, this treatment deals with a low voltage form of this mode, one that has been commonly known in the literature as the "low-voltage arc."*

The first-order discharge model is based upon the experimental observation that the stream of electrons entering the ball plasma from the cathode plasma is scattered so completely that the electrons within the ball have a Maxwellian energy distribution. The ion generation rate from the electrons in the high-energy tail of this distribution is computed and used with particle flow and continuity equations to derive expressions for the electron temperature, potential and density distributions, total ion generation, and the energy relations.

The analytical predictions are shown to be in good agreement with the results of probe measurements carried out with different noble gases, gas pressures, tube sizes, and anode current values. Of particular interest is the experimental verification of the prediction that the peak ball potential can be well below the ionization potential of the gas and, furthermore, need not be related in any essential way to the atomic excitation potentials.

Another result of interest is the verification of the conclusion from analysis that when the tube geometry exceeds a certain maximum size an anode glow and then a second ball plasma are formed. This second plasma is located adjacent to the anode surface and at a higher potential than the original plasma. The belief that this situation is the same as that involved in the formation of striations is discussed and supporting evidence is presented.

Introduction

In Part I[1] of this series the qualitative aspects of the various discharge modes were outlined. In Part II[2] the ion generation and loss relations of the anode-glow mode were derived and used in the

---

* Decimal Classification: R337.1.

[1] L. Malter, E. O. Johnson, and W. M. Webster, "Studies of Externally Heated Hot Cathode Arcs, I—Modes of the Discharge," *RCA Review*, Vol. XII, pp. 163-435, September, 1951.

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

analysis of the volt–ampere characteristics of the discharge. In Part III[3] the density and particle flow relations in the anode-glow plasma were treated in detail and criteria for the existence of the anode-glow mode were developed and compared with experiment. The transition from the anode-glow to the ball-of-fire mode was discussed qualitatively.

Described in this paper is an analytical and experimental investigation of one form of the ball-of-fire mode. This form operates with an anode potential that can be as low as a few volts and, in some cases, even zero or negative.[4] This discharge has been commonly known in the past as the "low-voltage arc." Numerous investigations and attempts to explain it have been made.[5] No comprehensive account of this unique discharge has yet been given, however. The outstanding problem has been to find the mechanism by which such low anode potentials can accomplish sufficient ionization to maintain the discharge. Since it appears these low anode potentials can occur without oscillations of the anode potential, early investigators were forced to conclude that a virtual anode must exist in the interelectrode space. Probe measurements indicated that this was indeed the case. These measurements indicated that the potential of this virtual anode was close enough to the lowest excitation potentials of the gas to make it appear possible that the ion generation could be accounted for by cumulative processes. This now seems to be the generally accepted view.[6,7] This view, however, runs into the difficulty of having to account for the flow of anode electron current out of the virtual anode against the strong retarding potential between it and the anode surface. The explanation that the effect of the retarding field can be compensated for by large plasma density gradients has been advanced.[6,7] To the best of the author's knowledge, however, no systematic attempt has been made to explain the presence of the required density gradients.

---

[3] E. O. Johnson and W. M. Webster, "Studies of Externally Heated Hot Cathode Arcs, III—Plasma Density Distributions in the Anode-Glow Mode," *RCA Review*, Vol. XVI, pp. 82-108, March, 1955.

[4] G. Medicus and G. Wehner, "On the Discharge Mechanism in Hot-Cathode Rare Gas Diodes," Paper C3, *Conference on Gaseous Electronics*, Schenectady, New York, October, 1951.

[5] M. J. Druyvesteyn, "Der Niedervoltbogen," *Zeit. f. Physik*, Vol. 64, p. 781, September, 1930.

[6] K. G. Emeleus, *The Conduction of Electricity Through Gases*, Third Edition, p. 51, John Wiley and Sons, New York, N. Y., 1951; p. 51.

[7] L. B. Loeb, *Fundamental Processes of Electrical Discharges in Gases*, John Wiley and Sons, New York, N. Y., 1939; pp. 636-639.

In this paper all of these problems are considered. It is found that: (1) the ion generation can be accounted for solely by energetic electrons in the tail of a Maxwellian energy distribution, (2) that the virtual-anode potential bears no direct relation to either the excitation or ionization potential of the gas since cumulative ionization is not at all necessary in the process of ion generation, and (3) that the flow of ion and electron currents is accounted for in a straightforward manner by the systematic application of flow and continuity equations.

The qualitative aspects of the discharge model are discussed and the analytical consequences of the discharge model are worked out. Expressions are derived for the electron temperature, energy balance, particle flow, the distribution of plasma density, the electric fields, and the potentials. Finally, these expressions are compared with experiment.

Preliminary Considerations

As with the studies of the anode-glow (AG) mode, the present investigation started out with a study of the ball-of-fire (BF) mode in diodes of cylindrical geometry.[1-3] It was soon realized, however, that progress would be difficult because of (1) the tendency of the ball to sidestep probes, and (2) the general instability of the BF mode in cylindrical diodes. What was needed, first of all, was a method to "anchor" the ball plasma geometrically. This was achieved by the simple expedient of providing a plasma sink in the cathode plasma that would force the ball to locate close by.[3] This sink was a third electrode, or grid, located across the cathode–anode current path, roughly halfway between the cathode and anode. The problem of discharge stability was conquered mainly by stabilizing conditions in the dark, or cathode, plasma which links the ball plasma with the cathode. Following the existence criteria for the AG plasma,[3] the cathode–grid spacing was made short, open, and direct.

In building up a picture for the potential distribution in the discharge consider the structure cross section shown in Figure 1a. If a tube with this structure is filled with a noble gas, such as xenon, to a pressure of about a millimeter, over certain current ranges a brilliant glow will only be observed in the grid–anode region. On the basis of observations such as this and probe measurements of the sort previously described,[1-2] there can be little doubt that the dark plasma in the cathode–grid region is the same as that which accompanies the AG discharge. That is, electrons enter this plasma from the cathode against a small retarding potential V_c, shown in Figure

1b, drift through this plasma by virtue of density gradients and small electric fields, and eventually enter the sheath near the plane of the grid wires. The electrons are accelerated in this sheath and eventually gain enough energy to ionize. In the discharge of Figure 1 the ions are generated in a plasma, the ball plasma, instead of in a sheath as is the case in the AG mode. Some of the ions generated in the ball plasma drift (because of the action of density gradients and small electric fields) into the sheath at the grid and fall into the dark plasma

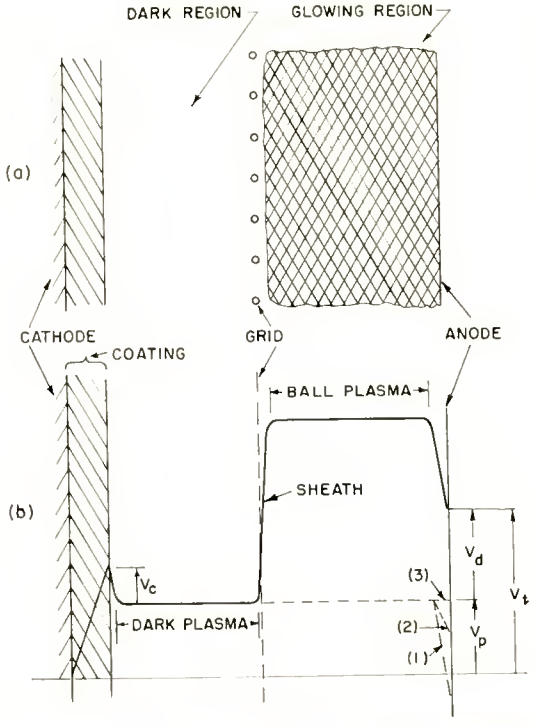


Fig. 1—(a) Cross section of tube structure; (b) Potential distribution for the ball-of-fire mode.

to sustain its density against diffusion losses of its particles to its boundaries. Since the dark plasma has already been treated in considerable detail in earlier papers of this series, attention will now be focused on the subject of this paper, the plasma in the grid–anode region.

With the discharge operating with a constant anode potential* well below the ionization potential of the gas, it is quite natural to assume that there must be a virtual anode out in the grid–anode space. A

* To within about 0.015 volt.

situation in conformity with this assumption and the above discussion is illustrated by the potential diagram of Figure 1b. In this discharge model the virtual anode is a plasma and, as is characteristic of such regions, should contain only small electric fields. Thus, almost the entire grid–anode region is represented as being at essentially the same potential. Close to the anode surface, however, is a sheath which "matches" the anode edge of the plasma to the anode. The necessity for this sheath will be explained in the analysis. The effective anode potential is the quantity V_a shown in the figure. The anode potential that would be measured with an external voltmeter is the quantity V_t. This potential is not really of fundamental importance since it is partially made up of the potential V_p which represents the voltage drop across the coating of the oxide cathode, minus the voltage V_c.* The value of V_p is usually found to lie between 1 and 5 volts.

The potential V_a is of fundamental importance because it is a direct measure of the energy consumed by the phenomena in the grid–anode region. This potential can be obtained from a simple experiment in which all of the plasma in the tube is generated by an auxiliary discharge. This type of operation, which has been previously described in considerable detail,[8] is characterized by the presence of only one plasma that extends from the cathode all the way to the anode. In the cathode–grid region the potential of this new plasma must correspond with that of the plasma present during the BF discharges; in the grid–anode region the new plasma is represented by the dotted lines shown in Figure 1b. The volt–ampere characteristic for this type of operation is shown by the solid line in Figure 2. The region of the characteristic from *a* to *b* has potential distribution typified by (1) in Figure 1b. Few electrons reach the anode from the plasma because of the retarding field in the sheath at the anode surface. The region of the characteristic from *b* to *c*, typified by (2) in Figure 1b, occurs when there is only a small retarding potential in the anode sheath. The knee in the characteristic, signified by point *c*, occurs when anode is at plasma potential as indicated by (3) in Figure 1b. The electron space current from the plasma is now completely saturated and any further increase in anode potential will have to be dissipated in an electron sheath at the anode. Hence the portion of the characteristic from *c* to *e* is flat. This behavior is always observed with helium. If, however, xenon or argon is used then the characteristic

* Contact potential differences can also add or subtract several volts.

[8] E. O. Johnson and W. M. Webster, "The Plasmatron, A Continuously Controllable Gas-Discharge Developmental Tube," *Proc. I.R.E.*, Vol. 40, pp. 645-659, June, 1952.

tends to turn up abruptly at some point *d* and trace out the path *df*, instead of *de*. Apparently, the electron sheath tends to form in the plane of the grid wires instead of at the anode surface.[3] The inference is that the potential *cd*, which is always observed to have a value between 2 and 3 volts, corresponds to the potential V_d in Figure 1b. The portion *df* of the characteristic is not in any way affected by the auxiliary discharge. However, if the auxiliary discharge is not used then the tube will trace out a volt–ampere characteristic corresponding

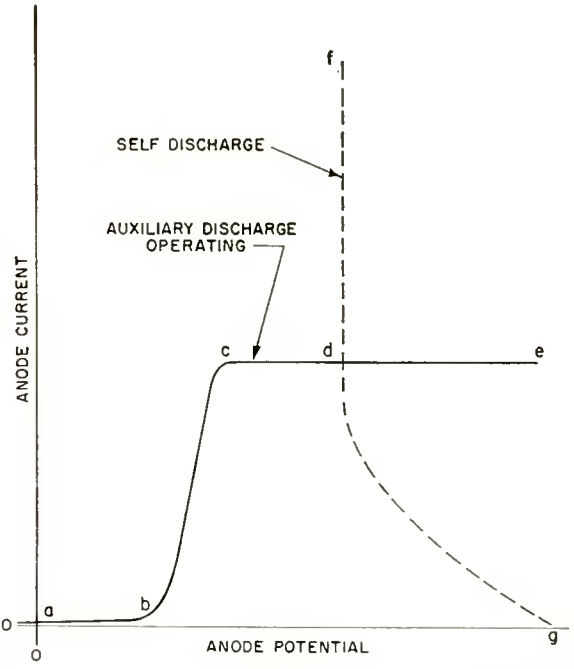


Fig. 2—Comparison of volt–ampere characteristics.

to the dotted line *gdf*, with the point *g* roughly corresponding to the ionization potential of the gas.

Oscilloscope photographs of the volt–ampere characteristic *abcdf* are shown in Figure 3. These were taken with the same sort of tube structure as that displayed in Figure 1a except that an auxiliary discharge assembly[8] was recessed into the anode surface. The lower characteristic was taken with 35, the middle one with 125, and the upper one with 215 milliamperes of auxiliary discharge current. The origin of each voltage–current axis is marked by the intersection of the vertical and horizontal lines. The small dash below each voltage axis, to the right of the origin, corresponds to two volts; the small

dash above each origin corresponds to two amperes. It is especially to be noted that at no time does the anode potential go appreciably beyond a point corresponding to d in Figure 2.* And, as can be seen in the photograph, the potential difference between the knee, and what corresponds to point d, is slightly in excess of two volts.** It can be inferred from these observations that the potential distribution in the tube rises as shown by the dotted lines in Figure 4. These lines emphasize that the potential of the whole grid–anode region must lift up uniformly, so that all appreciable potentials are absorbed by sheaths located near the grid and anode. Destruction of the plasma in the

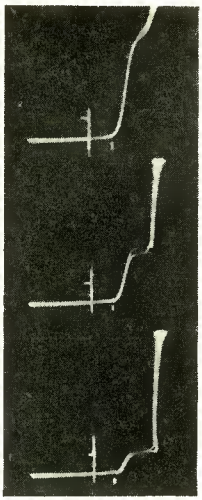


Fig. 3—Photograph of plasmatron volt–ampere characteristic.

grid–anode region would be required if the potential in this region did not lift up uniformly as shown. Such a catastrophe might be expected to take place only in the presence of very large potential differences between the electrodes. The appearance of a sheath at the grid, instead of one solely at the anode, can be accounted for by arguments previously advanced.[3] That is, an electron sheath will form where the product of the cross-sectional area of the electron current path and the electron space current is a minimum.

Besides reducing the cross-sectional area of the current path, the grid wires also are a plasma sink and must severely depress the plasma density in their vicinity. Hence, the current-carrying capability of

* The oscilloscope has a bandwidth ranging from d-c up to about 10 megacycles.

** At higher anode currents the knee of the characteristic becomes less well defined probably because of the increasing effect of cathode resistance.

the plasma in the plane of the grid wires must be reduced, along with the cross section of the current path, and it is not unexpected that a sheath should tend to form here. Experiments with tubes containing fewer grid wires indicate that this is the case. These tubes, for example, tend to trace out the path *cde*, instead of *cdf*, shown in Figure 2.

There is no reason to believe that the phenomena in tubes, such as those just discussed, are essentially any different than those in tubes studied by other investigators in the past. In particular, the electron sheath at the grid can be moved arbitrarily close to the cathode surface by simply moving over the grid by the requisite distance. If this is done then the potential distribution in the tubes described above

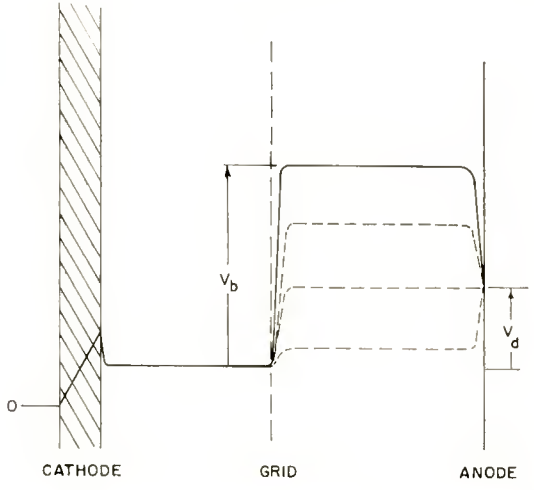


Fig. 4—Comparison of potential distributions.

will, in its general aspects, be the same as that described in the experiments of Druyvesteyn.[5] Rather than study the low-voltage arc under the conditions described by him, it seemed particularly convenient to treat the low-voltage arc as a modification of the plasmatron operation described in Reference (8), and the AG operation described in References (1)-(3).

Of particular interest in the analysis to follow will be the determination of the potentials V_b and V_d of Figure 4. For example, it will be shown, by both analysis and experiment, that the potential V_b for xenon is close to seven volts, and that the theoretical value for V_d turns out to be between 2 and 3 volts, just as found in the experiments described above.

Discharge Model

In this model the plasma in the grid–anode region is treated as a "mixing bowl." That is, electrons will be considered as entering this plasma in a monoenergetic stream that is immediately scattered, both in direction and energy. Specifically, it is assumed that this scattering is so complete that the electrons are forced into compliance with Maxwell-Boltzmann distribution laws. This, then, enables us to consider the electrons as having a temperature T_e. With this assumption as a basis, the ion generation function, arising from the electrons in the plasma having an energy greater than the ionization potential of the gas, is readily derived. Then, by combining the foregoing information with the flow and continuity equations, and certain boundary conditions, all the pertinent plasma parameters such as density, electron temperature, total ion generation, electric fields and potentials can be completely determined.

The ball plasma is considered as an entity wherein the energy of the incoming electrons is scattered randomly among the electrons already present. This sharing of energy enables every electron entering the ball to eventually reach the anode. However, on passing through the ball, electrons lose energy in generating ions and excited atoms and so, on the average, cannot possibly retain their initial energy. This requires that the point of exit from the ball be at a higher potential than the point of entrance, a condition specifically expressed by the potential difference V_d portrayed in the potential model shown in Figure 4.

The cardinal assumption of electron energy scattering in the ball is not an arbitrary one; it is supported by multitudinous observations made by many observers. Specifically, the factors upon which the scattering picture is based are:

1. A well-defined electron temperature is usually observed in almost any discharge even though the electron mean free paths are much too long to account for a sufficient degree of collision scattering.[9]
2. Direct measurements on the scattering of an electron beam in a plasma show that this scattering is very much greater than could be expected from any sort of simple collision picture.[9]
3. Analytical inference that some sort of cooperative action (such as plasma oscillations) could account for such scattering.[10]

---

[9] I. Langmuir, "Scattering of Electrons in Ionized Gases," *Phys. Rev.*, Vol. 26, pp. 585-613, 1925.

[10] D. Bohm and E. P. Gross, "Effects of Plasma Boundaries in Plasma Oscillations," *Phys. Rev.*, Vol. 79, pp. 992-1001, September, 1950.

4. Finally, as will be shown, actual probe temperature measurements of the ball very clearly give the impression that it has a definite electron temperature.

Whatever mechanism accomplishes this scattering it certainly does not seem to be simple coulomb interaction. This effect seems far too small for the plasma densities involved in the ball. For example, the Conwell-Weisskopf scattering formula[11]

$$1 - \cos\theta = 2\left(\frac{\mathcal{E}_1}{2\mathcal{E}}\right)^2 \ln\left[1 + \left(\frac{2\mathcal{E}}{\mathcal{E}_1}\right)^2\right]$$

where $\mathcal{E}_1$ is the potential energy at a distance $a/2$ from a target electron (or singly charged ion), indicates the average deflection angle θ suffered by an incoming electron of kinetic energy $\mathcal{E}$ when this electron travels a distance a in the vicinity of the target particle. If the plasma density n is 10^{15}, $a = n^{-1/3} = 10^{5}$, and if the kinetic energy of the incoming electron is 1 volt, then $1 - \cos\theta = 1.25 \times 10^{-7}$. That is, the incoming electron will suffer, on the average, a "deflection" amounting to 1.25×10^{-7}. In traveling one centimeter it will pass 10^{5} targets in a plasma density of 10^{15} particles per cubic centimeter and so be deflected by a factor 1.25×10^{-2}. This is trivial and would be even more so if the plasma density were 10^{13} and the kinetic energy 10 volts, values more in keeping with what is probably present in the tubes under consideration.

Generation Function

The following is a derivation[12] of the expression for the number of ions generated per second per electron, where the electrons have a Maxwellian energy distribution characterized by a temperature T_e, and the gas has a pressure p and an ionization potential V_i.

The number of ions N_p formed per centimeter of path by an electron of energy V volts is

$$N_p = pa\,(V - V_i), \tag{1}$$

where a is the differential ionization coefficient normalized to a pressure of 1 millimeter. If the electron has a velocity of v centimeters

[11] See W. Shockley, *Electrons and Holes in Semiconductors*, D. Van Nostrand, New York, N. Y., 1950; pp. 258-264.

[12] This follows that given by T. J. Killian, "The Uniform Positive Column of an Electric Discharge in Mercury Vapor," *Phys. Rev.*, Vol. 35, pp. 1238-1252, May, 1930.

per second, then the number of ions formed per second is $N_p v$ and the ion current I_p, in amperes, which arises from each electron, is given by

$$I_p = N_p v e, \tag{2}$$

where e is the electronic charge in coulombs. Now, if we have a total collection of N electrons with a Maxwellian distribution of velocities, the number of electrons having velocities between v and $v + dv$ is[13]

$$Nf(v)\,dv = \frac{4N}{\sqrt{\pi}}\left(\frac{m}{2kT_e}\right)^{3/2} v^2 \exp\left(\frac{-mv^2}{2kT_e}\right) dv. \tag{3}$$

In this expression m is the electron mass in grams, k is Boltzmann's constant in ergs per degree Kelvin, and T_e is the electron temperature in degrees Kelvin. By multiplying Equations (2) and (3), and making use of Equation (1), an expression is derived for the total ion current generated by all the electrons in the velocity range v to $v + dv$. It is

$$dI_p = \frac{2eNap}{\sqrt{\pi}}\left(\frac{m}{2kT_e}\right)^{3/2}\left(\frac{2e'}{300m}\right)^2 V(V - V_i)\exp\left(-\frac{eV}{kT_e}\right)dV,$$

or, when the constants are evaluated,

$$dI_p = 1.34 \times 10^{-5}\frac{Nap}{T_e^{3/2}} V(V - V_i)\exp\left(-\frac{eV}{kT_e}\right)dV. \tag{4}$$

In this expression the relation

$$v = \sqrt{\frac{2e'V}{300m}}$$

was used to replace the velocity v with the voltage V. The quantity e' is the electronic charge in electrostatic units.

The total ion current generated by N electrons is found by integration of Equation (4) for all electrons having an energy greater than V_i. That is,

$$I_p = 10^{-13}\, T_e^{\frac{1}{2}}\, Nap\left[V_i + \frac{2kT_e}{e}\right]\exp\left(\frac{-eV_i}{kT_e}\right).$$

[13] See, for example, J. D. Cobine, *Gaseous Conductors*, McGraw-Hill Book Co., New York, N. Y., 1941; pp. 8-15.

And, finally, the quantity z, the ions generated per second per electron, is

$$z = 0.625 \times 10^6 \, T_e^{\frac{1}{2}} \, ap \left[V_i + \frac{2kT_e}{e} \right] \exp\left(\frac{-eV_i}{kT_e} \right). \tag{6}$$

Plasma Analysis

The relations which govern the behavior of the charged particles in the ball plasma must be the flow equations,

$$\Gamma_e = -n u_e E - D_e \operatorname{grad} n \tag{7}$$

$$\Gamma_p = -n u_p E - D_p \operatorname{grad} n, \tag{8}$$

and the continuity equations,

$$\operatorname{div} \Gamma_e = z n^* \tag{9}$$

$$\operatorname{div} \Gamma_p = z n. \tag{10}$$

With the subscripts e and p referring to electrons and ions, respectively, Γ is the flow density rate, u is the mobility, E is the electric field, and D is the diffusion coefficient. As is usual in a first-order plasma analysis, the ion and electron densities are assumed to be equal and are represented by the quantity n. By eliminating the electric field E from Equations (7) and (8), making use of the relation

$$D_a = \frac{\mu_e D_p + \mu_p D_e}{\mu_e + \mu_p} \cong \frac{\mu_e D_p + \mu_p D_e}{\mu_e}, \tag{11}$$

and introducing Equations (9) and (10), there is obtained

$$\nabla^2 n + Gn = 0, \tag{12}$$

wherein $G = z/D_a$. In the simplest possible case, the infinite plane-parallel, the solution to Equation (12) is

$$n = C_1 \cos\left(\sqrt{G}x + C_2\right). \tag{13}$$

The infinite plane-parallel case is a fair approximation for the "ball" in the tubes under consideration since the anode region in these has

* The analysis is simplified, and at the same time trivially affected, if div $\Gamma_e = 0$.

lateral dimensions which are several times larger than the grid–anode distance. In any event, as will be seen, this simplifying assumption will not affect the basic conclusions derived from the analysis.

In evaluating the constants C_1 and C_2 the boundary conditions at the cathode edge of the ball are used. These are

1. $\frac{\Gamma_e}{\Gamma_c} \cong -\sqrt{\frac{m}{M}} \cong -\frac{\mu_p}{\mu_e}$, which must be satisfied if there is to be an essentially zero electric field at both ends of the double sheath[3] which connects the ball with the dark, or cathode, plasma.

2. $E = -\frac{kT_e}{e\lambda_e}$. This is a convenient way to define the cathode edge of the plasma in a first order treatment. It, in effect, stipulates that plasma conditions end where the electric field is so large that it gives an electron an energy, over a mean free path, that is equal to its mean thermal energy. When the electric field is this large, diffusion and mobility relations start to break down and the electron energy distribution will tend to depart from the Maxwellian.

By eliminating E from Equations (7) and (8) there is obtained for the plane parallel case

$$\frac{dn}{dx} = -\frac{\Gamma_e}{D_a}\left[\frac{\mu_p}{\mu_e} + \frac{\Gamma_p}{\Gamma_e}\right]. \tag{14}$$

The use of the first boundary condition with this equation shows that, at $x = 0$, $dn/dx = 0$, and from Equation (13) it is seen that $C_2 = 0$. By eliminating grad n from Equations (7) and (8) it is found that

$$n = \frac{1}{E}\frac{kT_e}{e}\frac{\Gamma_e}{D_a}\left[\frac{\Gamma_p}{\Gamma_e} - \frac{\mu_p}{\mu_e}\frac{T_p}{T_e}\right], \tag{15}$$

if use is made of the Einstein relation $D = \mu\, kT/e$. If the first and second boundary conditions are used in Equation (15) then, at the cathode edge of the ball plasma

$$n = \frac{\lambda_e\Gamma_e}{D_a}\frac{\mu_p}{\mu_e}\left[1 + \frac{T_p}{T_e}\right] = \frac{\lambda_e\Gamma_e}{D_a}\frac{\mu_p}{\mu_e} = C_1. \tag{16}$$

The term in brackets is negligibly greater than unity because $T_e \gg T_p$. Thus the density distribution for the infinite plane parallel case is

$$n = \frac{\lambda_e \Gamma_e}{D_a} \frac{\mu_p}{\mu_e} \cos \sqrt{G}x. \tag{17}$$

This relation will now be used to find the total ion generation, the ion current flow, the electric field in the plasma, and also the potentials within the plasma.

By substituting Equation (17) into Equation (10), and integrating, it is found that the total ion generation per unit cross section is

$$\sum \Gamma_p = z \int_0^d n dx = \sqrt{G}\, \lambda_e \Gamma_e \frac{\mu_p}{\mu_e} \sin \sqrt{G}\, d. \tag{18}$$

By equating the gradient of Equation (17) to its equal from Equation (14) and solving for the ion flow density Γ_p, there is obtained

$$\Gamma_p = \frac{\mu_p}{\mu_e} \Gamma_e \left[\lambda_e \sqrt{G} \sin \sqrt{G}x - 1\right]. \tag{19}$$

By substituting Equations (19) and (17) into (15), the electric field E is found to be

$$E = \frac{kT_e}{e\lambda_e} \left[\frac{\lambda_e \sqrt{G} \sin \sqrt{G}x - 1}{\cos \sqrt{G}x}\right]. \tag{20}$$

This relation can be integrated to find the potential inside the plasma with respect to the cathode edge. One obtains

$$V = \frac{-kT_e}{e} \left[\frac{1}{\lambda_e \sqrt{G}} \ln \tan \frac{1}{2}\left(\sqrt{G}x + \frac{\pi}{2}\right) + \ln (\cos \sqrt{G}x)\right]. \tag{21}$$

Before plasma conditions are completely determined by Equations (17)-(21) it is necessary to impose a boundary condition at the anode edge of the plasma so that the quantity G can be made determinate. As with the corresponding boundary condition at the cathode edge, this is most naturally achieved by stipulating, at $x = d$, that $E = kT_e/e\lambda_e$. Equation (20) then requires that

$$\lambda_e \sqrt{G} \sin \sqrt{G}\, d = 1 + \cos \sqrt{G}\, d. \tag{22}$$

This relation fixes G, and hence the electron temperature T_e, in terms of the gas constants and the dimension d of the ball plasma. The parameter d is either directly fixed by tube geometry (the grid–anode spacing, for example) or by a simultaneous solution of conditions both in the ball and in the dark plasma. The former case is the only one of concern in this analysis; the latter case is discussed briefly in Part III of this series of papers. The only physically meaningful solutions of Equation (22) require that $0 < \sqrt{Gd} < \pi/2$. The solutions to Equation (22) can be found from the curve in Figure 5. If this curve is entered with a value for λ/d, which is fixed by geometry and gas pressure, a value θ for $\sqrt{G}d$ is determined. Consequently, $G = \theta^2/d^2$. Making use of $G = z/D_a$, as previously defined, and the value for z given from Equation (6), there is derived

$$\sqrt{\frac{eV_i}{kT_e}}\left(\frac{eV_i}{kT_e}+2\right)^{-1}\exp\left(\frac{eV_i}{kT_e}\right)=\frac{6.8\times 10^7}{\theta^2}C^2p^2d^2. \quad (23)$$

The quantity C, which is given by

$$C=\sqrt{\frac{a\sqrt{V_i}}{\mu_p}}$$

is dependent solely upon the type of gas used. Hence Equation (23) completely defines the electron temperature in terms of the characteristics of the gas, its pressure, and the geometry of the tube. The value of C for the various gases is given in Table I and several solutions for Equation (22) are plotted in Figure 6.

Table I

Gas	C
Helium	4.68×10^{-3}
Neon	1.2×10^{-2}
Argon	5.19×10^{-2}
Xenon	8.1×10^{-2}
Mercury	1.1×10^{-1}

The value for mercury was taken from Table 8.8 in Cobine p. 241; the other values were computed from the values of V_i and a, given elsewhere in Cobine, and more recent published values of the mobility μ_p.[14]

It is interesting to compare the preceding analysis with the classical

[14] See the series of papers published by the MIT group, for example, M. A. Biondi and S. C. Brown, "Measurements of Ambipolar Diffusion in Helium," *Phys. Rev.*, Vol. 75, pp. 1700-1705, June, 1949.

one for the positive column. Since the same generation and loss mechanisms have been used in both cases it is not surprising that Equation (23), wherein θ has the value 2.405, is identical with the expression* derived from the radial conditions in the positive column. Even if $\theta \neq 2.405$, the temperatures derived from either analysis would be closely the same. In other words, the electron temperature is very insensitive to the imposed boundary conditions. Of much greater significance than the similar end result of the temperature analysis, however, is the basic difference between the use of the two analyses. Specifically, the present analysis considers both radial** and longitudinal plasma conditions in one homogeneous treatment; the classical positive column analysis derives the electron temperature from radial conditions and then proceeds to use this result in a different manner to derive the longitudinal conditions. In particular, this longitudinal analysis consists of (1) finding the electric field necessary to maintain the electron temperature in the presence of energy losses due to elastic and unelastic collision, and (2) finding the electron density n_0 along the axis of the column necessary to generate ions as fast as they are lost to the walls. This analysis specifically does not consider the longitudinal flow relations that must exist between the ions and electrons and, furthermore, utilizes the very arbitrary assumption that the plasma density along the axis of the column is essentially constant. This, in the author's opinion, is a serious error. As will be seen, the inevitable variation of longitudinal plasma density leads directly to a situation which calls for the rupture of the plasma at periodic intervals. Even though the present analysis is an approximate one, because it does not consider all the details of plasma behavior, it certainly seems to be a more basic approach to the behavior of a current-carrying plasma than the less homogeneous classical analysis that has been used with the positive column.

There is still one more boundary condition to consider. This has to do with the manner in which the plasma "connects" to the anode. If the plasma electron space current at the anode edge of the ball exceeds the anode current density called for by the external circuit, that is, if

$$\frac{n_a \bar{c}_e}{4} > i_e ,$$

* See, for example, Equation (8.69), p. 239 in Reference (13). Equation (8.70) is only correct if the pressure term is deleted.

** The radial conditions could easily have been included by solving Equation (12) for three dimensions.

where n_a is the electron density at the anode edge of the ball, and $\bar{c}_e$ is the mean thermal velocity of the electrons, then a sheath with a retarding field for electrons will form and it will have to absorb a potential

$$V_s = \frac{kT_e}{e} \ln \frac{n_a \bar{c}_e}{4\Gamma_e}, \tag{24}$$

a necessary consequence of Boltzmann's relation. This potential will serve to reduce the electron space current density, "boiling" out of the plasma, to the correct value at the anode surface. In this way the anode edge of the plasma is "matched" to the anode potential.

Plasma Size

It is appropriate to point out that the boundary conditions at both cathode and anode edges of the ball plasma could have been treated in a more general manner. That is, the specifications

$$E = \psi_c \frac{kT_e}{e\lambda_e}$$

and

$$E = \psi_a \frac{kT_e}{e\lambda_e},$$

with ψ_c and ψ_a being arbitrary factors, could have been made at the cathode and anode edges, respectively. If this were done, Equation (22) is modified and different assumed values for ψ_c and ψ_a would produce a family of curves in Figure 5, instead of the single curve shown. The effect of different assumed values of ψ_c and ψ_a can be briefly summarized as follows:

1. The value of T_e is only slightly affected by the assumed value of either ψ_c or ψ_a.
2. The plasma density and total ion generation given by Equations (17) and (18) are modified by the factor $1/\psi_c$; ψ_a has only a slight effect on these equations.
3. For the limiting case, wherein $\sqrt{G}d = \pi/2$, the value assumed for ψ_a has no effect upon λ_e/d. This means that the electric field at the anode edge of the plasma can change both in value and sign without affecting other discharge conditions.

With the above facts as a background the effect of plasma size, that is, the parameter d, can now be discussed. As exemplified by the

single curve in Figure 5, the complete family of curves relating λ_e/d with $\sqrt{G}d$, show that for a given gas pressure and hence λ_e, the quantity $\sqrt{G}d$ increases as d increases. Also, from Equation (17) it is seen that the plasma density at the anode must decrease, and with it, the quantity V_S in Equation (24). When V_S has decreased to zero the electron current carrying capability of the plasma at the anode edge has been saturated so that the plasma space current density J

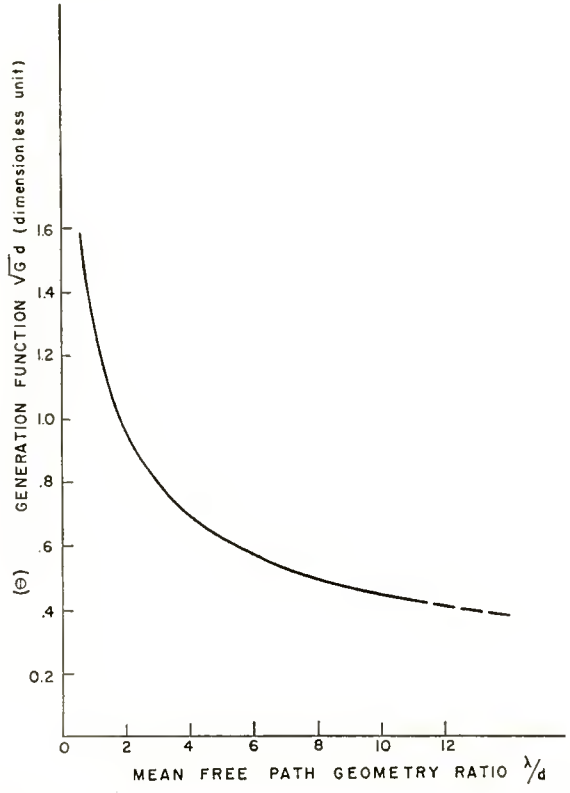


Fig. 5—Curve showing characteristic solutions for plasma generation function.

is equal to the anode current density demanded by the external circuit.[8] If this situation occurs for a value of d, say d_1, then the potential distribution should be as shown in Figure 7a, where V_d is determined by the factors described in the next section. As the grid–anode spacing is increased still further, V_d and d_1 should not change. However, the added spacing should be adsorbed by an electron sheath, see Figure 7b, whose thickness $(d_2 - d_1)$ is given by the 3/2 power law

$$(d_2 - d_1)^2 = \frac{P\,V^{3/2}}{J},$$

where P is 2.33×10^{-6} for the plane parallel case, J is the saturated electron current density as noted previously, and V is the potential across the sheath. The potential V should appear as an increase in the applied anode potential. When the value of $d_2 - d_1$ is such that V^* approximates the excitation and ionization potentials of the gas, an anode glow[15] (along with ion generation) should appear in the sheath.

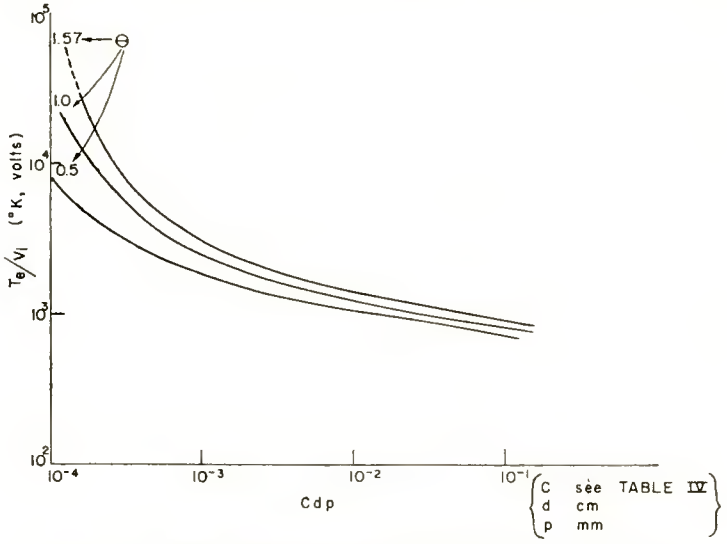


Fig. 6—Curves relating electron temperature to gas and geometric parameters.

A still further increase in the grid–anode distance should result, as shown in Figure 7c, in the formation of another plasma. The reasons for this have been expounded in Reference (3). This new plasma should be subject to roughly the same general conditions as the first one, and hence the whole process should repeat again and again as the grid–anode spacing is increased indefinitely. The increase in the anode voltage, each time a new plasma is added, should be between about 3 volts and the ionization potential of the gas. The actual value will depend upon the factors discussed in the next section.

* Because the electrons enter the sheath with appreciable energy, appreciable excitation and ionization will occur when V is less than either the excitation or ionization potentials of the gas.

[15] Such a phenomenon has already been reported, see G. Medicus, "A Ball-of-Fire Discharge with a Pronounced Current Saturation," *Jour. Appl. Phys.*, Vol. 24, p. 233, February, 1953.

Because of the form of the plasma density distribution this concatenation of plasmas should have a visual appearance like that of striations, that is, each plasma should glow brilliantly at its cathode end, and less brilliantly toward its anode edge. The distance between different striations should be the same because the phenomena in each plasma should be the same. The plasmas at the cathode and anode ends of the column could be an exception to this because of the pos-

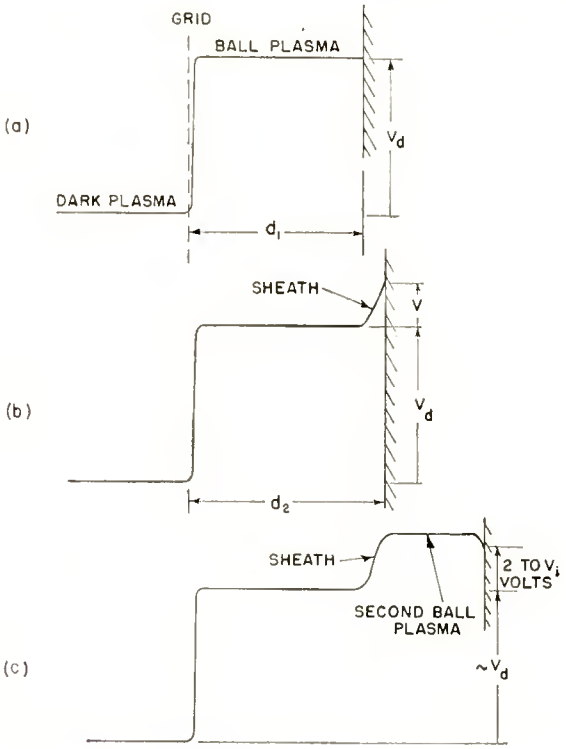


Fig. 7—Potential distributions for the ball-of-fire mode.

sibility of different boundary conditions. As noted earlier, however, the location of the sheaths connecting the plasmas is quite subject to instability. Thus it is not unexpected that striations should often have a fuzzy appearance, except under ideal conditions. Perhaps it is not too farfetched to picture an instability in one sheath propagating through the plasma to other sheaths, causing the traveling striations that are frequently observed.[16]

[16] M. J. Druyvesteyn and F. M. Penning, "Electrical Discharges in Gases," *Rev. Mod. Phys.*, Vol. 12, pp. 166-169, April, 1940.

difficult to assess because of the paucity of information on the generation probabilities of the various excited states in any of the noble gases other than helium, a gas which is of no interest in the present investigation. However, it is probably safe to say that the potential corresponding to the energy dissipated in all the excited states is of the order of ten times greater than that dissipated in ionization, or roughly 0.6 volt. Hence,

$$V_d \cong 0.6 + \frac{2\,kT_e}{e}. \tag{26}$$

If a more precise answer is desired, then the equation

$$I_x = 10^{-13}\, T_e^{\frac{1}{2}}\, N a_x p \left[V_x + \frac{2\,kT_e}{e} \right] \exp\left(-\frac{eV_x}{kT_e} \right)$$

must be used for each excited state x. This equation is Equation (5) with a_x and V_x substituted in place of a and V_i.

As will be shown in the next section, the experimental conditions make kT_e/e close to one volt. Thus Equation (26) predicts that $V_d \cong 2.5$ volts, a value that is in reasonable accord with the observed values discussed earlier. Apparently then, for quite typical conditions, most of the energy applied to the grid–anode region results in heating the anode by electron impact; at most one quarter as much energy goes into atomic processes, with only a fraction of this being involved in actual ion generation.

Cumulative Ionization

In the preceding analysis all ion generation is assumed to arise from a single-stage process. It is, however, appropriate to consider the importance of ionization arising from a two-stage process, namely, cumulative ionization, the mechanism widely accepted[5-7] as the source of ionization in the low-voltage arc. The following calculations are introduced to show that such a mechanism may indeed be the dominating one, under certain conditions.

On the average, the total population N_x of excited state x in the ball plasma is

$$N_x = \tau_x N\, z_x \tag{27}$$

where τ_x is the average lifetime of state x, N is the total electron population, and z_x is the number of atoms of state x generated per second per electron. The quantity z_x can be obtained from Equation

(6) wherein a and V_i are replaced by a_x* and V_x, respectively. In calculating the ionization rate z_{xi} arising from atoms in state x we can again employ Equation (6), wherein a_{xi} and V_{xi} $(= V_i - V_x)$ are substituted. However, in place of the neutral atom pressure factor p, the equivalent pressure p_x, or statistical weight, for the population N_x must be used. This equivalent pressure is simply

$$p_x = \frac{N_x}{3.55 \times 10^{16} S}, \tag{28}$$

where 3.55×10^{16} is the number of neutral atoms per cubic centimeter at a pressure of 1 millimeter, and S is the plasma volume. Using Equations (27) and (28), and also Equation (6) as noted, it is found that

$$Z_{xi} = \frac{(0.625 \times 10^6)\ 10^{-16}}{3.55\ S} T_e\ (a_{xi}a_x)\ \tau_x Np \left[V_x + \frac{2\,kT_e}{e} \right] \left[(V_i - V_x) + \frac{2\,kT_e}{e} \right] \exp\left(-\frac{eV_i}{kT_e} \right). \tag{29}$$

What is of real interest, however, is the ratio z_{xi}/z. This gives a direct measure of the relative importance of the two ionization processes. This ratio follows directly from Equations (6) and (29) and is

$$\frac{Z_{xi}}{Z} = 0.176 \times 10^{-10} \frac{T_e^{\frac{1}{2}}}{S} \left(\frac{a_{xi}a_x}{a} \right) \tau_x N \left[V_x + \frac{2\,kT_e}{e} \right] \left[V_i - V_x + \frac{2\,kT_e}{e} \right] \left[V_i + \frac{2kT_e}{e} \right]^{-1} \tag{30}$$

First, consider this ratio for the case of ordinary excited states where τ_x is of the order of 10^{-8} second. For the typical experimental conditions to be described in following sections $N = 10^{13}$ electrons per cubic centimeter, $T_e = 10^4$ degrees Kelvin, and $S = 10$ cubic centimeters. For xenon $V_x = 8.5$, and $V_i = 11.5$ volts. In ignorance of the true values of a_{xi} and a_x it will be assumed that $a_{xi} \cong a_x \cong a$. When these values are used in Equation (30) it is found that $z_{xi}/z = 10^{-4}$, indicating that a two-step process involving ordinary excited states

* If the cross-section for excitation peaks at potentials only slightly greater than the threshold potential, then the quantity a_x must be replaced by the function for the particular excitation cross section.

is of no importance in the discharges being considered.

For metastable states, where τ_x might be of the order of 10^{-4} second, it is seen that the two-stage process of cumulative ionization could be as important as the single-stage process.[17] It is difficult to get a better evaluation of the role of the two-stage process in the present studies because of ignorance of the generation probabilities a_{xi} and a_x. Regardless of the absolute values of these quantities, however, it is clear that the two-stage process must eventually dominate as N is increased.

The previously discussed discharge model can readily be modified to accommodate any combination of the one- and two-stage ionization processes. Inclusion of the two-stage process, for example, introduces a second-power plasma density term in Equation (12). This, nevertheless, does not change the general character of the density function. That is, this function must still decrease monotonically* as the anode is approached. Therefore, the general conclusions relative to the maximum plasma size, are still valid. Furthermore, because of the dominating character of the factor

$$\exp\left(\frac{-eV_i}{kT_e}\right),$$

which occurs in both generation functions, the electron temperature will not be very sensitive to the type of ionization mechanism that is present. However, as the two-stage process increases in importance, the computed electron temperature will decrease from the value computed on the basis of the single stage process only.

Comparison with Experiment

Experiments were carried out with a tube of 10-ampere size and also with one of 0.1-ampere size. The salient features of the geometry in both tubes were similar to those sketched in Figure 1a. The specific details of the geometry of the 0.1-ampere tube are shown in Figure 8; most of the dimensions of the 10-ampere tube were roughly four times as large. The probes in both sizes of tubes were similarly located in the center of the anode surface and were used in the conventional manner.

---

[17] C. Kenty, "Production of 2537 Radiation and the Role of Metastable Atoms in an Argon–Mercury Discharge," *Jour. Appl. Phys.*, Vol. 21, pp. 1309-1318, December, 1950, gives a detailed account of cumulative processes in a positive column of a mercury–argon discharge.

* Assuming that boundary condition (1), page 510, is still used.

The characteristic low-voltage operation was observed with both argon and xenon. Neon was not tried, and discharges in helium had anode potentials that were only a few volts below the ionization potential of the gas. Furthermore, when helium was used, the anode potential always had five to ten volts of fluctuation amplitude. In all experiments with argon and xenon the anode potential was always monitored with an oscilloscope. With the discharge conditions to be described, the maximum fluctuation amplitude (incoherent) of the anode potential was about 0.015 volt. When the cathode temperature

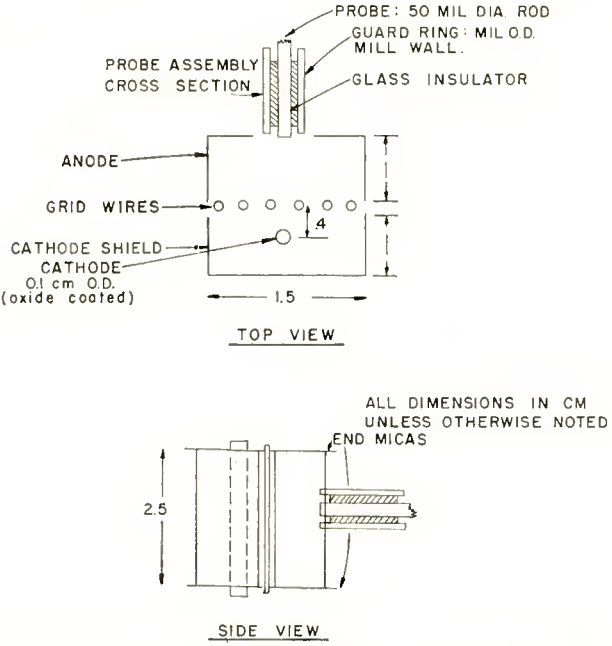


Fig. 8—Experimental triode structure for electron temperature measurements.

was increased from 825°C to about 925°C this fluctuation amplitude was reduced one half. Since the reduction continued as the cathode temperature was increased even further, it seems quite likely that this fluctuation is associated with the cathode itself.

Other principal precautions taken were: (1) probe currents were always kept below the value at which changes were noted in the anode voltage and current, (2) each run of probe measurements was consummated as rapidly as possible to minimize the effects of drifts in the probe work function.[18]

[18] M. A. Easley, "Probe Techniques for the Measurement of Electron Temperature," *Jour. Appl. Phys.*, Vol. 22, pp. 590-593, May, 1951.

Electron Temperature

Some typical plots of probe current and voltage for the 0.1-ampere tube are shown in Figure 9. The slope of these curves was not affected by the anode current. The electron temperature, as derived from the slope of these curves, was noted for gas pressures ranging from one to four millimeters, for xenon, and five to ten millimeters for argon.

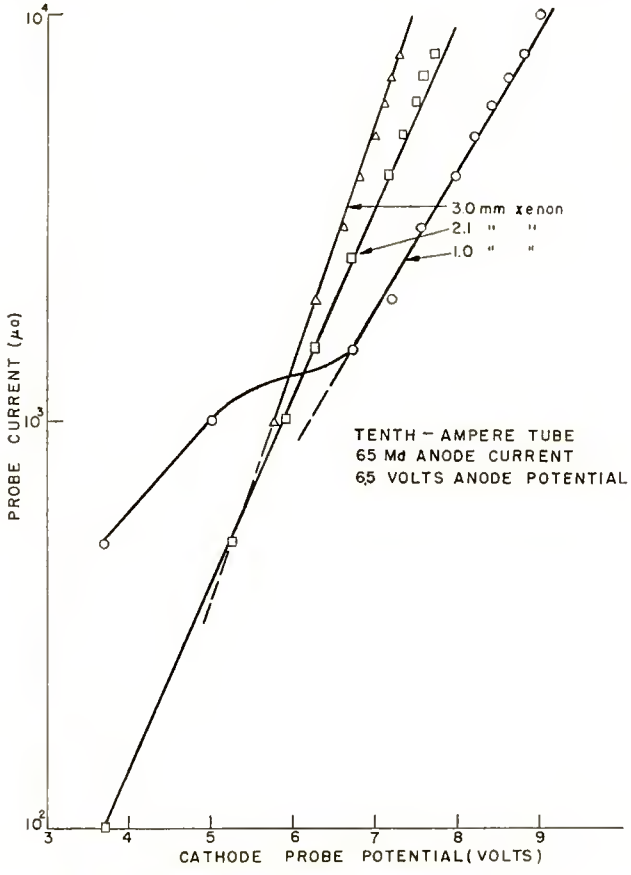


Fig. 9—Probe plots from experimental tube.

With the ten-ampere tube, the anode current was held at 8 amperes and the electron temperature was noted at only one gas pressure, 0.1 millimeter for one gas, xenon.

The results of these measurements are tabulated in Table II.

The values of C were obtained from Table I. The values of the electron mean free path λ_e were obtained from Cobine, p. 31. The value of d for the large tube was actually one centimeter. Although the small

Table II

Tube	Gas	C	Pressure p (mm)	λ_e (cm)	d (cm)	λ_e/d	(θ)	Cpd	Electron Temp. °K obs.	Electron Temp. °K theor.
10 amp	xenon	8.1×10^{-2}	0.1	3.33	1.0	3.33	0.75	8.1×10^{-3}	13350	13800
0.1 amp	xenon	8.1×10^{-2}	1.0	0.66	.5	1.32	1.2	4.1×10^{-2}	13900	11500
0.1 amp	xenon	8.1×10^{-2}	2.1	.33	.5	0.66	1.57	8.5×10^{-2}	11100	10300
0.1 amp	xenon	8.1×10^{-2}	3.0	.22	.5	0.44	(1.57)	12.1×10^{-2}	8700	9800
0.1 amp	xenon	8.1×10^{-2}	4.0	.15	.5	0.30	(1.57)	16.2×10^{-2}	8000	9200
0.1 amp	argon	5.2×10^{-2}	5.0	.066	.5	0.132	(1.57)	13.0×10^{-2}	13900	16200
0.1 amp	argon	5.2×10^{-2}	10.0	.033	.5	0.066	(1.57)	25.9×10^{-2}	13300	12000

tube also had a value of one centimeter, the sides of the anode probably acted to reduce the effective value of d. Hence, the rough value of 0.5 centimeter was taken. Only for the first three cases is the value of λ_e/d such that θ comes out to a value existing on the curve of Figure 5. This is not a catastrophe, however, since this difficulty arises directly from the approximation made in assuming that the cathode edge of the plasma is terminated where $E = -kT_e/e\lambda_e$. If the boundary value of E is made about ten times smaller, then the difficulty vanishes.* As can be seen from Figure 6, the actual value of the geometrical factor θ does not have very much effect on T_e, particularly at large values of the parameter Cpd. The theoretical value of T_e was found from the curves in Figure 6 by using the values of Cpd and θ listed in Table II. It can be seen that the agreement between theory and experiment is within about 15 per cent. This is about as good as can be expected. It is all the more significant in that it was obtained over a considerable range of tube sizes, gas pressures, anode current values, and with two types of gas.

The presence of good linearity on the semilog probe plots is, in itself, good justification for the basic assumption upon which the whole analysis was based. Only in the case where the xenon pressure was as low as 1 millimeter (see Figure 9) was there any indication of nonlinearity in the probe plots at probe currents below those clearly affecting discharge conditions.

Plasma Potentials

The potentials inside the plasma can be obtained with Equation (21). These potentials are related to the anode by Equation (24). The anode potential is related to the dark plasma potential by means of either Equation (25) or (26). The dark plasma is about two volts positive with respect to the cathode.

Several solutions of Equation (21) for different values of the parameter λ_e/d are shown in Figure 10. For convenience, the dimensionless quantity eV/kT_e is used as ordinate. It can be seen that the peak plasma potential is of the order of several tenths of the mean electron energy.

Applying Equation (24) to the 10-ampere tube, which gives results typical of the smaller tube, it is found that[19]

---

* What the scattering mechanism does to λ_e is not known. The value of λ_e may be greatly different from the literature values which were all taken under non-plasma conditions.

[19] The value of the electron space current $ne\bar{c}_e/4$ can be conveniently obtained from the nomograph on p. 202 of L. Malter and W. M. Webster, "Rapid Determination of Gas Discharge Constants from Probe Data," *RCA Review*, Vol. XII, pp. 191-210, June, 1951.

$$V_S = \frac{13{,}800}{11{,}600} \ln \frac{30}{1} = 4 \text{ volts},$$

if the value of $eV_e = J_e = 1$ ampere/cm^2 is used for the anode current, and the ion density of approximately 1.4×10^{13}, as computed in one of the following sections, is used. This means that the peak plasma potential is slightly greater than 4 volts with respect to the anode. Since the anode is 2 to 3 volts higher in potential than the dark plasma, the peak plasma potential is about 6 to 7 volts higher in potential than the dark plasma.

It was found that the probe in the 10-ampere tube floated at a

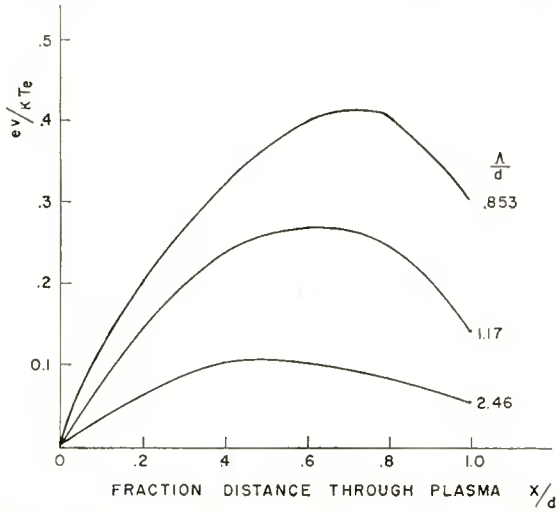


Fig. 10—Curves showing plasma potential distributions.

potential 4.8 volts below that of the anode. Using the curves in Figure 11, from which the conventional floating potential V_w of a probe with respect to a plasma can be conveniently obtained, it is found for xenon with $T_e = 13{,}800°\text{K}$ and $T_p = 300°\text{K}$ that $2.32 \times 10^4 V_w/T_e = 16.2$. Hence $V_w = 9.65$ volts* and we see that the plasma potential should be 4.45 volts above that of the anode. This agrees very well with the 4 plus volts computed from the theory.

The use of a hot probe to check plasma potential, while attractive at first sight, runs into the difficulty that it will inject copious quantities of thermal electrons into the "hot" plasma and, no doubt, greatly affect conditions therein. An objection can also be raised about the

* If V_w is computed by making use of Equation (19) for the ion current, and the usual Boltzmann relation for the electron current, a comparable value is obtained.

effect that a cold probe would have on the discharge if inserted out into the middle of the plasma. It would certainly affect the discharge geometry. As used in the present experiments, the cold probe, by being a part of the anode surface, acts just like a section of the anode and, thus, should have a minimum perturbing effect on the discharge.

Total Ion Generation

This can be computed by the use of Equation (18). A more convenient form of this equation is

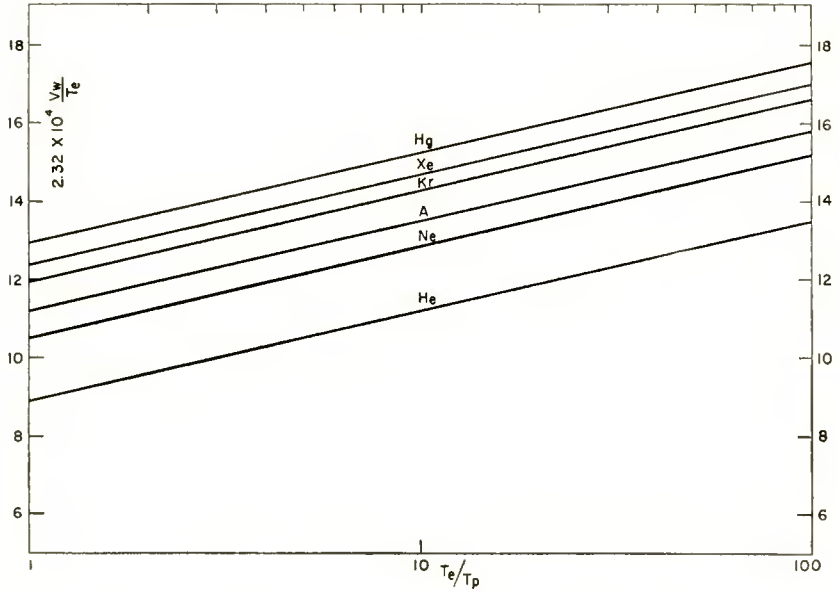


Fig. 11—Curves for calculation of probe floating potential.

$$\frac{\sum \Gamma_p}{\Gamma_e} = (\sqrt{G}d)\left(\frac{\lambda_e}{d}\right)\frac{\mu_p}{\mu_e}\sin\sqrt{G}d.$$

The 10-ampere tube mentioned in Table II gave results which are typical of those of the smaller tube so that only the former need be discussed. Using the values of $\sqrt{G}d$ and λ_e/d given from Table V and the value of 2.7×10^{-3} for μ_p/μ_e* it is computed that the ratio $\sum \Gamma_p/\Gamma_e$

* $\frac{\mu_p}{\mu_e} \cong 1.18\frac{\lambda_p}{\lambda_e}\sqrt{\frac{T_e}{T_p}}\sqrt{\frac{m}{M}} = 1.18\left(\frac{1}{5.6}\right)(6.3)\frac{1}{490} = 2.7 \times 10^{-3}$

for xenon, where the classical value for λ_p/λ_e has been used in ignorance of the true value.

should have a value of about 4.6×10^{-3}. The measured value, obtained by noting the ion currents to negative electrodes and making an estimate of the unmeasurable ion currents, has a value of approximately 2×10^{-2}. These values are within an order of magnitude of each other and, considering the difficulty in getting a correct value for the actual ion currents and approximations made in the analysis, the agreement is as good as can be expected. The results with the other tubes are of a comparable nature.

Plasma Density

This can be computed from Equation (17). For the 10-ampere tube the value of the ambipolar diffusion coefficient, measured with xenon at 0.1 millimeter pressure and normalized to 1.0 millimeter pressure, is given by[14]

$$D_a = D_p\left(1 + \frac{T_e}{T_p}\right) = 10\ (1 + 40) = 400\ \text{cm}^2/\text{sec}.$$

The value of μ_p/μ_e is as given in the previous subsection, and the value of the electron mean free path is, from Table II, about 0.33 centimeter at 1.0 millimeter pressure. With an anode current density of roughly 1 ampere per square centimeter, the maximum plasma density as computed from Equation (17) is 1.4×10^{13} particles per cubic centimeter.

The probe in the 10-ampere tube drew a saturated ion current density of about 25 milliamperes per square centimeter. Using this current density, and a presumed ion temperature of 300°C, the standard equation for plasma density

$$\frac{ne\bar{c}_p}{4} = J_p$$

gives a plasma density of about 2.2×10^{13}. The fact that this agrees with the above computed value within a factor of ten is probably all that can be deemed significant with respect to a check of the theory.

Critical Plasma Size

Experiments on this quantity were carried out with a tube of the same construction as that shown in Figure 8, except that it contained no probe and the grid–anode distance could be adjusted at will by virtue of a movable anode. Movement of the latter was accomplished by means of a magnet. The grid–anode distance could be varied from 0.2 to 3.0 centimeters.

The variation of anode potential as a function of grid–anode

distance in a tube containing argon is typified by the curve shown in Figure 12. The anode potential is seen to remain fairly constant with the initial increase in distance. Then, as can be seen, the anode potential rises abruptly by about ten volts. As noted in the figure, the appearance of a glowing layer on the anode surface accompanies this abrupt rise in anode potential. The change in spacing, over which this abrupt rise in potential takes place, is about five times larger than that computed from the 3/2 power law, wherein zero initial velocities of the electrons are assumed.*

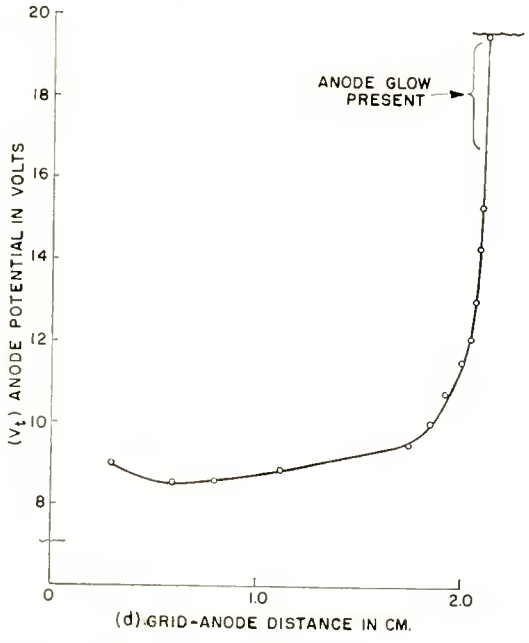


Fig. 12—Variation of anode potential with grid–anode distance.

When the spacing is increased beyond that noted in the figure the glowing layer is replaced by one or more glowing balls that are attached to the anode surface. When these appear, the anode potential becomes unstable and the average applied anode potential assumes a value that is only several volts in excess of the initial anode potential.

The behavior of the pressure and critical grid–anode spacing for different gases is shown in Table III. These values were obtained with a discharge current of 75 milliamperes. As this current is increased the critical value of spacing was found to increase slightly. It will be

* Both the initial velocities of the electrons entering the sheath and positive ions in the sheath help account for this discrepancy.

noticed that the product pd_1 is constant except when d_1 is large in value. Large values of d_1, particularly when comparable with tube diameter (about 4.0 centimeters) as might be expected, cause the critical spacing to be smaller than it should be on the basis of the plane parallel model. The value of the quantity λ_e/d suggests that the proper value for ψ_c is close to 0.1.

Table III

Gas	Gas Pressure p (mm)	Electron Mean Free Path[20] λ_e (cm)	Critical Distance d_1 (cm)	$\frac{\lambda_e}{d}$	pd_1
xenon	1.0	0.33	3.5	0.94	3.5
	2.0	.165	2.4	.069	4.8
	3.0	.11	1.7	.065	5.1
	4.85	.068	1.0	.068	4.85
argon	2.6	0.127	3.2	.040	8.3
	5.2	.064	2.5	.025	13.0
	8.0	.041	2.3	.018	18.5
	10.8	.031	1.7	.018	18.4
	13.0	.025	1.4	.018	18.2

In summary, experiment and analysis are in excellent semi-quantitative accord with respect to the limitations on plasma size and the phenomena which take place when these limits are exceeded.

Conclusions

This paper describes an analytical and experimental treatment of an externally heated hot-cathode discharge. This discharge is unique and of special interest since it operates with an anode potential that can be very small compared to the ionization potential of the gas.

This discharge is analyzed in a systematic manner by the application of particle flow and continuity relations and the notion that all ion generation stems directly from energetic electrons in the tail of a Maxwellian energy distribution. Expressions for electron temperature, plasma potential, plasma density, total ion generation, total energy consumption, and limitations on plasma size are derived.

Conventional probe measurements and experiments based upon comparisons with other types of discharges, already quite well understood, indicate that the analysis is in reasonable accord with reality.

[20] Reference (13), p. 31.

The electron temperature, energy balance, and plasma potentials are quite closely accounted for.

Of particular interest is the verified prediction that the plasma wherein the ion generation takes place cannot be larger than a certain critical size. The onset of an anode glow and a second plasma, when this dictate is not in accord with tube dimensions, is also correctly predicted. An argument is advanced to show that it is quite likely that this behavior is similar to that which takes place in the phenomenon of striations.

TRANSISTORIZED SYNC SEPARATOR CIRCUITS FOR TELEVISION RECEIVERS*

By

HUNTER C. GOODRICH

RCA Victor Television Division,
Camden, N. J.

Summary—This paper presents transistorized sync separator circuits which appear to meet commercial requirements for operation under both good and adverse conditions. Tests on a limited number of transistors indicate good interchangeability and satisfactory performance up to 60°C. The impulse-noise immunity of sync separators is discussed. A method of using diode switching to control the time constant of a single-transistor separator for optimum impulse-noise performance is given.

Transistor collector voltage ratings must be at least equal to the level of sync output voltages required. High-frequency response on the order of that of the experimental SX-160[1] is needed in the horizontal separator for good sync rise time. Low I_{eo} and I_{co} are desirable, but most of the circuits described will tolerate high-temperature leakages of the order of 50 to 100 microamperes.

Because of the abrupt low-voltage "knee" of the junction transistor, one transistor will produce double-clipped sync. Since this function is usually accomplished with a two-stage vacuum-tube amplifier, the use of a transistor results in increased circuit simplicity.

GENERAL DISCUSSION

FEW circuits in a television receiver have shown as much variety as those for sync separation and automatic gain control (a-g-c). This has resulted from the efforts of many engineers to devise economical circuits which will perform reliably under a wide variety of conditions. Some of these conditions encountered in the field are listed below:

1. A signal strength range on the order of 100,000 to 1.
2. Variations in sync percentage from the nominal 25 per cent.
3. Variation in tube (or transistor) characteristics including initial characteristics, aging, changes with temperature.
4. Variation in gain of associated circuits such as the r-f, i-f, and video stages.
5. Variations in line voltage.

---

* Decimal Classification: R583.13.

[1] C. W. Mueller and J. I. Pankove, "A P-N-P Triode Alloy Junction Transistor for Radio-Frequency Amplification," *RCA Review*, Vol. XIV, pp. 586-598, December, 1953.

6. Impulse noise of various types.
7. Airplane flutter.
8. Reflections and multipath reception.
9. Various degrees of picture contrast desired by user.

Vacuum-tube circuits have been developed which meet these conditions in a reasonably satisfactory manner. Although such circuits may be relatively uncomplicated in structure, they are the result of much engineering and field testing, and make optimum use of the characteristics of the particular tubes employed.

Transistors differ from vacuum tubes in many important respects. As a result, most of the vacuum-tube circuits will not work effectively with transistors. The purpose of this study has been to investigate the characteristics of transistors which are relevant to sync applications and to develop circuits which will utilize these characteristics to provide optimum performance at reasonable cost. Certain transistor characteristics such as low input impedance, uncontrollable forward collector conductivity, temperature sensitive leakage, and limited voltage ratings generally represent disadvantages for sync applications. On the other hand, the high transconductance, sharp cutoff and saturation characteristics, and the absence of heaters, represent definite advantages for transistors. Also, the existence of both p-n-p and n-p-n transistor types offers increased flexibility in devising circuits. Thus the application of transistors to sync circuits presents both new problems and new opportunities to the circuit designer.

The Junction Transistor as a Pulse Amplifier

All of the three basic transistor amplifier types—common emitter, common base, and common collector—can be used as sync amplifiers. The common emitter connection gives the poorest rise and fall times for a given transistor. However, since it is the only one giving both voltage and current gain, both of which are usually required, it is the most useful. Most of the following discussion of pulse amplifier characteristics is therefore devoted to the common emitter type.

Figure 1a illustrates the common emitter circuit and Figure 1b shows one form of its equivalent circuit.[1] Equivalent-circuit parameters are given for the 2N34 and the SX-160, the latter being an experimental alloy junction r-f transistor described by Mueller and Pankove.[1] Since the vertical straightness of the received picture depends on the timing accuracy of the sync pulses, reasonably good rise time of the horizontal sync amplifier is highly desirable. An important limitation on rise time is set by the time constant of $r_{bb'}$ and $c_{b'e}$. A low-frequency transistor such as the 2N34 may have a

rise time actually in excess of the 5-microsecond width of horizontal sync when operated with common emitter.

The 10 to 90 per cent rise and decay times of the medium-frequency experimental SX-160 with a 10,000-ohm collector load are of the order of 2 microseconds. Both rise and decay times increase with increasing collector load resistance as indicated in Figure 2. The rise time can be reduced greatly by overdriving, as will be discussed later. The equivalent circuit of Figure 1b indicates that any driving source resistance will lengthen the rise time by effectively adding to $r_{bb'}$. Tests on an SX-160 substantiated this as shown in Figure 3.

Figure 4 illustrates the collector family of curves for a junction transistor. A fortunate aspect of this characteristic for sync separators is the abrupt knee which occurs near zero collector voltage.

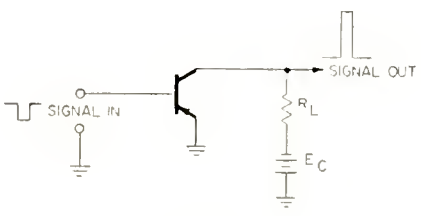


FIG. 1a

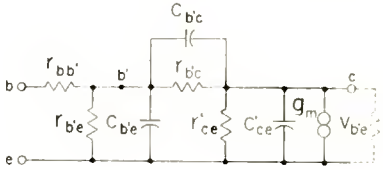


FIG. 1b

TYPICAL VALUES FOR I_C = 1 MA.

		2N34	SX-160
$r_{bb'}$	OHMS	250	75
$r_{b'e}$	OHMS	1000	500
$C_{b'e}$	UUF	10,000	1000
$r_{b'c}$	MEG	1	0.1
$C_{b'c}$	UUF	35	10
g_m	MHOS	.032	.032

Fig. 1—(a) Common-emitter circuit; (b) pi-equivalent circuit.

Thus an amplifier with the proper load may be driven into increasing collector conduction until abruptly limited by collector voltage saturation. Such an amplifier, when driven from cutoff to saturation, produces a double-clipped output with an amplitude only a few tenths of a volt less than the collector supply voltage. The relatively flat collector characteristic above the knee is a result of the fact that the current carriers passing through the base layer to the collector travel by diffusion instead of being attracted by an electrostatic field as is the case in a vacuum tube. The value of the collector voltage thus has little effect on the collector current as long as it is sufficient to attract the carriers that have reached the collector junction.

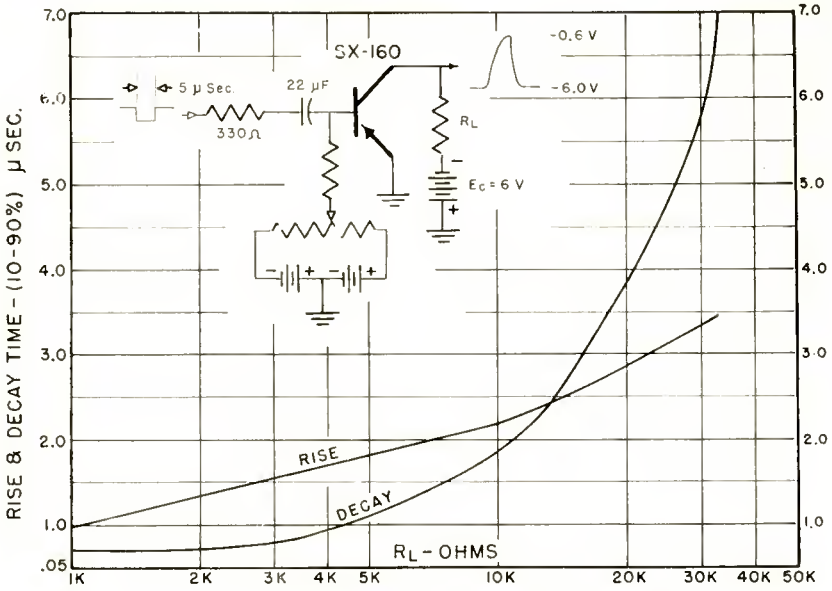


Fig. 2—Rise and decay time versus collector load.

Pulse-driving a transistor into collector saturation modifies both the rise and fall times at the output. The application of a pulse which drives the base–emitter junction farther in a forward direction than is necessary just to saturate the collector will be called *overdriving* the transistor. Overdriving produces a marked decrease in rise time.

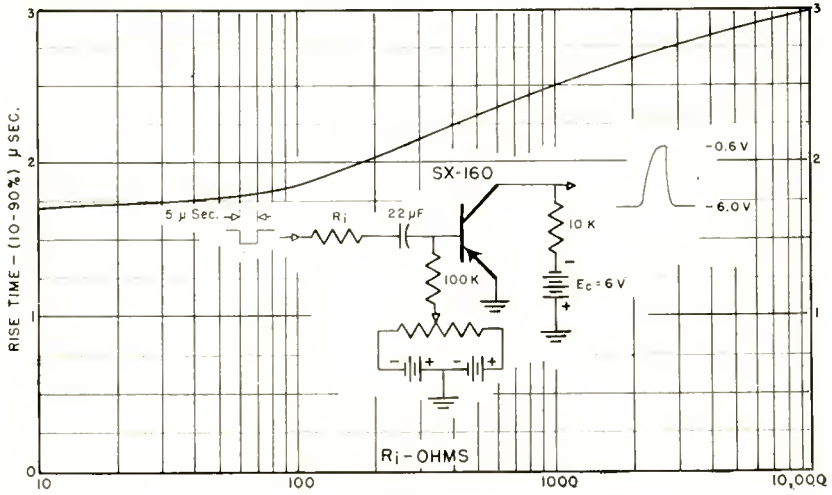


Fig. 3—Rise time versus external base resistance.

As an example, applying an input signal voltage double that required just to saturate the collector of an SX-160 reduced the rise time from 2 to 0.2 microsecond. Lesser amounts of overdriving will produce lesser degrees of pulse steepening. The variation in rise time as a function of base-driving current in excess of that required to saturate the collector on sync tips is given in Figure 5 for an SX-160 and a 2N34. Since the output transistor of a sync separator will normally be overdriven to some extent to achieve double clipping, a considerable degree of pulse steepening will result. However, the use of an inherently low-frequency transistor which requires a large degree of over-

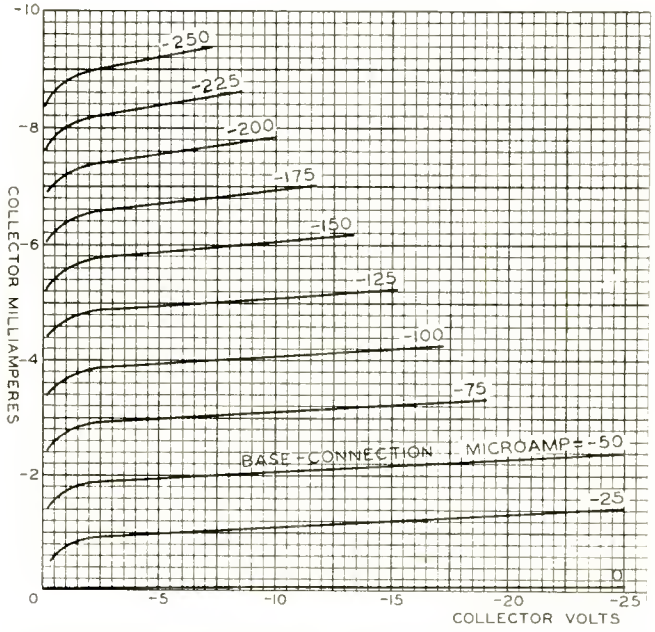


Fig. 4—Typical junction-type transistor characteristics.

driving to produce an acceptable rise time is wasteful of gain and has a deleterious effect on the trailing edge of the output.

The further effect of overdriving a junction transistor is that the trailing edge of the output pulse may be delayed, i.e., the pulse may be effectively widened. This effect has been called "back-porch" effect. Back-porch results from the fact that, while collector saturation limits the collector current to a value equal to the supply voltage divided by the load resistance, overdriving produces a minority carrier current flow into the base in excess of this limit. Some of the excess carriers are thus stored in the base and continue to flow to the collector after the input signal has returned to zero or reversed. Serious back-porch

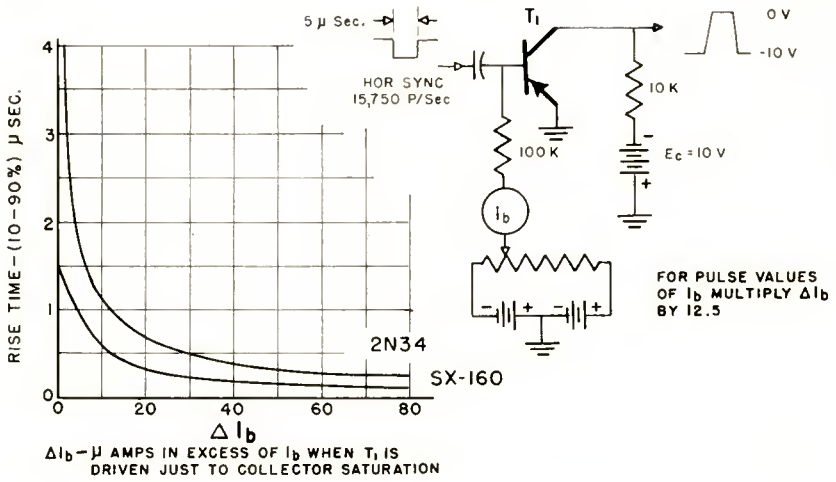


Fig. 5—Rise time with overdriving.

effect shifts the average timing of the sync output pulses and will cause a phase shift in many types of horizontal phase detectors. Other types of phase detector work primarily from the front portion of the sync pulses and will be little affected. In any case it is possible with a medium-frequency transistor such as the SX-160 to get adequate overdriving for good limiting and rise time without producing serious back-porch effect.

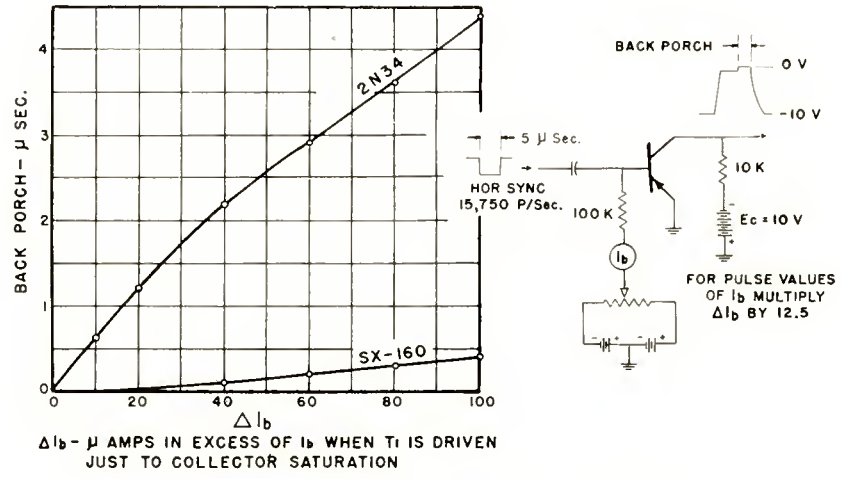


Fig. 6—Back-porch width with overdriving.

Driving Source

Transistors operated with common emitter or base have a low input resistance. The particular value of resistance varies considerably depending on transistor characteristics, circuit parameters, and operating current. An average SX-160 as typically operated in a sync separator with common emitter and an 8200-ohm collector load has an input resistance of the order of 500 to 1000 ohms within its amplifying range. Beyond the saturation point the input resistance falls to a fraction of this value; below cutoff it may rise to over a megohm. It is evident that a low-impedance video source is desirable to drive transistorized sync and a-g-c circuits most effectively. Since at low power levels a transistor is essentially a current-driven device, a high source impedance with a signal voltage proportionally increased to supply the required signal current might also be satisfactory. However, there are two limitations to the use of such a high-impedance source: (1) the high video amplitude existing between sync pulses while the transistor is at cutoff will cause increased base-to-emitter leakage, thus disturbing the bias, and (2) the source impedance adds to the base resistance to produce, in conjunction with $C_{b'e}$, a longer rise time of the output sync pulses.

It seems basic to a-g-c circuits that the video signal must be d-c coupled to the a-g-c rectifier or amplifier. Otherwise a sudden increase in signal strength can overload the i-f amplifier, reduce the a-c video component, and block the receiver as a result of the faulty information derived from the a-c signal. Since transistors will not withstand the normal plate voltages encountered in a television receiver, this means that the video signal applied to an a-g-c or combined sync and a-g-c transistor circuit must come from a source close to ground potential.

One obvious source of such a video signal is the second detector load. Since the transistors tested had a base-to-emitter capacity on the order of 15 micromicrofarads in the cutoff condition, such a transistor must be tapped down at least to the mid-point of the detector load to prevent undue capacitive loading. Unfortunately, the detector signal level in terms of video current available to the separator is not sufficient to operate the more economical separator circuits. A further disadvantage of this signal source is that any impulse noise present has not yet been clipped by the first video amplifier.

A second possible video source is from a resistive divider from video plate to ground as shown at point A of Figure 7a. For some separator and a-g-c circuits this source is quite satisfactory. However, since the impedance of the divider must be high compared with the plate-to-B+ load to maintain the necessary video plate voltage, only

a fraction of the total video plate current is available to the transistor circuits. In cases where the signal source need not be near d-c ground, as in a capacitance-coupled separator, the full video signal current is available from a tap (B) on the normal video plate load.

A third video source, which eliminates most of the disadvantages listed above, is shown in Figure 7b. The voltage developed across R_1, the normal cathode resistor, is not usually sufficient to supply the sync and a-g-c circuits. An additional resistor, R_4, is added to increase the video voltage at the cathode. This does not introduce degeneration since it is not in series with the grid-to-cathode voltage applied from the second detector load. Taking the signal from point C, the source impedance is low and grounded and carries the full video current in V_1. Optionally, a variable resistance R_5 can be inserted as a contrast control. As the resistance of R_5 is increased, the voltage across it

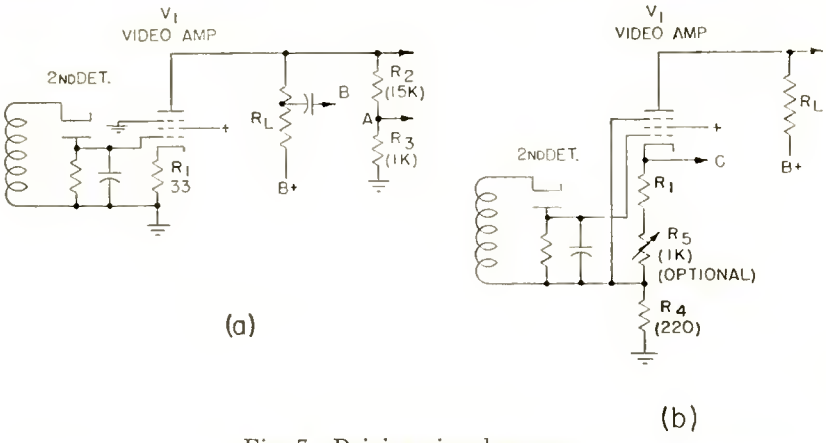


Fig. 7—Driving signal sources.

increases, but the voltage across R_4 decreases as degeneration decreases the gain of V_1. The result is that the voltage to the separator can be kept relatively constant. If the resistance of R_5 is made too great, the current available to the separator may be insufficient, but a fair range of contrast control may be achieved. In this circuit the capacitance of the V_1 grid lead to ground is effectively increased by the added video voltage developed across R_4. This capacitance should therefore be kept as low as possible.

Two-Stage Sync Separators

A critical factor in a transistorized sync separator is maintaining the proper separation bias level. Even with a good a-g-c system the video level applied to the separator will change considerably with

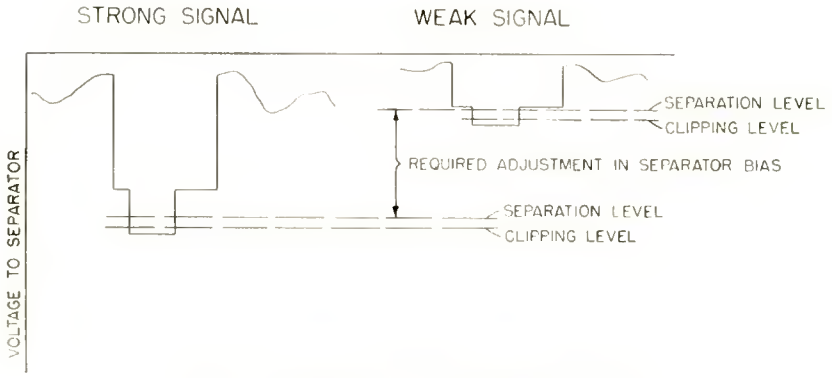


Fig. 8—Adjustment of separation level.

changes in line voltage, a-g-c control setting, and the extreme range in signal levels which may be encountered. Figure 8 illustrates how the separation level, i.e., the level where the separator begins to conduct, must change with signal level. This obviously requires the use of self-adjusting rather than fixed bias.

Figure 9 illustrates one form of two-stage sync separator. The emitter current flowing through R_1 establishes an emitter bias which biases the base–emitter junction of T_1 in the reverse direction except during sync. Collector current flows during sync to produce sync output pulses across R_2. E_2 and R_2 are proportioned to avoid collector saturation, so that the pulse amplitude across R_2 varies somewhat with variations in input amplitude. The emitter current likewise varies with signal amplitude so that the bias is kept at the proper level for separation on all signals. This circuit is analogous to the cathode-biased vacuum-tube separator frequently used.

The first separator in Figure 9 is relatively noncritical of transistor characteristics, assuming that the transistor has adequate high-

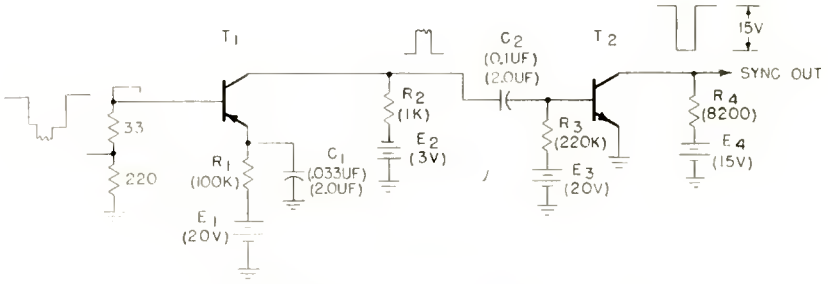


Fig. 9—Two-stage sync separator.

frequency response to give the required rise time. The collector current versus base-to-emitter voltage characteristic at low currents has been found fairly constant from unit to unit, at least for a given transistor type. The input resistance varies considerably, particularly with variations in α, but this does not affect the circuit operation if the signal has a sufficiently low impedance. An advantage of placing bias resistor R_1 in the emitter circuit is that if the d-c impedance of the base circuit is low, the collector-to-base leakage current, I_{co}, does not appreciably affect the bias. Between sync pulses the base–emitter junction is biased in the reverse direction, and any leakage at this junction, I_{eo}, will flow through R_1. In a transistor whose characteristics are satisfactory for the circuit, I_{co} should be small compared with the average emitter current, and hence have only a minor effect on bias.

In the circuit of Figure 9 a second transistor, T_2, is used to amplify and limit the amplitude of the pulses across R_2. T_2 is operated so that it is driven to saturation during sync by any usable signal. This provides sync of uniform amplitude and cuts off any noise pulses at sync level. Since the sync has been completely separated from the blanking pulses by T_1, it would be possible to operate T_2 with fixed bias or with direct coupling. However, the bias would have to be set so that T_2 would saturate on the sync of weak signals; this means that on a strong signal T_2 would be heavily overdriven. The result is a considerable widening of horizontal sync or strong signals due to back-porch effect. Such heavy overdriving is avoided by using self-bias on T_2, as is shown in Figure 9. Bias E_3 is adjusted so that T_2 will saturate on very weak sync. A transistor driven to collector saturation ceases to give normal transistor action since the collector current is fixed. The base-to-emitter characteristic under this condition is like that of a junction diode and has a low forward resistance. The effect of saturation on input resistance is shown in Figure 10. The input circuit of T_2 thus functions as a clamp circuit, producing sufficient reverse bias so that T_2 is always overdriven but never to the extent of producing excessive back-porch widening of the sync output. The average base current I_b due to conduction during sync, charges C_2 and must equal the currents discharging C_2:

$$I_b = I_3 + I_{co} + I_{eo},$$

where $I_3 =$ current through bias resistor R_3,

$I_{co} =$ collector junction reverse leakage current,

$I_{eo} =$ emitter junction reverse leakage current.

The base current actually flowing during sync will be the average I_b divided by the duty factor of sync; this is about 12 I_b. Variations of I_3 with signal level will be minimized by making E_3 relatively large compared with the amplitude of sync applied to T_2. The primary effect of elevated temperature on T_2 is to increase the base current flow

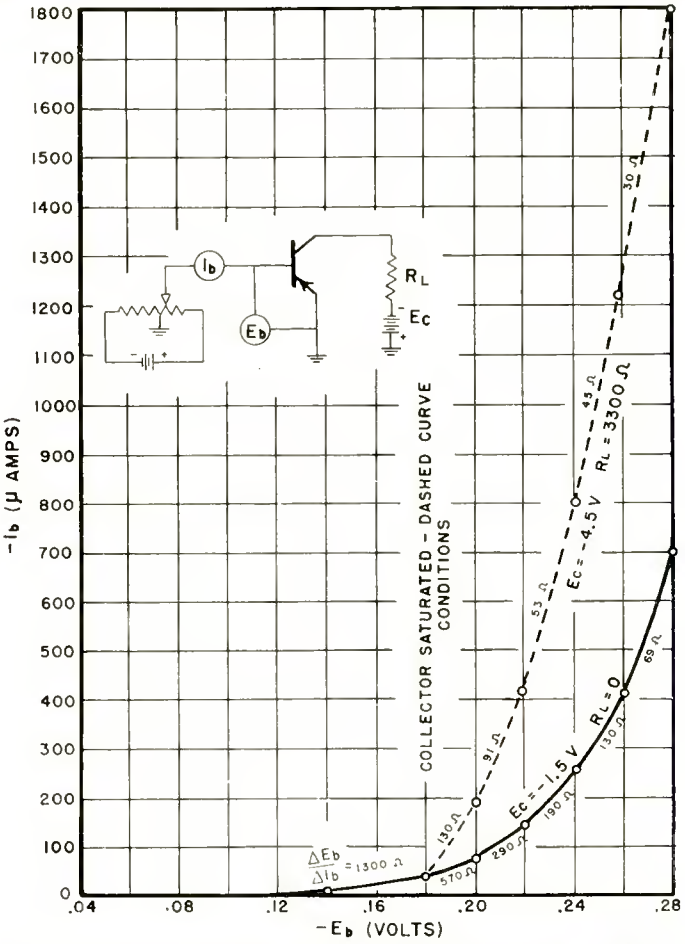


Fig. 10—Input resistance of common emitter stage (SX-160 for saturated and unsaturated collector).

during sync as a result of increased I_{co} and I_{eo}. With an experimental SX-161 transistor (the n-p-n counterpart of the experimental SX-160) an induced change in I_3 of 100 microamperes changed the output pulse width 0.5 microsecond. A similar increase in I_{co} and I_{co} would have the same effect. The 2N35 exhibited a greater storage effect, the output pulse being widened 2 microseconds by the same change in I_3.

The two-transistor separator of Figure 9, driven by the circuit of Figure 7b, was connected to a receiver. The sync output was applied to the balanced horizontal transistor phase detector.[2] An SX-160 and an SX-161 were used as T_1 and T_2 respectively. When the smaller values of C_1 and C_2 were used for separation of horizontal sync only, the performance of the circuit was quite satisfactory. Sync held over the full range of usable signal levels and produced no shift in raster when the sync in the signal was reduced to 12 per cent. Transistor interchangeability was good and simulated increases of 100 microamperes in I_{co} and I_{eo} in both stages produced only minor changes in centering. When the larger values of C_1 and C_2 are used to permit separation of both horizontal and vertical sync, performance is similar except for the decrease in horizontal impulse noise immunity resulting from the longer time constants. This effect is discussed in a later section.

The circuit of Figure 9 produces negative sync output which, as mentioned, is suitable for driving the transistor phase detector described in Reference (2). The sync polarity may be reversed by operating T_1 as a common-collector amplifier or by operating T_2 as a common-emitter amplifier. Figure 9 is presented as an example of a two-stage separator, and the discussion thereof explains some of the design factors involved. Many variations are possible. Since for each of the two stages there are three possible amplifier types (common emitter, base, or collector) two basic transistor types (p-n-p or n-p-n) and three means of biasing (self-base bias, self-emitter bias, or direct coupling), the number of possible combinations is in the hundreds. Many such combinations have obvious disadvantages, but many others would be workable circuits. Although it has not been possible to test all of these, the circuit of Figure 9 is believed to be among the better two-stage separators.

One-Stage Sync Separators

The two functions of sync separation and sync clipping are usually performed by separate vacuum tubes in commercial television receivers. A two-stage transistor separator was discussed in the previous section. However, the sharp cutoff and saturation characteristics of the junction transistor make possible its use as a one-stage separator, in which the same stage both separates and clips the sync pulses.

---

[2] B. Harris and A. Macovski, "A Symmetrical-Transistor Phase Detector for Horizontal Synchronization," *RCA Review*, Vol. XV, pp. 18-26, March, 1955.

The one-stage separator is shown in Figure 11. The bias developed by the emitter current flow through R_1 permits T_1 to conduct only during sync. At this time T_1 is driven to collector saturation, thus clipping the tops of sync as well as any noise that may be present. The output of T_1 is thus double clipped and will have a fixed output amplitude approximately equal to E_2. A one-stage separator can also be built with the self-bias elements R_1 and C_1 in the base circuit, but this has the disadvantage that collector leakage current I_{co} passes through the bias resistor.

There are certain differences in the self-biasing action between Figure 11 and T_1 of Figure 9. In Figure 9, T_1 is not driven to saturation. The average collector current will thus vary with signal level. Since the collector current is the major component of the emitter current, it will cause adjustments in emitter bias with signal level. The change in base current required to produce a given change in

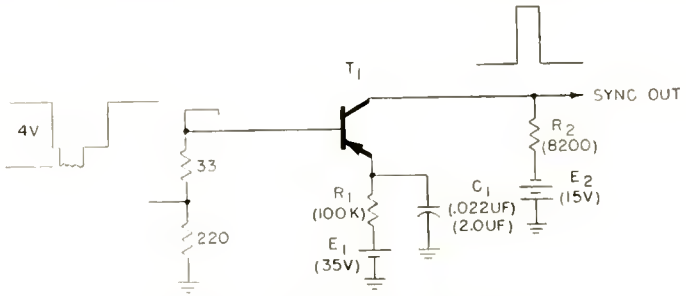


Fig. 11—One-stage sync separator.

bias current is low because of the base-to-emitter current gain of the transistor.

In Figure 11, since the collector is driven to saturation during sync at all signal levels, the collector current cannot increase with signal level. If the emitter bias is to adjust properly to varying signal levels, the adjustment must then come from changes in base-to-emitter current flow. As previously noted, after collector saturation the base–emitter junction takes on the characteristic of a forwardly biased diode. The fall in input resistance beyond saturation is shown by the change in slope of the dotted curve of Figure 10. The resistance falls well below 100 ohms. In this circuit the rectified input signal itself must provide the variations in bias, without benefit of the current gain of the transistor. The relationship between the required input signal level and source impedance and the variations encountered in signal amplitude and I_{eo} will be illustrated by an example. The separator circuit of Figure 11, driven by the circuit of Figure 7a, was

installed in a receiver. The source impedance to the separator was 250 ohms and the absolute level of sync tips varied from +1 to +4 volts above ground for maximum combined excursions of a-g-c control setting and signal strength. Since the emitter bias closely follows the level of sync tips, this means a 3-volt variation across R_1. This change in current through R_1 due to change in signal level will be called ΔI_s, in this case it is equal to 33 microamperes.

Let us allow a 50-microampere variation in I_{co} with temperature and unit-to-unit differences. Now equating ΔI_b, the maximum resultant variation in current charging C_1 during each sync pulse, to the variations in currents discharging C_1:

$$\Delta I_b = \Delta I_s + \Delta I_{co} = 83 \text{ microamperes.}$$

Dividing the above average value by the duty factor of sync to get the base current flowing during each sync pulse gives about 1 milliampere. The normal video input signal is about 4 volts across the 250-ohm load. Sync normally represents 1 volt of this signal, so that 4 milliamperes of sync current flows in the source load. Since the input resistance of the transistor at saturation is very low, most of this current is available as input current to the transistor. Thus there is a safety factor to take care of weak signals or low percentage sync. The performance of the system under adverse signal conditions and normal temperature variations is good; details are given in the following section where a more refined version of the circuit is given. It may be noted here that increasing R_1 will decrease that part of ΔI_b due to signal level variations, but will not change the effect of temperature changes. If the allowance made for I_{co} could be reduced, the signal source impedance could be raised.

It is also possible to build a common-base single-transistor separator as shown in Figure 12. This arrangement produces a shorter rise time than would be achieved with the same transistor operated with common emitter. The biasing action is very similar; however, since the video source must supply the sync output current directly, this circuit requires a greater video input current for proper separation.

Sync Separator Noise Immunity

The synchronizing ability of a television receiver in the presence of impulse noise is a function of the noise characteristics of both the sync and the a-g-c circuits. This discussion is devoted to the former. One characteristic of a sync separator which aids noise immunity

is the ability to clip off noise so that it never exceeds the level of the sync output. The separators described in the previous sections do this. Another very important factor is that the separator not block or cut off for long periods following noise pulses. Impulse noise disturbances are normally of relatively short duration and low duty factor. The loss of sync information during the actual noise pulses usually does little harm. However, if the separator has long time constants that can charge on the noise pulses and keep the separator cut off for a long time thereafter, a serious loss of sync information results.

A long separation time constant is necessary for the separation of vertical sync. If horizontal sync is taken from the same circuit, the time constant must be on the order of 100 times that required for horizontal separation alone. This seriously affects horizontal noise

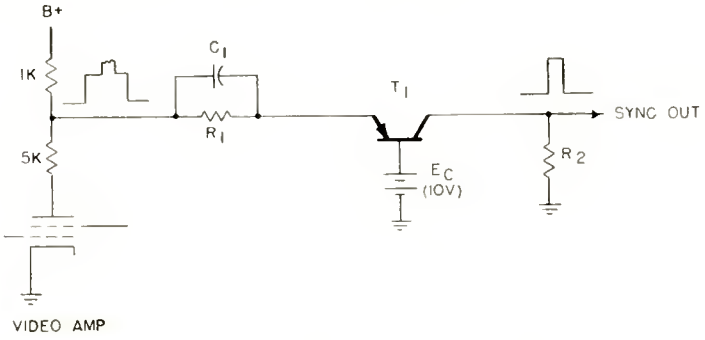


Fig. 12—Common-base one-stage sync separator.

immunity as shown by a comparison of lines 2 and 4 in Table I. One of the most effective solutions to the problem is the use of two separators, one optimized for horizontal and one for vertical. This represents an added expense, although the practicality of the one transistor separator makes it less so.

Another popular solution in vacuum-tube circuits is the use of a double time constant in the grid circuit of the separator. The application of this principle to a transistor circuit is shown in Figure 13.

In vacuum-tube circuits the equivalent in the grid circuit of R_3, which is part of the short-time-constant combination, is made very large so that it limits the degree of charging possible in the large-time-constant circuit. The presence of this series resistance does not seriously reduce the tube gain to vertical sync because of the very high input resistance of the vacuum tube. The low input resistance of a transistor amplifier limits the usefulness of such a circuit.

Table I—Comparative Sync Separator Noise Immunity

Sync Circuit	Minimum signal level giving good horizontal synchronization in presence of impulse noise	
	2-Volt Detector Level	6-Volt Detector Level
1. Typical vacuum-tube circuit with separate horizontal and vertical sync separators and a common sync output amplifier	380 μv	200 μv
2. Figure 11, one-stage transistor separator	3000	1200
3. Figure 13, above with double time constant	800	250
4. Figure 11 with $C_1 = 0.022$ microfarads for horizontal separation only	70	70
5. Figure 14, transistor with diode-controlled time constant	300	200

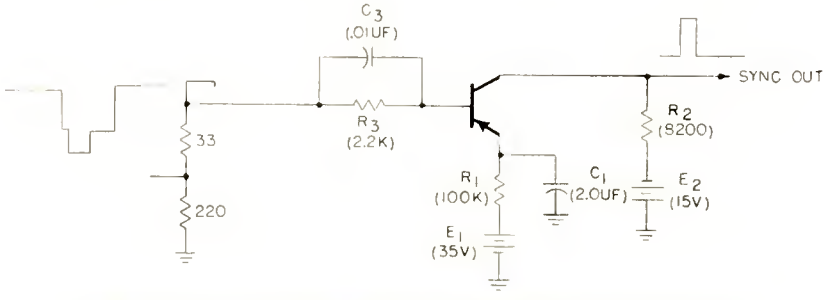


Fig. 13—One-stage separator with double time constant.

In this case, R_3 reduces the separator gain to vertical sync, and hence must be restricted in value. The values shown in Figure 13 reduced the vertical gain of the circuit by about one half. Comparison of lines 2 and 3 shows about a 3 to 1 improvement in horizontal noise immunity.

The common channel separator circuit of Figure 14 gives a marked improvement in noise immunity without appreciable loss in gain. This circuit is basically the common-emitter single-transistor separator previously described, adapted for positive power supply operation. However, it has two *R-C* time constant combinations in the emitter circuit which are brought into play at the proper time by diode D_1.

During each sync pulse, the emitter current of T_1 biases D_1 in the forward direction, placing the long time constant of R_3-R_4-C_3 in the bias circuit. This will maintain the emitter bias for the duration of

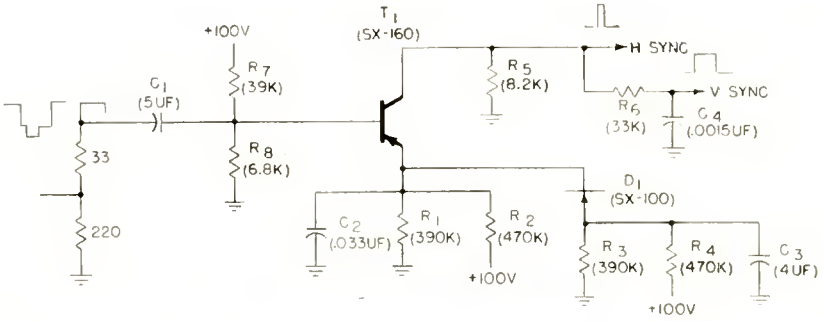
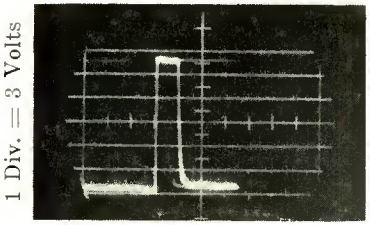
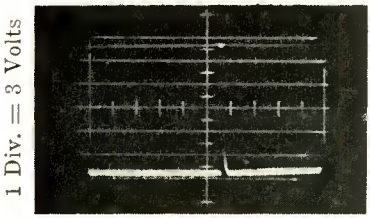


Fig. 14—One-stage sync separator with diode-controlled time constant.



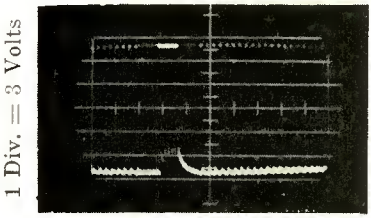
1 Div. = 5 μ Sec

(a) Horizontal rate.



1 Div. = 1200 μ Sec

(b) Field rate.



1 Div. = 240 μ Sec

(c) Expanded vertical sync.

Fig. 15—Sync output waveforms (one-transistor separator of Figure 14 with off-the-air signal).

the vertical sync pulse, permitting its proper separation. Any noise pulses will also cause D_1 to conduct, charging C_3. However, after the noise pulse the excess charge on C_3 will keep D_1 open, and horizontal separation can resume as soon as the potential across the short time constant combination R_1-R_2-C_2 reaches the normal value. Comparison of lines 2 and 5 shows an average 8 to 1 improvement in horizontal noise immunity for the double time constant circuit.

Results

The one-stage separator with diode-controlled time constant shown in Figure 14, driven by the circuit of Figure 7b, was installed in a receiver and checked against commercial requirements. It was found that the weakest intelligible signal could be held in sync and that no blanking appeared in the sync output until the sync was reduced below 15 per cent. The sync output had a 0.4-microsecond rise time and a 1-microsecond decay as shown in Figure 15. The resulting raster had no noticeable bends. In the presence of impulse noise the vertical and horizontal synchronization was comparable with that of a standard commercial receiver. The disturbance due to airplane flutter was somewhat less than in the standard receiver. Transistor interchangeability among the SX-160 transistors was good, the transistors being selected only for a V_{ce} greater than 30 volts. Heating the transistors to 60°C gave no appreciable change in performance. Of the transistors so tested, the highest I_{eo} was 80 microamperes at 60°C. The circuit showed good line-voltage tolerance, giving good separation from 85 to 130 volts. Thus it is possible to meet normal commercial performance standards with a transistor sync separator, and to do so with somewhat less circuitry than is usually employed with vacuum tubes.

INTERCEPTION NOISE IN ELECTRON BEAMS AT MICROWAVE FREQUENCIES*

By

W. R. Beam

RCA Laboratories,
Princeton, N. J.

Summary—*A theory is derived for interception-produced partition noise current and velocity in an electron beam of finite or infinite cross section. By considering the case of nonuniform interception probability across the cross section, expressions for partition current and velocity are obtained:*

$$\overline{i^2}_{partition} = 2\,e\,I_0\,\Delta f\,\Omega_2$$

$$\overline{v^2}_{partition} = \frac{2e\,\Delta f}{I_0}\left(\frac{kT_c}{mv_0}\right)^2 \frac{\Omega_2}{\Omega_1^2}$$

where e,m *are the electronic charge and mass,*

I_0 *is the d-c current incident on the plane of interception,*

Δf *is the measurement bandwidth,*

k *is Boltzmann's constant,*

T_c *is the cathode temperature in °K,*

v_0 *is the average axial velocity of the electrons,*

$\Omega_1 I_0$ *is the current transmitted through the interceptor,*

Ω_2 *is a partition parameter, calculated from the interception probability, and equal to* $\Omega_1(1-\Omega_1)$ *in the case of uniform probability.*

A suitable model, based on thermal velocities and confined flow in the beam, is used as a basis for calculation of Ω_1 *and* Ω_2 *in interception at a round aperture. Curves of* Ω_1 *and* Ω_2 *are given. Experimental confirmation of the accuracy of the model is shown in experimental data.*

A simplified analysis of the reduction of noise by interception of the edge of an electron bunch shows that reduction of noise current by this mechanism is generally much smaller than increase caused by partition noise.

For the case of interception by a nearly ideal mesh grid, considering only a single space charge wave mode, the fundamental noise parameter S is computed, showing that under the assumed conditions the noisiness of a heavily intercepted beam may not be considerably greater than that of a beam emerging from a cathode.

Introduction

THE increased noise in screen-grid tubes caused by interception of cathode current at the screen grid was analyzed in 1940 by North.[1] This analysis treated random, equally probable intercep-

* This work was supported in part by the U. S. Army Signal Corps under contract DA-36-039-sc-64443.

[1] D. O. North, "Fluctuations In Space-Charge-Limited Current at Moderately High Frequencies, Part III," *RCA Review*, Vol. V, pp. 244-260, October, 1940.

tion of all electrons on the screen grid, and did not include any velocity fluctuation effects. Since the invention of the traveling-wave tube, a noise theory[2-4] has been developed in which velocity fluctuations play an essential role. The interaction of electrons in "space-charge waves" leads to expressions for noise figure which include both current and velocity fluctuations introduced at the cathode, as well as constants of the region between cathode and radio-frequency interaction region.

In this study the theory of partition noise is extended to the case of an electron beam which is intercepted after passing from the cathode through a field region which establishes the character of the noise space-charge waves. Since interception in such electron beams is almost invariably the case of round beams passing through round apertures rather than wire mesh, an interception noise theory must be developed to handle such conditions. The analysis assumes a linearization whereby the partition noise and modified original noise of the beam are independent. This allows the effects due to partition to be separated from the modified original noise waves.

In order to calculate the magnitude of the noise due to the random interception process, some model must be assumed in which each electron is assigned a probability of interception. The model chosen has proved by experiment to be a realistic one. The partition noise is shown to contain a random velocity component in addition to the partition current. This velocity may under some conditions become important in determining noise behavior. The theory herein follows a method used by Rack,[5] where description of the electron current is in terms of velocity groups; it makes clearer from a statistical point of view the consequences of linearizing the fluctuations, and quickly establishes the independence of the several fluctuation components. A general expression for total noise current and noise velocity is readily obtained.

In the theories of minimum noise figure of microwave amplifiers[2-4] there appears a parameter, the "noisiness," which is invariant for a non-intercepted beam. For interception by a fine mesh grid, this noisiness value can be shown to be a minimum, when there is no inter-

---

[2] S. Bloom and R. W. Peter, "A Minimum Noise Figure for the Traveling-Wave Tube," *RCA Review*, Vol. XV, pp. 252-267, June, 1954.

[3] J. R. Pierce and W. E. Danielson, "Minimum Noise Figure of Traveling-Wave Tubes with Uniform Helices," *Jour. Appl. Phys.*, Vol. 25, pp. 1163, September, 1954.

[4] H. A. Haus and F. N. H. Robinson, "The Minimum Noise Figure of Microwave Beam Amplifiers," *Proc. I.R.E.*, Vol. 43, pp. 981-991, August, 1955.

[5] A. J. Rack, "Effect of Space Charge and Transit Time on Shot Noise from Diodes," *Bell Sys. Tech. Jour.*, Vol. 17, pp. 592, 1938.

ception, and to increase to a maximum near 50 per cent interception, but to decrease again at large interception. Expressions are derived for the noisiness, the noise-current standing-wave ratio, and position of the current maximum after interception. It is impossible to predict the exact nature of higher order space-charge waves in the beam, but these are included in the expressions.

Current and Velocity Fluctuations: Random Interception

The simplest physical model for interception noise is a filamentary electron stream which is partially intercepted in a random manner at an interception plane P. The average current in the stream before interception is I_0, and after interception, $k_0 I_0$.

The beam may be described, in the manner of Rack,[5] in terms of velocity classes: the entire spectrum of velocity is divided into a number of classes v_m (where v_m includes a range of velocities complementary to the ranges v_{m-1}, v_{m-2}, $\cdots v_{m+1}$, $v_{m+2}\cdots$. In each class m, the incident beam contains n_m electrons which approach the plane of interception during a time interval Δt. Because of the random nature of electron emission, n_m measured during different Δt intervals will take on different values; the variation may be expressed as

$$n_m = n_{0m} + \delta_m, \tag{1}$$

where δ_m has zero average value and n_{0m} is constant.

The average current approaching the interception plane during an interval Δt is

$$I_1 = \frac{e}{\Delta t}\sum_m (n_{0m}+\delta_m) = \frac{e}{\Delta t}\left[\sum_m n_{0m} + \sum \delta_m\right] = I_0 + I_{1ac}. \tag{2}$$

At the plane of interception, the fraction of electrons *not* intercepted can be expressed as $(k_{0m}+k_{1m})$, where k_{0m} is the average transmission and the average of k_{1m}, over all Δt is zero. (The value k_{0m} is independent of velocity class (v_m), for truly random interception.) The transmitted current may be expressed as

$$I = \frac{e}{\Delta t}\sum_m (n_{0m}+\delta_m)\ (k_0+k_{1m}) = \frac{e}{\Delta t}\sum_m (n_{0m}k_0+n_{0m}k_{1m}+\delta_m k_0+\delta_m k_{1m}). \tag{3}$$

If δ is small compared to n_0, the final term may be dropped.* The first term is simply the average transmitted current, so the a-c part of transmitted current is:

$$I_{2ac} = \frac{e}{\Delta t} \left(\sum_m \delta_m k_0 + \sum_m n_{0m} k_{1m} \right). \tag{4}$$

The k_{1m} and δ_m are independent, so that the currents of the two parts of I_{2ac} will add as their squares. The first part is simply $I_{1ac} \times k_0$, the a-c incident current multiplied by the transmission coefficient. The second part is the random interception current, as found by North.[1] The total mean-square current within a frequency band Δf can be shown† to be

$$\overline{I^2_{2ac}} = k^2_0 \, \overline{I^2_{1ac}} + 2eI_0 \Delta f (k_0 - k^2_0). \tag{5}$$

If $I_{1ac}{}^2$ is made equal to Γ^2 $(2eI_0\Delta f)$, the equation becomes identical to North's[1] Equation (7), upon identifying I_n/I with k_0.

Identification of the first part of Equation (4) with the incident a-c current is obvious, but it is important to realize that the δ's in the incident current may not be random; that is, the space-charge-wave interaction in the electron gun makes the current and velocity fluctuations statistically different from the values at the cathode. If the δ's are assumed random, and independent of m, the result will always be full shot-noise current, while it is well established[6] that lower values exist, due to space-charge-wave interaction, at frequencies where space-charge smoothing[7] is not thought to occur.

The velocity fluctuations in the beam may be established using a method similar to that of Rack.[5] The average electron velocity over an interval Δt is given by

---

* This is equivalent to saying that the *deviations* from average interception of the current $(e/\Delta t)\Sigma\delta_m$ constitute a negligibly small a-c current. The result, Equation (4), is found to be equivalent to considering partition noise and modified initial noise completely independent. An analysis presented in Appendix II supports this assumption.

† See Appendix I.

[6] C. C. Cutler and C. F. Quate, "Experimental Verification of Space Charge Transit Time Reduction of Noise in Electron Beams," *Phys. Rev.*, Vol. 80, pp. 875-878, December, 1950.

[7] D. O. North, "Fluctuations in Space-Charge-Limited Currents at Moderately High Frequencies, Part II," *RCA Review*, Vol. IV, pp. 371-388, January, 1940.

$$v_2 = \frac{\sum_m (n_{0m}+\delta_m)\ v_m\ (k_0+k_{1m})}{\sum_m (n_{0m}+\delta_m)}, \tag{6}$$

in which interception has been considered.

After manipulation in which cross products of δ and k_1 are neglected, and the constant part dropped,

$$v_{2ac} = \frac{\sum_m \delta_m k_0 (v_m - \bar{v}) + \sum_m n_{0m} k_{1m}\ (v_m - \bar{v})}{\sum_m k_0 n_{0m}}, \tag{7}$$

where $$\bar{v} = \frac{\sum_m n_{0m} v_m}{\sum_m n_{0m}}.$$

The first part of this expression is simply v_{1ac}, the velocity fluctuation in the incident beam.

The second part, v'_2, (the partition velocity) may be evaluated using the method outlined in Appendix 1:

$$\overline{(v'_2)^2} = \frac{2\Delta f \Delta t\ \overline{[\sum_m n_{0m} k_{1m}\ \ (v_m - \bar{v})]^2}}{[k_0 \sum_m n_{0m}]^2}. \tag{8}$$

Because of the independence of the k_{1m},

$$\overline{[\sum_m n_{0m} k_{1m})\ (v_m - \bar{v})]^2} = \sum_m \overline{(n_{0m} k_{1m})^2\ (v - \bar{v})^2}$$

$$= \sum_m (v_m - v)^2\ \overline{(n_{0m} k_{1m})^2} = \sum_m (v_m - \bar{v})^2\ (k_0 - k_0^2)\, n_{0m};$$

$$\overline{v'^2_2} = \frac{2\Delta f \Delta t\,(k_0 - k_0^2)}{k_0^2\ (\sum_m n_{0m})^2} \sum_m (v_m - \bar{v})^2 n_{0m};$$

$$\overline{v'^2_2} = \frac{2e\Delta f}{I_0} \frac{k_0 - k_0^2}{k_0^2}\ \overline{(v - \bar{v})^2}, \tag{9}$$

where the v_m may be replaced by a continuous velocity distribution. The quantity $\overline{(v-\bar{v})^2}$ should *not* depend on the space charge interaction

occurring before interception, since it is a measure of the temperature of the electron gas. Robinson[8] has found, at potentials greater than a few times kT_c/e that $\overline{(v-\bar{v})^2} = (kT_c/mv_0)^2$, where $mv_0^2 = 2eV_0$. V_0 is the potential at the plane of interception with respect to the cathode.

The independence of the new fluctuation components may be established by calculation of $\overline{v_{ac}I_{ac}}$ after interception. Taking the product of Equations (4) and (7),

$$v_{2ac}I_{2ac} = \frac{e}{\Delta t} \frac{\sum_m (\delta_m k_0 + n_{0m}k_{1m})\,]\,[\sum_m \delta_m k_0 (v_m - \bar{v}) + \sum_m n_{0m}k_{1m}(v_m - \bar{v})\,]}{\sum_m k_0 n_{0m}} \tag{10}$$

Since the δ_m and k_{1m} are independent of one another, the only terms which contribute to the correlation are

$$\frac{e}{\Delta t}\left[\frac{(\sum_m \delta_m k_0)\,[\sum_m \delta_m k_0\,(v_m-\bar{v})\,] + (\sum_m n_{0m}k_{1m})\,[\sum_m n_{0m}k_{1m}\,(v_m-\bar{v})\,]}{\sum_m k_0 n_{0m}}\right].$$

The first term is simply $k_0 v_{1ac} I_{1ac}$, and the second may be shown to reduce the zero, since

$$\sum_m n_{0m}k_{1m} \sum_p n_{0p}k_{1p}\,(v_p - \bar{v}) = \sum_m (n_{0m}k_{1m})^2\,(v_m - \bar{v}).$$

The mean-square value of $n_{0m}k_{1m}$ is proportional to n_{0m}, and

$$\sum_m n_{0m}\,(v_m - \bar{v}) = 0.$$

The cross-correlation function after interception is simply k_0 times the value before interception, which establishes the independence of the newly introduced current and velocity fluctuations.

In final terms

$$\overline{I_{2ac}^2} = k_0^2 \overline{I_{1ac}^2} + 2eI_0\,(k_0 - k_0^2)\,\Delta f;$$
$$\overline{v_{2ac}^2} = \overline{v_{1ac}^2} + \frac{2e\Delta f}{I_0}\left(\frac{1-k_0}{k_0}\right)\left(\frac{kT_c}{mv_0}\right)^2. \tag{11}$$

[8] F. N. H. Robinson, "Space-Charge Smoothing of Microwave Shot Noise in Electron Beams," *Philosophical Magazine*, Vol. 43, pp. 51-62, January, 1952.

Extension of Theory to a Finite Beam, with Nonuniform Interception Probability

To attack the problem of noise in a beam of extended cross section, having different interception probabilities at each point on the cross section, it is useful to consider it as a multitude of small beams, each of which can be described as having a fixed k_0. If the average transmission for that beam element characterized by subscript α be $k_{0\alpha}$, and the δ, I_0, n, etc., are likewise characterized by α, the total (summed) a-c current in the composite beam after interception is

$$I_{2ac} = \frac{e}{\Delta t} \sum_{\alpha} \sum_{m} (\delta_{m\alpha} k_{0\alpha} + n_{0m\alpha} k_{1m\alpha}) . \tag{12}$$

It will be observed that the part of I_{2ac} representing the modified initial current fluctuations over the beam is just

$$\frac{e}{\Delta t} \sum_{\alpha} \sum_{m} \delta_{m\alpha} k_{0\alpha} .$$

This is simply the sum of the initial fluctuation currents, reduced by the transmission factor.

In dealing with the space-charge-wave modes in an electron beam, Rowe[9] established the statistical independence of the various noise wave modes at the cathode plane. Were the mode functions identical at all points along the axis of flow, the modes would continue to be independent. In actuality, the mode functions change shape as beam velocity, beam diameter, and diameter of surrounding walls vary along the axis. This results in an interchange between the modes which makes them slightly correlated to one another. Were this not so, it would be theoretically possible to consider separately the modes and arrive at an expression for the modification of noise in terms of the initial mode current distribution. It is perhaps simplest to express the first part of Equation (12) as $I_{1\ \text{modified}}$; exactly how the various noise wave modes will be modified after interception can only be calculated if all mode coefficients are initially known. In the case of interception by a mesh grid whose holes are relatively finely spaced, say 1/10 the beam diameter or less, the various mode currents are modified approximately in the ratio of transmitted current to incident current, or

[9] H. E. Rowe, "Shot Noise in Electron Beams at Microwave Frequencies," Sc.D. Dissertation, Massachusetts Institute of Technology (1952), pp. 161-167; Unpublished.

$$I_{1\ \text{modified}} = I_{1ac} \times \frac{\text{Average Current Transmitted}}{\text{Average Incident Current}}.$$

The partition noise current, $(e/\Delta t) \sum_{m,\alpha} n_{0ma} k_{1ma}$, is independent of the velocity class m, so that if $n_{0ma} = n_{0m}$ (the average current in each element the same), the partition current squared is the sum of the partition currents squared in each element α. Hence

$$\overline{I_{2ac}^2} = \overline{I_1^2{}_{\text{modified}}} + \sum_{\alpha} (k_{0\alpha} - k_{0\alpha}^2)\, 2e\Delta f I_{0\alpha}\ ; \tag{13}$$

in a similar manner, making the same assumptions,

$$\overline{v_{2ac}^2} = \overline{v_1^2}_{\text{modified}} + \frac{\sum_{\alpha} (k_{0\alpha} - k_{0\alpha}^2)\, I_{0\alpha} 2e\Delta f \left(\dfrac{kT_c}{mv_0}\right)^2}{\left(\sum_{\alpha} k_{0\alpha} I_{0\alpha}\right)^2}\ . \tag{14}^*$$

These two expression can be put into integral form, in which $j_0 dA$ represents $I_{0\alpha}$.

$$\overline{I_{2ac}^2} = \overline{I_1^2}_{\text{modified}} + \left\{ 2e\Delta f \int (k_0 - k^2{}_0)\, j_0 dA \right\} \tag{15}$$

$$\overline{v_{2ac}^2} = \overline{v_1^2}_{\text{modified}} + \left\{ 2\Delta f \left(\frac{kT_c}{mv_0}\right)^2 \frac{e \int (k_0 - k^2{}_0)\, j_0 dA}{\left(\int k_0 j_0 dA\right)^2} \right\} \tag{16}$$

where j_0 is the beam current density and k_0 the probability of electrons of a given point in the beam's cross section to be transmitted through

---

* The quantity v_1 modified is, as I_1 modified, strictly dependent on the mode situation in the beam. The calculated partition components of Equations (13) and (14) are uniformly weighted over the beam cross section, and thus do not exactly represent the fundamental mode excitation, since even the fundamental space-charge-wave functions vary over the cross section. At low interception, where $I_{\text{modified}} = I_{1ac} \times \dfrac{\text{Transmitted current}}{\text{Initial current}}$, $v_{1\ \text{modified}} = v_{1ac}$ may be shown. This again is the result of the assumption $\overline{\delta_{m\alpha}\delta_{p\beta}} = \overline{\delta_m \delta_p}$.

the plane of interception. The bracketed terms in Equations (15) and (16) are the partition-noise components. The integral is over the beam cross section. The integrals can be normalized:

$$\Omega_1 = \frac{\int k_0 j_0 dA}{\int j_0 dA}\ ;\quad \Omega_2 = \frac{\int (k_0 - k^2_0)\, j_0 dA}{\int j_0 dA}\ ;\quad I_0 = \int j_0 dA. \tag{17}$$

Ω_1 can be recognized as the transmission factor for the entire beam, analogous to k_0 of the uniform probability case. Ω_2 is likewise analogous to $k_0(1-k_0)$.

Then,
$$\overline{I_{2ac}{}^2} = \overline{I_1{}^2}_{\text{modified}} + 2eI_0\Delta f\Omega_2\ ;$$
$$\overline{v_{2ac}{}^2} = \overline{v_1{}^2}_{\text{modified}} + \frac{2e\Delta f}{I_0}\left(\frac{kT_c}{mv_0}\right)^2 \frac{\Omega_2}{\Omega_1{}^2}. \tag{18}$$

When interception is at a mesh electrode, the aforementioned approximation gives:

$$\overline{I_{2ac}{}^2} = \Omega_1{}^2\,\overline{I_{1ac}{}^2} + 2eI_0\Delta f\Omega_2\ ,$$
$$\overline{v_{2ac}{}^2} = \overline{v_{1ac}{}^2} + \frac{2e\Delta f}{I_0}\left(\frac{kT_c}{mv_0}\right)^2 \frac{\Omega_2}{\Omega_1{}^2}. \tag{18a}$$

The results of this analysis should not be too casually compared with experimental data, for in the usual noise measurements, the probe (a cavity or helix) is not equally sensitive to currents at every point in an extended beam. This calculation of partition current does not take that fact into account, and also does not consider the fact that partition noise current excites higher order space-charge waves, which may also be picked up by the probe, and result in somewhat confusing measurements of noise current versus distance from the plane of interception.

Noisiness and Noise Standing-Wave Ratio after Interception (Interception at Ideal Grid or High-Transmission Aperture)

When noise characteristics of a beam are determined by one space-charge-wave mode, it has been shown by Haus and Robinson[4] that the noise waves possess two invariants, S and Π where

$$S^2 = \overline{V^2}\,\overline{I^2} - \overline{[\mathrm{Im}(VI^*)]}^2;$$
$$\Pi = \mathrm{Re}(\overline{VI^*}), \tag{19}$$

where Re denotes the real part, Im denotes the imaginary part, I^* is the complex conjugate of I, and $\overline{V^2}$ is defined as $(v_0^2 m^2/e^2)\ \overline{v^2}$. $\overline{\mathrm{Im}(VI^*)}$ and $\overline{\mathrm{Re}(VI^*)}$ measure the correlation between velocity (voltage) and current fluctuations; $\overline{\mathrm{Re}\ (VI^*)}$ is zero unless the fluctuations are correlated at the potential minimum; and correlation is thought by some writers,[2, 10] to be nonexistent. The noise figure of an amplifier which couples to these space-charge waves of noise has a minimum noise value[4] of the form

$$F = 1 + (S - \Pi) \times (\text{constant}).$$

It is obvious that the application of S or Π limits analysis to a single space-charge wave mode. It is therefore not possible to obtain expressions for S and Π when mode transformations take place. If (1) the interception occurs at a mesh, and (2) the higher order space-charge-wave modes are neglected, *or* the beam is infinite in cross section, Equations (11) or (18a) may be used. Using Ω_1 alternatively to represent k_0 in the ideal grid case, and Ω_2 to represent $k_0(1-k_0)$, an expression for S may be derived, recognizing that

$$\begin{matrix}\mathrm{Re} \\ \mathrm{Im}\end{matrix}\ (VI^*)_{\substack{\text{after} \\ \text{interception}}} = \Omega_1 \begin{matrix}\mathrm{Re} \\ \mathrm{Im}\end{matrix}\ (VI^*)_{\substack{\text{before} \\ \text{interception}}} \dagger$$

$$\Pi_2 = \Omega_1 \Pi_1$$

$$S_2^2 = \{\Omega_1^2 \overline{V_1}^2 \overline{I_1}^2 - \Omega_1^2 [\overline{\mathrm{Im}(V_1 I_1^*)}]^2\} + [\overline{V_1^2}\ \overline{I'^2} + \Omega_1^2\, \overline{I_1^2}\ \overline{V'^2}] + \overline{V'^2\, I'^2}, \tag{20}$$ ‡

where $\overline{V'^2}$ and $\overline{I'^2}$ are the partition fluctuations:

$$\overline{V'^2} = \frac{2\Delta f}{eI_0} (kT_c)^2 \frac{\Omega_2}{\Omega_1^2}; \quad \overline{I'^2} = 2e\Omega_2 I_0 \Delta f. \tag{21}$$

When, as in the present case, only a single mode of space-charge-wave interaction is considered, the mean-square current and "voltage" as a function of axial position measured in plasma-wavelength angular units ϕ always obey relations of the following form, in a drift region:

---

[10] D. A. Watkins, "Noise at the Potential Minimum in the High Frequency Diode," *Jour. Appl. Phys.*, Vol. 26, pp. 622-624, May, 1955.

† When the plane of interception corresponds to a current maximum or minimum of the noise wave, $\overline{\mathrm{Im}(VI^*)} = 0$ and $S = \overline{V^2}\,\overline{I^2}$.

‡ Here, subscripts 1 refer to values prior to interception, and subscripts 2, after interception.

$$\overline{I'^2} = \frac{S}{W} \frac{(1+R\cos 2\phi)}{\sqrt{1-R^2}},$$

$$\overline{V'^2} = SW \frac{(1-R\cos 2\phi)}{\sqrt{1-R^2}}, \tag{22}$$

where $W =$ "beam impedance" $= \dfrac{2V_0}{I_0} p \dfrac{\omega_0}{\omega}$,

$$R = \frac{\eta^2-1}{\eta^2+1},$$

$$\eta = \frac{I_{max}}{I_{min}} = \text{noise current standing wave ratio},$$

$$\omega_0 = \sqrt{\frac{\overline{\rho_0 e}}{m\epsilon}} = \text{plasma frequency},$$

$p =$ plasma frequency reduction factor.[11]

If, now, the substitutions

$$W'^2 = \frac{\overline{V'^2}}{I'^2} \Omega_1^2 = \left(\frac{kT_c}{eI_0}\right)^2, \tag{23}$$

$$S_0^2 = \overline{V'^2 I'^2} \frac{\Omega_1^2}{\Omega_1^2} = (2kT_c\Delta f)^2$$

are made in Equation (20), S_2^2 may be expressed as

$$S_2^2 = \Omega_1^2 S_1^2 + \frac{\Omega_2 S_0 S_1}{\sqrt{1-R_1^2}} \left[\left(\frac{W'}{W_1} + \frac{W_1}{W'}\right) + R_1 \cos 2\phi_1 \left(\frac{W'}{W_1} - \frac{W_1}{W'}\right)\right] + \frac{\Omega_2^2}{\Omega_1^2} S_0^2, \tag{24}$$

where ϕ_1 is the angle between a current maximum of the incident noise wave, and the plane of interception.

Since the $\cos 2\phi_1$ term is the only one which can be negative, minimum S_2 always occurs at $\phi_1 = 0$, or *at the current maximum*, when

[11] D. A. Watkins, "Traveling-Wave Tube Noise Figure," *Proc. I.R.E.*, Vol. 40, pp. 65-70, January, 1952.

$W_1 < W'$ (which is true where the potential of the intercepting plane is greater than $kT_c/e \simeq .1$ volt).

The noisiness after interception at the current maximum becomes

$$S_2^2{}_{\min} = \Omega_1^2 S_1^2 + \Omega_2 S_0 S_1 \left(\frac{W'}{W_1} \eta_1 + \frac{W_1}{W'} \frac{1}{\eta_1} \right) + \frac{\Omega_2^2}{\Omega_1^2} S_0^2. \qquad (25)$$

When interception probability is constant over the beam,

$$\Omega_1 = k_0, \quad \Omega_2 = k_0 (1-k_0),$$

$$S_2^2{}_{\min} = k_0^2 S_1^2 + k_0 (1-k_0) S_0 S_1 \left(\frac{W'\eta}{W_1} + \frac{W_1}{\eta W'} \right) + (1-k_0)^2 S_0^2. \qquad (26)$$

Following the assumptions of Bloom and Peter,[2]

$$S_1^2 = 4 (4-\pi) (kT_c \Delta f)^2;$$

Equation (26) reduces to the absolute minimum

$$\frac{S_{2\min}}{S_1} = \frac{(1-.074k_0)}{.925}, \quad \text{when} \frac{W'\eta_1}{W_1} = 1. \qquad (27)$$

This interesting result indicates that noisiness of such an intercepted beam does not need to be much higher than that of a beam from a cathode. If II were appreciable, minimum noise figure, which is proportional to S-II, might be changed considerably by interception. Realizing the condition $W'\eta/W_1 = 1$ requires a very high standing-wave-ratio before the interception plane, which is neither readily attainable, except at low potentials, nor generally desirable in a low-noise tube.

In the region after interception, the Equations (22) are again applicable, with subscripts 2 in the appropriate positions. By combining these relations to eliminate $\cos 2\phi_2$,

$$R_2^2 = \left(\frac{\eta_2^2-1}{\eta_2^2+1} \right)^2 = 1 - \frac{S_2^2}{\overline{V_2^2}\,\overline{I_2^2}} \left[1 - \left[\frac{\overline{V_2^2} - W_2^2 \overline{I_2^2}}{\overline{V_2^2} + W_2^2 \overline{I_2^2}} \right]^2 \right], \qquad (28)$$

$$\cos 2\phi_2 = -\frac{1}{R_2} \left[\frac{\overline{V_2^2} - W_2^2 \overline{I_2^2}}{\overline{V_2^2} + W_2^2 \overline{I_2^2}} \right], \qquad (29)$$

which when solved for ϕ_2 gives the position of the standing waves at the place of interception with respect to the current maximum.

The expressions for S_2, R_2, and $\cos 2\phi_2$ must not be applied indiscriminately, since they apply only to mesh interception, and to a single mode of space-charge-wave interaction.

Interception at a Round Aperture

Undoubtedly the most interesting interception problem from a practical point of view is the interception of part of a solid beam of circular cross-section by a round aperture. In this type of interception, as in interception by an ideal grid, random processes play a certain part, as well as processes which are completely predetermined. In handling this problem, if the statistical processes can be separated from the fixed ones, Equations (15) and (16) may be called upon to derive the partition current and velocity.

The source of any random interception in the beam must lie in those radial thermal velocities which are uncorrelated to the longitudinal motion of the electrons.* These velocities are controlled by the axial focusing magnetic field, so that at infinite field no radial velocity remains, and interception involves no statistical "random" process which is independent of the axial fluctuations.

The model to be used in calculations of Ω_1 and Ω_2 is derived from the following reasoning: All of the electrons emitted from a particular point on the cathode will move in circular trajectories in the transverse plane and (in ideal geometries and in the absence of space charge) pass through a particular point in the beam once each cyclotron wavelength. This picture is much too idealized, so we feel safe to assume that after transit through a practical electron gun, in the presence of space charge, the electrons are fairly well mixed, and do not image the cathode.

For simplicity, the beam will be assumed to be of (nominal) radius r_b, such that for every point P on the cross section, $0 \leq r \leq r_b$, $0 \leq \theta \leq 2\pi$, there are electrons which move in the transverse plane in circular paths having P as center. The radii of the circular trajectories are distributed in a way determined by the distribution of transverse velocities at the cathode (i.e., the path radius being $\rho = mv_T/(Be)$, the number of electrons between radii ρ and $\rho + d\rho$ is the same as the number of electrons with total transverse velocity between v_T and $v_T + dv_T$, where $\partial\rho/\partial v_T = m/(Be)$. In this model in confined flow, the

* In a later section, effects of correlated components of radial and axial fluctuations will be discussed.

net effect of each electron on the longitudinal space charge wave interaction is the same as if it were located at P, instead of circling P. This is the most important concept in the theory, as it provides the key to assigning statistical independence to the interception process. The model is, of course, most accurate if $\omega_{\text{cyclotron}} \gg \omega_{\text{plasma}}$. The interception of electrons "belonging to" point P is assumed to be random; the average transmission can be found by integrating the probability distribution around P over those regions which are not covered by the anode aperture. In this fashion, a $k_0(r)$ can be found, i.e., a value of interception probability, for each point $0 \leq r \leq r_b$. Axial symmetry is assumed, for simplicity. From the value of $k_0(r)$, the partition current and velocity can be calculated by integration over the cross-section $0 \leq r \leq r_b$.

From elementary statistics, the number of electrons emitted, out of a total of N_0, with transverse velocity between v_T and $v_T + dv_T$ is

$$dN = C_1 \exp\left[-\frac{mv_T^2}{2kT_c}\right] v_T dv_T , \tag{30}$$

$$N_0 = \int_{v_T^2=0}^{\infty} dN = C_1 \int_0^{\infty} \exp\left[-\frac{mv_T^2}{2kT_c}\right] v_T dv_T ,$$

$$\therefore C_1 = \frac{mN_0}{kT_c} ,$$

$$dN = \frac{mN_0}{kT_c} \exp\left[-\frac{mv_T^2}{2kT_c}\right] v_T dv_T . \tag{31}$$

Now, since the radius is given, from elementary considerations, by $\rho = mv_T/(Be)$, where B is the axial magnetic field,

$$dN = \frac{mN_0}{kT_c} \exp\left[-\frac{(Be\rho)^2}{2mkT_c}\right] \left(\frac{Be}{m}\right)^2 \rho d\rho .$$

In terms of cartesian coordination, $\rho^2 = x^2 + y^2$.

$$dN = \frac{mN_0}{kT_c} \exp\left[-\frac{(Be)^2}{2mkT_c}(x^2+y^2)\right] \left(\frac{Be}{m}\right)^2 \frac{dxdy}{2\pi} . \tag{32}$$

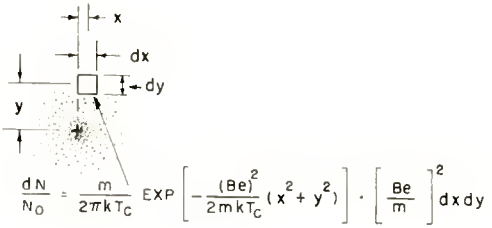
Figure 1a illustrates this probability distribution. Consider this distribution located about a point P at radius r, from the axis of the beam,

with the edge of the aperture at radius r_a. For all *significant* radii r, the edge of the aperture may be considered to be a straight line (Figure 1b). The fraction of the N_0 electrons *not* intercepted is

$$\frac{N_T}{N_0} = \frac{m}{kT_c}\left(\frac{Be}{m}\right)^2 \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{(r_a-r)} \exp\left[-\frac{(Be)^2}{2mkT_c}(x^2+y^2)\right]dxdy,$$

which reduces to

$$k_0(r) = \frac{N_T}{N_0} = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\frac{Be}{m}\sqrt{\frac{m}{2kT_c}}(r_a-r)} e^{-u^2}du,$$



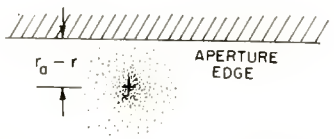


Fig. 1—Model assumed for interception at aperture edge.

where $u = \frac{Be}{m}\sqrt{\frac{m}{2kT_c}}\,r$.

By the definition of the error function,

$$\text{erf } y = \frac{2}{\sqrt{\pi}}\int_0^y e^{-u^2}du,$$

$$k_0(r) = \frac{1}{2}\left(1+\text{erf}\left[\frac{Be}{\sqrt{2mkT_c}}(r_a-r)\right]\right). \qquad (33)$$

With this value of $k_0(r)$, Ω_1 and Ω_2 can be computed from Equation

(17). While the theory, within its limitations, allows j_0 to vary in any desired manner, for computation here it will be assumed that j_0 equals a constant $(0 \leq r \leq r_b)$, and $j_0 = 0$, $(r > r_b)$. The actual beam current will extend beyond r_b, and Ω_1 is the actual transmission coefficient (which is *not* unity for $r_b < r_a$).

In practical beams, the radii of paths taken by electrons about their centers (P) may be somewhat greater than given by the thermal velocities alone, because of perturbations in focusing, which tend to "wash out" after some distance of travel because of initial differences of longitudinal velocity. We shall not attempt to explain quantitatively effects due to this additional "smearing out" of the cross section of the beam.

When $k_0(r)$ is substituted into Equation (17), with j_0 constant as specified, we obtain

$$\Omega_1 = \int_0^1 (1+\text{erf}\,[\alpha r_a - \alpha u r_b])\, u\,du,$$

$$\Omega_2 = \frac{1}{2}\int_0^1 (1+\text{erf}\,[\alpha r_a - \alpha u r_b])\,(1-\text{erf}\,[\alpha r_a - \alpha u r_b])\, u\,du, \tag{34}$$

where $$\alpha = B\sqrt{\frac{e}{m}}\left(\frac{kT_c}{e}\right)^{-\frac{1}{2}} = 4.52 \times 10^3\, B_{\text{gauss}} T_c^{-\frac{1}{2}}\ (\text{meters}^{-1}).$$

In Figures 2 through 7, Ω_1 and Ω_2 are plotted versus $\alpha(r_b - r_a)$, for a number of values of αr_b. It will be observed that for a certain Ω_1, Ω_2 is a decreasing function of αr_b, indicating that the partition components, proportional to Ω_2, decrease with increasing magnetic field. It is also interesting that, for a given αr_b, the value of Ω_2 first increases as interception is increased, then slowly decreases; the reason for this is that, once enough of the beam is intercepted, the "region of uncertainty" reaches maximum area, and on further interception, the region of uncertainty can only decrease, as the perimeter of the intercepting aperture becomes smaller. The correctness of this result has been verified experimentally, as described later. Examination of Equation (17) shows that the greatest contribution to partition noise is in those regions where k_0 is near 1/2. In strong magnetic fields, these regions are indeed small in area, which explains clearly the behavior of Ω_2. In Figure 3, there is also plotted $\Omega_1(1-\Omega_1)$, which would be the value of Ω_2 if the assumption of uniform interception probability were made.

This curve, compared with actual Ω_2, illustrates graphically the error which can be made in this assumption. Another point of interest is that the actual value of Ω_2 approaches $\Omega_1(1-\Omega_1)$ for sufficiently small interception, even in this aperture case.

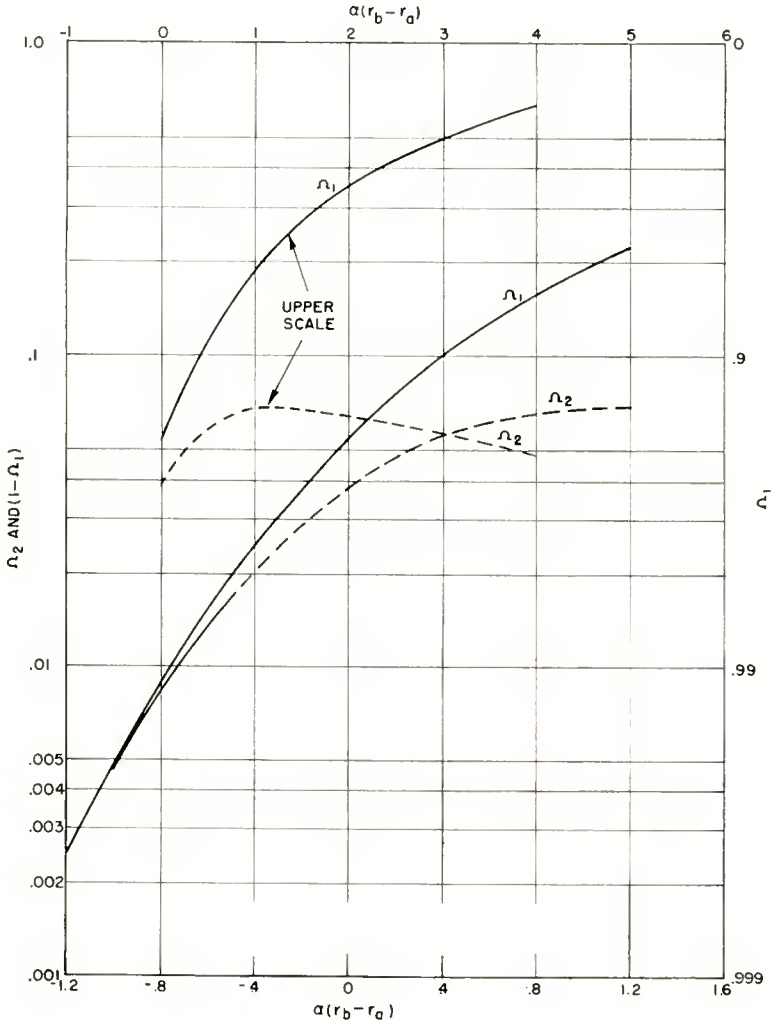


Fig. 2—Interception noise parameters versus differential radius ($\alpha r_b = 10$).

In many practical beams, the current density as a function of radius will *not* vary as assumed. This would indicate need for a different choice of $j_0(r)$. This is beyond the scope of the paper, but is not necessary for understanding of the processes taking place.

The extension of the probability distribution of Equation (32) to the problem of interception at a mesh grid of finite wire size, for example, while not particularly simple, is direct. For each point P within the limits of one grid module, a value k must be calculated;

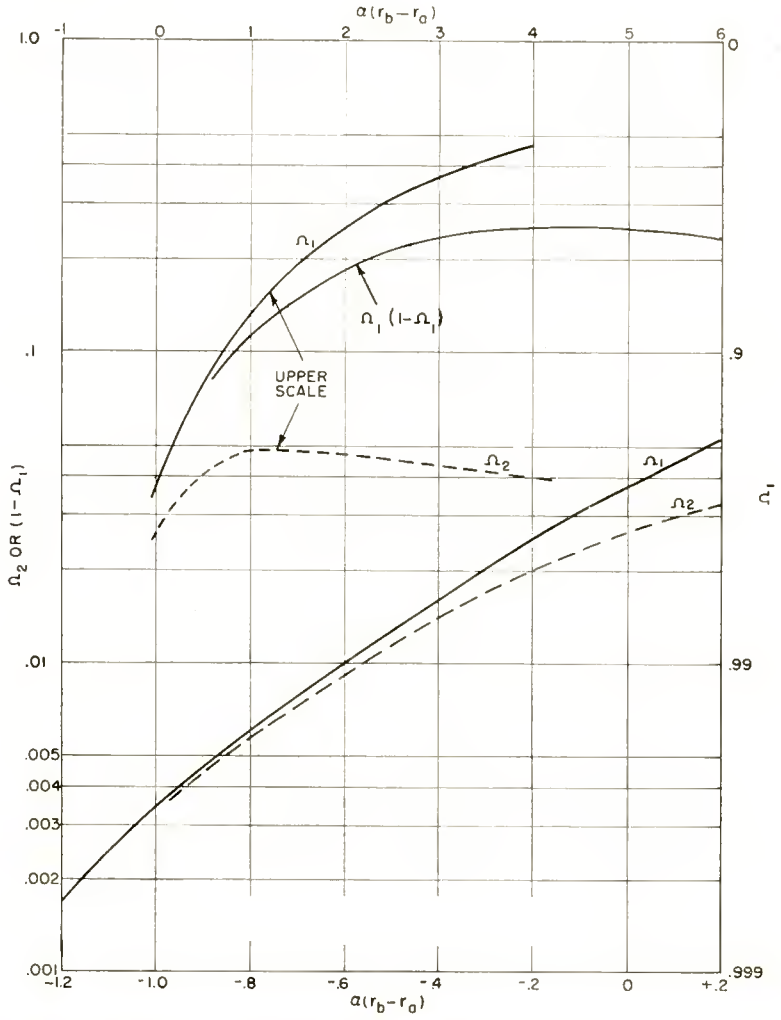


Fig. 3—Interception noise parameters versus differential radius ($\alpha r_b = 15$).

then k_0 and $k_0(1-k_0)$ may be integrated to determine Ω_1 and Ω_2. The mesh may be so fine that k_0 is essentially constant over the grid module; the partition noise will approach the North[1] value, if the smaller dimension of the module, let us call it h, satisfies the relation

$\alpha h \ll 1$. This means simply that the electron distribution about each point must extend over a number of meshes.

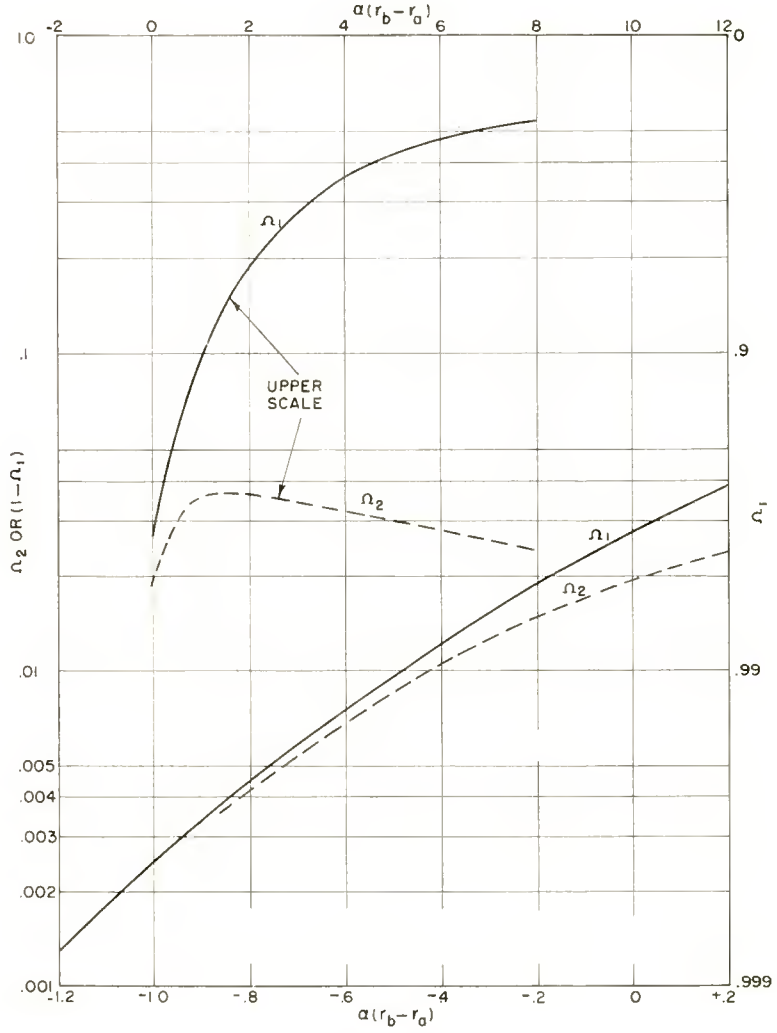


Fig. 4—Interception noise parameters versus differential radius ($\alpha r_b = 20$).

Interception Due to Transverse Motion of the Bunched Beam

Any electron beam with space charge density modulation may be expected to "bulge" in diameter at places where the charge density is greater than average, and shrink in diameter at points of smaller than

average charge density. The actual picture is not quite so simple, because of the many modes of space-charge interaction which can be supported by a beam in a finite magnetic field. Nonetheless, any such radial expansion and contraction correlated with the longitudinal space-charge-

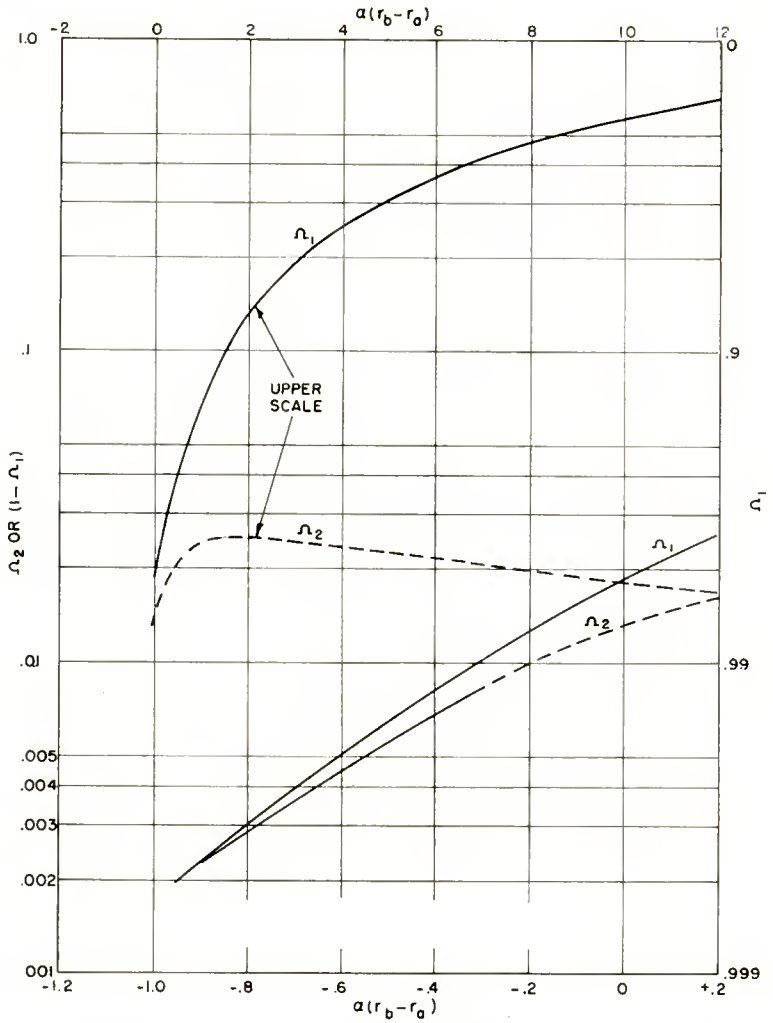


Fig. 5—Interception noise parameters versus differential radius ($\alpha r_b = 30$).

waves may cause unusual effects if part of the beam is intercepted, for since the beam swells in size when the total current is going through a maximum value, the interception will increase, reducing the fluctuation current after interception.

To gain some idea of the possible order of magnitude of this effect, it is necessary to determine approximately how much increase of beam diameter will accompany a given level of fluctuation current.

Hahn[12] has derived an expression for the variational part of the beam radius, r_1, where $r = r_0 + r_1$. It is, however, difficult using his expressions to arrive at a simple expression for r_1 in terms of I_1, the beam perturbation due to axial "bunching". One may calculate the increase of equilibrium radius[13] (r_{eq}) when the average current in the beam is increased by the value of the fluctuations. The equilibrium radius of a beam in parallel combined flow is given by

$$r_{eq}^2 = \frac{r_c^2}{2}\left[\left(\frac{B_0}{B_z}\right)^2 + \sqrt{\left(\frac{B_0}{B_z}\right)^4 + 4}\,\right]. \tag{35}$$

r_c is the cathode radius, B_z the focusing field, and B_0 the field required to focus the same diameter, beam of the same current in shielded (Brillouin) flow.

For cases of practical confined flow, $\left(\frac{B_0}{B_z}\right)^4 \ll 4$, and

$$r_{eq} \simeq r_c \sqrt{1 + \frac{1}{2}\left(\frac{B_0}{B_z}\right)^2}. \tag{36}$$

Since B_0 is proportional to beam current I_0,

$$2r_{cq}dr_{cq} = \frac{r_c^2}{2B_z^2} \times 2B_0 dB_0 = \left(\frac{B_0}{B_z}\right)^2 \frac{dI}{I_0} r_c^2.$$

Since $r_{cq} \simeq r_c$,

$$\frac{dr_{eq}}{r_{eq}} \simeq \frac{1}{2}\left(\frac{B_0}{B_z}\right)^2 \frac{dI}{I_0}. \tag{37}$$

(dI represents the fluctuation in current.)

If the current density of the beam at the edge of the aperture is J_{AP},

[12] W. C. Hahn, "Small Signal Theory of Velocity-Modulated Electron Beams," *General Electric Review*, Vol. 42, pp. 258-270, June, 1939.

[13] C. C. Wang, "Electron Beams in Axially Symmetrical Electric and Magnetic Fields," *Proc. I.R.E.*, Vol. 38, pp. 135-147, February, 1950.

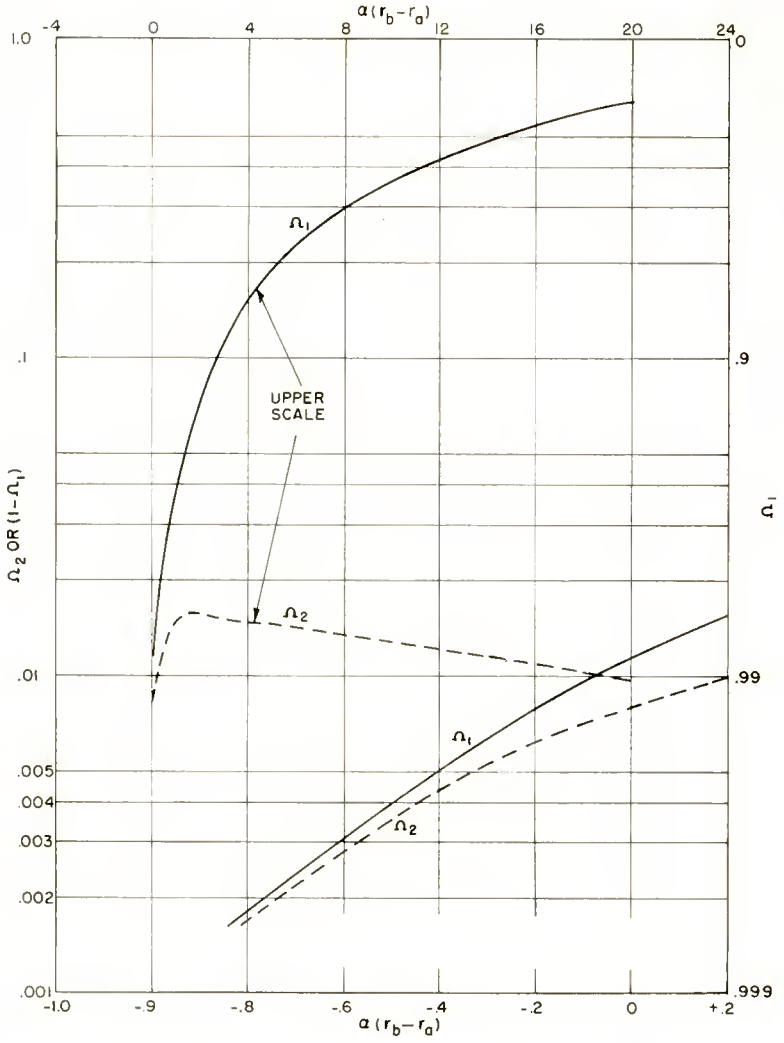


Fig. 6—Interception noise parameters versus differential radius ($\alpha r_b = 50$).

an increase dr in beam diameter alone will cause interception current to increase by an amount

$$dI_{\text{int}} = J_{AP}(2\pi r_a)\,dr,$$

where dr is measured at $r = r_a$. Now

$$\left.\frac{dr}{r}\right|_{r_a} = \left.\frac{dr}{r}\right|_{r_{eq}} \tag{38}$$

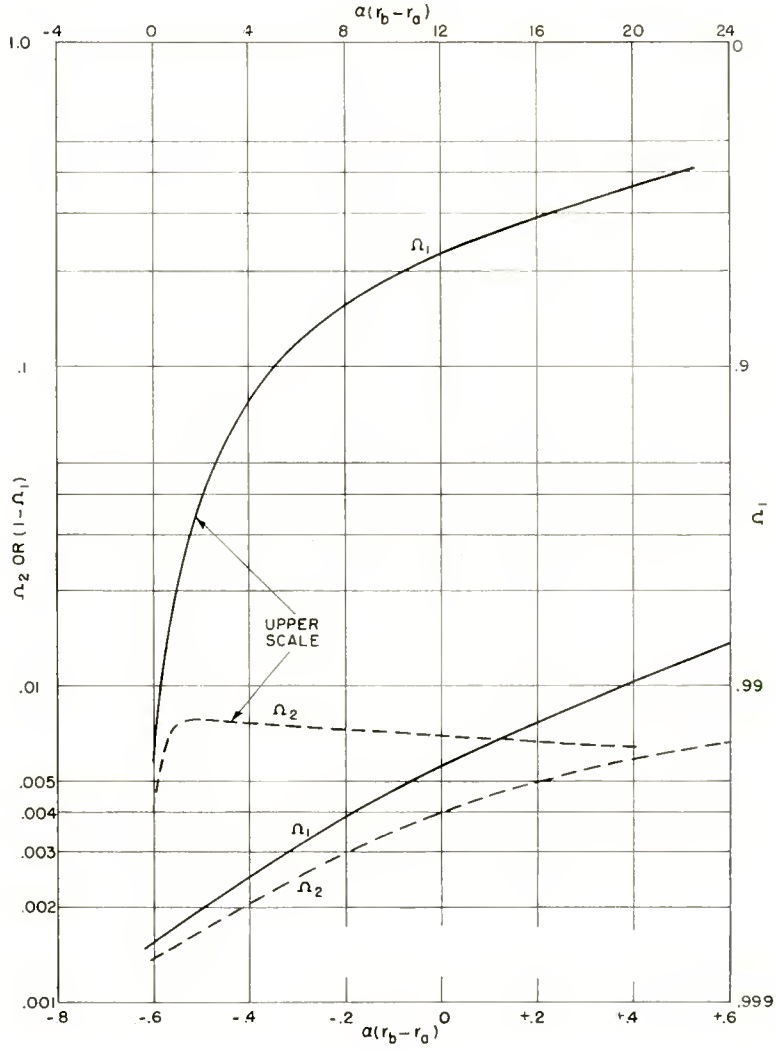


Fig. 7—Interception noise parameters versus differential radius ($\alpha r_b = 100$).

so that, obtaining dI from Equation (37)

$$\left.\frac{dI_{\text{int}}}{dI}\right|_{\substack{\text{due to radial}\\ \text{expansion}}} = \frac{J_{AP}}{J_{AVG}}\,\frac{r_a^2}{r_{eq}^2}\left(\frac{B_0}{B_z}\right)^2, \tag{39}$$

(r_{eq} may be identified as r_b)

where J_{AVG} is the average current density over the beam radius r_b. In many low-noise amplifiers, B_0/B_z is of the order of .1 or .2, giving an upper limit to dI_{int}/dI of about .01 or .04. This represents possible reduction of noisiness of a few tenths of a decibel, but of course there is always an increase due to partition.

In Brillouin flow conditions, effects will be larger, but then so will the partition current, since radial excursions of the electrons will be greater. It appears unlikely that improvement in noise figure can be gained in this manner.

Significance and Experimental Verification

The significance of the analytical results derived in this investigation may be summed up as follows:

1. Interception of an electron beam by a random process introduces both a current and a velocity fluctuation component, which are uncorrelated to one another or to the fluctuations originally in the beam.

2. The magnitude of the partition noise components depends upon the extent of the areas in which electrons have almost equal a priori chances of interception or transmission. The partition noise approaches the North[1] value when interception probability is small at all points (as when only the outermost parts of the beam are intercepted by an aperture, for example) or nearly the same at all points.

3. The partition noise of a beam focused by an axial field is approximately inversely dependent upon the field strength. At the limit of infinite field, there is no partition noise component.

4. While the effects of interception or noisiness are always deleterious, the degradation is not as serious as may be predicted from use of the North partition noise formula.

In order to verify the significance of this theory, an experiment has been carried out in a demountable noise-measuring apparatus. The noise current in the beam was measured by a cavity located immediately following an apertured disk. The noise current in the partially intercepted beam was measured as a function of magnetic field strength and interception. The results are illustrated in Figure 8.

The points representing experimental data are seen to fall near the curve showing the theoretical variation of partition noise with magnetic field, except at low fields, where natural changes of beam diameter cause interception to increase appreciably. The theoretical curve, representing the partition component of noise current (which should be the only field-dependent part) has been shifted vertically until it fits best the experimental data. The line marked "theoretical asymptote at $B = \infty$" thus represents the modified original noise current.

Conclusions

Partial interception of an electron beam has been found to give rise to two new independent fluctuation sources, a velocity fluctuation and a current fluctuation. The magnitude of these components depends upon the degree of uncertainty of interception. In the case of an electron beam in confined flow, the uncertainty is produced by the transverse

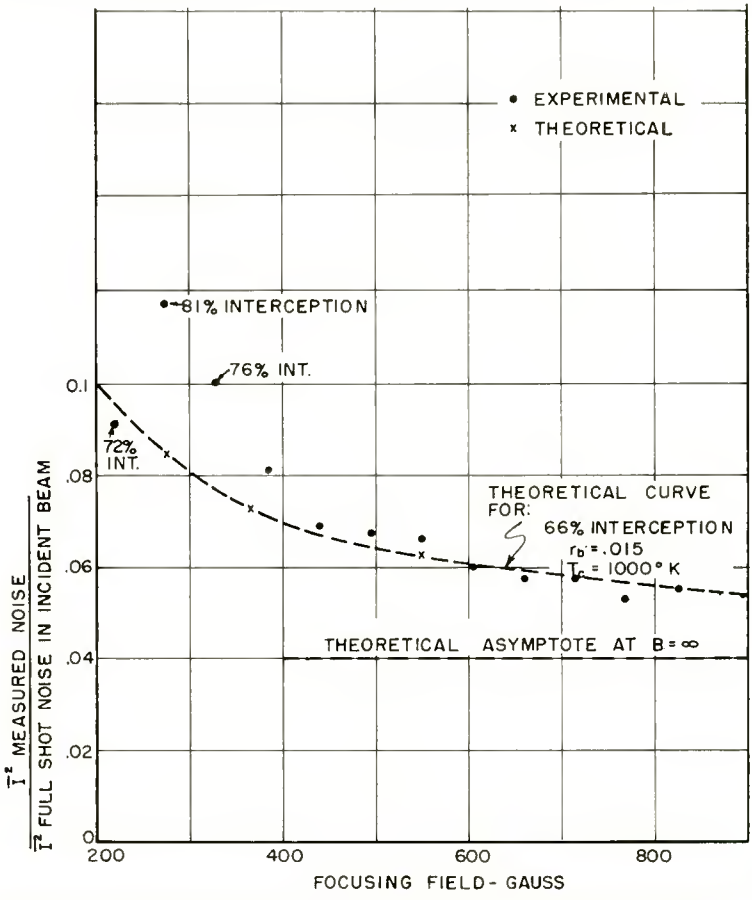


Fig. 8—Noise current versus magnetic field after .020-inch-diameter aperture.

motion of the electrons uncorrelated with the axial space-charge waves. A theory has been derived, in which the partition noise components are determined from integrals over the cross section of the beam. Values of partition noise are graphed for the case of interception by a round aperture. The theory has been qualitatively verified by experiment.

In applying this theory to problem of large interception by apertures, it is not possible to predict the total noise after interception, because of the unknown effects due to higher order space-charge wave interaction.

Appendix I — Mean-Square Values of Random Functions

Let us suppose that there exists a random series (which may be a voltage, current, velocity, etc. function) which has an average value h_q in the time interval Δt_q. For convenience, all Δt_q are assumed to be of equal duration. Such a function might be the random partition fluctuation current of Equation (4) or the velocity fluctuation of Equation (6), which are functions known to have no prediscrimination of frequency. (This is not true of the space-charge-wave current or velocity fluctuations in the incident beam.)

Using a trick familiar in Fourier theory, we may imagine the time series to be periodic, of period $\tau \gg \Delta t$. The Fourier spectrum of a pulse of average height h_q and width Δt can be expressed as

$$H = \sum_{l=1}^{\infty} C_l \cos\left(\frac{2\pi l t}{\tau} + \theta_l\right). \tag{40}$$

The coefficient $C_l = 2h_q\Delta t/\tau$, and θ_l depends on the location of the pulse. The mean-square value of one component of the "signal" is $\frac{1}{2}\ (2h_q\Delta t/\tau)^2$. The number of frequency components, each of the same mean square, within a band of width Δf cycles, is $\Delta t/\tau$.

The total mean-square signal within the band is obtained by adding up $\Delta t = \tau\Delta f$ terms of magnitude $2(h_q\Delta t/\tau)^2$ each, for a total of $2\Delta f(h_q\Delta t)^2/\tau$. This is the mean-squared signal of a single pulse. Actually, in the period τ, there are many pulses of each height, specifically: $\tau/\Delta\tau \times$ (probability of given height). Because of the random phases of the pulses, mean-squared values may be added.

For heights $h_q = R$, let the probability be p_R. Then, the total mean-squared signal will be

$$2\Delta f \sum_R \left(\frac{R^2\Delta t^2}{\tau} \cdot p_R \cdot \frac{\tau}{\Delta t}\right) = 2\Delta f \sum_R R^2 p_R \Delta t, \tag{41}$$

where the sum is over all *possible* values of R.

Example: If the number h_q represents, for example,

$$I' = \sum_m \frac{(n_{0m}k_{im})\,e}{\Delta t},$$

(the interception current fluctuations), then the mean-square a-c value will be

$$I'^2 = 2\Delta f \sum_{R=1}^{n_n} \left\{ \frac{e^2}{\Delta t^2} R^2 \Delta t \,[\text{probability that } \sum_m n_{0m}k_{im} = R] \right\}; \qquad (42)$$

a more formalized method of expression is

$$I'^2 = \frac{2e^2\Delta f}{\Delta t} \overline{(\sum_m n_{0m}k_{im})^2}, \qquad (43)$$

where the barred quantity is called (in mathematical probability) the "expectation" or "average value" of $(\sum_m n_{0m}k_{im})^2$.† It can be shown that since the numbers k_{im} are statistically independent,

$$\overline{(\sum_m n_{0m}k_{im})^2} = \sum_m \overline{(n_{0m}k_{im})^2}.$$

The quantity $\overline{(n_{0m}k_{im})^2}$ is the average value of the square of the deviation from average interception $(k_0 n_{0m})$, and can be expressed as

$$(n_{0m}k_{im})^2 = \sum_{p=1}^{n_{0m}} k_0{}^p (1-k_0)^{n_{0m}-p} \frac{n_{0m}!}{p!\,(n_{0m}-p)!} [p-k_0 n_{0m}]^2.$$

Uspensky[14] shows, by evaluating this expression, that

$$\sum_m \overline{(n_{0m}k_{im})^2} = \sum_m n_{0m}(k_0-k_0{}^2). \qquad (44)$$

† The average implied is not (necessarily) a time average, but an average over-all possible values the function may assume.

[14] J. V. Uspensky, *Introduction to Mathematical Probability*, McGraw-Hill Book Company, New York, N. Y., p. 116.

Therefore,

$$I'^2 = \frac{2e^2\Delta f}{\Delta t}(k_0 - k_0^2)\sum_m n_{0m},$$

but since

$$\frac{e}{\Delta t}\sum_m n_{0m} = I_0,$$

then

$$I'^2 = 2eI_0\Delta f(k_0 - k_0)^2. \tag{45}$$

Since the result does not depend on the number of velocity classes, or the number in each class, it may be applied to the practical problem, wherein the velocity classes are not discrete.

Appendix II — Accuracy of Linearization Assumption

The question of the accuracy of the assumption that the cross-product terms caused by the random interception of the initial fluctuation current may be answered satisfactorily by determining the magnitude of the fluctuations introduced as the result of random interception of the a-c quantity δ.

$$I_{\substack{\text{2nd order}\\ \text{fluctuation}}} = \frac{e}{\Delta t}\sum_m \delta_m k_{im}. \tag{46}$$

For simplicity, let us assume that δ_m represents a random quantity (we neglect the auto-correlation introduced by space-charge interaction). In carrying out the procedure described in Appendix I, it is necessary to calculate the quantity

$$I^2_{\text{2nd order}} = \frac{2e^2\Delta f}{\Delta t}\overline{(\delta k_0 - g)^2}$$

in which g represents the number of electrons (out of the δ) transmitted. The average must be taken over all g, *as well as all* δ. In the symbolic form, let $P(j)$ be the probability that $\delta = j$, and let $Q(g,j)$ be the probability that g electrons are transmitted. Then

$$\overline{I^2}_{\text{2nd order}} = \frac{2e^2\Delta f}{\Delta t}\sum_{\delta=j=-n_0}^{\infty}\sum_{g=0}^{jk_0}(jk_0 - g)^2 P(j)Q(g,j),$$

which may be rewritten

$$\overline{I^2}_{\text{2nd order}} = \frac{2e^2\Delta f}{\Delta t} \sum_j P(j) \left[\sum_g (jk_0 - g)^2 Q(g,j) \right];$$

the bracketed part of the expression may be recognized as the same expression found in Equation (44).

$$\overline{I^2}_{\text{2nd order}} = \frac{2e^2\Delta f}{\Delta t} \sum_j P(j) j(k_0 - k_0^2) = \frac{2e^2\Delta f}{\Delta t} k_0(1 - k_0) \left[\sum_j jP(j) \right];$$

in this expression, the bracketed quantity represents the average value of δ, which is *zero*. The additional mean-squared current due to the random interception is therefore zero, hence the omission of the cross-product term $k_1\delta$ seems justifiable; even though the δ is not truly random, it can still be demonstrated that there is no second-order mean-square fluctuation current.

TRANSIENT RESPONSE OF DETECTORS IN SYMMETRIC AND ASYMMETRIC SIDEBAND SYSTEMS

By

T. Murakami and R. W. Sonnenfeldt

RCA Victor Television Division,
Camden, N. J.

Summary—*Detector response in symmetric- and asymmetric-sideband systems is analyzed in this paper. It is shown that the limitations of the envelope detector are eliminated when a synchronous detector is used. Unlike the envelope detector, the synchronous detector gives an output linearly related to the original modulation, so long as the transmission system is linear. Curves and equations are given to develop this result in general, and then in particular for the flat, staggered-triple bandpass filter for various modulation levels. The improved transient performance of the synchronous detector is shown, experimentally. A simple form of synchronous detector is analyzed. Detailed conclusions are drawn from the analysis and experimental results.*

Introduction

THE envelope of an amplitude-modulated wave is distorted by asymmetric-sideband transmission. This system of transmission is used in television receivers and causes distortion and loss of picture fidelity.[1-5] The two main effects of asymmetric sideband transmission are: (1) nonlinear envelope distortion, and (2) phase modulation of the carrier. Nonlinear envelope distortion can, however, be eliminated by use of a detector which responds only to the envelope of the in-phase component, or to a linear combination of the envelopes of the in-phase and quadrature components of the modulated carrier.

---

[1] W. J. Poch and D. W. Epstein, "Partial Suppression of one Sideband in Television Reception," *RCA Review*, Vol. 1, pp. 19-35, January, 1937.

[2] S. Goldman, "Television Detail and Selective Sideband Transmission," *Proc. I.R.E.*, Vol. 27, pp. 725-732, November, 1939.

[3] R. D. Kell and G. L. Fredendall, "Selective Sideband Transmission in Television," *RCA Review*, Vol. 4, pp. 425-440, April, 1940.

[4] E. C. Cherry, "The Transmission Characteristics of Asymmetric Sideband Communication Networks," (Part 1) *Jour. I.E.E.*, Vol. 89, Part III, pp. 19-42, March, 1942, and (Part 2) *Jour. I.E.E.*, Vol. 90, Part III, pp. 75-88, June, 1943.

[5] J. Avins, B. Harris and J. S. Hovarth, "Improving the Transient Response of Television Receivers," *Proc. I.R.E.*, Vol. 42, pp. 274-284, January, 1954.

It will be shown that the in-phase and quadrature components of the detected output are linear functions of the depth of modulation for any linear system of carrier-wave transmission.

It is next shown that when a modulated carrier is transmitted through a bandpass system the synchronously detected output is equivalent to transmission of the modulation alone through a low-pass analog filter. A general derivation is given for the transfer function of the low-pass analog filter, which corresponds to any given bandpass filter transfer characteristic.

A detailed study is made of the step-carrier response of a flat-staggered triple-tuned circuit of ten per cent bandwidth. Results are given for envelope detection and for in-phase, quadrature, and off-axis synchronous detection. Experimental data is included to verify the theoretical results.

A practical synchronous detector circuit is analyzed and it is shown how this circuit can be used for in-phase, quadrature and off-axis detection.

Transmission of Amplitude-Modulated Carrier in Linear Systems

To understand what happens when an amplitude-modulated carrier is transmitted through a linear filter, consider the modulated carrier given by

$$F(t) = [1 + mf(t)\ U(t)] \cos \omega_0 t \tag{1}$$

where m is the depth of modulation, $f(t)$ the modulating function, ω_0 the carrier angular frequency, and $U(t)$ is the unit step function. Treating the steady-state and transient portions of the input given by Equation (1) separately, the Fourier transform of the transient part of $F(t)$ is

$$\begin{aligned} G(\omega) &= \int_{-\infty}^{\infty} m f(\lambda)\ U(\lambda) \cos \omega_0 \lambda\ e^{-\omega \lambda}\, d\lambda \\ &= \frac{m}{2} \int_0^{\infty} f(\lambda)\ [e^{i\omega_0 \lambda} + e^{-i\omega_0 \lambda}]\, e^{-i\omega \lambda}\, d\lambda \\ &= \frac{m}{2} \int_0^{\infty} f(\lambda)\ [e^{i(\omega_0 - \omega)\lambda} + e^{-i(\omega_0 + \omega)\lambda}]\, d\lambda. \end{aligned} \tag{2}$$

Let the filter to which the modulated carrier is applied have a transfer

function given by $T(\omega)\ e^{i\theta(\omega)}$. The transient component of the output of the filter is given by[6]

$$R(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} T(\omega)\ G(\omega)\ e^{i[\omega t + \theta(\omega)]}\, d\omega\,. \tag{3}$$

Substituting the value of $G(\omega)$ as given by Equation (2) in Equation (3) gives

$$R(t) = \frac{m}{4\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda)\ e^{i[\omega t + \theta(\omega) + (\omega_0 - \omega)\lambda]}\, d\lambda$$

$$+ \frac{m}{4\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda)\ e^{i[\omega t + \theta(\omega) - (\omega_0 + \omega)\lambda]}\, d\lambda. \tag{4}$$

Let $\omega = -\omega'$ in the second integral of Equation (4), then

$$R(t) = \frac{m}{4\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda)\ e^{i[\omega t + \theta(\omega) + (\omega_0 - \omega)\lambda]}\, d\lambda$$

$$+ \frac{m}{4\pi} \int_{-\infty}^{\infty} T(-\omega')\ d\omega' \int_{0}^{\infty} f(\lambda)\ e^{i[-\omega' t + \theta(-\omega') - (\omega_0 - \omega')\lambda]}\, d\lambda\,. \tag{5}$$

Since ω' is a dummy letter of integration, it can now be changed back to ω. If the filter is physically realizable, the filter transfer function must have $T(\omega) = T(-\omega)$ and $\theta(-\omega) = -\theta(\omega)$, so Equation (5) becomes

$$R(t) = \frac{m}{4\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda)\ e^{i[\omega t + \theta(\omega) + (\omega_0 - \omega)\lambda]}\, d\lambda$$

$$+ \frac{m}{4\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda)\ e^{-i[\omega t + \theta(\omega) + (\omega_0 - \omega)\lambda]}\, d\lambda$$

[6] M. S. Corrington, T. Murakami, and R. W. Sonnenfeldt, "The Complete Specification of a Network by a Single Parameter," *RCA Review*, Vol. XV, p. 392, September, 1954.

$$= \mathrm{Re}\ \frac{m}{2\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda)\ e^{i[\omega t + \theta(\omega) + (\omega_0 - \omega)\lambda]}\ d\lambda, \tag{6}$$

where Re means the real part of the quantity. Adding the steady-state component of the response to Equation (6), the total response of the filter to the modulated carrier $[1 + m f(t)\ U(t)] \cos \omega_0 t$ becomes

$$R(t) = \mathrm{Re} \left\{ T(\omega_0)\ e^{i[\omega_0 t + \theta(\omega_0)]} + \frac{m}{2\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda)\ e^{i[\omega t + \theta(\omega) + (\omega_0 - \omega)\lambda]}\ d\lambda \right\}. \tag{7}$$

Extracting the carrier frequency from both terms in Equation (7), we have

$$R(t) = \mathrm{Re} \left\{ e^{i[\omega_0 t + \theta(\omega_0)]} \left[T(\omega_0) + \frac{m}{2\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda)\ e^{i[(\omega - \omega_0)(t - \lambda) + \theta(\omega) - \theta(\omega_0)]}\ d\lambda \right] \right\}$$

$$R(t) = \mathrm{Re} \left\{ e^{i[\omega_0 t + \theta(\omega_0)]}\ [P(t) + i\,Q(t)] \right\}, \tag{8}$$

where

$$P(t) = T(\omega_0) + \frac{m}{2\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda) \cos [(\omega - \omega_0)(t - \lambda) + \theta(\omega) - \theta(\omega_0)]\ d\lambda, \tag{9}$$

$$Q(t) = \frac{m}{2\pi} \int_{-\infty}^{\infty} T(\omega)\ d\omega \int_{0}^{\infty} f(\lambda) \sin [(\omega - \omega_0)(t - \lambda) + \theta(\omega) - \theta(\omega_0)]\ d\lambda. \tag{10}$$

Envelope of Response

Equation (8) can be rewritten as

$$R(t) = \mathrm{Re} \left[\{\cos [\omega_0 t + \theta(\omega_0)] + i \sin [\omega_0 t + \theta(\omega_0)]\} \{P(t) + iQ(t)\} \right]$$

$$= P(t) \cos [\omega_0 t + \theta(\omega_0)] - Q(t)\ [\sin \omega_0 t + \theta(\omega_0)]. \tag{11}$$

Equation (11) shows that $P(t)$ is the envelope of the in-phase component and $Q(t)$ the envelope of the quadrature component of the carrier after passage through the filter. It is to be noted that $P(t)$ and $Q(t)$ as given by Equations (9) and (10) are linear functions of m, the depth of modulation.

The carrier envelope, $E(t)$, is found by rewriting Equation (11) as

$$R(t) = \{[P(t)]^2 + [Q(t)]^2\}^{\frac{1}{2}} \cos\left[\omega_0 t + \theta(\omega_0) + \tan^{-1}\frac{Q(t)}{P(t)}\right]$$

$$= E(t) \cos[\omega_0 t + \theta(\omega_0) + \psi(t)], \tag{12}$$

where $$E(t) = \{[P(t)]^2 + [Q(t)]^2\}^{\frac{1}{2}} \tag{13}$$

and $$\psi(t) = \tan^{-1}\frac{Q(t)}{P(t)}. \tag{14}$$

The carrier envelope given by Equation (13) has nonlinear distortion, since $E(t)$ is a nonlinear function of m, the depth of modulation. The term $\psi(t)$, which is the phase modulation of the carrier, is zero when the quadrature component $Q(t)$ is suppressed.

Detection of P(t) and Q(t)

The envelopes of the in-phase and quadrature components, $P(t)$ and $Q(t)$, of the modulated carrier at the output of the filter can be recovered by means of a product-type detector.

Let Equation (11) be multiplied by $2\cos[\omega_0 t + \theta(\omega_0) + \phi]$. This is the steady-state carrier at the output of the filter, with an arbitrary phase angle ϕ. The result is

$$\begin{aligned} 2R(t) \cos[\omega_0 t + \theta(\omega_0) + \phi] &= 2P(t) \cos[\omega_0 t + \theta(\omega_0)] \cos[\omega_0 t + \theta(\omega_0) + \phi] \\ &\quad - 2Q(t) \sin[\omega_0 t + \theta(\omega_0)] \cos[\omega_0 t + \theta(\omega_0) + \phi] \\ &= P(t)\{\cos\phi + \cos[2\omega_0 t + 2\theta(\omega_0) + \phi]\} \\ &\quad + Q(t)\{\sin\phi - \sin[2\omega_0 t + 2\theta(\omega_0) + \phi]\}. \end{aligned} \tag{15}$$

After passage through a low-pass filter which removes the $2\omega_0 t$ terms, the output e_0 becomes

$$e_0 = P(t) \cos\phi + Q(t) \sin\phi. \tag{16}$$

When $\phi = 0$, only the in-phase component $P(t)$ is detected; when $\phi \neq 0$, the detection is off-axis, the effect of which is discussed in a later section. When $\phi = \pm\pi/2$, only the quadrature component is detected. The detected output from an off-axis detector is seen to be a linear function of the depth of modulation, m, since the output is a linear combination of $P(t)$ and $Q(t)$.

Low-Pass Equivalent for Asymmetrical Bandpass Filter

The study of the transient response of bandpass systems is facilitated by use of equivalent low-pass analogs.[7-14] A more general derivation valid for all networks and waveforms, will be given here. Low-pass analogs for asymmetrical bandpass filters exist for detection of either the in-phase or quadrature components of the modulated carrier. If a modulated carrier is transmitted through an asymmetric bandpass filter, as shown by Figure 1a, and the output is detected synchronously, either the in-phase component, $P(t)$, or the quadrature component, $Q(t)$, can be obtained with suitable choice of the reference carrier phase, ϕ.

For the equivalent low-pass filter to exist, an input corresponding to the modulating function should produce the output $P(t)$ or $Q(t)$, in the equivalent low-pass filter. This is shown by Figure 1b for the in-phase equivalent filter and by Figure 1c for the quadrature equivalent filter.

Low-Pass Equivalent for In-Phase Detection

The low-pass equivalent for the in-phase component of the asym-

---

[7] H. Nyquist, "Certain Topics in Telegraph Transmission Theory," *Trans. A.I.E.E.*, Vol. 47, pp. 636-639, April, 1928.

[8] Murray F. Gardner and John L. Barnes, *Transients in Linear Systems*, Vol. 1, pp. 247-249, John Wiley & Sons, Inc., New York, N. Y., 1942.

[9] P. R. Aigrain, B. R. Teare, Jr., and E. M. Williams, "Generalized Theory of the Band-Pass Low-Pass Analogy," *Proc. I.R.E.*, Vol. 37, pp. 1152-1155, October, 1949.

[10] Colin Cherry, *Pulses and Transients in Communication Circuits*, Dover Publication, Inc., pp. 236-259, 1950.

[11] H. M. James, N. B. Nichols, and R. S. Phillips, *Theory of Servo-Mechanisms*, pp. 117-124, McGraw-Hill Book Co., Inc., New York, N. Y., 1947.

[12] F. W. Budd, Jr., "A Network Theorem and Its Application," *Proc. I.R.E.*, Vol. 39, pp. 685-688, June, 1951.

[13] J. S. S. Kerr, "Transient Response in a Color Carrier Channel with Vestigial Sideband Transmission," *I.R.E. Convention Record*, pt. 4, pp. 18-23, 1953.

[14] S. S. L. Chang, "On the Separability of Laplace Transform Variable and Its Application in Carrier Systems," *I.R.E. Convention Record*, pt. 2, pp. 27-33, 1955.

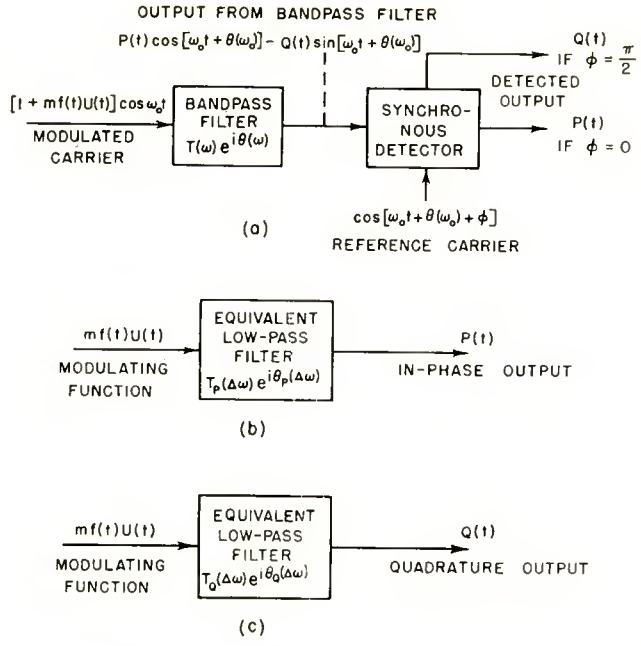


Fig. 1—Diagram of bandpass system and equivalent low-pass system.

metrical bandpass filter can be obtained by use of Equations (9) and (10). By splitting the integral term into two equal parts, Equation (9) can be rewritten as

$$P(t) = T(\omega_0) + \mathrm{Re}\left\{\frac{m}{4\pi}\int_{-\infty}^{\infty} T(\omega)\,d\omega \int_{0}^{\infty} f(\lambda)\, e^{i[(\omega-\omega_0)(t-\lambda)+\theta(\omega)-\theta(\omega_0)]}\,d\lambda\right\} + \mathrm{Re}\left\{\frac{m}{4\pi}\int_{-\infty}^{\infty} T(\omega)\,d\omega \int_{0}^{\infty} f(\lambda)\, e^{i[(\omega-\omega_0)(t-\lambda)+\theta(\omega)-\theta(\omega_0)]}\,d\lambda\right\}. \tag{17}$$

Let $\omega = \omega_0 + \omega'$ in the first integral of Equation (17),

$\omega = \omega_0 - \omega'$ in the second integral of Equation (17).

then

$$P(t) = T(\omega_0) + \mathrm{Re}\left\{\frac{m}{4\pi}\int_{-\infty}^{\infty} T(\omega_0+\omega')\,d\omega' \int_{0}^{\infty} f(\lambda)\, e^{i[\omega'(t-\lambda)+\theta(\omega_0+\omega')-\theta(\omega_0)]}\,d\lambda\right\}$$

$$+ \operatorname{Re}\left\{ \frac{m}{4\pi} \int_{-\infty}^{\infty} T(\omega_0 - \omega')\, d\omega' \int_0^{\infty} f(\lambda)\, e^{i[-\omega'(t-\lambda) + \theta(\omega_0-\omega') - \theta(\omega_0)]}\, d\lambda \right\}. \tag{18}$$

In Equation (18) $P(t)$ is written in terms of the frequency characteristics above and below the carrier frequency, ω_0, such as in Figure 2. Equation (18) may be rearranged to give

$$P(t) = T(\omega_0) + \operatorname{Re}\left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} \tfrac{1}{2} \left[T(\omega_0 + \omega')\, e^{i[\theta(\omega_0+\omega') - \theta(\omega_0)]} \right.\right.$$

$$\left.\left. + T(\omega_0 - \omega')\, e^{-i[\theta(\omega_0-\omega') - \theta(\omega_0)]} \right] e^{i\omega' t}\, d\omega' \int_0^{\infty} m f(\lambda)\, e^{-i\omega'\lambda}\, d\lambda \right\} \tag{19}$$

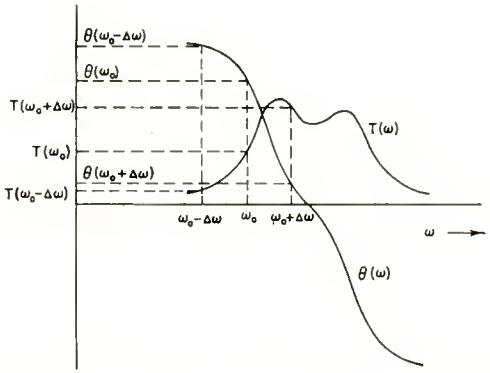


Fig. 2—Transfer characteristics of bandpass filter.

since the sign of the imaginary component can be changed without affecting the real component.

The second integral of Equation (19) represents the transform of the driving force $m f(t)$, so the transfer function of the equivalent low-pass analog for the in-phase component, $P(t)$, is given by

$$T_P(\Delta\omega)\, e^{i\theta_P(\Delta\omega)} = \tfrac{1}{2} \left[T(\omega_0+\Delta\omega)\, e^{i[\theta(\omega_0+\Delta\omega) - \theta(\omega_0)]} \right.$$

$$\left. + T(\omega_0-\Delta\omega)\, e^{-i[\theta(\omega_0-\Delta\omega) - \theta(\omega_0)]} \right], \tag{20}$$

where $T_P(\Delta\omega)$ is the amplitude characteristic and $\theta_P(\Delta\omega)$ is the phase characteristic of the equivalent low-pass filter, and $\Delta\omega$ is the deviation from the carrier frequency, ω_0.

To simplify the notation let

$$\begin{aligned} T_P(\Delta\omega) &= T_P, \theta_P(\Delta\omega) = \theta_P, \\ T(\omega_0+\Delta\omega) &= T_U, \theta(\omega_0+\Delta\omega) - \theta(\omega_0) = \theta_U, \\ T(\omega_0-\Delta\omega) &= T_L, \theta(\omega_0-\Delta\omega) - \theta(\omega_0) = \theta_L. \end{aligned} \tag{21}$$

Then Equation (20) becomes

$$T_P\, e^{i\theta_P} = \tfrac{1}{2}\,[T_U\, e^{i\theta_U} + T_L\, e^{-i\theta_L}]. \tag{22}$$

Equating real and imaginary parts in Equation (22) gives

$$T_P \cos\theta_P = \tfrac{1}{2}\,[T_U \cos\theta_U + T_L \cos\theta_L], \tag{23}$$

$$T_P \sin\theta_P = \tfrac{1}{2}\,[T_U \sin\theta_U - T_L \sin\theta_L]. \tag{24}$$

Solving Equations (23) and (24) for T_P and θ_P gives the amplitude and phase characteristics of the low-pass analog, for in-phase detection, in terms of the characteristics of the bandpass filter.

$$T_P = \tfrac{1}{2}\,[T_P^2 + T_L^2 + 2T_U\, T_L \cos\,(\theta_U + \theta_L)\,]^{\frac{1}{2}}, \tag{25}$$

$$\theta_P = \tan^{-1}\frac{T_U \sin\theta_U - T_L \sin\theta_L}{T_U \cos\theta_U + T_L \cos\theta_L}. \tag{26}$$

Low-Pass Equivalent for Quadrature Detection

Similarly the equivalent low-pass analog of Figure 1c which gives $Q(t)$ as the response when the modulating function $f(t)$ is applied, can be found by starting with Equation (10). If the transfer characteristic for the quadrature component filter is denoted by $T_Q(\Delta\omega)\, e^{i\theta_Q(\Delta\omega)}$ then using the same notation as before for the bandpass filter and with $T_Q(\Delta\omega) = T_Q$ and $\theta_Q(\Delta\omega) = \theta_Q$, the relations,

$$T_Q\, e^{i\theta_Q} = \frac{1}{2i}\,[T_U\, e^{i\theta_U} - T_L\, e^{-i\theta_L}], \tag{27}$$

$$T_Q = \tfrac{1}{2}\,[T_U^2 + T_L^2 - 2T_U\, T_L \cos\,(\theta_U + \theta_L)\,]^{\frac{1}{2}}, \tag{28}$$

$$\theta_Q = \tan^{-1}\frac{T_L \cos\theta_L - T_U \cos\theta_U}{T_L \sin\theta_L + T_U \sin\theta_U}, \tag{29}$$

can be derived in a similar manner.

Equations (22) and (25) through (29) give the relations between

the bandpass filter and the corresponding in-phase and quadrature low-pass analogs.

Relations Between Vestigial- and Double-Sideband Systems

In the case of a symmetrical bandpass filter, with the modulated carrier tuned to the center of the band, $T_U = T_L$ and $\theta_U = -\theta_L$. Substituting these into Equation (22),

$$T_P e^{i\theta_P} = T_U e^{i\theta_U}. \tag{30}$$

This means that when an in-phase component detector is used an asymmetric bandpass filter and a symmetrical bandpass filter can have the same low-pass equivalent, and the same transient response. The bandwidth of the symmetrical filter will be nearly twice that of the corresponding vestigial-sideband filter. Since the symmetrical bandpass filter does not have a quadrature component, when the carrier is in the center, the envelope detector and the synchronous detector give the same response.

Equations (25) and (26) show that the amplitude and phase characteristics of the low-pass analog are each functions of both the amplitude and phase of the bandpass filter. These equations also show that there is an infinite number of asymmetrical bandpass amplitude and phase characteristics for a given low-pass analog transfer characteristic.

Figure 3a shows the amplitude and phase characteristics of a flat-staggered-triple tuned circuit. The corresponding amplitude and phase curves for the in-phase and quadrature low-pass analogs are shown by Figures 3b and 3c respectively, for the carrier at the 50 per cent point on the low side of bandpass characteristic.

Summary of Theoretical Results

It is useful at this point to summarize the results obtained.

1. In any linear system, symmetrical or asymmetrical, the waveshape of the transient output of a synchronous detector is independent of the depth of modulation. An important consequence is that transient waveshape correction is possible in the transmitter, r-f or i-f sections of the receiver, or in the video amplifier, by linear passive networks.

2. The transient waveshape of the output from an envelope detector is nonlinearly dependent on the depth of modulation. This means that the waveshape will be altered as the depth of modulation is changed. Compensation by linear passive networks is not possible, except for very small depths of modulation.

3. Any bandpass system (symmetry is not required) followed by

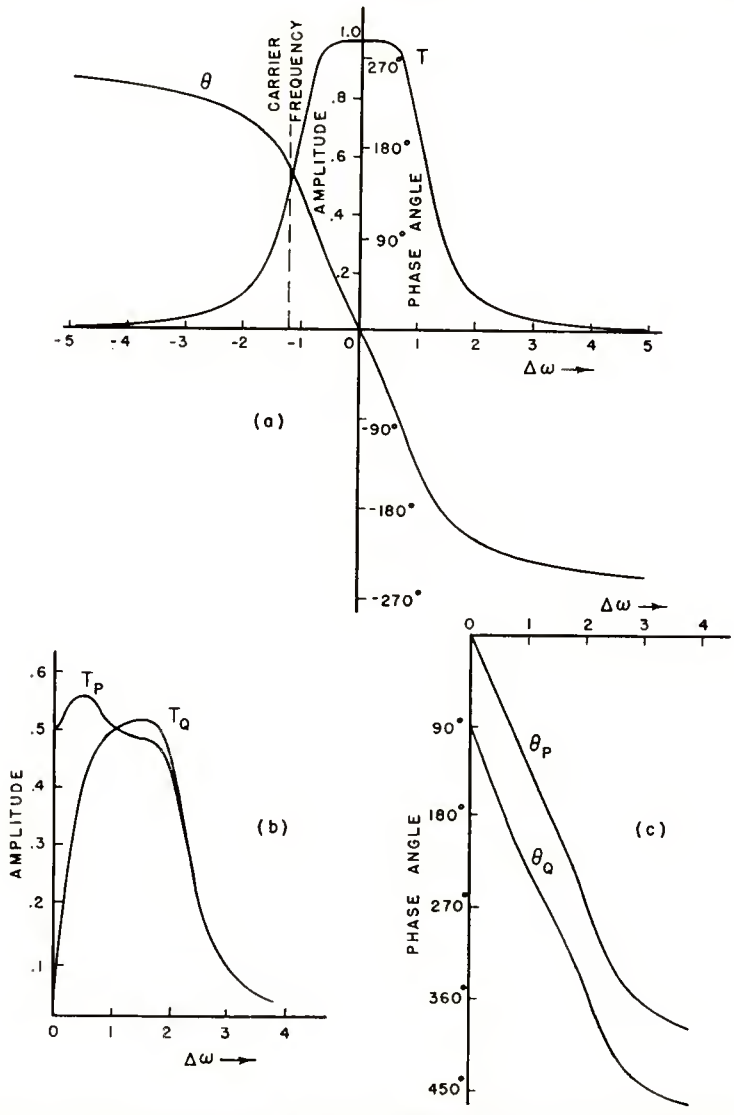


Fig. 3—(a) Bandpass filter transfer characteristic; (b) low-pass analog amplitude characteristics; (c) low-pass analog phase characteristics.

a synchronous detector can be replaced by a completely equivalent low-pass system. The analog relationship between the systems is such that transmission of the modulation alone through the analog low-pass filter is equivalent to synchronous detection of the amplitude modulated carrier after transmission through the corresponding bandpass system. This relationship is valuable because it reduces problems involving

modulated carriers in bandpass systems to the solution of the simpler low-pass system.

4. An infinite number of asymmetric sideband systems can be found with the same in-phase low-pass analog as a given asymmetric sideband system. This means that the same transient fidelity can be obtained from single- and vestigial-sideband systems using an in-phase component detector as can be obtained from corresponding symmetric-sideband systems with envelope detection. The saving of bandwidth is a maximum of two to one.

5. The amplitude response of the low-pass analog depends on the amplitude and phase characteristics of the bandpass system. Therefore, a knowledge of both the sweep response of the bandpass system and the phase characteristic is necessary to determine the frequency response of the detector output for amplitude modulated carrier inputs.

Transient Response of Practical Circuits

It is assumed that the bandpass system to be considered experimentally can be represented by a maximally flat triple-tuned circuit which has been described by Wallman.[15] The response to a stepped carrier wave will be calculated for the double- and vestigial-sideband systems of transmission. To study the effect of the method of detection on the transient, the response will be found for both envelope and in-phase component detection. The Laplace transform method will be used in the derivation of the equations for the transient response.

Double-Sideband System, Exact Analysis

The exact equation for the response of a flat-staggered triple-tuned circuit, with 10 per cent bandwidth, to a unit step of carrier voltage applied to the center frequency, is given by Equation (31) (See Appendix I for derivation).

$$S(\alpha) = \sin 40\alpha + 0.5768\, e^{-1.0433\alpha} \cos 41.7573\alpha \qquad (31)$$
$$- 0.5768\, e^{-0.9567\alpha} \cos 38.2928\alpha - e^{-2\alpha} \sin 39.9500\alpha,$$

where $\alpha = .025\ \omega_c t$, $\omega_c =$ center frequency.

This equation can be rewritten as

$$S(\alpha) = \sqrt{[P(\alpha)]^2 + [Q(\alpha)]^2} \sin \left[40\alpha + \tan^{-1} \frac{Q(\alpha)}{P(\alpha)} \right], \qquad (32)$$

[15] G. E. Valley and H. Wallman, *Vacuum Tube Amplifiers*, McGraw-Hill Book Co., Inc., New York, 1948, p. 176.

where

$$P(\alpha) = 1 - 0.5768\, e^{-1.0433\alpha} \sin 1.7573\alpha - 0.5768\, e^{-0.9567\alpha} \sin 1.7071\alpha - e^{-2\alpha} \cos .0500\alpha, \tag{33}$$

and

$$Q(\alpha) = 0.5768\, e^{-1.0433\alpha} \cos 1.7573\alpha - 0.5768\, e^{-.9567\alpha} \cos 1.7071\alpha + e^{-2\alpha} \sin .0500\alpha. \tag{34}$$

The envelope of the response is given by $\sqrt{[P(\alpha)]^2 + [Q(\alpha)]^2}$ and the in-phase and quadrature components by $P(\alpha)$ and $Q(\alpha)$ respectively. Figure 4 shows the envelope, in-phase and quadrature components of the detected modulation. It is noted from this exact analysis that the quadrature component in a double sideband bandpass system with 10 per cent bandwidth is negligible; the in-phase and envelope response are nearly identical.

Double Sideband System Approximate Analysis

The approximate method of analysis which is based on arithmetic symmetry, as first given by Wheeler,[16] gives very accurate results for the double sideband case. The approximate equation for the envelope of the output from a flat staggered triple when a unit step of carrier is applied at the center frequency is given by, (see Appendix II, Equation (69) for derivation),

$$E(\alpha) = 1 - e^{-2\alpha} - \frac{2}{\sqrt{3}} e^{-\alpha} \sin \sqrt{3}\alpha, \tag{35}$$

where $\alpha = \frac{1}{2} \Delta\omega_c t$, $\Delta\omega_c = \frac{1}{2}$ bandwidth at 70 per cent points.

The envelope response corresponding to a square-wave-modulated carrier, applied to a flat-staggered triple-tuned circuit at the center frequency, is shown in Figure 5 for 70 and 100 per cent modulation. It is seen that envelope and in-phase detection are identical so long as the modulation percentage is not so great as to produce rectification ($\alpha = 9$ to 11 for $m = 1$) in the envelope detector. The square-wave response of the double-sideband system is seen to have the same shape for upward and downward modulation.

---

[16] H. A. Wheeler, "The Solution of Unsymmetrical Sideband Problems with the Aid of Zero-Frequency Carrier," *Proc. I.R.E.*, Vol. 29, pp. 446-458, August, 1941.

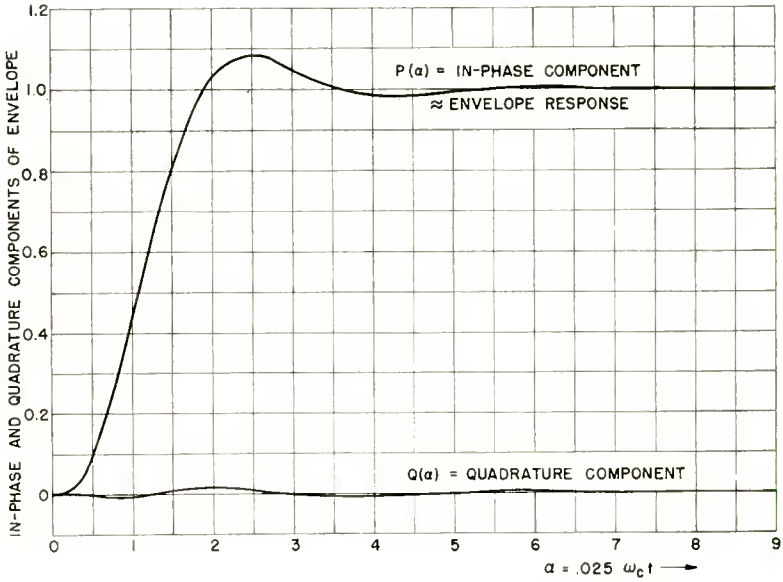


Fig. 4—Transient response of flat staggered triple to a unit step of carrier at center frequency (exact analysis).

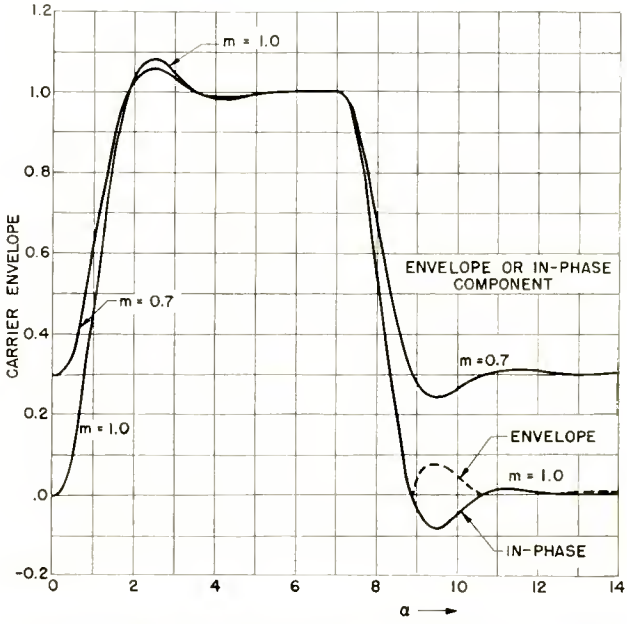


Fig. 5—Square-wave response of flat staggered triple; carrier at center frequency (approximate analysis).

Vestigial Sideband System

Envelope Response: Comparison of the exact and approximate solutions for several cases shows that the maximum error involved in the approximate method of analyzing the transient response of a flat-staggered triple with 10 per cent bandwidth is less than 3 per cent for the vestigial case. This simplified method will be used in the calculation of the curves to follow. The approximate analysis for the envelope response of a flat-staggered triple, to a suddenly-increased carrier at the 50 per cent response point on the selectivity curve, is given in Appendix II, Equation (70) as

$$E(\alpha) = [1 + m^2X^2 + m^2Y^2 - 2m(X \cos 2\omega_0\alpha + Y \sin 2\omega_0\alpha)]^{\frac{1}{2}}, \quad (36)$$

where

$$X = (1 - \omega_0{}^2)\, e^{-2\alpha} + \frac{1}{\sqrt{3}}(1 - 2\omega_0{}^2)\, e^{-\alpha} \sin\left(\sqrt{3}\alpha - \frac{\pi}{3}\right) + e^{-\alpha} \cos\left(\sqrt{3}\alpha - \frac{\pi}{3}\right),$$

$$Y = \omega_0 e^{-2\alpha} + \sqrt{3}\, \omega_0 e^{-\alpha} \sin\left(\sqrt{3}\alpha - \frac{\pi}{3}\right) + \omega_0 e^{-\alpha} \cos\left(\sqrt{3}\alpha - \frac{\pi}{3}\right),$$

$$\omega_0 = \sqrt[6]{3}, \quad \alpha = t,$$

and $m =$ depth of modulation.

The equation for a sudden reduction in carrier amplitude of unit amplitude, to one of amplitude $1 - m$, is obtained from Appendix II, Equation (71), and is

$$E(\alpha) = [(1 - m)^2 + m^2(X^2 + Y^2) + 2m(1 - m)(X \cos 2\omega_0\alpha - Y \sin 2\omega_0\alpha)]^{\frac{1}{2}}, \quad (37)$$

where the symbols have the same meaning as before.

In-Phase and Quadrature Response: The corresponding in-phase and quadrature components are given by the expressions

$$P(\alpha) = 1 - m\ (X \cos 2\omega_0\alpha + Y \sin 2\omega_0\alpha), \quad (38)$$

$$Q(\alpha) = m\ (X \sin 2\omega_0\alpha - Y \cos 2\omega_0\alpha), \quad (39)$$

for the suddenly applied sine wave.

Similarly

$$P(\alpha) = 1 - m + m(X \cos 2\omega_0\alpha + Y \sin 2\omega_0\alpha), \tag{40}$$

$$Q(\alpha) = -m(X \sin 2\omega_0\alpha - Y \cos 2\omega_0\alpha), \tag{41}$$

for the suddenly decreased sine wave.

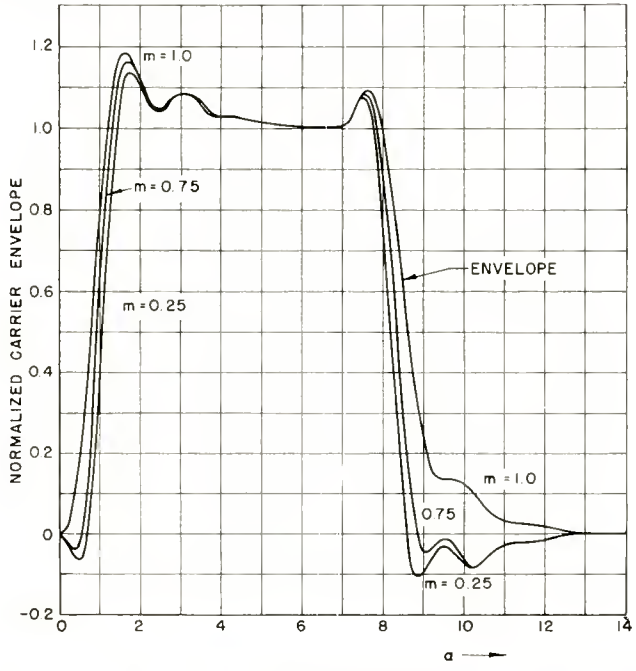


Fig. 6—Square-wave response of flat staggered triple; carrier at 50 per cent on selectivity curve.

A normalized plot of Equations (36) and (37) is shown in Figure 6 for various depths of modulation. The response is seen to change in shape with the depth of modulation due to the nonlinear relation between $E(\alpha)$ and m in Equation (36). For small depths of modulation, where $m \ll 1$, Equation (36) for $E(\alpha)$ can be expanded in a binomial series. Neglecting terms in m^2 and considering only the terms in m, the result is the same as Equation (38) for $P(\alpha)$. The contribution of $Q(\alpha)$ becomes negligible, since it is proportional to m^2. This means that the envelope and in-phase component detectors give the same output for small ($m < .1$) depths of modulation.

Normalized Curves: The normalized in-phase and quadrature response curves for square wave modulation of the carrier are shown in Figure 7. They were obtained from Equations (38) to (41) which are linear functions of m. It is noted that the in-phase component has the same shape for upward and downward modulation, for all depths of modulation. The rising and decaying transients of the envelope response are considerably different for large depths of modulation, as shown by Figure 6.

A comparison of Figures 6 and 7, for large depths of modulation

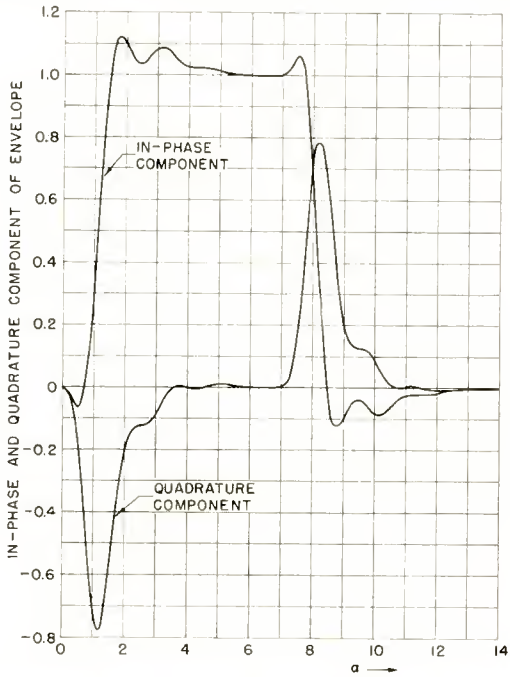


Fig. 7—Normalized square-wave response of flat staggered triple; carrier at 50 per cent.

shows that the in-phase component has a 20 per cent better decay time than the envelope response for 75 per cent modulation. The rise times are nearly equal.

Off-Axis Detection

So far we have considered envelope detection and detection of the in-phase and quadrature components. Now we shall consider the third method of detection, previously discussed, termed off-axis detection. In this method a linear combination of the in-phase and quadrature components is obtained. This is not the same as envelope detection

where the square root of the sum of the squares of the in-phase and quadrature components is detected.

Consider the vector diagram shown by Figure 8. The vector indicating the envelope $E(\alpha)$ is shown with the in-phase and quadrature components $P(\alpha)$ and $Q(\alpha)$. If the detector detects along a vector displaced by the angle ϕ from the in-phase component it becomes an off-axis detector. The new output $E'(\alpha)$ at angle ϕ will be given by

$$\begin{aligned} E'(\alpha) &= E(\alpha) \cos(\psi - \phi) \\ &= E(\alpha) \cos\psi \cos\phi + E(\alpha) \sin\psi \sin\phi \\ E'(\alpha) &= P(\alpha) \cos\phi + Q(\alpha) \sin\phi, \\ \frac{E'(\alpha)}{\cos\phi} &= P(\alpha) + Q(\alpha) \tan\phi, \end{aligned} \tag{42}$$

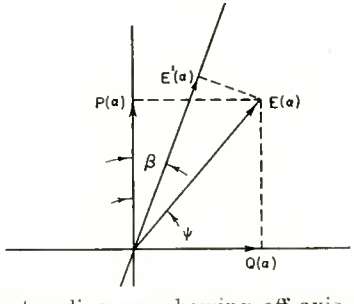


Fig. 8—Vector diagram showing off-axis detection.

since

$$P(\alpha) = E(\alpha) \cos\psi,$$

and

$$Q(\alpha) = E(\alpha) \sin\psi.$$

Figure 9 is a plot of Equation (42) and shows the effect of off-axis detection on the response of a flat-staggered triple-tuned bandpass circuit for various detection angles, ϕ. A 70 per cent square-wave modulated carrier is applied at the 50 per cent point on the selectivity curve in the computation of these curves. It is noted that the rise and decay times of the response remain equal for each ϕ. The precursory and following overshoots are identical for upward and downward modulation for any given detection angle. This means that off-axis detection can improve the transient response, with no change in the network. Inspection of Figure 9 shows that a phase angle ϕ of 0° (in-phase detection) does not necessarily give the best balance between the precursory ripple and the following overshoot.

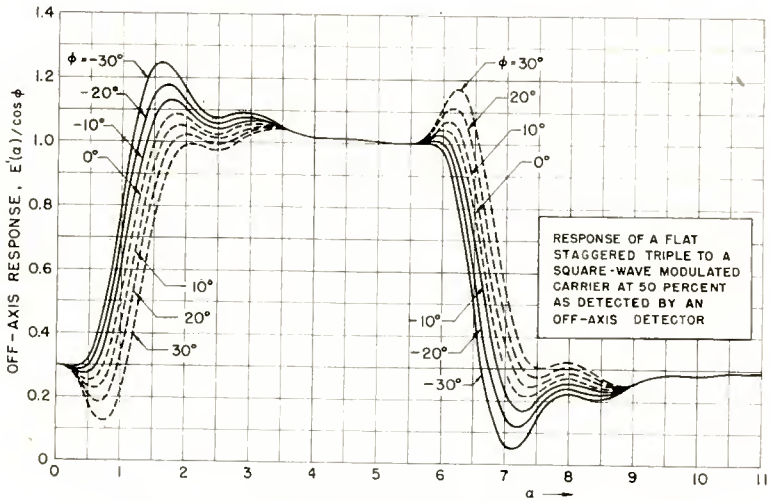


Fig. 9—Square-wave response with off-axis detection.

In-Phase Detector Circuit

The detected output given by Equation (16) was seen to be the result of multiplication of the output response of a bandpass filter by a carrier frequency term, the terms of double carrier frequency being removed by a low-pass filter. If the carrier frequency is available, this multiplication can be obtained from a variety of linear modulator circuits, operated as so-called product-detectors. A circuit of considerable interest is shown in Figure 10.

It is shown in Appendix III that the circuit performs the multiplication required if $e_2/e_1 \ll 1$. When the incoming signal is

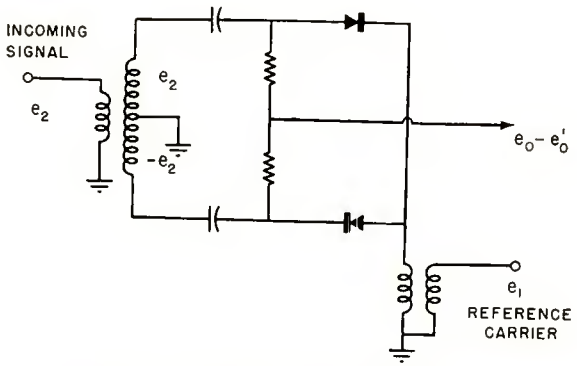


Fig. 10—In-phase component detector.

$$e_2 = E(t) \cos [\omega_0 t + \theta(\omega) + \psi(t)], \tag{43}$$

and the reference carrier is

$$e_1 = E_1 \cos [\omega_0 t + \theta(\omega_0) + \phi], \tag{44}$$

the output is

$$e_0 - e_0' = 2 [P(t) \cos \phi + Q(t) \sin \phi], \tag{45}$$

which gives the off-axis detection of Equation (16).

Experimental Results

The transient response of a flat-staggered triple-tuned bandpass filter to vestigial sideband transmission was measured experimentally, using both envelope and in-phase detection. The bandwidth of the filter was four megacycles. The upper part of Figure 11a shows the sweep-frequency response of an in-phase component detector, with frequency increasing to the right of the notch. The lower oscillogram of Figure 11a shows the response of an in-phase component detector to an 85 per cent square-wave modulated carrier applied at the 50 per cent response point on the selectivity characteristic. The square-wave repetition rate is 100 kilocycles. The corresponding frequency and transient responses of an envelope detector for the same bandpass filter are shown in Figure 11b. By comparison of Figures 11a and 11b it is noted that the transient symmetry is considerably improved by in-phase detection. The frequency responses are nearly the same. The details of the dissymmetry in the envelope detector can be more clearly seen in Figure 12a where the corners of the square-wave have been greatly expanded in time. The upper curve is the rising transient and the lower is the falling one. Figure 12b shows the comparatively symmetrical corners of the output from the in-phase detector. The dots are time markers and are spaced 0.05 microsecond apart. Figure 13 gives a comparison of rise time and overshoot. Figure 13a shows an expanded view of the envelope detector response. Figure 13b is the corresponding in-phase response. Figures 14a and 14b show the responses of envelope and in-phase component detectors respectively, when the repetition frequency of the square-wave modulation is increased to one megacycle.

Conclusions

Video Equivalent

A linear video equivalent exists when an amplitude-modulated car-

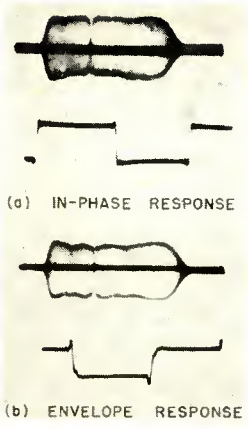


Fig. 11—Output from in-phase and envelope detector.

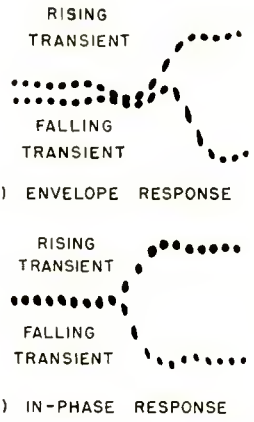


Fig. 12—Detail of envelope and in-phase detector output.

rier is transmitted through an asymmetrical sideband system and the output is detected by a synchronous detector. To define the bandpass system with its modulated carrier in terms of a linear video equivalent and the same modulating function, it is necessary to define two low-pass equivalent circuits. One of the low-pass equivalents will give the envelope of the in-phase component of the carrier as the output when the modulation is applied at the input. The other will give the envelope of the quadrature component of the carrier for the same input. The low-pass equivalent transfer characteristics change with the position

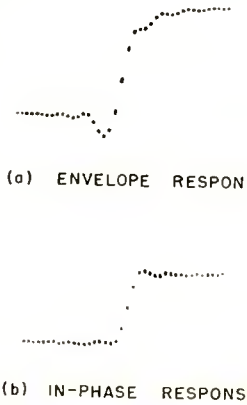


Fig. 13—Output from envelope and in-phase detector showing improvement in rise time.

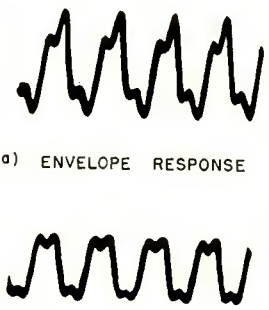


Fig. 14—1-megacycle square-wave response of envelope and in-phase detector.

of the carrier in the bandpass system. These linear video equivalent transfer characteristics are simply related to the asymmetrical bandpass transfer characteristics.

Envelope Detectors

When an amplitude-modulated carrier is transmitted through a linear system with sideband asymmetry, and an envelope detector is used, there will be nonlinear distortion in the detected output. The output waveform of an envelope detector changes with the depth of modulation. When the modulation contains jumps or steps, as in television or other pulse systems, the rise and decay times will be unequal. The precursory and following overshoots associated with the rise and decay of the envelope of the carrier, will not be identical for the increasing and decreasing carriers. These distortions are most serious when the depth of modulation is large and are undesirable in monochrome and color television.

Synchronous Detector

Nonlinear distortion in the output is eliminated by synchronous detection of an amplitude-modulated carrier transmitted through a linear system. The output waveform of a synchronous detector is independent of the depth of modulation. When the modulation contains pulses or other steep-sided waveforms, the rise and decay times will be the same. The precursory and following overshoots, associated with the rise and decay of the envelope of the in-phase component of the carrier, will be identical for the increasing and decreasing carriers. If necessary, an appropriate passive filter can be used to correct the amplitude and phase of the system, either before or after detection, to obtain a more symmetrical response.

Off-Axis Synchronous Detection

When a synchronous detector is used, the detection may be either in-phase with the carrier, or off-axis. With off-axis detection an improvement in the symmetry of the response to a square-wave can often be obtained. In television, where traps are used in the i-f amplifier, off-axis detection can be used to correct for part of the phase nonlinearity introduced by the traps.

In monochrome television worthwhile improvements in the transient fidelity and picture quality are obtained by use of a synchronous detector. In color television, identical rising and decaying transients have the additional advantages of giving sharper and cleaner color transitions.

Appendix I —

Exact Transient Analysis of a Flat Staggered Triple

As given by Wallman,[17] the exact equation of the normalized operational impedance for the maximally-flat staggered-triple tuned-bandpass circuit can be written as

$$\frac{Z}{R} = \frac{1}{\left[d_1 + \frac{s}{\alpha_1} + \frac{\alpha_1}{s} \right] \left[d_1 + s\alpha_1 + \frac{1}{s\alpha_1} \right] \left[\delta + s + \frac{1}{s} \right]}, \quad (46)$$

where

$$\delta = \text{over-all per cent bandwidth } \frac{\Delta\omega}{\omega_c},$$

$$\omega_c = \text{center frequency},$$

$$d_1^2 = \frac{4 + \delta^2 - \sqrt{16 + 4\delta^2 + \delta^4}}{2},$$

$$\left(\alpha_1 - \frac{1}{\alpha_1} \right)^2 = \delta^2 - d_1^2.$$

There are two stages staggered at $\omega_c\alpha_1$ and ω_c/α_1 of dissipation factor d_1 and one stage centered at ω_c of bandwidth $\Delta\omega$. Equation (46) can be written in factored form as

$$\frac{Z}{R} = \frac{s^3}{(s+a-ib)\,(s+a+ib)\,(s+c-id)\,(s+c+id)\,(s+f-ig)\,(s+f+ig)}, \quad (47)$$

where

$$a = \frac{\alpha_1 d_1}{2}, \quad b = \frac{a_1}{2}\sqrt{4 - d_1^2}, \quad c = \frac{d_1}{2\alpha_1},$$

$$d = \frac{1}{2\alpha_1}\sqrt{4 - d_1^2}, \quad f = \frac{\delta}{2}, \quad g = \frac{\sqrt{4 - \delta^2}}{2}.$$

[17] Reference (15), pages 184-185.

Expanding Equation (47) into partial fractions we have

$$\frac{Z}{R} = \frac{k_1}{s+a-ib} + \frac{k_1^*}{s+a+ib} + \frac{k_2}{s+c-id} + \frac{k_2^*}{s+c+id} + \frac{k_3}{s+f-ig} + \frac{k_3^*}{s+f+ig}, \tag{48}$$

where

$$k_1 = \frac{3ab^2-a^3-i(b^3-3a^2b)}{2ib\,[(c-a)^2-b^2+d^2+i2b(c-a)]\,[(f-a)^2-b^2+f^2+i2b(f-a)]}$$
$$= x_1 + iy_1,$$

$$k_2 = \frac{3cd^2-c^3-i(d^3-3c^2d)}{2id\,[(a-c)^2-d^2+b^2+i2d(a-c)]\,[(f-c)^2-d^2+g^2+i2d(f-c)]}$$
$$= x_2 + iy_2,$$

$$k_3 = \frac{3fg^2-f^3-i(g^3-3f^2g)}{2ig\,[(c-f)^2-g^2+d^2+i2g(c-f)]\,[(a-f)^2-g^2+b^2+i2g\,(a-f)]}$$
$$= x_3 + iy_3,$$

and * denotes the conjugate.

The response to a suddenly applied sine wave will be found by taking the inverse Laplace transform of $\left[\frac{Z(s)}{R}\,\frac{\omega_0}{s^2+\omega_0^2}\right]$

For the first two terms of Equation (48) the response to a suddenly applied sine wave will be given by

$$S_1(t) = 2\mathrm{Re}\left\{\mathcal{L}^{-1}\left[\frac{x_1+iy_1}{s+a-ib}\,\frac{\omega_0}{s^2+\omega_0^2}\right]\right\} \tag{49}$$

$$= 2\mathrm{Re}\left\{\omega_0(x_1+iy_1)\left[\frac{e^{-(a-ib)t}}{(a-ib)^2+\omega_0^2} + \frac{\sin(\omega_0 t-\psi)}{\omega_c\,[(a-ib)^2+\omega_0^2]^{\frac{1}{2}}}\right]\right\}, \tag{50}$$

where

ω_0 = carrier frequency,

$$\psi = \tan^{-1} \frac{\omega_0}{a - ib}.$$

Re means the real part of the quantity.

Rewriting Equation (50) gives

$$S_1(t) = \frac{2\omega_0 e^{-at}}{u_1^2 + v_1^2} [(x_1u_1 - y_1v_1) \cos bt - (x_1v_1 + y_1u_1) \sin bt]$$

$$+ \frac{2}{u_1^2 + v_1^2} \{[a(x_1u_1 - y_1v_1) + b(x_1v_1 + y_1u_1)] \sin \omega_0 t$$

$$- \omega_0 (x_1u_1 - y_1u_1) \cos \omega_0 t\}, \qquad (51)$$

where

$$u_1 = a^2 - b^2 + \omega_0^2, \qquad v_1 = 2ab.$$

By appropriate change of letters in Equation (51), expressions for the other pairs of terms of Equation (48) can be written. When added together the total response can be written as

$$S(t) = \frac{2\omega_0 e^{-at}}{u_1^2 + v_1^2} [A_1 \cos bt - B_1 \sin bt]$$

$$+ \frac{2}{u_1^2 + v_1^2} [(aA_1 + bB_1) \sin \omega_0 t - \omega_0 A_1 \cos \omega_0 t]$$

$$+ \frac{2\omega_0 e^{-ct}}{u_2^2 + v_2^2} [A_2 \cos dt - B_2 \sin dt]$$

$$+ \frac{2}{u_2^2 + v_2^2} [(cA_2 + dB_2) \sin \omega_0 t - \omega_0 A_2 \cos \omega_0 t]$$

$$+ \frac{2\omega_0 e^{-ft}}{u_3^2 + v_3^2} [A_3 \cos gt - B_3 \sin gt]$$

$$+ \frac{2}{u_3^2 + v_3^2} [(fA_3 + gB_3) \sin \omega_0 t - \omega_0 A_3 \cos \omega_0 t], \qquad (52)$$

where

$$A_1 = x_1u_1 - y_1v_1, \qquad B_1 = x_1v_1 + y_1u_1,$$
$$A_2 = x_2u_2 - y_2v_2, \qquad B_2 = x_2v_2 + y_2v_2,$$
$$A_3 = x_3u_3 - y_3v_3, \qquad B_3 = x_3v_3 + y_3v_3,$$
$$u_2 = c^2 - d^2 + \omega_0^2, \qquad u_3 = f^2 - g^2 + \omega_0^2,$$
$$v_2 = 2cd, \qquad v_3 = 2fg.$$

For the double sideband case $\omega_0 = 1.0$; for a 10 per cent bandwidth the constants are

$$a = 0.02608 \qquad x_1 = -25.71826$$
$$b = 1.04393 \qquad y_1 = -15.67779$$
$$c = 0.02391 \qquad x_2 = -24.27309$$
$$d = 0.05731 \qquad y_2 = \ \ 13.18140$$
$$f = 0.050 \qquad x_3 = \ \ 49.99118$$
$$g = 0.99875 \qquad y_3 = \ \ 2.50271$$

The equation for the response becomes

$$S(t) = 0.5768\, e^{-.0261t} \cos 1.0439t - 0.5768\, e^{-.0239t} \cos 0.9573t - e^{-.05t} \sin .9987t + \sin t. \tag{53}$$

Normalizing by $\alpha = .025t$

$$S(t) = 0.5768\, e^{-1.0433\alpha} \cos 41.7573\alpha - 0.5768\, e^{-.9567\alpha} \cos 38.298\alpha - e^{-2\alpha} \sin 39.9500\alpha + \sin 40\alpha, \tag{54}$$

where

$\alpha = .025\ \omega_0 t$ for midband frequencies other than unity.

Appendix II — Approximate Transient Analysis of Flat-Staggered Triple

The low-pass analog of the bandpass system can be used when the fractional bandwidth is small, and arithmetic symmetry can be assumed. When the carrier is in the center of the passband, Wheeler has shown[18] that in such a narrow-band system both the modulated carrier driving

[18] Reference (16), p. 454.

force and the system transfer function can be shifted to zero frequency and the response of the low-pass analog to the modulation is identical to the envelope response of the bandpass system. When the applied carrier is not at the center frequency of the bandpass system as shown in Figure 15a, the frequency of the modulated carrier applied to the low-pass analog is equal to the difference between the center frequency and the applied frequency of the bandpass system. This is shown by ω_d in Figure 15b. If the shifted modulated carrier, $m(t)$ sin $m_d t$, is applied to a bandpass system with a low-pass analog transfer function $G(s)$, it can be shown that the envelope of the output is given by[19]

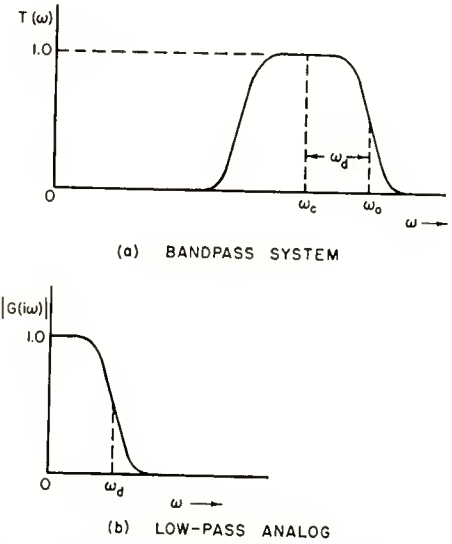


Fig. 15—Bandpass system with low-pass analog.

$$E(t) = \mathcal{L}^{-1}\,[G(s + i\omega_d)\; M(s)\,], \tag{55}$$

where

$$M(s) = \mathcal{L}\;[m(t)\,].$$

For the flat-staggered triple-tuned circuit the normalized impedance equation for the arithmetically symmetrical case with a bandwidth of two units is given by

$$\frac{Z}{R} = \frac{1}{(s^2 + s + 1)\;(s + 1)}. \tag{56}$$

[19] Reference (8), p. 248.

From Equation (55) the transform of the envelope response will be

$$\mathcal{L}\ [E(t)] = \frac{1}{s\left[(s + i\omega_d + \frac{1}{2})^2 + \frac{3}{4}\right]\left[s + i\omega_d + 1\right]}, \tag{57}$$

so

$$E(t) = \frac{1}{(1 + i\omega_d)\ (1 - \omega_d{}^2 + i\omega_d)} - \frac{e^{-(1 + i\omega_d)t}}{(1 + i\omega_d)} + \frac{2e^{-(\frac{1}{2} + i\omega_d)t}}{\sqrt{3}\left[1 - \omega_d{}^2 + i\omega_d\right]^{\frac{1}{2}}} \sin\left(\frac{\sqrt{3}}{2}t - \psi\right) \tag{58}$$

where

$$\psi = \tan^{-1}\frac{\sqrt{3}/2}{-(\frac{1}{2} + i\omega_d)} + \frac{\pi}{3}.$$

This can be rewritten as

$$E(t) = \frac{1}{(1 + i\omega_d)\ (1 - \omega_d{}^2 + i\omega_d)}\left[1 - X\cos\omega_d t - Y\sin\omega_d t + i(X\sin\omega_d t - Y\cos\omega_d t)\right], \tag{59}$$

where

$$X = (1 - \omega_d{}^2)\ e^{-t} + \frac{1}{\sqrt{3}}(1 - 2\omega_d{}^2)\ e^{-t/2}\sin\left(\frac{\sqrt{3}}{2}t - \frac{\pi}{3}\right) + e^{-t/2}\cos\left(\frac{\sqrt{3}}{2}t - \frac{\pi}{3}\right), \tag{60}$$

and

$$Y = \omega_d e^{-t} + \sqrt{3}\,\omega_d e^{-t/2}\sin\left(\frac{\sqrt{3}}{2}t - \frac{\pi}{3}\right) + \omega_d e^{-t/2}\cos\left(\frac{\sqrt{3}}{2}t - \frac{\pi}{3}\right). \tag{61}$$

The phase modulation on the carrier will be

$$\phi(t) = \tan^{-1} \frac{X \sin \omega_d t - Y \cos \omega_d t}{1 - X \cos \omega_d t - Y \sin \omega_d t}. \tag{62}$$

If the amplitude of the response is normalized by dividing by the final value

$$\frac{Z}{R} = \left| \frac{1}{(1 + i\omega_d)\ (1 - \omega_d{}^2 + i\omega_d)} \right| = \frac{1}{[1 + \omega_d{}^6]^{\frac{1}{2}}} \text{ at } \omega = \omega_d,$$

the in-phase and quadrature components are given by

$$P(t) = 1 - X \cos \omega_d t - Y \sin \omega_d t, \tag{63}$$

$$Q(t) = X \sin \omega_d t - Y \cos \omega_d t. \tag{64}$$

The envelope response is

$$E(t) = [1 + X^2 + Y^2 - 2(X \cos \omega_d t + Y \sin \omega_d t)]^{\frac{1}{2}}. \tag{65}$$

For the double sideband case $\omega_d = 0$ and X and Y in Equation (59) reduce to

$$X = e^{-t} + \frac{1}{\sqrt{3}} e^{-t/2} \sin\left(\frac{\sqrt{3}}{2} t - \frac{\pi}{3}\right) + e^{-t/2} \cos\left(\frac{\sqrt{3}}{2} t - \frac{\pi}{3}\right),$$
$$Y = 0, \tag{66}$$

so that the envelope response becomes

$$E(t) = 1 - e^{-t} - \frac{1}{\sqrt{3}} e^{-t/2} \sin\left(\frac{\sqrt{3}}{2} t - \frac{\pi}{3}\right) + e^{-t/2} \cos\left(\frac{\sqrt{3}}{2} t - \frac{\pi}{3}\right)$$
$$= 1 - e^{-t} - \frac{2}{\sqrt{3}} e^{-t/2} \sin \frac{\sqrt{3}}{2} t. \tag{68}$$

For a bandwidth of $2\Delta\omega_c$ megacycles

$$E(t) = 1 - e^{-\Delta\omega_c t} - \frac{2}{\sqrt{3}} e^{-\Delta\omega_c t/2} \sin \sqrt{3} \frac{\Delta\omega_c t}{2}, \tag{69}$$

with t in microseconds. The factor $\Delta\omega_c$ is introduced due to the scale change. For an increase in carrier amplitude, Equation (65) can be written as

$$E(t) = |1 - m(X \cos \omega_d t + Y \sin \omega_d t) + im(X \sin \omega_d t - Y \cos \omega_d t)|, \quad (70)$$

where m is the increase of carrier amplitude as compared to unity as shown in Figure 16.

The response to a sudden reduction in carrier amplitude of amount m is given by

$E(t) =$ steady-state $- m$ (response to a suddenly applied sine wave),

$$E(t) = |1 - m + m\ (X \cos \omega_d t + Y \sin \omega_d t) - im(X \sin \omega_d t - Y \cos \omega_d t)|. \quad (71)$$

Appendix III—Analysis of Synchronous Detector Circuit

Consider the detector circuit shown in Figure 17a with two signals

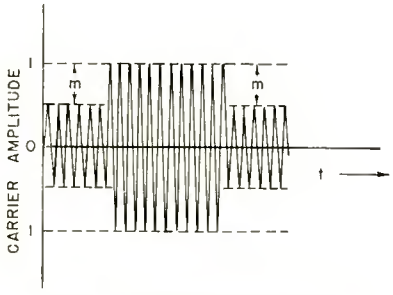


Fig. 16—Modulated carrier.

applied to the diode as indicated, if $e_1 \gg e_2$, the diode will be gated on the peaks of e_1, hence the term synchronous detector. Let the reference carrier be $e_1 = E_1 \cos [\omega_0 t + \theta(\omega_0) + \phi]$ and $e_2 = E(t) \cos [\omega_0 t + \theta(\omega_0) + \psi(t)]$ be the incoming signal. The rectified output voltage e_0 can be shown to be the envelope of $e_1 + e_2$ which is equal to

$$e_0 = [E(t)^2 + E_1^2 + 2E(t)\ E_1 \cos \beta]^{\frac{1}{2}}, \quad (72)$$

where $\beta = \psi(t) - \phi$, from the vector diagram of Figure 17b. If a balanced detector is used, the voltage across the second diode will be

$$e_0' = [E(t)^2 + E_1^2 - 2E(t)\ E_1 \cos \beta]^{\frac{1}{2}}. \quad (73)$$

The resultant output voltage from the balanced detector of Figure 10 is

$$e_0 - e_0' = [E(t)^2 + E_1^2 + 2E(t)\,E_1 \cos\beta]^{\frac{1}{2}} - [E(t)^2 + E_1^2 - 2E(t)\,E_1 \cos\beta]^{\frac{1}{2}}. \quad (74)$$

e_0 in Equation (72) can be written as

$$e_0 = E_1\,[1 + h^2 + 2h\cos\beta]^{\frac{1}{2}}, \quad (75)$$

where $h = \dfrac{E(t)}{E_1}$.

Expanding Equation (75) in a power series we have

$$e_0 = E_1\left\{1 + \tfrac{1}{2}\,[h^2 + 2h\cos\beta] - \frac{1}{8}\,[h^2 + 2h\cos\beta]^2 + \frac{3}{48}\,[h^2 + 2h\cos\beta]^3 - \frac{5}{128}\,[h^2 + 2h\cos\beta]^4 + \ldots\right\}$$

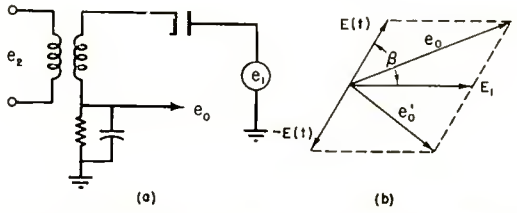


FIG. 17 DETECTOR WITH ASSOCIATED VECTOR DIAGRAM

Fig. 17—Detector with associated vector diagram.

$$e_0 = E_1\left\{1 + h\cos\beta + \tfrac{1}{2}\,h^2\sin^2\beta - \tfrac{1}{2}\,h^3\cos\beta\sin^2\beta + \tfrac{1}{2}\,h^4\sin\beta\left[1 - \frac{5}{4}\sin^2\beta\right] + \ldots\right\}. \quad (76)$$

Similarly

$$e_0' = E_1\left\{1 - h\cos\beta + \tfrac{1}{2}\,h^2\sin^2\beta + \tfrac{1}{2}\,h^3\cos\beta\sin^2\beta + \tfrac{1}{2}\,h^4\sin\beta\left[1 - \frac{5}{4}\sin^2\beta\right] + \ldots\right\}. \quad (77)$$

The resultant output voltage will then be given by

$$e_0 - e_0' = 2E_1 h \cos\beta \left[1 - \frac{h^2}{2} \sin^2 \beta \right], \tag{78}$$

neglecting terms containing powers of h higher than the 4th. If the assumption is made that $h \ll 1$, this condition being valid when the injected reference carrier is made large compared to the incoming signal, then

$$e_0 - e_0' = 2E(t) \cos\beta = 2E(t) \cos[\psi(t) - \phi] \tag{79}$$

$$= 2[E(t) \cos\psi(t) \cos\phi + E(t) \sin\psi(t) \sin\phi]$$

$$e_0 - e_0' = 2[P(t) \cos\phi + Q(t) \sin\phi]. \tag{80}$$

Equation (80) is seen to be the same as Equation (16), so that this detector is seen to perform the required operation.

EXPERIMENTAL HIGH-TRANSCONDUCTANCE GUN FOR KINESCOPES

By

F. H. Nicoll

RCA Laboratories,
Princeton, N. J.

Summary—Use of electroformed fine mesh on the control grid aperture of an experimental electron gun gives an order of magnitude reduction in required video drive. Focus and current characteristics are acceptable for many applications.

Introduction

In the past many attempts have been made to reduce cutoff and increase the transconductance of electron guns. The use of a fine mesh near the cathode, as in other radio tubes, would seem to be an obvious means, but early attempts to use such a mesh failed because of poor focus and the great loss of current. The success of the present work is due in part to the availability of fine flat electroformed mesh of high transparency and in part to the way the mesh is used so as not to alter the predominant focusing field.

In this work the fine mesh is placed over the cathode side of the grid aperture. Such an arrangement means minimum complication, and is therefore attractive where large numbers of tubes may be involved.

Assembly Problems

All guns described here have been made with an electroformed 200-mesh nickel screen.* This is a good compromise between a coarse grid giving poor transconductance, and a finer grid requiring unattainably small cathode–grid spacings. A photomicrograph of the mesh over a .035-inch grid aperture is shown in Figure 1. This shows the high transparency (60-75 per cent) and centering.

In order to obtain reasonable field penetration through the 200-mesh screen to the cathode surface, closer grid–cathode spacings are required than is standard in the cathode-ray tube art. An improved assembly was developed in which each individual grid–cathode assembly was spaced and measured before insertion in the gun. The cathode is

* This mesh is fabricated by the RCA Tube Division, Lancaster, Pa.

rigidly mounted with a ceramic support in a nichrome cylinder. This cylinder is a push fit in the apertured grid cylinder. Spacing of grid to cathode is adjusted by sliding these two members in a special jig until the correct value is obtained, at which point the two are welded. This unit, with correct grid–cathode spacing, is then mounted in the correct position in a push fit cylinder on the gun. Final operating spacings between grid and cathode were usually 0.001 to 0.0015 inch. In addition to the special cathode assembly described above, an extra-fine-grain cathode spray was employed. With these precautions, close spacings were mechanically well maintained even under conditions of ordinary tube shipment and shaking tests.

In the first experiments the mesh was attached to the .035-inch apertured grid cup by means of spot welding. This caused many failures due to weld splash, and incomplete welding of all the wires often gave

Fig. 1—Photomicrograph of 200 mesh per inch screen stretched across grid aperture.

poor mechanical support to the mesh near the edges of the aperture. A more satisfactory method was to fire the nickel mesh and nichrome cup while they were held in close contact by the jig shown in Figure 2. This consists of a lava piece (1) which is held in close contact with the mesh and grid cup (2) by the recessed cap (3) which is screwed to the base (4). The assembly is fired in a covered boat in dry hydrogen at 1065°C for 10 minutes. This method gave excellent thermal and electrical contact up to the very edge of the aperture.

Gun Design

The experiments were done by modifying a type 5TP4 gun. This gun is easily adapted for either electrostatic or magnetic focus (including an ion trap) and can be used in a wide variety of cathode-ray tube

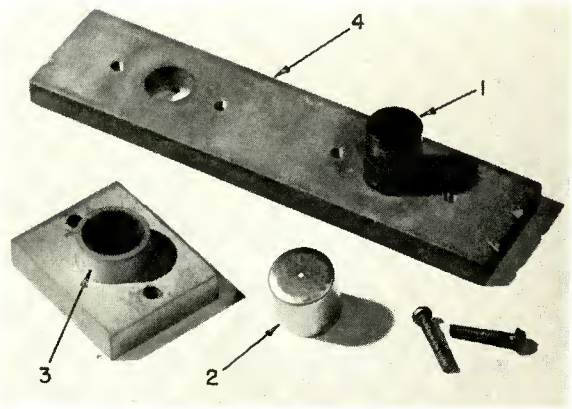


Fig. 2—Assembly jig for attaching screen to grid aperture by firing procedure.

applications, including the direct-view kinescope.

Figure 3 shows a cross-section of the G_1, G_2 and A_1 region of the modified 5TP4 gun. The grid mesh is placed on the cathode side of the grid aperture, and is very close to the cathode. As a result, the electrons emerge with low velocity and are focussed to a small cross-over by the curved field produced by field penetration into the grid aperture. The beam diverges again into the first anode from which it can be focused either electrostatically or magnetically. By this construction, focusing

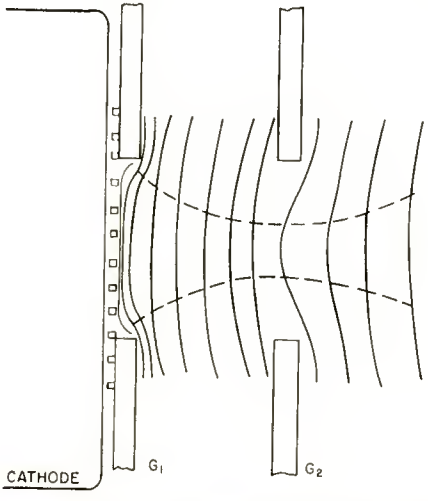


Fig. 3—Cross section of G_1–G_2 region of 5TP4 type gun with screen over grid aperture.

action is essentially maintained as it would be without the mesh. However, when the mesh is used, the thickness of the grid aperture becomes a much more important design parameter, since its thickness and the G_1–G_2 spacing determine the eventual spot size of the beam. A satisfactory combination uses a 0.005-inch-thick grid aperture with a 0.035-inch-diameter opening, spaced 0.020-inch from the second (screen) grid.

Gun Performance

Figure 4 shows photographs of the mesh image which is produced on the face of the kinescope with no final focusing and with the 1st anode and final cone at 12 kilovolts. This image is produced by the lens action of the field in the grid screen-grid region alone. The image is inverted and the photographs show the result for three different

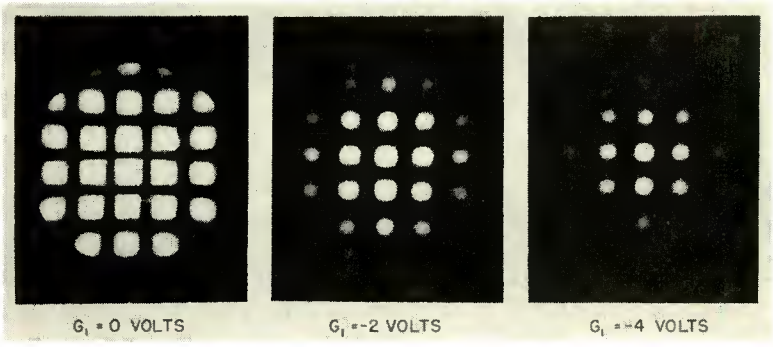


Fig. 4—Electron-produced image of grid aperture with screen.

grid biases. As can be seen, when the negative bias is increased the number of spots is reduced and also the area of the spots. The change in focused spot size as a function of bias with these guns is les than in the conventional grid–cathode arrangement. This tends to reduce the difference in focus between low lights and high lights, which is desirable in the interests of eliminating the line structure of a television picture.

During operation of the tube with magnetic or electrostatic focusing, no image of the mesh is visible when the focused spot is produced. However, for guns incorporating incorrect aperture thicknesses and spacing, an image of the mesh may become visible at focus, making the result unusable.

The grid-control characteristic of the experimental 5TP4 with mesh grid is compared with the same gun with no mesh in Figure 5. About

a tenfold increase in sensitivity is obtained with the result that a 5-volt video signal can produce a satisfactory modulation. It can be seen that similar peak currents can be obtained and also that the fraction of current collected by the aperture in the anode shown in dotted lines is somewhat less for this particular mesh-grid gun.

The resolution obtainable with these guns in tubes has been measured by observation of a resolution pattern in a 16GP4 type bulb (70 degrees deflection) and by direct measurement of the resolution from a flying-spot tube using this gun. These tests have given a resolution figure of 325-350 lines at 300 microamperes. This resolution can be

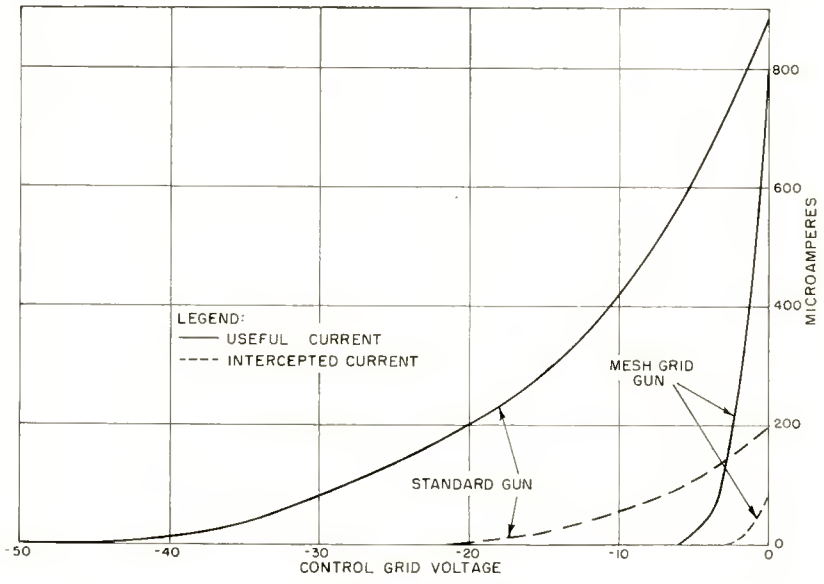


Fig. 5—Graph showing improved control characteristic of new high-transconductance gun.

further improved by the usual methods of increasing gun length or increasing deflection angle for a given size of picture.

Applications

The high-transconductance gun described was found useful in an experimental television kinescope designed to be driven by transistor video circuits. Experimental radar display tubes have also been made for use with especially high video frequencies, so that the video drive power could be kept to a minimum. The close tolerances and critical nature of the gun design make the construction relatively costly com-

pared with conventional guns, so that the uses so far have been confined to experimental tubes.

ACKNOWLEDGMENTS

The author would like to acknowledge the help of M. Colacello and P. J. Messineo in the assembly problems connected with this development, and to thank V. K. Zworykin and D. W. Epstein for their continued interest in the work.

AN ACCURATE ELECTRONIC MULTIPLIER†

By

SIDNEY STERNBERG

RCA Laboratories,
Princeton, N. J.

Summary—*An accurate analog electronic multiplier has been developed. It is a time division multiplier, the basic principles of which have been previously reported.*[1] *This paper emphasizes the dynamic considerations affecting the accuracy of multiplication. A novel dynamic balancer for switch tubes is described which makes it possible to obtain and maintain a static accuracy of .01 per cent at a switching rate of 10 kilocycles. The dynamic performance of the multiplier is specified by the gain–phase characteristics of a filter which is an integral part of any time division multiplier. A filter is described which maintains a very small phase shift at the computing frequencies and also eliminates the requirement for pulse handling d-c amplifiers. An accuracy of .02 per cent in amplitude with a phase shift of .0016 milliradian at 20 radians per second has been obtained. A small dynamic systems synthesizer has been constructed which contains eleven multiplier chassis. The performance of this multiplier in terms of combined accuracy and speed of response is better than any other known to the author.*

INTRODUCTION

ANALOG computers have, in the past, been limited with regard to speed, reliability, and accuracy by the multiplier component. Where accuracy was not a requirement, electronic multipliers have been built which were fast and reliable. However, accuracy could be obtained only with mechanical or electromechanical devices which were slow and subject to short-term wear. It is the purpose of this paper to describe the development of a high-accuracy electronic multiplier which uses a time-division principle. A multiplier of this type was first reported in 1952.[1] The work therein reported was done under contract with the Office of Naval Research as part of the Typhoon analog computer project. The resulting equipment was a laboratory model which proved the basic time-division concepts and demonstrated an electronic switch with excellent characteristics. In the early part of 1953 work began on a dynamic systems synthesizer. This was to be a large-scale electronic analog differential analyzer with the prime

† This work was done under contract with the U. S. Air Force, Wright Air Development Center Armament Laboratory as part of the Dynamic Systems Synthesizer project.

[1] E. A. Goldberg, "A High Accuracy Time Division Multiplier," *RCA Review*, Vol. XIII, pp. 265-274, September, 1952.

purpose of solving six degrees of freedom guided missile systems problems. Since the basic concept of the synthesizer was to build a machine without mechanical moving parts,[2] a highly accurate electronic multiplier was required. It was decided to use the time division principle and attempt to provide for long-term stable accuracy and increased speed of multiplication. This paper describes the results of the development work and the final multiplier design.

Review of Basic Principles

In a time-division multiplier a waveform is generated whose amplitude is proportional to one variable and whose time division is proportional to the other. The average value of the waveform represents the multiplication. The basic component of the multiplier is an electronic switch. This switch should be sufficiently accurate that the accuracy of the multiplication depends only on the combined stability of the wirewound resistors used in the signal paths and stability of power-supply voltages. The switch must operate at a sufficiently high frequency so that the computing signal frequencies are not affected by the required filtering operation.

Figures 1a and 1b show the systems used for varying amplitude and time division respectively. The output of the amplitude channel is $e_o = e_x e_y / k$, where k is the standard reference voltage.

The electronic switch used in this system is a current-switching device. The switched current is regulated by the high-gain amplifier #1. Amplifier #2 is a current polarity reversing device making possible algebraic or bipolar multiplication. Amplifier #3 in the amplitude channel (Figure 1a) sums the plus and minus currents and converts to a low-impedance voltage output which represents the product. In the time-division channel (Figure 1b) amplifier #3 provides high loop gain assuring accurate setting of the d-c bistable multivibrator to a time-division linearly proportional to e_x.

The summation of currents at (a) in Figure 1b is equal to zero.

$$\sum i_a = 0. \tag{1}$$

Assume the conductances equal, then

$$kt - k(T - t) - Te_x = 0,$$

[2] J. Lehmann, "Development of a New Large-Scale Three-Dimensional Simulator," paper presented at Typhoon Symposium III, University of Pennsylvania, Philadelphia, Pa., October 13, 1953.

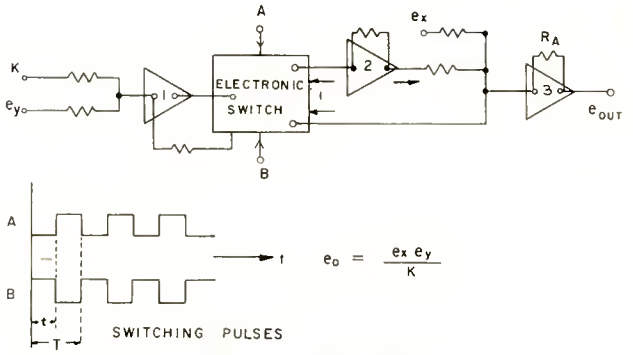


Fig. 1a—Amplitude channel.

$$t = \frac{T}{2}\left(\frac{e_x}{k} + 1\right). \tag{2}$$

The output of the amplitude channel (Figure 1a) is

$$e_o = \frac{(e_y + k)}{T}\left[t - (T - t)\right] - e_x. \tag{3}$$

Substituting Equation (2) for t,

$$e_o = \frac{e_y e_x}{k}. \tag{4}$$

All of the above amplifiers are wide-band electronic chopper-stabilized d-c amplifiers.[3]

The bulk of this paper is concerned with two basic components:

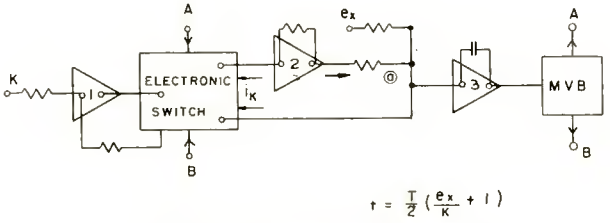


Fig. 1b—Time-division channel.

[3] S. Sternberg, "Electronic Zero Stabilization of DC Amplifiers," paper presented at the Typhoon Symposium III, University of Pennsylvania, Philadelphia, Pa., October 13, 1953.

the electronic switch and the filter. The accuracy of the multiplication is therein decided assuming the amplifiers and resistors are not limiting factors.

Electronic Switch

The electronic switch, which is the heart of the multiplier, must perform the function of switching an electronic signal without affecting the information contained in that signal.

Figure 2 is a schematic diagram which exhibits the current feature of the switch. Tube V-2 is the electronic switch which is in a series circuit with a current regulator tube V-1 and two resistors in parallel, 0.15 and 1 megohm. The voltage at point C, e_c, is directly proportional to the switched current, i. It is also directly proportional to $(e_y + k)$. The voltage e_c is the output voltage of a high-gain d-c feedback

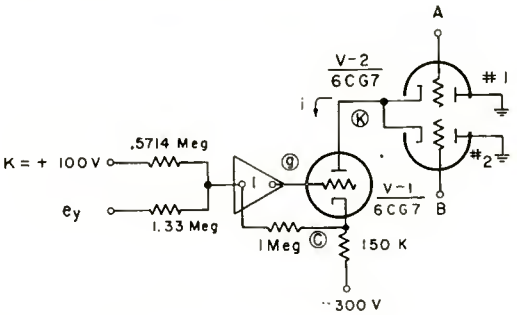


Fig. 2—Unbalanced electronic switch.

summing amplifier which includes in the forward section of the loop amplifier #1 and V-1. The switch current, i, is therefore directly proportional to $(e_y + k)$. Changes in plate resistance of either V-1 or V-2 are compensated for by changes in e_{yc}, always maintaining the relationships

$$i = k_1 e_c = k_2 (e_y + k), \tag{5}$$

$k = +100$ volts,

e_y varies between $+100$ and -100 volts,

e_c varies between -250 and -100 volts,

i varies between .08 and 1.23 milliamperes.

The grounded plates of V-2 shown in Figure (2a) are, in practice, current summing points of d-c feedback amplifiers.

From the above discussion it would appear that the requirements

of the switch have been fulfilled. However, one must consider the dynamics of the system and how they affect the accuracy of multiplication. As was pointed out earlier, the actual signal representing the multiplication is the average value of the pulse waveform.

Thus, we may write

$$e_o = \frac{R_A}{T}\left[\int_o^t i_1\,dt - \int_t^T i_2\,dt\right] \tag{6}$$

$$e_o = \frac{R_A}{T}\left[Q_1 - Q_2\right] \tag{7}$$

where T = waveform period in seconds,

R_A = feedback resistor of amplifier #3 (Figure 1a),

i_1 and i_2 = currents switched in Figure 3,

e_o = output voltage, Figure 1a.

Figure 3 is an equivalent circuit for Figure 2. Amplifier #1 and the current regulator tube V-1 is represented as an infinite impedance current source. The electronic switch tube V-2 is shown schematically as two mechanical switches. Point k is the common cathode of V-2. e_1 and e_2 are the voltages appearing at k depending upon which switch is closed. The voltages are different for the following reasons. The grids of the 6CG7 V-2 are pulsed 180° out of phase so that only one section of the tube is conducting at any time. The "on" section acts as a cathode follower. The common cathode point, k, assumes a voltage level which is the voltage on the grid plus the grid bias voltage required to pass through the tube the current driven from an infinite-impedance source. These voltages are not the same for both sections of the duo-triode. It is difficult to find two tubes whose plate current versus grid bias characteristics are the same and will remain so.

The voltage at point k (Figure 3) will vary between e_1 and e_2 at a frequency equal to the switching frequency. The condenser C must charge and discharge due to the changing voltage. Thus, if $e_1 > e_2$, less current passes through switch #1 than through switch #2 during Δt required to change the voltage across C.

$$i_1 = i - i_C\,,$$

$$i_2 = i + i_C\,,$$

$$\Delta Q = 2\,i_c \Delta t = 2C\Delta e_{kb}\,,$$

C = parasitic capacity from common cathode to ground,

$e_{kb} = e_1 - e_2$.

ΔQ is the unbalance charge which produces an error in multiplication.

In Equation (7)

$$e_o \text{ (full scale)} = 100 \text{ volts}$$

$$100 = \frac{R}{T} Q_{fs}$$

$$Q_{fs} = \frac{100T}{R}.$$

In the multiplier $R \approx .25$ megohm,

$$Q_{fs} = 4T \times 10^{-4} \text{ coulombs.}$$

$$\text{Per cent error of full scale} = E = \frac{\Delta Q}{Q_{fs}} = \frac{\Delta Q f}{4} \times 10^6,$$

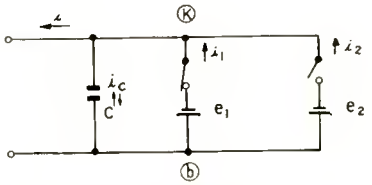


Fig. 3—Equivalent circuit for Figure 2.

$$E = \frac{C \Delta e_{kb} f}{2} \tag{8}$$

where Δe_{kb} is in volts, f is in cycles per second, and C is in microfarads.

If some representative values are chosen,

$$C = 50 \text{ micromicrofarads},$$

$$\Delta e_{kb} = 0.3 \text{ volt},$$

$$f = 10 \text{ kilocycles};$$

the error is in the order of 0.1 per cent. For the purposes of the dynamic systems synthesizer this is not tolerable.

In order to reduce this error, one or all of three factors C, Δe_{kb}, or f, must be reduced. The switching frequency, however, should be

made as large as possible since reducing it would only transfer the problem to the filter which is required to detect the average value or product. Reduction in C consists of careful chassis wiring. The largest development effort was in the direction of reducing Δe_{kb}. Figure 4 is a schematic of the first part of a dynamic switch tubes balancer. The voltage swing which charges C_k and causes a charge unbalance appears at point k (Figure 4). This signal is detected by an a-c amplifier with a gain $-G$. The output of the a-c amplifier is used to set the "turn on" level of the grid voltages for both sections of the switch tube. In effect, the voltage swing which had appeared at point k has now been transferred to the grids, where it appears as a differential between the "turn on" levels of the grid voltages. The voltage swing at k is reduced by the ratio of $1/G$. In practice, the value of G for the a-c amplifier is approxi-

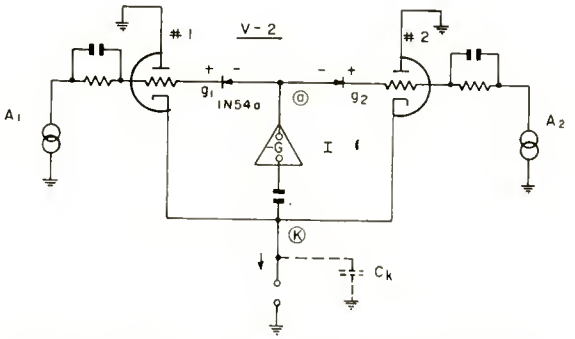


Fig. 4—Dynamic switch tubes balancer, Part I.

mately 100. 1N54A germanium crystals are used to limit the voltage level on the grids during their positive swings to a value set at the output (a) of the a-c amplifier. The bandwidth of the a-c amplifier must be broad enough to accommodate changes in unbalance of the two sections of the switch due to changes in current during computation. In practice, the correction was accomplished on a pulse-to-pulse basis with a bandwidth greater than the switching frequency. The amplifier consists of a single tube, a 6U8.

In the above discussion only the cathode-to-ground capacitance has been considered. Actually, grid-to-cathode capacitance must be considered although its contribution to switch unbalance is much less serious than the former. If for the moment we consider $C_{g1k} = C_{g2k}$ then an identical analysis is applicable in this case and Equation (8) may be applied.

It is obvious that transferring the voltage unbalance from cathode

to grid does not remove the charge unbalance due to cathode-to-grid capacitance.

The voltage swing across C_{g1k} and C_{g2k} must be made equal. Figure 5 is a schematic of the final balancer which includes the second part of the dynamic switch tube balancer. The physical equipment consists of a reversing amplifier, 12AT7, and another pair of 1N54A germanium diodes. The purpose of this part of the balancer is to set the "turn off" level of the grids voltages for both sections of the switch tube.

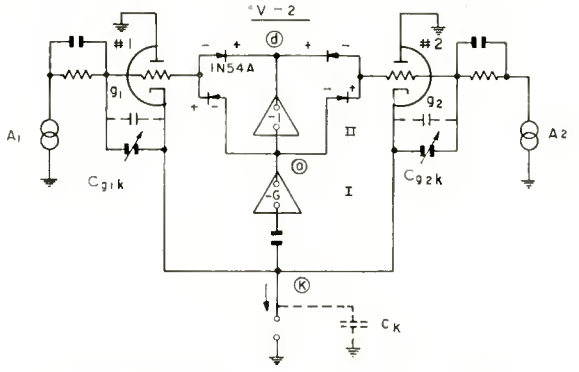


Fig. 5—Complete dynamic switch tubes balancer.

Refer to Figure 6 and compare the resultant voltages on the grid of each switch section.

$$e_{g1} = (-52 + \Delta e_{a1}) - (-62 + \Delta e_{b1}),$$

$$e_{g2} = (-52 - \Delta e_{a2}) - (-62 - \Delta e_{b2}),$$

$$\Delta e_{a1} = \Delta e_{b2}, \qquad \Delta e_{b1} = \Delta e_{a2},$$

$$e_{g1} = e_{g2}.$$

Thus this second charge unbalance condition is corrected.

One further effect remains to be considered. C_{g1k} generally does not equal C_{g2k}. This also produces a charge unbalance in accordance with Equation (8). The only difference is that now capacitance is the differential term and the voltage is total swing on the grids. It is felt that this effect is a static one and will not vary appreciably with signal level and time. A static correction, the setting of a small differential capacitor, is considered adequate, (Figure 5).

The differential capacitor is tube mounted. All 6CG7 tubes stored for use in the dynamic systems synthesizer are checked and capacitance equalized. The net result is an electronic multiplier in which any tube

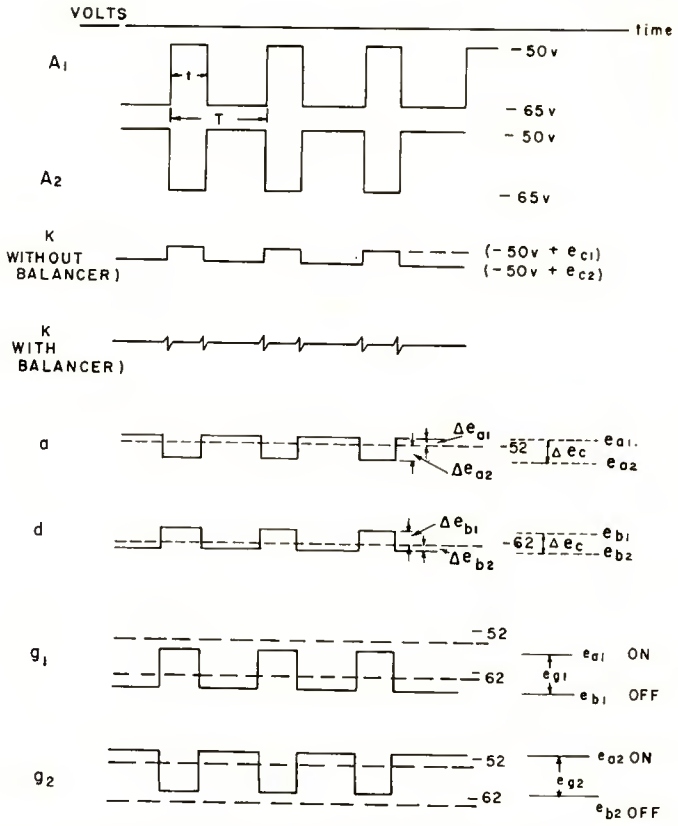


Fig. 6—Waveforms for Figure 5. Letters refer to points on Figure 5. Numbers 1 and 2 refer to sections 1 and 2 of the switch tube V2. The voltage ordinate scale factor is not the same for all waveforms and in some cases is exaggerated for clarity.

may be replaced, including the switch tube, without affecting the multiplier operation or requiring a manual adjustment. The schematic diagram of the over-all electronic switch is shown in Figure 7.

Filter

The product of two variables appears at the output (e_o in Figure 1a) as the average value of a pulse wave train. In some cases the information can be used in this form, relying on smoothing in other parts of the computer as a function of the equations to be solved, to reduce the a-c components of the pulses. However, in general this is not desirable. It is necessary to provide, as part of the multiplier, a low-pass filter for measuring the average value of the pulse wave train.

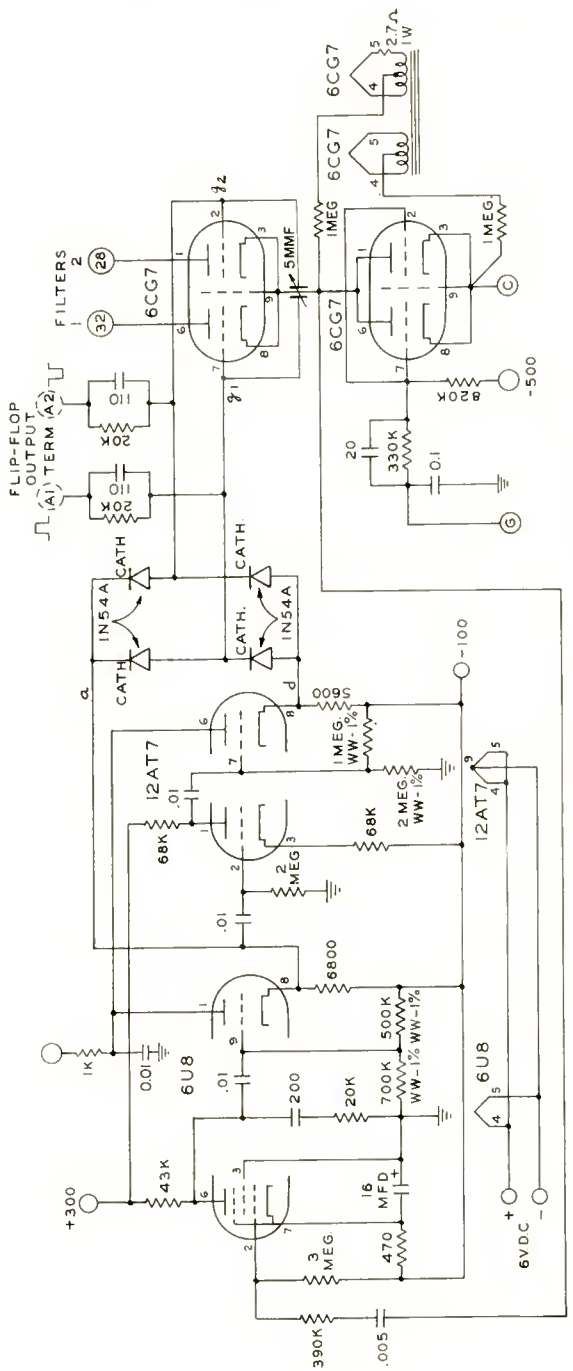


Fig. 7—Schematic diagram of complete electronic switch.

The design of any filter is always at best a compromise between more important and less important characteristics plus the ever-present economic considerations. The first decision made was to maintain a multiplier accuracy of .01 per cent at d-c. The accuracy specification leads to the choice of a switching-rate frequency (Equation (8)) which must be attenuated by the filter. Preliminary measurements on the bench indicated that 10 kilocycles gave reliable, long term, accurate multiplication. The requirement on attenuation at 10 kilocycles is 80 decibels. It is especially important to remove the switching frequency at the output of the multiplier if this output is multiplied again by the same time-division variable. The switch of the second multiplier will then tend to act as a synchronous rectifier and generate errors in computation. This situation can occur when attempting a series expansion where powers of the variable must be generated.

The maximum computing frequency was chosen as 16 cycles per second. This means that a product can be generated which varies sinusoidally plus and minus full scale at a frequency of 16 cycles per second. For a mechanical multiplier to have equivalent performance it would require an angular shaft rotational speed of 100 radians per second, and an angular acceleration of 10,000 radians per second per second. The only remaining specification for the filter is the gain–phase characteristic and the required configuration of elements.

It is relevant to digress at this point to consider the problem of amplifier design in the multipliers. In the amplitude channel there are two amplifiers, the current reversing amplifier and the output summing amplifier, which are required to handle high-level, high-frequency signals (Figure 1a). The power requirements for such an amplifier appear to be exorbitant for a computer of the size of a six-degree-of-freedom missile simulator. Ideally, pulses should appear only during the multiplication in the electronic switch and nowhere else in the system. For this reason it was decided that the filters should immediately follow the output of the electronic switch (Figure 8).

Application of standard filter-design procedures yields immediately the gain–phase characteristics associated with the above specifications. However, the phase shift at the computing signal frequencies, using for example a constant-k design, is too large for the specifications of the simulator. In Figure 8 it should be noted that the filter, by choice of placement, is driven from a very-high-impedance source and is terminated in a practically zero impedance, the summing point of a high-gain d-c amplifier. In a completely mismatched filter system, zero phase shift is obtainable in the transmission band with 180° phase shift in the attenuation region for a single LC section. The undesirable

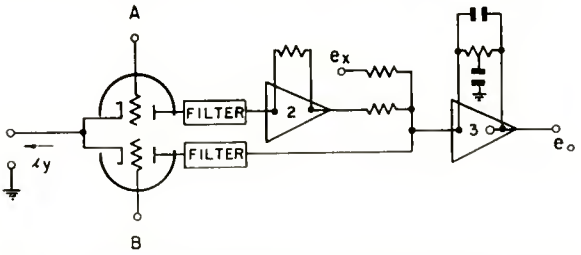


Fig. 8—Amplitude channel showing filter position.

feature of such a system is the existence of resonant gain peaks. In solving a set of simultaneous differential equations there will always be internal closed-loop systems. Closed-loop systems containing multipliers with the above characteristics are potentially unstable. A filter section was designed which preserves the small phase characteristics at low frequencies possible with a mismatched filter, but at the same time minimizes the effect of resonant peak gains. This is done at the cost of adding an additional inductance without realizing any increased attenuation in the attenuation band (Figure 9).

The transfer function for the section shown in Figure 9 is given in Equation (9), where $L_1 > L_2$:

$$\frac{i_o}{i_i} = \frac{1 + j\omega \dfrac{L_1 + L_2}{R}}{1 - \omega^2 L_1 C + j\left(\omega \dfrac{L_1 + L_2}{R} - \dfrac{\omega^3 L_1 L_2 C}{R}\right)} . \tag{9}$$

Figure 10 shows three gain curves for three different values of R. The dashed line is the locus of all possible peaks as a function of the value of R.

Figure 11 is a simplified schematic of the filter used in the multiplier.

The transfer function of the over-all filter shown in Figure 11 is given in Equation (10).

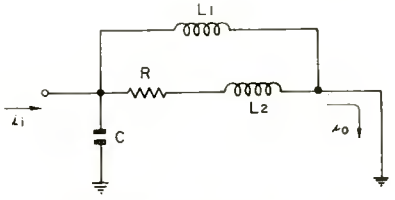


Fig. 9—Simplified filter.

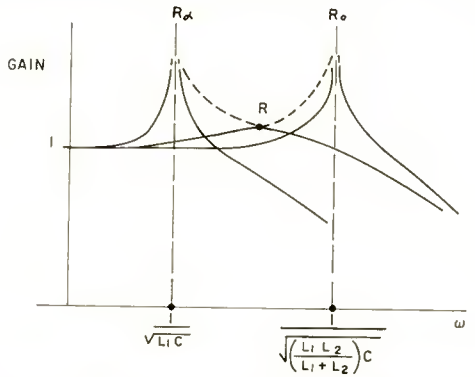


Fig. 10—Gain curves. R_∞ denotes open circuit; R_0 denotes short circuit; R denotes resistance for smallest peak.

$$Y = \frac{i_0}{i_i} = \left[1 + j\omega \frac{L_1 + L_2}{R}\right]\left[1 - \omega^2 C(2L_2 + L_1) + \omega^4 C^2 L_2 L_1 + j\left(\omega \frac{L_1 + L_2}{R} - \omega^3 \frac{L_2 C}{R}(3L_1 + 2L_2) + \omega^5 \frac{C^2 L_2^2 L_1}{R}\right)\right]^{-1} \quad (10)$$

The following values are used in the circuit of Figure 11:

$L_1 = 10$ henrys, $C = .02$ microfarad, $L_2 = 1$ henry, $R = 13{,}000$ ohms. For multiplier computing frequencies $0 \leq \omega \leq 100$ radians per second.

$$Y_C \doteq \frac{1 + j\omega T_1}{1 - \omega_2 T_2 + j\omega T_1}, \quad (11)$$

$$T_1 = \frac{L_1 + L_2}{R}, \quad T_2 = C(2L_2 + L_1).$$

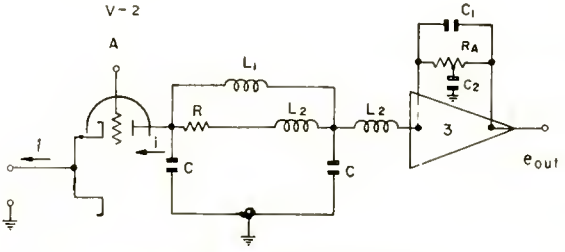


Fig. 11—Over-all filter.

An additional approximation yields

$$Y_C \doteq |Y_C|\, e^{j\phi} \doteq \frac{1 + \omega^2 (T_1^2 - T_2)}{1 + \omega^2 (T_1^2 - 2T_2)}\, e^{-j\omega^3 T_1 T_2} \tag{12}$$

$$\phi_{\omega=100} = -.0116 \text{ degree}$$
$$|Y_C|_{\omega=100} = 1.00239$$
$$\phi_{\omega=20} = -.000093 \text{ degree}$$
$$|Y_C|_{\omega=20} = 1.0001.$$

Computation of two points shows that the scale factor at $\omega = 20$ radians per second has increased by an amount equal to the error of multiplication at $\omega = 0$.

From Equation (10)

$$|Y|_{f=10\text{ kc}} = \frac{1}{5{,}548}.$$

The attenuation at the switching frequency of 10 kilocycles is approximately 75 decibels.

The above computations are for a filter with ideal inductors and capacitors. In practice, of course, the coils contain appreciable resistance. The resistance in the coils will modify T_1 and T_2 but not to any important extent. However, additional terms will be added to the transfer function which will be linear functions of ω and whose coefficients will contain r.

$$r_1 = 600 \text{ ohms, for } L_1,$$
$$r_2 = 60 \text{ ohms, for } L_2.$$

The lagging phase shift contributed by these additional terms is considerably larger in the multiplier computing frequency band than the ϕ of Equation (12). Since rC terms are linear in ω, it is possible to compensate for these lagging terms with a lead network. A lead–lag network is placed across the output amplifier #3 (Figure 11). The amplifier isolates the main body of the filter from the lead–lag network allowing a great deal of flexibility in the choice of elements.

The transfer function for the network in the feedback of amplifier #3 is given in Equation (13).

$$Y_A = \frac{1 + j\,\omega T_3}{1 - \omega^2\, T_3\, T_4 + j\,\omega T_4}, \tag{13}$$

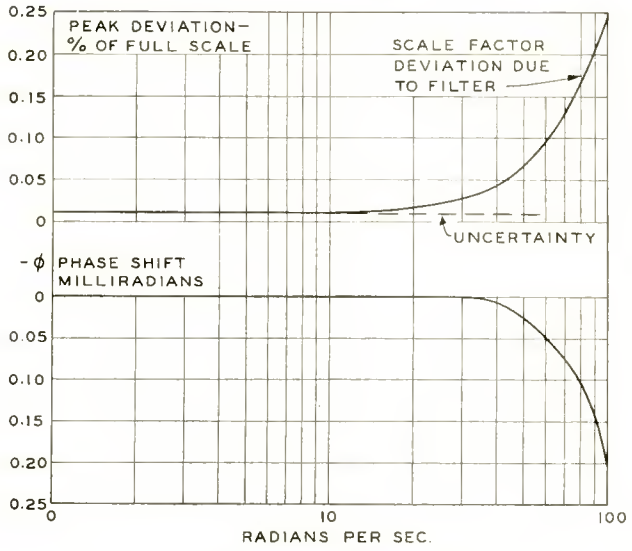


Fig. 12—Multiplier performance.

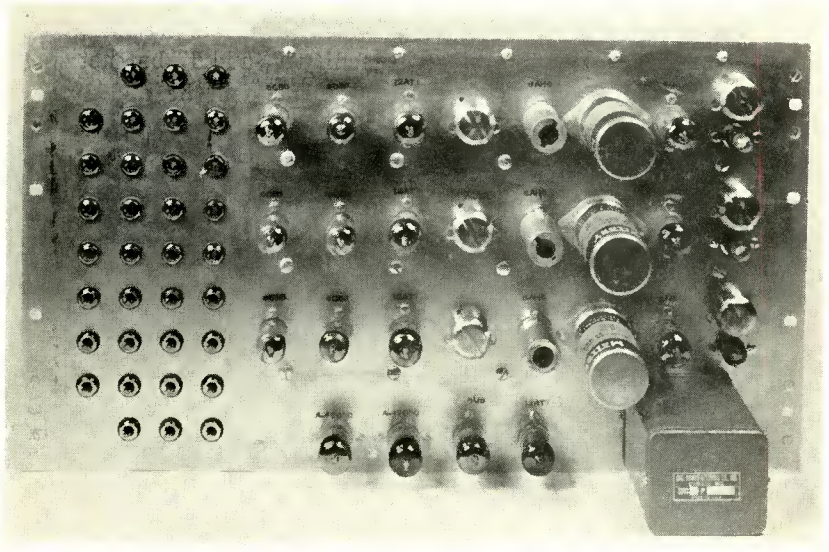


Fig. 13—Multiplier chassis.

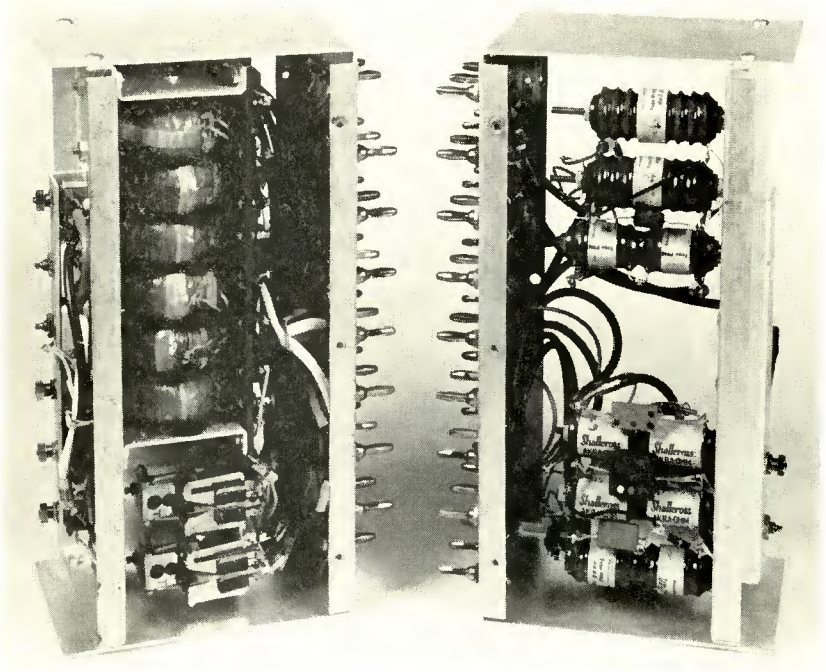


Fig. 14—Multiplier plug-in boxes.

where $T_4 = R_A C_1$,

$$T_3 = \frac{R_A C_2}{4}.$$

In this case, $C_1 = 1{,}000$ micromicrofarads, $C_2 = 5{,}000$ micromicrofarads, and $R_A = 244{,}000$ ohms.

T_3 and T_4 are adjusted to cancel the RC terms in the filter and as much of the nonlinear phase delay as is possible with a straight-line approximation to a cubic. In the attenuation region this network gives a lagging phase shift and at 10 kilocycles adds 5 decibels to the over-all attenuation characteristics.

The measured values of an operational multiplier (Figures 13 and 14) are in close agreement with the calculated quantities above. The gain–phase characteristics over the computing band are shown in Figure 12. The emphasis throughout the design of the filter, as previously stated, has been to minimize the phase shift. If in some specialized application an improved gain characteristic is desired, the

feedback network in the output amplifier may be changed from a phase to a gain equalizer. The gain peak occurs at 500 cycles and is equal to 2.7. Figure 13 is a photograph of a multiplier chassis, and Figure 14 is the plug-in box which contains the filters and wirewound resistors. It should be noted that by simply changing the plug-in box the multiplier chassis may be converted to three operational amplifiers for use as summers and integrators. An all-electronic dynamic systems synthesizer can therefore be built around a single basic unit.

ACKNOWLEDGMENTS

The writer wishes to acknowledge the important contributions and suggestions made during the period of this work by A. W. Vance, J. Lehmann, and E. A. Goldberg. The filter calculations were made by C. C. Shumard, and M. L. Wadlin was responsible for product design.

A NEW LOW-POWER SINGLE-SIDEBAND COMMUNICATION SYSTEM

By

E. A. Laport* and K. L. Neumann†

Summary—*A new single-sideband high-frequency telephone system is described for simplex or duplex operation. It is usable for telephony, manual telegraphy, and teleprinter operation over short and medium distances. It is primarily intended for suppressed-carrier operation, but can also be operated with carrier and one sideband for intercommunication with existing amplitude-modulated stations on the same working frequency. The system is adapted for use by nontechnical personnel for many of the simpler telecommunication requirements around the world, with utmost spectrum conservation.*

Introduction

THE extreme congestion of the high-frequency spectrum calls for leadership by engineers in effecting every possible means of better spectrum utilization. Electromagnetic spectrum space is a limited world resource that must be carefully used to meet constantly expanding radio needs. Conventional methods of double-sideband (AM) radio transmission are wasteful because they use much more bandwidth than is needed to convey the desired intelligence.

The high frequencies (from 3 to 30 megacycles) are particularly important in the world's radio pattern because long distances can be reached for telephone and telegraph with simple and relatively inexpensive equipment. This feature carries with it the equally easy creation of interference to others over long distances. At times a few watts of radiated power will communicate thousands of miles and interfere with other services over vast areas.

There are tens of thousands of high-frequency stations in use. A majority are for short-distance intermittent use for fixed point-to-point and land and marine-mobile private, public and government services. A large part of these services use AM telephony. Bandwidth reduction by one half, which would result if all AM systems were supplanted by single-sideband systems, would represent a substantial contribution to frequency conservation and interference reduction.

Since technical leadership in spectrum conservation must come from

* Radio Corporation of America, New York, N. Y.
† Radiomarine Corporation of America, New York, N. Y.

industry, the system to be described was conceived and developed to make more advanced methods of transmission economical enough and simple enough to compete with existing equipment and eventually to displace it.

Single-sideband (SSB) systems have been in use for three decades. They have been applied virtually exclusively for wire-line carrier systems and long-haul radiotelephone circuits. In the form now in use the cost and complexity has been beyond the means of the mass of high-frequency users. The novelty of the system to be described here (designated SSB-1) is three-fold: Single-sideband suppressed carrier techniques are provided at a cost that is of the same order as conventional amplitude-modulated equipment of comparable power; it can be installed and used by people with little or no technical skill; it is applicable to both simplex and duplex forms of communication. It therefore becomes practicable to use single-sideband transmission for short-haul systems wherever AM is now used.

Comparison of Single-Sideband Suppressed-Carrier and Amplitude Modulation

The advantages of single-sideband over amplitude modulation will be reviewed briefly to indicate the practical value of the new techniques to be described. The comparative spectra, drawn on the basis of frequency and power, are shown in Figure 1, indicating two principal advantages of SSB. These are:

1. Only half as much bandwidth is needed for intelligence transmission because the carrier wave and the redundant second sideband are not transmitted. In receiving, noise is reduced 3 decibels by virtue of one-half the bandwidth needed, and interference within the range of the eliminated sideband is likewise eliminated. Interference coming within the pass-band of the wanted sideband of course is received, but in the random interference situation generally existing, its probability is reduced by one-half.

2. All the power is transmitted in one sideband (Figure 1a) and is not divided between the carrier and the second sideband as in AM, represented in Figure 1b. This higher transmission efficiency amounts to a power gain of 4 (6 decibels) for the wanted single sideband. There is also the reduction of 3 decibels of noise contributed by the 50 per cent reduction in receiver bandwidth. Altogether, then, the power advantage of SSB over AM of the same rating is 8, or a system improvement in signal-to-noise ratio of 9 decibels in typical communication. Under practical conditions the full 9 decibels is usually realized as confirmed by many careful tests. The SSB-1 system with its peak

envelope output of 60 watts, is therefore equivalent in effectiveness to a 500-watt AM transmitter before counting the additional effective power gain of the 20-decibel speech compressor. Speech compression can also be used with an AM transmitter but its effectiveness is better realized with SSB during typical selective fading encountered in wave propagation.

The most important advantage of SSB, not evident from Figure 1, is:

3. Distortion in propagation is greatly reduced. AM transmission requires that rather strict phase and amplitude relations be held between the carrier and its side frequencies.[1] Deviations of phase and

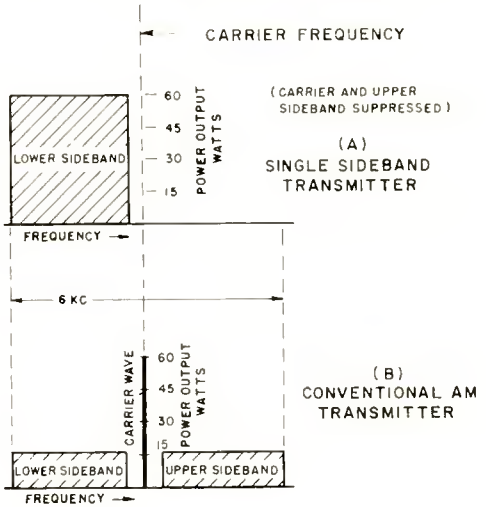


Fig. 1—Comparative frequency–power spectra.

amplitude in propagation introduce distortion and mutilation that is a well-known phenomenon in high-frequency broadcasting and telephony. Single-sideband eliminates this phase and amplitude dependency needed for AM, so that selective fading is much less destructive with SSB. The SSB system gain under this circumstance cannot be related in any way with power gain, because a mutilated signal, however strong, is useless.

There are some additional advantages of less importance, but nevertheless desirable:

4. Single-sideband realizes a considerable saving in power consumption, because power is radiated only during speech transmission.

[1] E. A. Laport, "Characteristics of Amplitude-Modulated Waves," *RCA Review*, Vol. 1, pp. 26-38, April, 1937.

Even on low-power equipment this has value since many users must generate their own power. Furthermore, the SSB-1 can be substituted for AM transmitters of much higher power.

5. Single-sideband has a certain degree of privacy because the carrier is suppressed. Common home receivers cannot resolve such a signal, though communication receivers equipped with beat oscillators often can. This eliminates vast numbers of casual listeners among the general population.

The SSB-1 System and its Adaptions

The production of a low-cost SSB system that is simple enough for use in a wide variety of elementary services required a new approach from the system viewpoint. Many of the customary practices in long-distance usage are unnecessary for single-channel telephone traffic. The question that was resolved in the development of the SSB-1 equipment was: "What is the simplest way in which single-sideband can be realized in a practical and reliable equipment for general use?"

These broad objectives guided the development:

1. Simplex (press-to-talk) telephony, and simplex telegraph. This allows many stations to intercommunicate on one working frequency.

2. Immediate switching to four fixed operating frequencies — transmitter and receiver to be switched together.

3. Operation in the 3 to 15 megacycle range, the frequencies most useful for the short and medium distances.

4. Power output of the order of 50-75 watts on one sideband, with carrier and the second sideband suppressed as completely as practicable.

5. Provision for transmitting carrier and one sideband when desired for communication with stations using AM equipment.

6. Manual synchronization of replacement carrier by vernier adjustment of a precision crystal oscillator.

7. Provision for working with a single antenna at all frequencies.

8. To be operationally as simple as existing AM equipment.

Conversion of communication systems to SSB cannot be expected to be immediate. Existing users, those most likely to appreciate and need the benefits of SSB, now use AM equipment. Therefore the SSB-1 system was designed to work with existing AM stations. To do this, the carrier can be transmitted or suppressed as desired by means of a single switch. When carrier and one sideband are transmitted, the SSB-1 emission can be received on any AM receiver tuned to its frequency. The SSB-1 receiver in turn receives an AM signal, rejects its carrier and one sideband, and processes the wanted sideband in the receiver circuits as though only the one sideband were received. This

realizes the advantage of reduced distortion which accrued from use of SSB.

The SSB-1 transmitter has an audio oscillator for on–off keyed telegraph transmission. The signal appears on the air as an on–off keyed continuous wave that can be received on a conventional communication receiver. When the carrier is not suppressed, a conventional AM receiver delivers the original audio tone to the operator or the transcribing equipment. To receive conventional telegraph from other stations, continuous-wave on–off telegraphy is converted by the SSB-1 to a keyed tone signal when the transmitting frequency of the distant station is within the pass-band of the SSB-1 sideband. If the distant station is frequency-shift keyed, a frequency-shifted tone is delivered by the SSB-1 receiver. The distant station may be carrier-keyed or tone-keyed AM and still be received if its carrier frequency is right.

A network of stations using the SSB-1 system can operate simplex telephony or telegraphy on one working frequency by means of quartz crystals that bring their carrier frequencies within about 75 cycles of each other. The SSB-1 has a vernier frequency control that permits each station on a network frequency to be "netted" to the frequency of the main station. Once so adjusted, all sets will remain virtually synchronized for long periods for both transmission and reception. A frequency discrepancy of up to 75 cycles does not adversely affect the intelligibility of transmitted speech, though the speaker sounds more natural when the replacement carrier is synchronous. Frequency trimming is a panel control that the receiving user can adjust as needed; but this is seldom necessary if the netting procedure is performed occasionally.

The SSB-1 system has four preset operating frequencies selectable by a single switch. Any one station therefore has a day and a night frequency for possible use on two networks.

The system plan adapts the SSB-1 for integration into existing AM networks, SSB networks, or combinations of both. Progressive conversion from AM to SSB can then take place at whatever rate the users desire. Furthermore, it can be used by nontechnical operators. Once adjusted to its working frequencies by a technician, anyone can use the equipment after simple instructions about press-to-talk manipulation, calling procedure, and even frequency changing at prescribed hours.

Description of the Equipment

The system arrangement is represented in Figure 2. The photographs, Figures 3 and 4, show the appearance of the SSB-1 equipment as packaged for fixed station services.

Figure 5 shows the functional relations within the SSB-1 transmitter and receiver. It will be seen from this how the frequency-generating and translating chain is made common to both transmitter and receiver. Figure 6 shows the radio and audio power chassis and Figure 7 shows a complete set of tubes for the equipment.

The SSB-1 is equipped with switching and signaling to three extension telephones that may be located at any distance from the radio set where it is feasible to run the interconnecting wires. When a call arrives for one of the extension telephones, the operator sets the switch to the proper extension. This lights a signal light on the selected extension handset. A separate buzzer or calling bell can also be used for audible signaling if desired. When the extension telephone

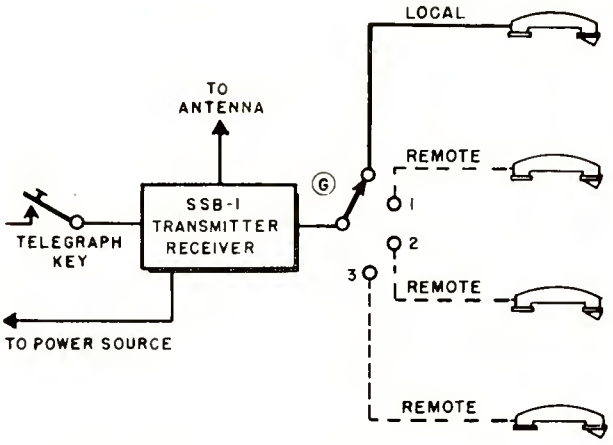


Fig. 2—General arrangement of SSB-1 system.

handset is lifted, a light appears above the extension selector switch indicating to the operator that the called telephone is responding. When a call is to be made from an extension telephone, the signal light on the SSB-1 shows the operator which one is calling. If the radio circuit is clear, the selector switch is set to the indicated extension and the call proceeds by radio. Between calls on the extensions, the operator guards the circuit using the local handset. When an extension is in use, the operator's handset is disconnected. When all extension lights are out, the operator switches to the local handset. There is a loudspeaker built in for standby for calls, and for monitoring when the transmitter is off. A switch transfers the receiver output from loudspeaker to handset.

To meet the needs of most users, controls are minimized as far as

possible and there are no meters. There is a gain control for input speech levels. A peak-radio-frequency indicating light tells the operator when peak levels are being reached. Adjustment is made so that it flashes on peaks of speech. Another gain control regulates received audio output levels. Except for the frequency-change switch, the carrier on-off switch, and the carrier synchronizing control, these are the only adjustments on the transmitter-receiver panel. On the power panel there are only switches for "Loudspeaker-Handset" transfer; switching from local to one of the three extension telephones; transfer to "Telegraph Transmit"; and three power switches for "Power," "Re-

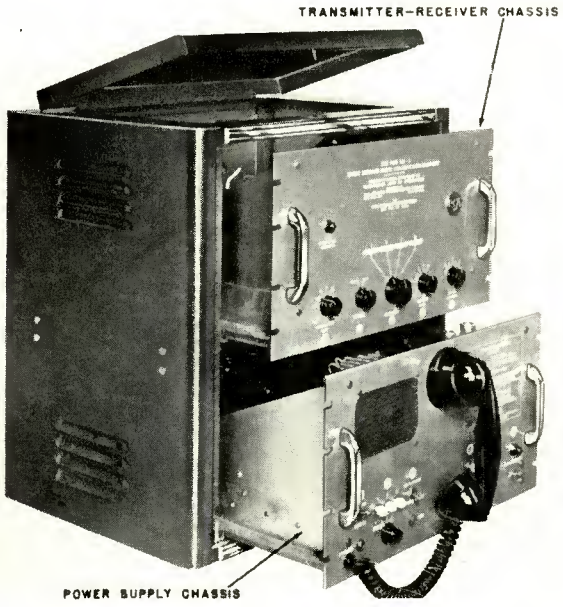


Fig. 3—Accessibility details for SSB-1 (a rear door is not visible).

ceiver" and "Transmitter" on–off. There are jacks for connecting head telephones, and for a telegraph key. These can be seen in Figure 4. All other fixed working adjustments are made internally. The two units are on 19-inch standard rack panels, and arranged for roll-out as seen in Figure 3. There are hinged doors on top and back to reach all tubes and crystals.

A complete SSB-1 equipment weighs 150 lbs, which, with its small size, makes it feasible to fly a new unit to a station requiring servicing and returning an inoperative unit to a central shop where competent servicing is available.

Performance

The peak envelope power output is from 58 to 65 watts depending upon the frequency and the range of antenna impedances. The carrier frequency is held to a tolerance within 0.0005 per cent by quartz crystals. The power input is 85 watts on receive only. On telegraphy at full power output the power input to the SSB-1 is 310 watts. The power input during voice transmission with a moderate value of speech compression is of the order of 240 watts, which is a small fraction of

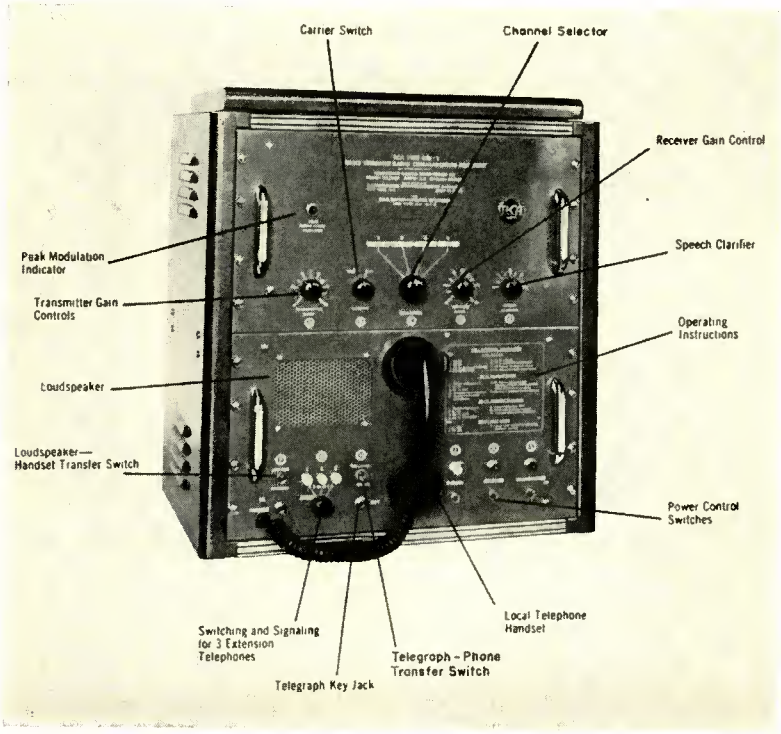


Fig. 4—Identification of essential exterior controls.

that needed for an AM transmitter of equal communication effectiveness.

The carrier frequency control, called a "voice clarifier," is adjustable over a range of 150 cycles at all working frequencies. When a network is correctly set up, this provides an adjustment of ± 75 cycles, a value found to provide fully intelligible speech anywhere within its range, but permitting adjustment to virtual synchronism for most natural reception. The distortion of a single-tone frequency at full modulation (no clipping) is 2.5 per cent at 1,000 cycles.

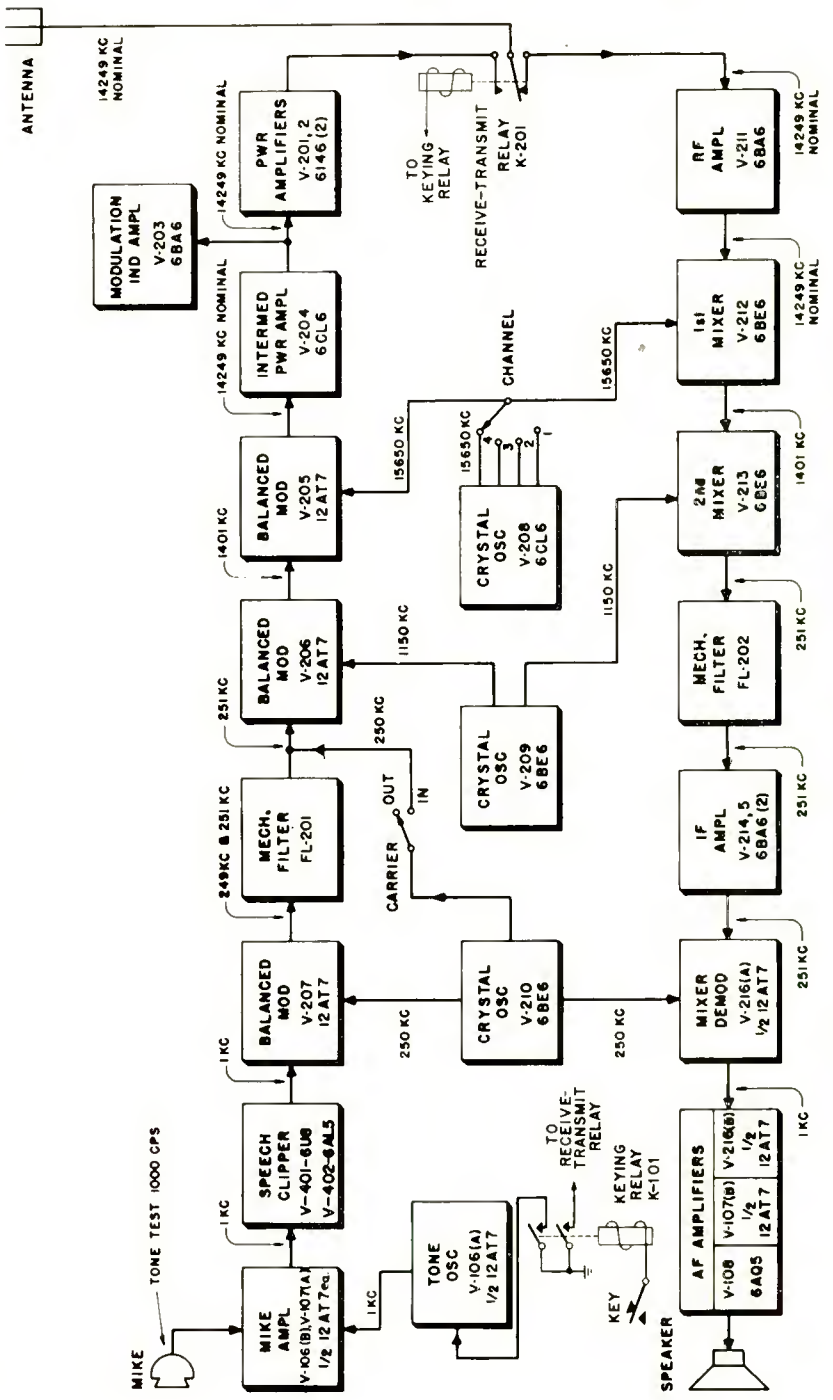


Fig. 5—SSB-1 transmitter, receiver and frequency-control arrangements.

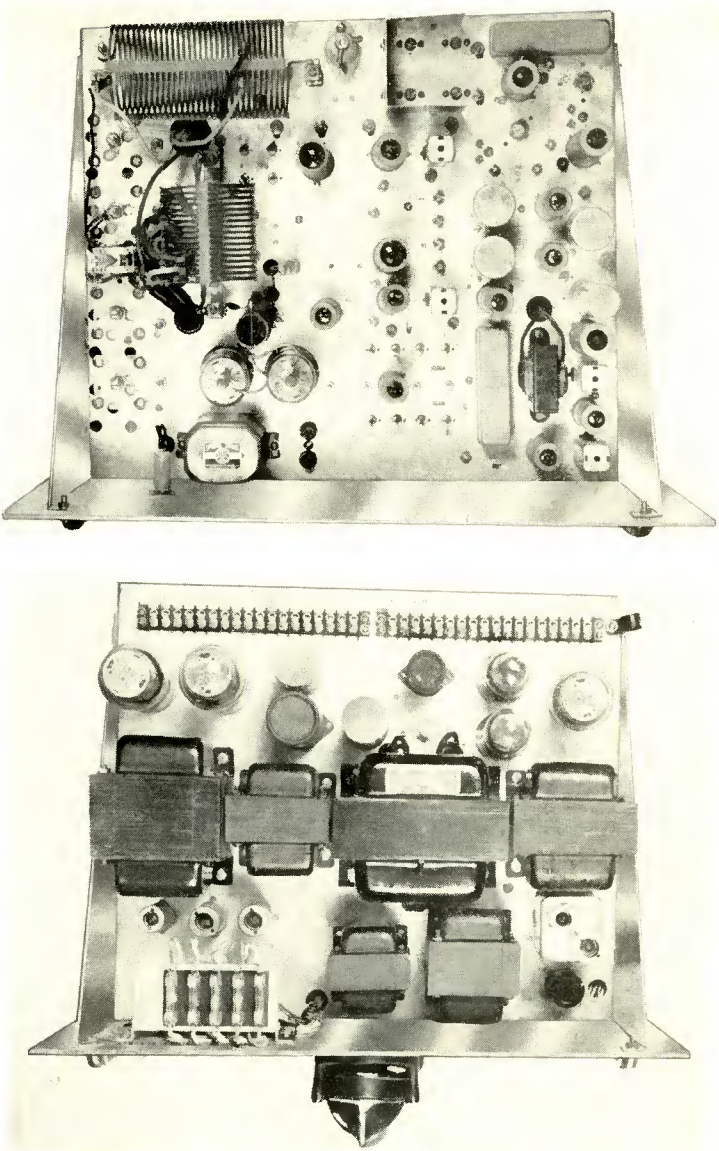


Fig. 6—SSB-1 radio-frequency chassis (top) and power/audio chassis (bottom).

On-line teleprinter operation is available in two ways:

(1) A send–receive teleprinter can be used to (on–off) key the tone oscillator in the SSB-1 transmitter for sending. For receiving, a tone rectifier is needed between the SSB-1 receiver and the teleprinter

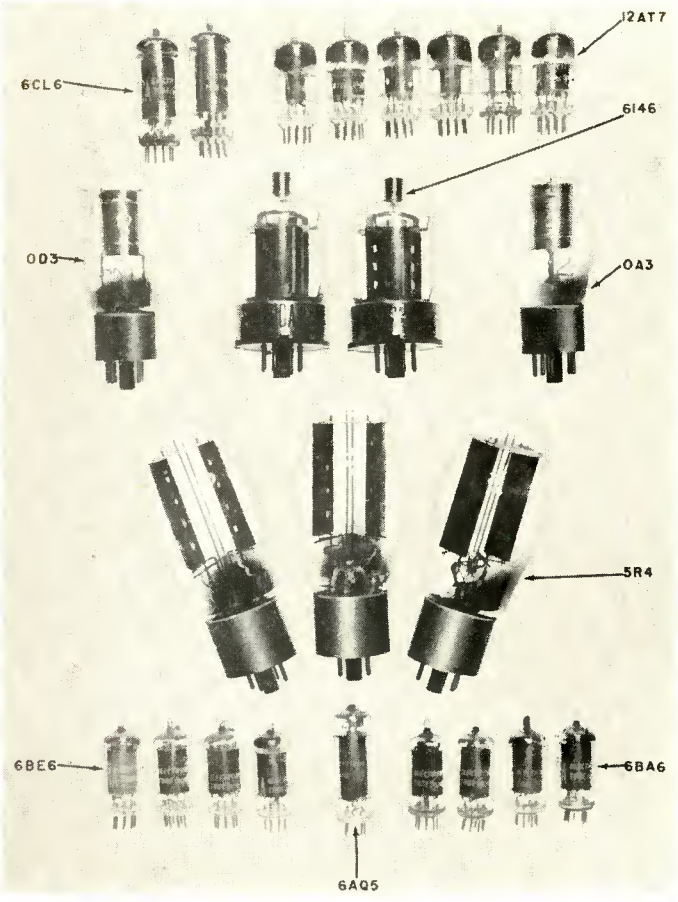


Fig. 7—Complete set of tubes for SSB-1.

to provide direct-current signals. A separate send–receive switch performs the equivalent function of the press-to-talk switch on the hand telephone.

(2) A voice-frequency carrier channel may be used to provide frequency-shift teleprinter signals for transmission, and reception, connected between the SSB-1 and the send–receive teleprinter. A send–receive switch is also used.

The output on transmission of carrier and one sideband is rated at 60 watts peak envelope power, giving 15 watts of carrier power and 15 watts peak of sideband. The carrier and sideband levels are adjusted to be of about equal amplitude on speech peaks for this form of emission, for reception by AM receivers. The unwanted sideband is

suppressed 50 decibels or more relative to the transmitted sideband.

On two-tone test at full modulation with carrier suppressed, a transmitter and receiver together give cross-modulation products about 26 decibels below full level for one tone.

Duplex Operation of Stations

Two SSB-1 sets may be used in 2-frequency duplex as shown in Figure 8, by using one always as the transmitter, and the other as the receiver. Output terminals from the receiver and input terminals to the transmitter are available at 600 ohms for the speech circuits to connect to appropriate telephone control terminal equipment for switching, calling and 4-wire to 2-wire conversion for coordination with normal telephone exchanges. Intercity telephone transmission is available in this way. Provisions for signaling between exchanges, or for intercity subscriber dialing, are external to the SSB-1 equipments.

Several channels of voice-frequency carrier can also be used for teleprinters over an SSB-1 duplex circuit.

Field Tests

Two separate field tests were made on the Mississippi River marine-mobile telephone service of the Radiomarine Corporation of America, using an SSB-1 at their station WCM near Pittsburgh, Pa. The tests used the assigned marine frequencies of 4067, 6455 and 8205.5 kilocycles. An SSB-1 was also installed at their station KE2XQH, 75 Varick Street, New York, N. Y., 320 miles distant. At WCM, switching was provided to connect a 50-watt AM transmitter, a 1,000-watt AM transmitter and the SSB-1 to the same antenna so they could be rapidly interchanged for comparison. With no speech compression in the SSB-1 it was about 10 decibels superior to the 50-watt AM set, and about 6 decibels inferior to the 1,000-watt AM set, as expected. With speech compression in the SSB-1, it is a few decibels superior to the 1,000-watt set, with better intelligibility and relative freedom from selective fading.

During this test period, WCM operators also used the SSB-1 with carrier and sideband to communicate with ships on the Mississippi and Ohio rivers. These ships are equipped with AM sets, a few of which had crystals giving frequencies near enough to that of the WCM SSB-1 to give satisfactory cooperation. Consistently satisfactory service, often superior to that from WCM's 1,000-watt transmitter, was obtained at all distances up to the available maximum when ships were in the Mississippi delta region, some 920 miles distant. This was

a practical test of the advantage of asymmetric sideband emission and the cooperation with existing amplitude-modulated equipment during three weeks' trials.

A later test of twenty days was made with an SSB-1 equipment installed on a river tug, and turned over to the ship's officers to use in their own way. Single-sideband suppressed-carrier operation on two-way simplex service in this test, at various distances in the river system, proved that the system and the equipment gave exceptionally

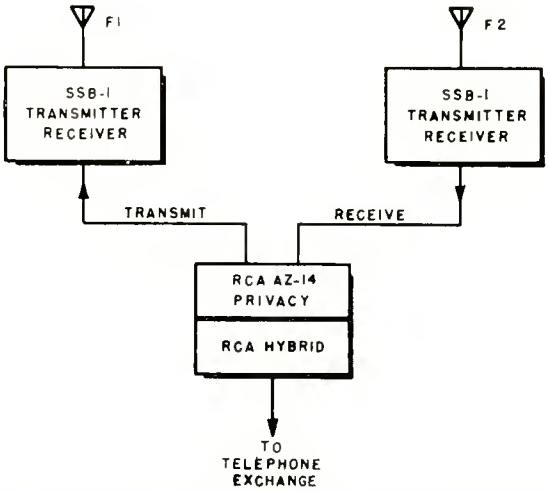


Fig. 8—External facilities for using 2 SSB-1's for duplex telephony.

satisfactory performance and that it was markedly superior to AM under every condition.

Acknowledgments

The authors wish to give credit to members of the Engineering Department of the Radiomarine Corporation of America who helped in the development of this equipment. N. L. Barlow, C. E. Schneider and I. Linker contributed to the electrical design; B. Autry contributed to the mechanical design; W. N. Hayton of Radiomarine station WCM at Pittsburgh and officers of the towboat *Valvoline* contributed to the experimental field tests.

A LOCKED-OSCILLATOR QUADRATURE-GRID FM SOUND DETECTOR

By

Jack Avins and Thomas Brady

RCA Laboratories,
New York, N. Y.

Summary—*This paper describes a quadrature-grid FM detector which employs a pentode tube having a sharp-cutoff suppressor characteristic. At high input signal levels AM rejection and static limiting are provided by grid damping and degeneration, and at low input signal levels by operation as a locked oscillator detector.*

The circuit is relatively simple and capable of performance comparable to that of a ratio-detector audio-amplifier combination.

General Description

THERE have been many attempts to design a simple FM detector circuit around a low-cost pentode. In the past these attempts were unsuccessful because of feedback difficulties and because of plate current rectification leading to poor AM rejection. These problems have been overcome by deliberately setting up the circuit so that it oscillates and thereby functions as a locked-oscillator–quadrature-grid detector for very weak signals. Intermediate and strong signals are demodulated directly with the self-oscillation automatically suppressed. This type of operation leads to improved sensitivity, greater AM rejection, and static limiting over a wider range than would otherwise be possible.

The basic circuit of the detector is shown in Figure 1. It uses a new pentode* having a sharp-cutoff suppressor-grid characteristic. The 4.5-megacycle input signal is applied to the control grid. Space-charge coupling develops a 4.5-megacycle voltage across a tuned circuit connected to the suppressor grid. This voltage is in quadrature with the control-grid signal when that signal is at its center or unmodulated frequency. As the modulation swings the frequency of the input voltage up and down from this center value, the phase of the suppressor-grid voltage swings back and forth from the quadrature position. The combined effect on the plate current of these control and suppressor-grid voltages results in frequency-modulation detection. This operation is similar to that of a conventional quadrature-grid detector.

* Type 6DT6 for 6.3-volt heater supply; type 3DT6 for series-string heaters (3 volts, 0.6 ampere).

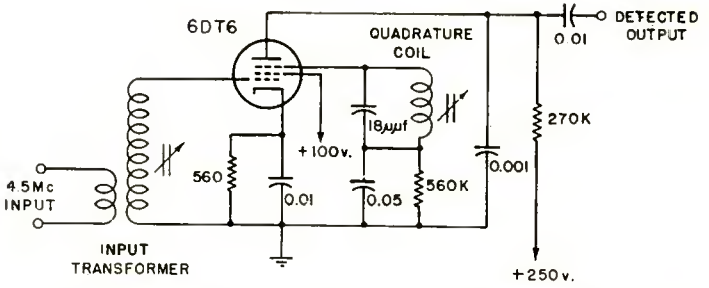


Fig. 1—Locked-oscillator quadrature-grid detector circuit.

Locked Oscillator Mode

At low signal levels the circuit functions as a locked-oscillator detector because of internal feedback between the suppressor and control grid circuits, as shown in Figure 2. The voltage gain from the control grid to the quadrature grid is 3, accompanied by a phase shift of 90 degrees. In the reverse direction there is coupling from the suppressor to the control grid through the $G_1 - G_3$ interelectrode capacitance. Sufficient energy is returned to the input to maintain oscillation at a level of about one volt r-m-s at the control grid and to develop the required suppressor bias. This action provides a 4.5-megacycle level substantially independent of reduction in the injected signal from the driver. A signal with an amplitude down to approximately one-third volt r-m-s is sufficient to lock the oscillator and pull it to full deviation.

Figure 3a shows the normal demodulated output of the detector operating in the locked oscillator mode. The output is linear over the

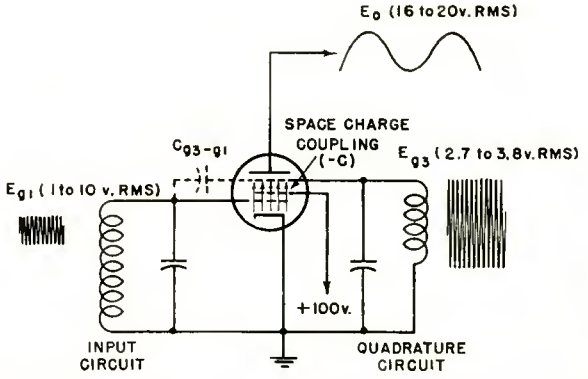


Fig. 2—Quadrature-grid space-charge coupling and G_3–G_2 feedback path. For high values of input, the signal levels at the input and suppressor grids increase as shown.

range in which the detector is locked, with a beat occurring outside of this range. The addition of 30 per cent amplitude modulation results in the output characteristic shown in Figure 3b. The AM rejection is high near the middle of the characteristic but falls off near the extremes as the input falls below the minimum locking level during the downward part of the AM cycle.

The sensitivity of the detector to small signals is dependent in part on the capacitance from G_1 to G_3. Increased capacitance makes the detector less sensitive so that a higher input is required to lock the oscillator over the full-frequency swing. Abnormally low capacitance will cause nonlinear operation and instability. In practice, the 0.1 micromicrofarad $G_1 - G_3$ capacitance supplies the major part of the required feedback capacitance. This capacitance is not critical and variations in tube capacitance cause negligible changes in performance.

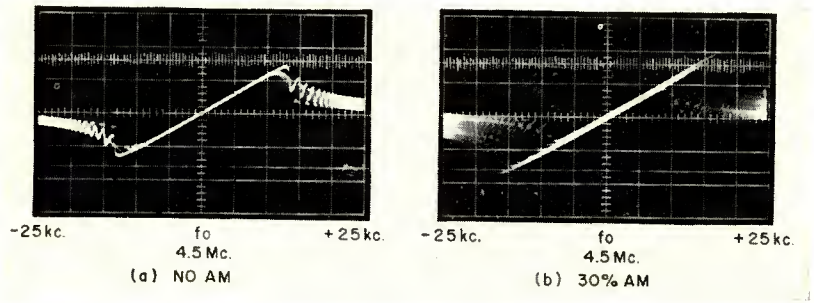


Fig. 3—Demodulated output below threshold of full-deviation sensitivity.

High-Level Operation

At intermediate and high signal levels, the input overrides the self-oscillation and the normal signal levels are supplied by the input rather than by self-oscillation, with no change taking place in the detected output. The detected output remains substantially constant because grid current loading of the input tuned circuit tends to keep the input signal constant, and cathode degeneration augments the bias on both the control and the suppressor grids.

Figure 4 shows how the output characteristic varies as a function of input signal level. The series of oscillograms in Figure 4a are for a relatively wide deviation of 50 kilocycles to show the region near the extremes of deviation. The oscillograms of Figure 4b are for 25-kilocycle deviation and 30 per cent amplitude modulation.

The linearity and AM rejection characteristics are apparent.

The damping effect of the grid current in tending to keep the input

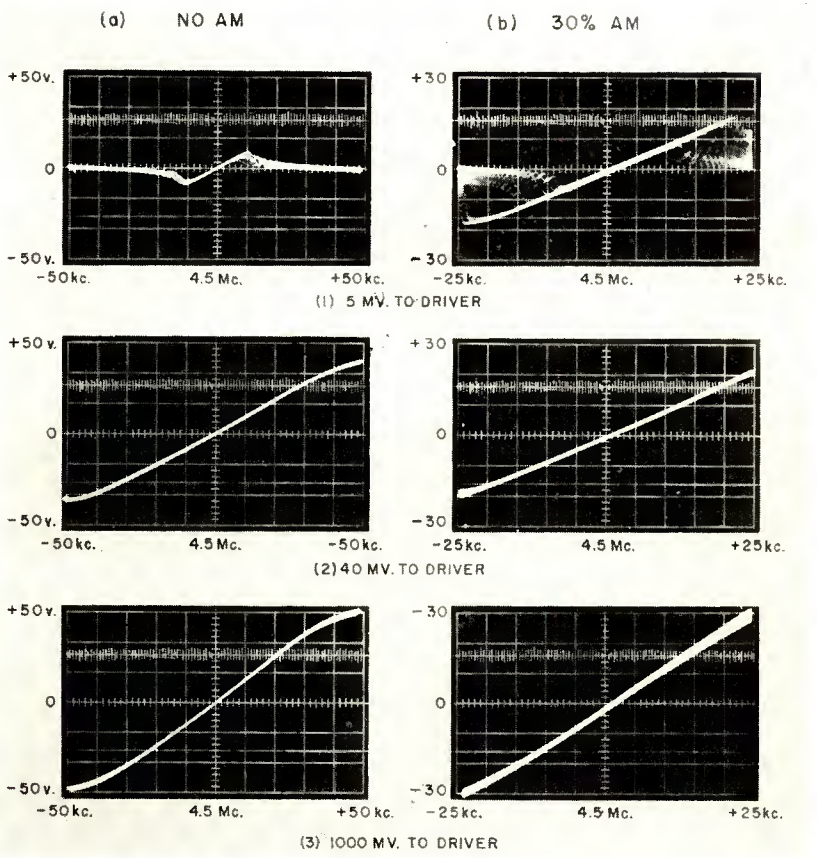
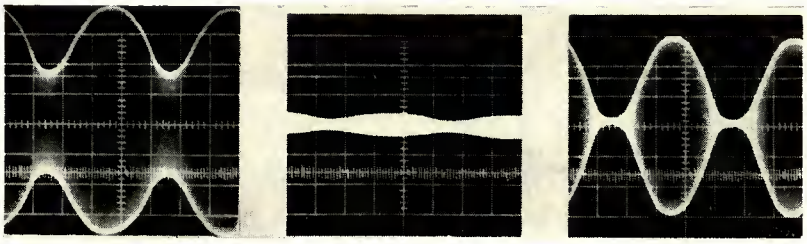


Fig. 4—Demodulated output as a function of input signal level. The input levels refer to the driver grid shown in Figure 7.

signal relatively constant is illustrated in Figure 5. These oscillograms were taken with relatively high driver input; Figure 5a represents a grid swing of approximately 25 r-m-s volts (measured with the detector tube cold) which is reduced by grid damping to about 3 r-m-s volts. The resulting reduction in the amplitude modulation is observed in the right-hand portion of the figure; this oscillogram was taken with the gain of the oscilloscope increased ten times.

Degeneration

The AM rejection at high levels is further augmented by cathode degeneration. As shown in Figure 6, a voltage is developed across the r-f bypassed cathode resistor corresponding to the amplitude-modulation envelope of the signal. This voltage affects the cathode-suppressor



Undamped input (Scope gain × 1) | Damped input (Scope gain × 1) | Damped input (Scope gain × 10)

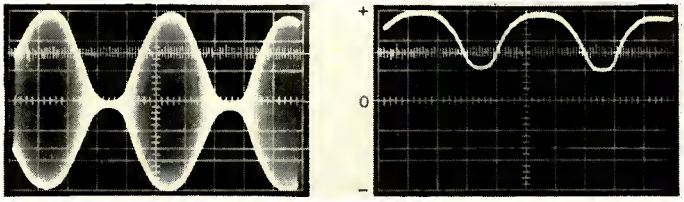
Fig. 5—Effect of grid-current damping on the input level and AM percentage.

potential so as to prevent changes in cathode current from affecting the plate current; the plate current remains fixed with variations of signal amplitude, while the screen current increases from about 3 to 6 milliamperes over the operating range. The cathode resistance required to obtain optimum rejection is not critical and experience with a large number of tubes indicates that a fixed resistor will provide good AM rejection.

Circuit Considerations

Efficient operation under weak signal conditions requires that a high impedance be presented to the input grid. This results in minimum G_1 to G_3 capacitance being required for maintenance of oscillation in the locked-oscillator mode. At the same time, high grid impedance contributes to improved limiting in the grid circuit at high input levels.

Either double-tuned or single-tuned transformer coupling can be used between the driver and detector. Somewhat better performance can be obtained with a double-tuned transformer; however, the degree



(a) Input signal, 4.5 megacycles, 100 per cent amplitude modulated.

(b) Cathode voltage.

Fig. 6—The component of cathode voltage corresponding to the input envelope, which augments the AM rejection.

of improvement obtained is not commensurate with the additional cost and greater alignment complexity.

The circuit shown in Figure 7 has been found to yield good performance. A single-tuned transformer provides a step-up between the driver plate and the detector input. The input is tuned with the tube and wiring capacitances in order to obtain a high input impedance. The turns ratio of the input transformer affects the sensitivity, AM rejection, and the low-level operation. An increase in the step-up ratio to the grid causes a loss in both sensitivity and AM rejection. A decrease in the step-up produces a relatively low grid impedance which in turn decreases the feedback; this adversely affects linearity and stability in the locked-oscillator region. The stability of the driver is also dependent upon the turns ratio, with an increased stepup resulting in improved stability.

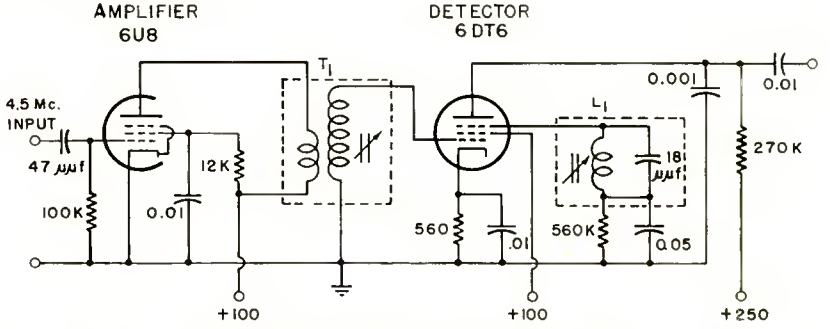


Fig. 7—Typical locked-oscillator and quadrature-grid detector circuit using a single-tuned input transformer.

A capacitance of 18 micromicrofarads is used to tune the quadrature-grid circuit. A slight improvement in AM rejection may be obtained by reducing this value, but this tends to decrease the linearity at high levels. An increased audio output may be obtained by an increase in the Q of the quadrature tuned circuit but at the expense of linearity, bandwidth, and sensitivity. The required suppressor-grid bias is developed across a 560,000-ohm resistor. Lower values of this resistor tend to cause nonlinearity due to suppressor-grid current loading of the quadrature tuned circuit, while higher values adversely affect operation at low levels. The suppressor-grid resistor is bypassed for the amplitude-modulation frequencies.

As discussed previously, the cathode resistor is not critical and a fixed value between 560 and 820 ohms may be used, depending on the other parameters of the circuit. The position of the AM crossover on the discriminator characteristic is determined by the cathode resistor and the value is chosen to center the crossover.

Alignment

The alignment of the detector may be readily accomplished with off-the-air signals. The quadrature grid is adjusted initially for maximum undistorted sound output at a normal signal level so that grid-current damping will mask the effects of control-grid circuit detuning. The signal is then reduced until distortion or noise occurs and the grid circuit and the 4.5-megacycle takeoff are then tuned to provide minimum distortion and noise, and maximum sound.

If a frequency-modulated signal generator and oscilloscope are used, the pattern of Figure 4b2 is obtained by adjustment of the quadrature-grid tuning. A reduction of the input will yield the pattern of Figure 3a which is centered about the operating frequency by means of the interstage transformer tuning. The driver input circuit is tuned to obtain maximum width of the lock-in region and the detector input grid readjusted if necessary to maintain a centered response.

Audio-Frequency Amplifier

It has frequently been considered desirable to provide a phono connection in a television receiver. This can be done with the new circuit by using the detector as an audio amplifier. By feeding the low-level audio signal to the suppressor grid, a gain of approximately 35 is obtained. Switching is relatively simple because the suppressor grid bypass condenser can be lifted off ground and used as the input coupling condenser. A resistor can be placed in series with this lead to suppress any tendency for oscillation at 4.5 megacycles.

Table I—Supply Requirements and Output Characteristics

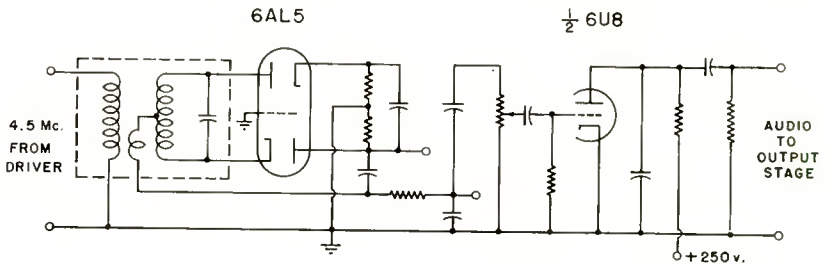
	Min. Signal	Max. Signal
Screen Current	3.8 ma.	5.7 ma.
Plate Current	0.3 ma.	0.3 ma.
Input Grid Current	0.1 ma.	0.5 ma.
Audio Frequency Output (25 kc deviation)		
$E_{bb} = 250$ v.; $R_L = 0.27$ megohm	16 v. r-m-s	21 v. r-m-s
$E_{bb} = 450$ v.; $R_L = 1.0$ megohm	28 v. r-m-s	36 v. r-m-s

Performance

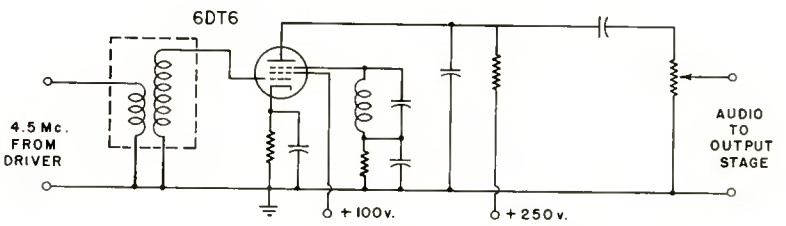
Figure 8 shows the circuit of the new detector in comparison with that of a ratio detector. Both circuits are driven from a similar driver tube, so that this part of the circuit is not shown. Similarly the output of each circuit drives the audio power output tube. The reduction in the number of components is apparent. The supply requirements and output characteristics are indicated in Table I.

Returning the plate load to the B boost supply and increasing its value to 1 megohm will increase the output, so that a 30 per cent modulated signal will provide a peak-to-peak swing of 30 volts on the grid of the output tube.

Performancewise, as installed in a typical television receiver, the new circuit has yielded quieting and maximum deviation sensitivities comparable to those obtained with the ratio detector. In general, the



(a) ORIGINAL SOUND DETECTOR



(b) DEVELOPMENTAL SOUND DETECTOR

Fig. 8—Comparison of the ratio detector and locked-oscillator detector circuits, as employed in a typical receiver.

maximum sensitivity tends to be improved with the new circuit because of the efficient operation provided by the locked-oscillator mode.

Extensive tests with typical impulse noise sources have shown the new circuit to provide improved quieting both for weak and strong signal conditions. The static limiting is particularly effective in preventing changes in volume with tuning and in preventing "blasting" effects such as are encountered under some multipath conditions.

RCA TECHNICAL PAPERS†

Third Quarter, 1955

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

"All-Pass Amplifier," H. J. Woll, *Electronics* (July) 1955
"Basic Problems in Film Pickup for TV Broadcasting," E. M. Gore, *Elec. Eng.* (July) 1955
"A Beam Power Tube for Ultra-High-Frequency Service," W. P. Bennett, *RCA Review* (September) 1955
"Class B Operation of Transistors—1. Design Considerations," K. E. Loofbourrow, *Electronic Design* (July) 1955
"Class B Operation of Transistors—II. Basic Circuits," K. E. Loof-Bourrow, *Electronic Design* (August) 1955
"Color Effects and How It Works," C. R. Monro, *Broadcast News* (August) 1955
"Color Television Luminance Detail Rendition," W. G. Gibson and A. C. Schroeder, *Proc. I.R.E.* (August) 1955
"Control of Light Intensity in Television Projectors," K. Sadashige and B. F. Melchionni, *Jour. S.M.P.T.E.* (August) 1955
"Design of D. C. Tractive Electromagnets," M. R. Alexy, *Electrical Manufacturing* (July) 1955
"Design Problems of VHF Repeater Stations," J. R. Neubauer, *Wire & Radio Communications* (August) 1955
"Differential Amplifier for Null Detection," W. E. Barnette and L. J. Giacoletto, *Electronics* (August) 1955
"Differential Gain Tests TV Color," J. O. Schroeder, *Electronics* (August) 1955
"The Direct Conversion of Radiation into Electrical Energy," J. J. Loferski, P. Rappaport, and E. G. Linder, *RCA Industry Service Laboratory Research Bulletin RB-3* (July 13) 1955
"Dislocations in Germanium," S. G. Ellis, *Jour. Appl. Phys.* (September) 1955
"The Effects of Junction Shape and Surface Recombination on Transistor Current Gain—Part II," K. F. Stripp and A. R. Moore, *Proc. I.R.E.* (July) 1955
"Electrolytic Etching at Small-Angle Grain Boundaries in Germanium," S. G. Ellis, *RCA Industry Service Laboratory Research Bulletin RB-1* (July 6) 1955
"Energy Levels of a Crystal Modified by Alloying or by Pressure," R. H. Parmenter, *Phys. Rev.* (September 15) 1955
"The Engineer and His Professional Society," B. E. Shackelford, *RCA Engineer* (August-September) 1955
"Extending UHF-TV with Booster Amplifiers," J. Epstein, W. C. Morrison, and O. M. Woodward, Jr., *Electronics* (July) 1955

† 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.

"144-Megacycle Transmitter Part II: Operation and Adjustment," R. M. Mendelson, *RCA Ham Tips* (August) 1955

"On the Problem of Optimum Detection of Pulsed Signals in Noise," A. H. Benner and R. F. Drenick, *RCA Review* (September) .. 1955

"Optical Projection of Radiant Power," L. T. Sachtleben, *RCA Engineer* (August-September) 1955

"Performance of a Radio-Frequency Alloy Junction Transistor in Different Circuits," L. J. Giacoletto, *RCA Industry Service Laboratory Bulletin LB-989* (September 2) 1955

"Periodic Magnetic Field Focusing for Low-Noise Traveling-Wave Tubes," K. K. N. Chang, *RCA Industry Service Laboratory Research Bulletin RB-11* (September 8) 1955

"Periodic Magnetic Field Focusing for Low-Noise Traveling-Wave Tubes," K. K. N. Chang, *RCA Review* (September) 1955

"Photoconductivity of the Sulfide, Selenide, and Telluride of Zinc or Cadmium," R. H. Bube, *RCA Industry Service Laboratory Research Bulletin RB-5* (August 9) 1955

"Plasma Oscillations at Extremely High Frequencies," M. A. Lampert, *RCA Industry Service Laboratory Research Bulletin RB-12* (September 14) 1955

"Plasma Oscillations in Electron Beams," K. G. Hernqvist, *Jour. Appl. Phys.* (August) 1955

"Point-Contact Transistor Amplifiers," R. F. Schwartz, *Audio* (July) 1955

"Radiative Transitions in Semiconductors," R. Braunstein, *Phys. Rev.* (Letter to the Editor) (September 15) 1955

"Radiative Transitions in Semiconductors," R. Braunstein, *RCA Industry Service Laboratory Research Bulletin RB-4* (August 4) 1955

"The RCA 'Ampliphase' Fifty," C. J. Starner, J. Q. Lawson, and C. D. Mulford, *Broadcast News* (August) 1955

"The RCA Color Corrector," L. Shapiro, *RCA Engineer* (August-September) 1955

"RCA Engineers Design First Air-Force Equipment Using Transistors," D. E. Shumaker, *RCA Engineer* (August-September) 1955

"Reliability Engineering," C. M. Ryerson, *Jour. S.M.P.T.E.* (July) . 1955

"The Role of the Mechanical Engineer in Electronic Equipment Design," T. G. Greene and P. C. Harrison, *RCA Engineer* (August-September) 1955

"Ser-Cuits," W. W. Cook, *Service* (July) 1955

"Simplified Design Procedures for Tuned Transistor Amplifiers," C. C. Cheng, *RCA Review* (September) 1955

"Some Aspects of Thermal Conversion in Germanium," H. Kroemer, *RCA Industry Service Laboratory Research Bulletin RB-2* (July 8) 1955

"Some Notes on Integration of Color Equipment and Existing Monochrome Facilities," A. H. Lind, *Broadcast News* (August) .. 1955

"Speculations on the Energy Band Structure of Zinc-Blende-Type Crystals," F. Herman, *RCA Industry Service Laboratory Research Bulletin RB-8* (August 23) 1955

"Stabilization of Pulse Duration in Monostable Multivibrators," A. C. Luther, Jr., *RCA Review* (September) 1955

"Surface Treatment of Silicon for Low-Recombination Velocity," A. Moore and H. Nelson, *RCA Industry Service Laboratory Bulletin LB-991* (September 23) 1955

"The Synchroguide: A Design History," S. Tourshou, *RCA Engineer* (August-September) 1955

"System Parameters Using Tropospheric Scatter Propagation," H. H. Beverage, E. A. Laport, and L. C. Simpson, *RCA Review* (September) 1955

"A Test Set for Transistor Performance Measurement at 455 Kc," T. M. Scott, L. A. Freedman, and D. D. Holmes, *RCA Industry Service Laboratory Bulletin LB-986* (August 2) 1955

"Three Vidicon Color Film System," K. H. Hardiman, N. L. Hobson, B. F. Melchionni, and A. Reisz, *Broadcast News* (August) .. 1955

"A Time-Sampling and Amplitude-Quantizing Tube," R. P. Stone, C. W. Mueller and W. M. Webster, *Proc. I.R.E.* (August) 1955

"Transistor Fabrication by the Melt-Quench Process," J. I. Pankove, *RCA Industry Service Laboratory Bulletin LB-985* (July 6) .. 1955

"Uniform Planar Alloy Junctions for Germanium Transistors," N. H. Ditrick and C. W. Mueller, *RCA Industry Service Laboratory Bulletin LB-990* (September 9) 1955

"Uniform Strains and Deformation Potentials," R. H. Parmenter, *Phys. Rev.* (September 15) 1955

"Whiskers on 'Gangs'—a Metallurgical Mystery," L. Pessel and W. D. Rhoads, *RCA Engineer* (August-September) 1955

"Wide-Range Electronic Tuning of Microwave Cavities," F. R. Arams and H. K. Jenny, *Proc. I.R.E.* (September) 1955

"WR-61A Color-Bar Generator Design," S. Wlasuk, *RCA Engineer* (August-September) 1955

Correction:

In the paper entitled "System Parameters Using Tropospheric Scatter Propagation," by H. H. Beverage, E. A. Laport, and L. C. Simpson, which appears on pages 432-457 of the September 1955 issue, the following corrections should be noted:

In the caption for Figure 14, the phrase which reads "diversity gain 10 decibels" should be changed to read "no diversity gain."

In the caption for Figure 15, the phrase which reads "50 decibels 50 per cent of the time" should be changed to read "50 decibels 99 per cent of the time."

Expression (8), page 447, should not be an equation; it should read

$$W_r - F - 10 \log_{10} I + 204 - 12$$

AUTHORS

JACK AVINS received the A.B. degree in Physics from Columbia University in 1932. From 1931 to 1934 he was an assistant in the Physics Department at Columbia. From 1935 to 1938 he was employed by John F. Rider, Publisher, Inc. and from 1939 to 1941 by Service Instruments, Inc., where he designed radio test equipment. In August 1941 he reported for active duty in the Signal Corps and during 1942 and 1943 he served as senior American radar instructor at the Military College of Science in England. In 1944 he became Chief of the Radar Division, Signal Corps Publications Agency, Fort Monmouth. In January 1946, he joined the Industry Service Laboratory, RCA Laboratories, New York, N. Y., where he has been engaged in receiver design. Mr. Avins is a Member of Phi Beta Kappa and a Senior Member of the Institute of Radio Engineers.

W. R. BEAM—(See *RCA Review*, Vol. XVI, No. 2, June 1955, p. 315.)

THOMAS J. BRADY, JR. received the B.E.E. degree from Manhattan College in 1954, and is taking graduate work at the College of the City of New York at present. Beginning in 1943 he served a total of 7 years as an electronic technician and as an instructor in the U. S. Navy. He worked in the quality control department of Loral Electronics, New York and joined the RCA Industry Service Laboratory in 1954. He is currently working on television receiver circuitry. Mr. Brady is a member of Eta Kappa Nu, the Institute of Radio Engineers, and the American Institute of Electrical Engineers.

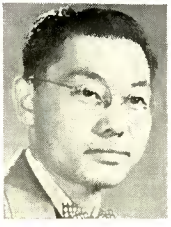
DAVID W. EPSTEIN received the B.S. degree in Engineering Physics from Lehigh University in 1930. In 1934 he received an M.S. and in 1937 a D.Sc. in Electrical Engineering from the University of Pennsylvania. He joined the Radio Corporation of America in 1930 and was associated with the Research Department of RCA Victor Company, Inc., from 1930-35, the Research Department of RCA Manufacturing Company from 1935-41, and RCA Laboratories from 1941 to date. Dr. Epstein is a Member of the American Physical Society, a Fellow of the Institute of Radio Engineers, and a Member of the Optical Society of America and Sigma Xi.

HUNTER C. GOODRICH received the B.A. degree in Physics from Wayne University in 1942. From 1942 to 1945 he was a civilian engineer with the Signal Corps Engineering Laboratory in Detroit working on radio interference measuring equipment. In 1945 he joined the Radio Corporation of America where he has worked on advanced development problems in automobile radio, color television, and transistorized television circuits. Mr. Goodrich is a member of Sigma Pi Sigma and a Senior Member of the Institute of Radio Engineers.

EDWARD O. JOHNSON—(See *RCA Review*, Vol. XVI, No. 1, March 1955, p. 173.)

PETER E. KAUS received the B.S. degree from the University of California in 1947, the M.S. degree from U.C.L.A. in 1952 and is now working toward his Ph.D. From 1951-1953 he was a teaching assistant at U.C.L.A. In 1954 he joined RCA Laboratories in Princeton where he has been working in the field of electron optics. He is a member of the American Physical Society and of the Society of Sigma Xi.

EDMUND A. LAPORT—(See *RCA Review*, Vol. XVI, No. 3, September 1955, p. 485.)



T. MURAKAMI received the B.S. degree in E.E. from Swarthmore College in 1944, and the M.S. degree from the Moore School of Electrical Engineering, University of Pennsylvania in 1947. From 1944 to 1946 he was an assistant and research associate in the Department of Electrical Engineering at Swarthmore College. Since 1946 he has been with the Advanced Development Section of the RCA Victor Television Division, Camden, N. J., working on radio frequency circuit development. Mr. Murakami is a Senior Member of the Institute of Radio Engineers and a member of Sigma Xi.

KARL L. NEUMANN received the B.S. degree in Electrical Engineering from New York University in 1949. He joined the Engineering Department of the Radiomarine Corporation of America in 1940. From 1943 to 1946 he served in the U. S. Coast Guard as an officer. At present, he is Supervisory Engineer of the Radiotelegraph Division. In this activity he has supervised development of both military and commercial high power telephone and telegraph transmitters as well as compulsory and emergency shipboard communications equipment. Mr. Neumann is a member of the Institute of Radio Engineers.

FREDERICK H. NICOLL received the B.Sc. degree in Physics from Saskatchewan University, Canada, in 1929 and the M.Sc. degree in 1931. He held an 1851 Exhibition Scholarship to Cambridge University, England for three years research and received the Ph.D. degree from that university in 1934. He was a research physicist with Electric and Musical Industries, Ltd. in London from 1934 to 1939. From 1939 to 1941 he was with the RCA Victor Division, Radio Corporation of America, at Camden, N. J. as a research engineer. Since 1941 he has been with the RCA Laboratories in Princeton, N. J. engaged in research on cathode-ray tubes and electron optics. Dr. Nicoll is a Senior Member of the Institute of Radio Engineers and a Member of the American Physical Society and of Sigma Xi.

RICHARD W. SONNENFELDT received the B.S. degree in Electrical Engineering in 1949 at Johns Hopkins University. From 1941-1943 and from 1946-1949 he designed control circuits and systems for the Charles Electric Company of Baltimore. In 1949 he joined the RCA Victor Division and has been engaged in advanced development work on monochrome and color television receivers and circuits. Mr. Sonnenfeldt is a member of Tau Beta Pi, Omicron Delta Kappa, and a Senior Member of the Institute of Radio Engineers.

SIDNEY STERNBERG received the B.S. degree in Physics in 1943 from the City College of New York. He studied Communications at M.I.T. in 1944 and received the Master's degree in Electrical Engineering in 1949 from New York University. From 1945 through 1946 he was working with military radar and communications equipment as an officer in the U. S. Navy. He was associated with the Office of Naval Research from 1947 through 1949 when he worked on simulators and computers, including research on general purpose digital computers. In 1950 he joined RCA Laboratories, where he has been engaged in analog computer research, feedback circuitry and general control problems. He is presently engaged in information systems' research and data-handling problems. Mr. Sternberg is a member of Sigma Xi.

D. D. VANORMER—(See *RCA Review*, Vol. XVI, No. 1, March 1955, p. 175.)

RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH • ENGINEERING

DECEMBER 1955

INDEX

VOLUME XVI

TABLE OF CONTENTS

March

June

AUTHORS, VOLUME XVI