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 XXI SEPTEMBER 1960 NUMBER 3

CONTENTS

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

RCA REVIEW

BOARD OF EDITORS

A COMPATIBLE STEREOPHONIC SYSTEM FOR THE AM BROADCAST BAND*

By

J. Avins, L. A. Freedman, F. R. Holt,
J. H. O'Connell, J. O. Preisig and R. N. Rhodes

RCA Laboratories
Princeton, N. J.

Summary—*A number of methods have been proposed for stereo broadcasting in the AM band. These are classified as either "multiplicative" or "additive," depending upon whether the quadrature sidebands representative of the stereo information are modulated by the envelope information or are essentially independent of the envelope modulation. The several system possibilities are evaluated with respect to compatibility, out-of-band radiation, stereo performance, signal-to-noise ratio, and receiver simplicity.*

The preferred system uses a pre-emphasized $L - R$ signal to frequency modulate the carrier; the resultant signal is amplitude modulated with $L + R$. Stereophonic field test transmissions over Radio Station WNBC, New York City, have demonstrated compatibility with existing monophonic receivers and excellent reception on stereophonic receivers of simple design.

Introduction

THE wide use of stereophonic recordings for home entertainment has led to interest in providing compatible stereophonic broadcasting in the AM broadcast band. In the evaluation of possible systems, consideration must be given to (1) *Stereo Performance*: Distortion (intermodulation and harmonic), crosstalk, and signal-to-noise ratio; (2) *Compatibility*: The performance of a monophonic receiver during a stereo transmission, distortion and signal-to-noise ratio, and the effects of misalignment and mistuning; (3) *Out-of-band Radiation*: Radiation in adjacent channels and the effect on receivers tuned to these channels; (4) *Transmitter Design*: Conversion of transmitters for stereophonic transmission; (5) *Receiver Design*: Cost, performance, and reliability.

This paper describes the more important of the various possible systems and discusses their performance with respect to the above factors. Basic System Concepts, Additive and Multiplicative Systems, Receiver Considerations, and a detailed discussion of the AM-FM System appear in the body of the paper. Appendixes A through C deal with Broadcast Standards, Audio Considerations, and Computer Calculations; Appendixes D through F describe details of the Quad-

* Manuscript received August 1, 1960.

rature and Independent Sideband, Modified Quadrature, and AM-FM/AM Systems.

As a result of the work described, the AM-FM System was found to be superior with respect to compatibility, signal-to-noise ratio (implying no reduction in station coverage), and receiver simplicity. This system was proposed to the United States Federal Communications Commission in November, 1959.*

Basic System Concepts

Many systems using two essentially full-bandwidth audio channels have been proposed over the years for accomplishing compatible AM stereo broadcasting. These systems have the common characteristic that the signal can be expressed in terms of the two forms of modulation which can be simultaneously applied to an r-f carrier: amplitude and angular modulation.

The term angular modulation as used here indicates that the phase angle of the transmitted signal is varied to provide the stereo information. This phase variation can be expressed in terms of either phase or frequency modulation. Depending upon the circuitry used to generate and detect the signal, either the phase modulation or the frequency modulation viewpoint may be preferable.

The most general form which the modulating signals, derived from two audio channels L (left) and R (right), can take is that in which the audio signal intended to amplitude modulate the carrier is a function $f_1(L, R)$ and the audio signal for angular modulation is some other function $f_2(L, R)$. Figure 1 shows the comparison between a monophonic and stereophonic signal in these terms.

It is generally accepted that $(L + R)$ is a compatible signal for monophonic reproduction. If the signal used for amplitude modulation $f_1(L, R)$ is simply $(L + R)$, then monophonic receivers will detect and reproduce $(L + R)$.

A commonly used component for transmission in the second audio-frequency path of a compatible system is $(L - R)$. Accordingly, the angular modulation function $f_2(L, R)$ may be arranged to be responsive to $(L - R)$. This is by no means a requirement, as other

* Petition of Radio Corporation of America for approval of standards for the RCA System of Stereophonic Broadcasting for Standard Broadcast Stations, November 12, 1959.

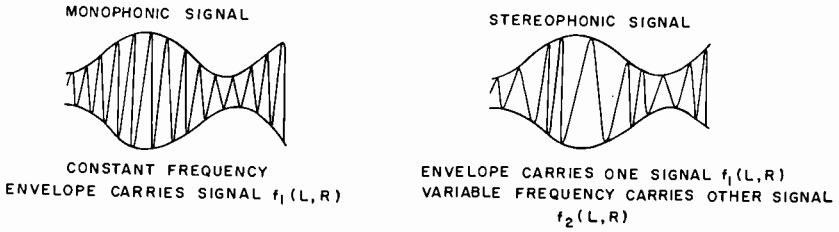


Fig. 1—Monophonic and stereophonic modulation of a carrier.

functions of L and R can be used to supply sufficient information to allow recovery of the original L and R signals.

It may be shown that there is no interaction between the two forms of modulation, provided the transmission channel has sufficient bandwidth to accommodate all the significant sidebands arising from the composite modulation. However, if any part of the sidebands is lost or subjected to a nonuniform amplitude or nonlinear phase response, then the two modulations may interact with one another to produce distortion and cross talk.

Transmitter Concepts

Since the signal for any system can be generated by suitable frequency and amplitude modulation of the carrier, a general method applicable to all systems may be used to convert an existing monophonic transmitter for stereophonic operation. The r-f source in a transmitter can be replaced by a frequency (or phase) modulator (Figure 2); the signal intended for amplitude modulation can be used as the signal source for the existing amplitude modulator.

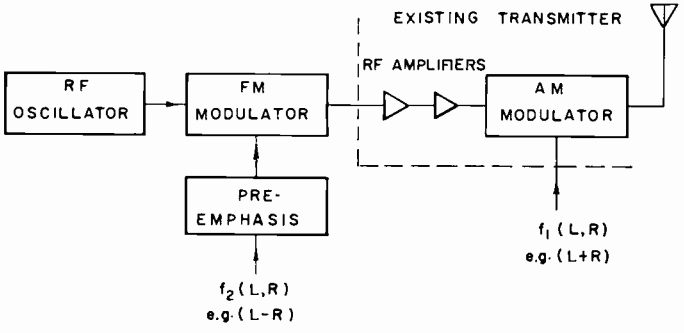


Fig. 2—A stereo transmitter. The r-f exciter of a conventional transmitter is replaced by an FM modulator.

The signals used in the high-power transmitter to obtain the required output can be obtained at a low level, as in Figure 3. The low-power r-f signal has its amplitude modulation removed by a limiter (care being taken not to introduce phase modulation in the process). This leaves the FM component which supplies the desired r-f excitation for the high-power transmitter. Detection of the envelope of the r-f signal provides the input to the AM modulator of the transmitter. Provided certain conditions are satisfied, the resulting transmitter output signal has the same frequency and amplitude components as the input signal produced by the "miniature transmitter."

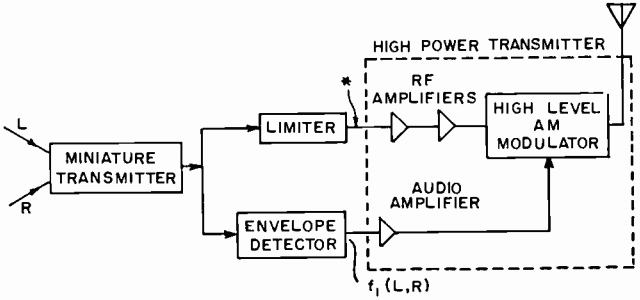
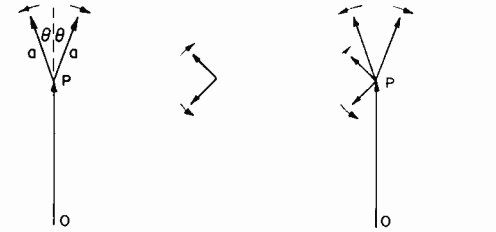


Fig. 3—A high-power transmitter employing a miniature transmitter to generate the amplitude and angular components.

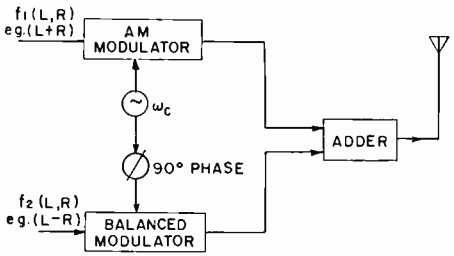
The conditions which must be satisfied include the provision of sufficient bandwidth to pass the wide band of frequencies involved in the limiter–exciter and envelope-detector–AM-modulator paths. It is also necessary that the relevant time delays in the r-f and AM channels be equalized and that the gains be adjusted for identical per cent modulation. Unless these conditions are satisfied, undesirable higher-order modulation components are transmitted; these components may produce out-of-band radiation components and impair both compatibility and stereo quality.

Additive and Multiplicative Systems

The ways in which an r-f carrier can carry two full-band information signals may be grouped under two headings: "additive" and "multiplicative." This classification is useful in showing the similarities and differences of the many possible systems, at both the transmitter and receiver. In the additive systems, sidebands are added to



(a) AM SIGNAL + (b) QUADRATURE = (c) ADDITIVE SIGNAL SIDEBANDS



(d) THE TRANSMITTER

Fig. 4—The amplitude and quadrature components of an additive signal and a typical stereo transmitter for an additive system.

the normal AM signal, as shown in Figure 4; in the multiplicative systems, the carrier and the FM sidebands containing the stereo information are all amplitude modulated (multiplied) by the normal monophonic signal, as shown in Figure 5.

The additive signal will be described first. Figure 4a shows the amplitude-modulated carrier, as in normal AM broadcasting. Figure

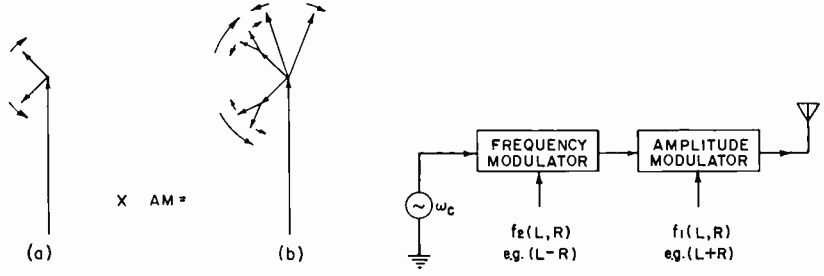


Fig. 5—An FM signal (a) followed by the sidebands which result from the multiplicative process (b). Any compatible stereo signal may be generated by the transmitter shown in (c).

4b shows the sidebands in quadrature, and the combined signal is shown in Figure 4c. The transmitter for generating the signal can be of the form shown in Figure 4d.

In the additive system, the in-phase sidebands are not affected by the presence of the quadrature sidebands, and vice versa. In the multiplicative systems, both the carrier and the frequency-modulation sidebands are modulated by that signal (e.g., $(L+R)$) which is to appear as amplitude modulation. Figure 5 shows how the signal is derived, and also a transmitter for such a system. This transmitter is similar to that shown in Figure 2. The use of the "miniature-transmitter" technique is usually not necessary in a multiplicative system.

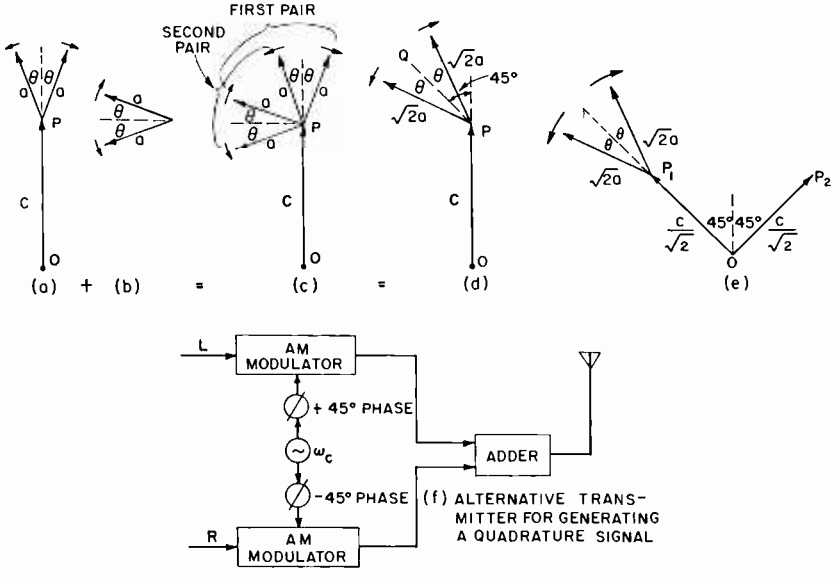


Fig. 6—The quadrature modulation system. The addition of the angular sidebands (b) to the carrier and AM sidebands (a) results in the modulated signal (c). The vectors may be combined (d) to illustrate the carrier components (e) and to suggest a transmitter form (f).

Additive Signal

Suppose the transmitter of Figure 4d is supplied with audio signals in which $f_1(L, R) = (L+R)$ and $f_2(L, R) = (L-R)$. This form of additive signal will be called a "quadrature modulation" signal. The vector diagrams for an L signal are shown in Figure 6.

In this instance, the vectors of the signal from the AM modulator (of Figure 4) are as shown in Figure 6a. Those from the balanced

modulator are as shown in Figure 6b. The vectors when added together as in Figure 6c can be combined into pairs about the line PQ, as shown in Figure 6d. The carrier may be considered to be the two components OP_1 and OP_2 shown in Figure 6e, so that the complete signal for L modulation is equivalent to amplitude modulation of OP_1, one of the two carriers phased at 90° to one another (which gives rise to the name "quadrature modulation system").

Similar analysis shows that an R signal corresponds to amplitude modulation of the other carrier (OP_2) and that complex signals may be considered as the sum of R and L modulations of their respective carriers.

A transmitter suitable for generating this signal is shown in Figure 6f. This differs from that shown in Figure 4d, but is equivalent to it if the modulations do not exceed certain values (Appendix D).

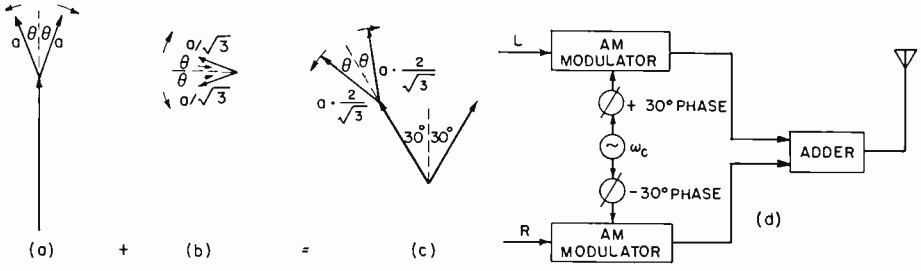


Fig. 7—A reduced-angle quadrature modulation system. This is directly analogous to Figure 6 with the carrier separation reduced to 60°.

There are several ways in which this simple additive system may be modified to produce subsystems. Detailed consideration of some of these is found in Appendixes D and E, but some basic variations will be mentioned now.

If the signals used in the transmitter of Figure 4d are given by $f_1(L, R) = (L + R)$ and $f_2(L, R) = (L - R)/k$, where k is larger than unity, then the sidebands in quadrature with the carrier are reduced in amplitude. In the alternative form of transmitter (Figure 7d), the angle between the L and R carriers becomes less than 90°. Such a signal shown in Figure 7 (drawn for 60° carrier separation) has the advantage of increased compatibility. This, however, results in decreased $(L - R)$ signal-to-noise ratio.

Another modification leads to a more basic difference. Suppose the audio signal $f_2(L, R)$ is given by $(L - R)$ phase-shifted through 90°. Then (as shown in Figure 8), the resulting signal, for an L-only

input, is given by a single-sideband signal. An R-only signal (i.e., $(L+R) = -(L-R)$) gives a single-sideband signal on the other side of the carrier. In any combination of L and R signals, the sidebands containing L information will be on one side of the carrier, and those containing R information will be on the other. This is often called the "independent sideband" system.

If, in a manner similar to that described for the quadrature signal, it is desired to reduce the energy content of the quadrature sidebands, $f_2(L, R)$ of Figure 4d will be given by $(L-R)/k$, phase shifted 90°. The locus of an L signal is then given by the ellipse of Figure 8e, instead of the circle of Figure 8d, and the L sidebands are no longer disposed on one side of the carrier.

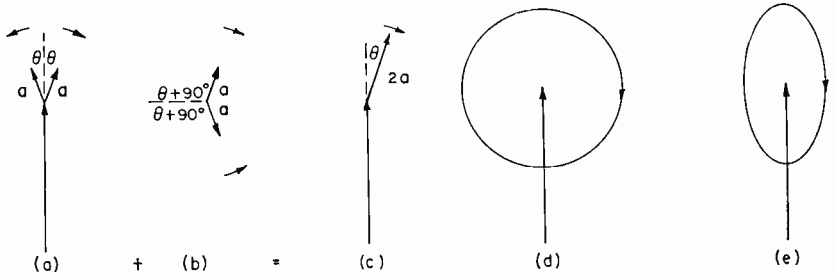


Fig. 8—The independent sideband system. For an L-only signal the in-phase sidebands (a) are added to the quadrature sidebands (b) to form the L signal shown in (c). The circular locus of the L-signal vector (d) becomes the ellipse (e) for a reduced-amplitude quadrature signal.

Multiplicative Signal

A direct method of obtaining a multiplicative signal is to use the transmitter of Figure 5c, with f_2 (L, R) fed to the FM modulator. The possibility then exists of modulating the phase angle of the carrier by many radians, without impairing compatibility. This simple multiplicative system will be called the AM-FM system. Included in the category of AM-FM are systems in which the $(L-R)$ signal is pre-emphasized. If the pre-emphasis extends to very low frequencies, the signal may then become essentially a phase-modulated signal in which the angular excursion tends to be independent of modulating frequency.

Functions of L and R other than $(L-R)$ can be chosen to give angular modulation of the carrier. One possible signal which is discussed in Appendix F is given by $f_2(L, R) = (L-R)/(1+L+R)$. This system is called "AM-FM/AM"; the signal used to angle modulate

the carrier, instead of being $(L - R)$ as in the simple AM-FM system, is $(L - R)$ divided by the amplitude of the AM envelope.

Another multiplicative signal may be derived from a signal containing carrier and quadrature sidebands by leaving these sidebands unchanged but amplitude modulating the carrier with the function required to give the desired compatible envelope. This "modified quadrature" system is described in Appendix E.

Receiver Considerations

All the systems convey $(L - R)$ information by angular modulation of the carrier. Because of this, the receiver best suited for one system will give an output related to the transmitted stereo program of any system. However, in order to avoid excessive distortion it is necessary to choose a particular type of receiver for the system being considered. Before the best system can be selected, it is necessary to know the problems associated with the receivers.

The stereo receiver must perform all the functions of a monophonic receiver in addition to providing means whereby a stereophonic audio output may be obtained. Accordingly, its fundamental performance characteristics, such as sensitivity, selectivity and signal-to-noise ratio should be comparable with those of monophonic receivers. Furthermore, the cross talk between the two stereophonic channels must be low over most of the tuning range, so that the listener will not be subjected to the confusion caused by shifts in auditory perspective.

Stereo receivers may be classified according to the detectors used. The detectors (either separate or combined) must give two outputs, which can be matrixed, if necessary, to provide signals for the L and R speakers.

An envelope detector and an FM detector may be combined in a stereo receiver as shown in Figure 9a. In some system applications, the FM detector is designed to be amplitude insensitive only at the center frequency.

An exalted carrier or synchronous detection receiver is shown in Figure 9b. Alternatively, an envelope detector and a single-output exalted carrier detector may be used.

Synchronous Detector and Exalted Carrier Receivers

In these receivers the carrier is effectively increased relative to the sidebands before detection. A typical synchronous detector circuit is shown in Figure 10. The locally generated or amplified carrier source must be accurate in frequency and phase, and the information

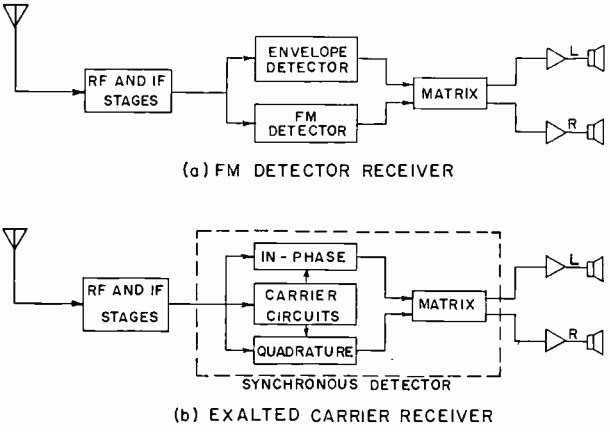


Fig. 9—Stereo receivers: (a) multiplicative system receiver with envelope and FM detectors and (b) additive system receiver employing synchronous detection.

to ensure this must be obtained from the incoming signal. The circuit shown has conventional r-f and i-f stages, with the synchronous detector operating at the i-f frequency of the receiver. Circuits should preferably be included to suppress the audio whistle which results during tuning from an unlocked local oscillator.

Another method is to separate the carrier from the sidebands by means such as a narrow-band crystal filter. The problems of this exalted carrier circuit are similar to those of a synchronous detector receiver. Both receiver types are suited to an additive signal.

Receivers containing an envelope detector and a single output synchronous detector have a better tuning characteristic. They are

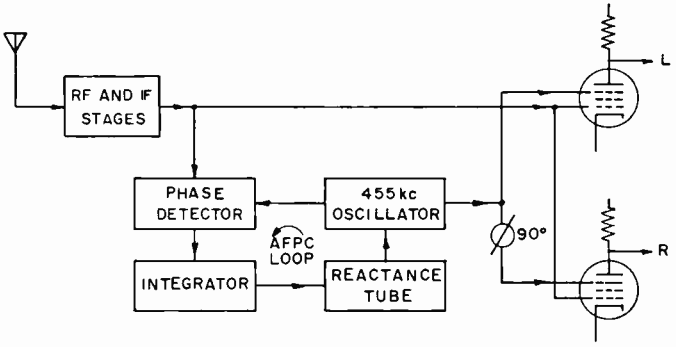


Fig. 10—Synchronous receiver showing signal path and AFPC loop. Injected carriers have relative 90° phase for additive signal recovery.

best suited, not to an additive signal, but to the previously mentioned Modified Quadrature System which is described in Appendix E.

An exalted carrier or synchronous detector receiver has potential advantages for monophonic AM reception as well as for reception of some stereo signals, and considerable effort has been spent over the years in attempts to design a receiver suitable for monophonic reception. The fact that no design suitable for commercial production has been found points up the difficulties involved.

FM Detectors with Limiters

A receiver that requires no carrier injection is made possible by the use of an envelope detector and an amplitude-insensitive FM detector (Figure 9a). This is suited to a compatible stereo signal in which the carrier is frequency-modulated with $(L - R)$ and then amplitude modulated with $(L + R)$. The detector output is relatively insensitive to the effects of receiver mistuning, and does not require the complication of an a-f-c (automatic-frequency-control) circuit.

If a static limiter is used, the output of the $(L - R)$ detector is independent of the received signal level. Therefore, if the a-g-c (automatic-gain-control) characteristic is not flat, it is necessary to provide means for making the output of the $(L - R)$ detector "track," i.e., be proportional to the mean level of the signal reaching the $(L + R)$ detector.

A stereo receiver of this type can be designed to tune in a manner which is similar to that of a monophonic receiver, and there is no significant distortion in a slightly detuned receiver.

FM Detectors without Limiters

A balanced FM detector (not preceded by a limiter) gives an output proportional to the frequency deviation multiplied by the instantaneous amplitude of the signal. One of several possible circuits is shown in Figure 11. In this circuit, outputs may be obtained as shown, giving

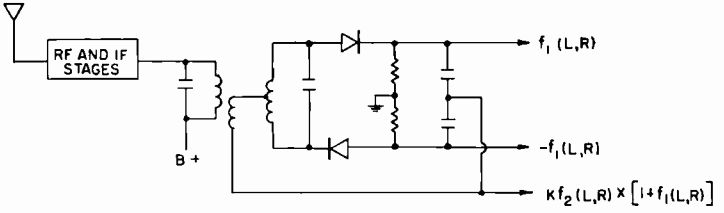


Fig. 11—Balanced discriminator circuit arranged to provide amplitude [e.g., $\pm (L + R)$] and angular [e.g., $k(L - R)$] components.

two phases of the AM modulation and a signal proportional to the product of the envelope and the FM modulation. The functions of L and R are shown for the signals produced by the transmitter of Figure 5c. If $f_1(L, R)$ is $(L+R)$ and $f_2(L, R)$ is $(L-R)/(1+L+R)$, a balanced FM discriminator acts as a multiplier of the AM and FM modulations. As a result, the output of the balanced discriminator is independent of the AM modulation and equal to $(L-R)$ provided the carrier frequency is at the null of the discriminator. Since the FM modulation required for this desirable result is of the form $(L-R)/(1+L+R)$, this type of signal and the associated system are referred to as "AM-FM/AM."

System Designation	Receiver Detectors		Signal	
	"AM"	"Angular"	"AM"	"Angular"
Quadrature and Independent Sideband	Synchronous (In-Phase)	Synchronous (Quadrature)	(L+R) In-Phase	(L—R) in Quadrature
Modified Quadrature	Envelope	Synchronous (Quadrature)	(L+R) as AM	(L—R) in Quadrature
AM-FM/AM	Envelope	Discriminator (No Limiter)	(L+R) as AM	(L—R)/ (1+L+R) on FM
AM-FM	Envelope	Discriminator (Limiter)	(L+R) as AM	(L—R) on FM

Fig. 12—Listing of stereo systems with corresponding receiver detectors for best reception of the indicated modulation components.

This type of receiver has a poor tuning characteristic. If the receiver is detuned, so that the carrier is not at the discriminator null, the AM-to-FM cross talk increases rapidly with detuning. This makes it necessary to incorporate some form of a-f-c in the receiver—applied either to the local oscillator or to the discriminator. The circuit then becomes relatively complicated.

System Considerations

The selection of a system requires consideration of the possible receiver types and the signals best suited to them. This leads to a tabulation such as that shown in Figure 12. Receivers are classified in accordance with the types of detector used in the AM and angular-modulation channels, and the signal make-up is given in terms of the modulations previously described. The particular receiver and signal

which lead to minimum distortion are listed in each row of the table along with the appropriate system designation.

The individual systems are compared in the table of Figure 13. Here, four systems are evaluated with respect to the following factors: compatibility, loss of (AM) signal-to-noise ratio, stereo performance, receiver simplicity, and out-of-band radiation. The evaluations given are based on extensive analytical and experimental work which is described in the remainder of this paper. As indicated in Figure 13, it was found that out-of-band radiation is negligible with all four systems. The AM-FM System is superior with respect to compatibility, signal-to-noise ratio (implying no reduction in station coverage), and receiver simplicity.

System Designation	Compatibility	Loss of Signal/Noise	Stereo Performance	Receiver Simplicity	Out-of-band Radiation
Quadrature and Independent Sideband	Fair	3 db (approx)	Good	Complex circuits necessary to derive carrier and suppress squeal	Negligible
Modified Quadrature	Good	3 db (approx)	Good	Complex circuits necessary to derive carrier	Negligible
AM-FM/AM	Good	2 db (approx)	Good	AFC required	Negligible
AM-FM	Very good	< 1 db	Good	Simple FM detector	Negligible

Fig. 13—Comparison of system characteristics. Receiver detectors for each system are given in Figure 12.

The AM-FM System is described in the next section. Descriptions of the Quadrature and Independent Sideband, Modified Quadrature, and the AM-FM/AM Systems are contained in the Appendixes.

THE AM-FM SYSTEM

In the AM-FM system, pre-emphasized $(L - R)$ frequency modulates the carrier; $(L + R)$ amplitude modulates the resultant. Some limitations of this system are first treated in a qualitative manner.

The choice of parameters is then described followed by transmitter and receiver design considerations.

Distortion

The effect of the bandwidth restriction introduced by reception of stereo signals on conventional monophonic receivers as well as on narrow-band stereo receivers is important. The spectrum of an FM signal consists of the carrier frequency and sidebands spaced from it by the audio frequencies, harmonics of the audio frequencies, difference frequencies between them, and difference frequencies between the

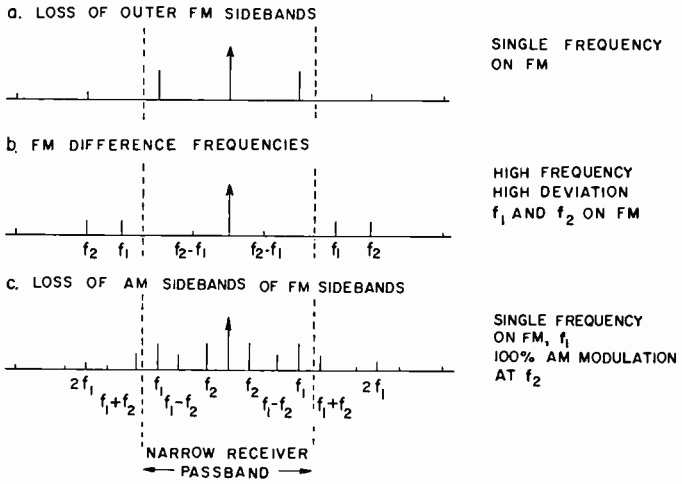


Fig. 14—Effect of bandwidth restriction on spectrum components of an AM-FM signal for three modulation conditions.

harmonics. An AM-FM signal has all these FM sidebands for the $(L-R)$ signal and in addition AM sidebands, corresponding to the $(L+R)$ signal, around each of them.

As an example, consider an FM signal with a single frequency of modulation (Figure 14a). Loss of the higher sidebands will lead to output from an FM detector at these frequencies, which correspond to harmonics of the modulating frequency. In addition, there will be AM components introduced into the signal.

Another effect is shown in Figure 14b. Here the FM signal consists of two high-frequency tones, f_1 and f_2, so that the transmitted signal contains components at f_2-f_1, $2f_2-f_1$, $2f_1-f_2$, and all other combinations. If the bandwidth is restricted, these components may result

in audio outputs from the receiver at these difference frequencies. The major component, at $f_2 - f_1$, leads to an undesired output.

Another type of distortion is shown in Figure 14c. The FM sidebands (shown here for a single-frequency audio signal, f_1) will each be surrounded (as will the carrier) by the AM sidebands, f_2. Loss of these outer sidebands, as shown, will leave the inner ones to give difference-frequency outputs from the receiver.

Distortion can result from amplitude (or phase) asymmetry in the receiver passband. For example, when an AM signal is applied, one of the sidebands will be amplified more than the other, thereby introducing a phase-modulation component. This phase modulation occurs principally during the troughs of the AM signal. The resulting AM-to-FM conversion increases with increasing AM modulation percentage. In like manner, passband asymmetry causes conversion of FM-to-AM.

The low-frequency deviation and pre-emphasis have been chosen so that these distortions are not significant. Thus the deviation is sufficiently small so as to make negligible sum-and-difference frequency components, FM-to-AM conversion, and out-of-band radiation. At the same time the deviation is large enough to make negligible the conversion of AM-to-FM in typical receiver passbands and to yield a high signal-to-noise ratio.

Field-test observations made on both stereo and conventional monophonic transmitters have indicated that adjacent-channel interference and noise tend to capture the FM channel during the troughs of envelope modulation as the carrier approaches zero. Incidental phase modulation also increases near the troughs in typical transmitters and contributes distortion. These causes of distortion, interference, and noise are greatly reduced by placing a limit of 90 per cent on the permissible downward AM modulation.

Selection of Deviation and Pre-Emphasis

The selection of deviation versus modulating frequency is complicated by conflicting trends. Thus, increasing the deviation produces these desirable effects:

(a) Improves the signal-to-noise ratio,

(b) Reduces AM-to-FM cross talk resulting from asymmetry in receiver passband,

(c) Reduces effect of incidental phase modulation in transmitters,

(d) Simplifies the stereo receiver (less receiver gain required).

But increasing the deviation also produces these undesirable effects:

(a) Increases the sum-and-difference intermodulation distortion components (Figure 14b).

(b) Increases FM-to-AM cross talk,

(c) Increases out-of-band radiation,

(d) Decreases compatibility on narrow-band receivers.

These conflicting factors can be reconciled by using a relatively low deviation with appropriate pre-emphasis as shown in Figure 15. These are the system standards used in the field test over Radio Station WNBC in December 1959 and January 1960. The low-frequency deviation is 1000 cycles. A double time constant circuit pre-emphasizes the $(L-R)$ audio signal with a time constant of 100 microseconds; a continuing rise in deviation at the higher audio frequencies is prevented by the de-emphasis of the second time constant which is 25 microseconds. This arrangement gives relatively high deviation at the frequencies from 1 to 5 kilocycles at which noise is most audible; at the same time the pre-emphasis is reduced at the higher frequencies which might otherwise supply energy to produce low-frequency intermodulation components. The maximum deviation is limited in this way to 4 kilocycles. The complementary receiver de-emphasis is shown.

In terms of the equivalent phase modulation, the double time constant circuit of Figure 15 produces an angular excursion for 100 per cent modulation of approximately 1 radian at 1 kilocycle decreasing to 0.5 radian at 5 kilocycles.

The performance obtained with these system constants in the laboratory and during field tests indicated the possibility of increasing the low-frequency deviation 25 per cent to 1250 cycles and simplifying the double time constant to a single 100-microsecond time constant, in order to achieve a better balance between the conflicting requirements previously given. Laboratory measurements as well as experience with receivers designed for this increased deviation with a single 100-microsecond time constant have confirmed the desirability of this simplification.

Delays

On going through the complete stereo system from program source to loudspeakers, the signals will be subject to appreciable delays. Care must be taken to prevent differences in the delays in the two channels. These channels can be L and R, or $(L+R)$ and $(L-R)$. For L and R signals, a relative delay of up to milliseconds can be tolerated, but for the $(L+R)$ and $(L-R)$ signals the tolerable difference is only

5 microseconds.[1,2] Both the transmitter and receiver introduce different delays into the $(L+R)$ and $(L-R)$ channels, and these require equalization.

To simplify the receiver, it is desirable to pre-compensate at the transmitter for the average delay of a typical receiver. Early tests showed that a typical receiver circuit introduced 25 microseconds more delay in the $(L-R)$ channel than in the $(L+R)$ channel, and so in subsequent field and laboratory tests a compensating equal delay was

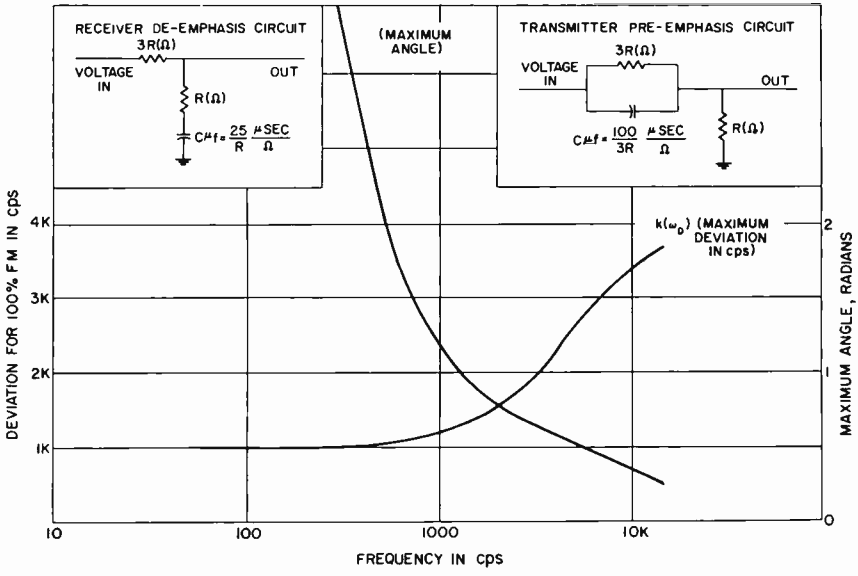


Fig. 15—Double time constant pre-emphasis characteristic employed during field tests. Corresponding receiver de-emphasis circuit is shown.

introduced into the transmitter circuits of the $(L+R)$ channel. If such a compensating delay is specified in the system standards, receivers can be designed to match it, whereas to match a zero relative delay would require a more complicated circuit.

WNBC Transmitter Modification

The basic transmitter shown in Figure 2 is directly applicable to the AM-FM system. The incoming L and R audio signals are matrixed

[1] NSRC Report of Panel 2, March 16, 1960.

[2] D. E. L. Shorter and G. J. Phillips, "A Summary of the Present Position of Stereophonic Broadcasting," BBC Engineering Monograph No. 29, April 1960.

to give $(L + R)$ and $(L - R)$; $(L - R)$ is used to frequency modulate a carrier source and this FM wave replaces the crystal frequency source at the transmitter. $(L + R)$ is applied to the AM modulator in the normal manner.

Experimental transmissions were made over Radio Station WNBC in New York which uses a 50-kilowatt RCA Model 50B transmitter operating at 660 kilocycles. The 5-kilowatt grid-modulated stage is followed by a 50-kilowatt class-B linear amplifier. A block diagram of the stereo installation is shown in Figure 16. The four experimental additions are shown in heavy outline. A photograph of this equipment is shown in Figure 17.

The stereo program material originated in Radio City, New York. The L and R signals were transmitted over two program lines to the transmitter at Port Washington, Long Island, a distance of approximately 20 miles.

Bridging amplifiers are used to feed the L and R signals to the matrix unit without disturbing the program-line connections to the normal audio equipment. The matrix unit adds and subtracts the incoming signals to produce the $(L + R)$ signal, which is used for amplitude modulation of the transmitter, and the $(L - R)$ signal, which contains the stereo information for modulation of the FM exciter. The matrix unit also contains an adjustable delay line which delays the $(L + R)$ signal to compensate for delays in the FM exciter unit.

The $(L + R)$ and $(L - R)$ signals are fed through two limiting amplifiers before they modulate the transmitter. The $(L + R)$ limiting amplifier operates in a normal manner, the $(L - R)$ limiting amplifier is arranged so that any gain reduction of the $(L + R)$ signal will cause the same amount of gain reduction of the $(L - R)$ signal.

The FM exciter (Figure 17) contains a 110-kilocycle temperature-controlled crystal, audio amplifier, pre-emphasis circuit, serrasoid modulator, a frequency multiplier, and an output amplifier. The level of the $(L - R)$ input signal to the FM exciter is monitored by a VU meter calibrated to indicate percentage of FM modulation. The FM exciter feeds the transmitter through a *stereo-mono* switch at the same point that the normal crystal exciter feeds the transmitter and provides the signal level required to drive the transmitter to normal power output. The transmitter is neutralized and tuned to minimize conversion of AM modulation to FM modulation.

To aid in compatibility testing, an automatic timer was connected (as shown) to the $(L - R)$ input terminals on the FM exciter to

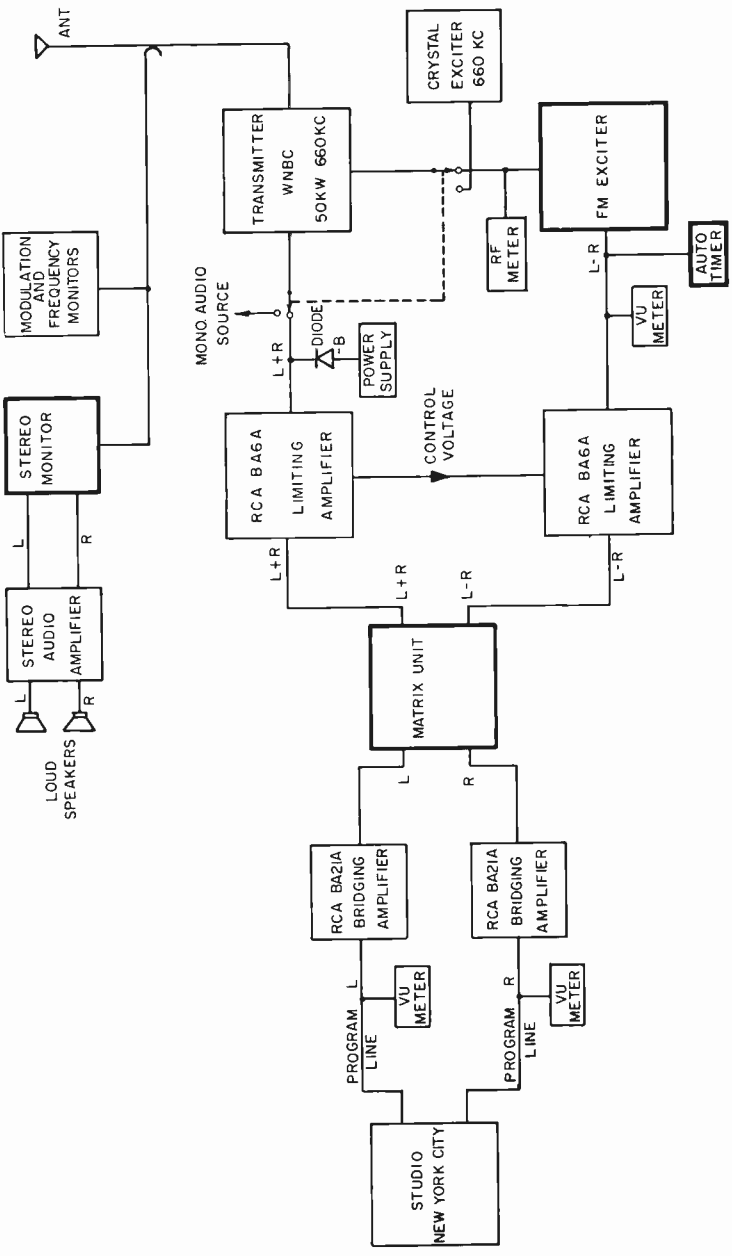


Fig. 16—Block diagram of stereo installation at radio station WNBC. Heavy outline indicates laboratory-built items.

alternately short and unshort these terminals every five seconds when the timer was turned on.

The $(L + R)$ signal from the limiting amplifier feeds the transmitter through a *stereo-mono* switch at the same point that normal audio signals are fed to the transmitter. The level and compression of the $(L + R)$ signal are the same as for normal operation of the transmitter, to provide a high average percentage modulation without

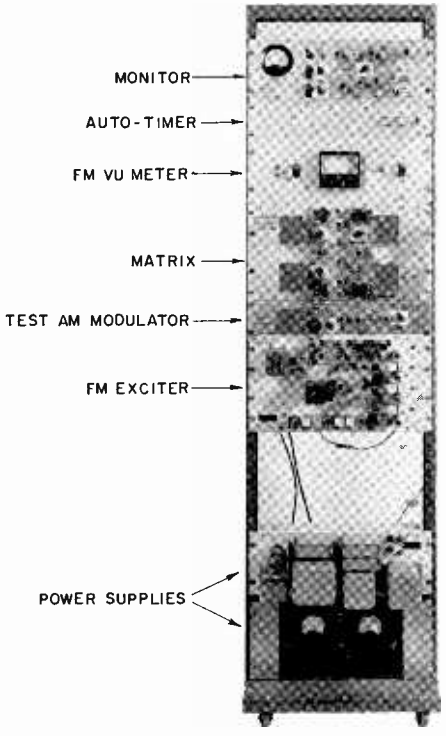


Fig. 17—Laboratory equipment used during field tests to convert radio station WNBC to stereo operation.

overmodulation of the transmitter for stereo operation. A diode and power supply are installed at the output of the $(L + R)$ limiting amplifier to clip all negative audio peaks which would cause more than 90-95 per cent negative modulation of the transmitter.

The percentage modulation and modulation peaks of the transmitter are indicated on a standard modulation meter. A standard frequency monitor measures the transmitter frequency.

In addition to the standard monitors, a stereo monitor receiver

was constructed and installed at the transmitter to measure the stereo performance. The monitor, tuned to 660 kilocycles with a bandwidth of 30 kilocycles, consisted of an AM detector, a limiter, FM discriminator, matrix and output amplifiers. The monitor detects and combines the AM and FM signals to give the original L and R signals. These are fed to a stereo audio amplifier and loudspeakers, to permit monitoring of the broadcast stereo signal.

The proper levels of the signals were normally checked by putting tones on the program lines and observing the modulation meters. The procedure used to make this check was to put a tone at 50 per cent on program line L and adjust the $(L + R)$ and $(L - R)$ gain controls on the matrix unit to obtain 50 per cent modulation of the AM transmitter and FM exciter. Similarly, a 50 per cent tone on line R gave 50 per cent modulation when the gain of the R bridging amplifier was adjusted. Two in-phase tones of 100 per cent should cause 7 decibels gain reduction of the $(L + R)$ limiting amplifier and 90 per cent AM modulation with no FM modulation. Two out-of-phase tones at 50 per cent on the program lines should give 100 per cent modulation of the FM exciter and no AM modulation. A tone at 90 per cent on either L or R will cause 90 per cent AM modulation and 90 per cent FM modulation. Final checks of the limiting amplifiers were made with program material.

The over-all performance of the transmitter was measured with the aid of the stereo receiver monitor and a spectrum analyzer. The curves shown in Figure 18 show the cross talk between L and R as a function of frequency. For comparison, the channel separation on a laboratory transmitter and field test receiver is also shown.

The spectrum of the transmitted signal was observed at the transmitter output. The spectrum analyzer photographs in Figure 19 were taken with a 7.5-kilocycle tone for the following conditions: (a) a monophonic transmission with 50 per cent AM modulation and (b) a stereophonic transmission with 50 per cent AM and 50 per cent FM modulation.

Incidental phase modulation resulting from amplitude modulation of the transmitter was investigated in some detail. Such phase modulation produces spurious output in the FM channel of the receiver in the form of cross talk and distortion components.

The four oscillograms in Figure 20 were obtained by amplitude modulating the transmitter at the indicated percentage of modulation. For each modulation frequency and for each percentage modulation, the incidental phase modulation is shown below the envelope modulated waveform. The incidental phase modulation is shown in terms

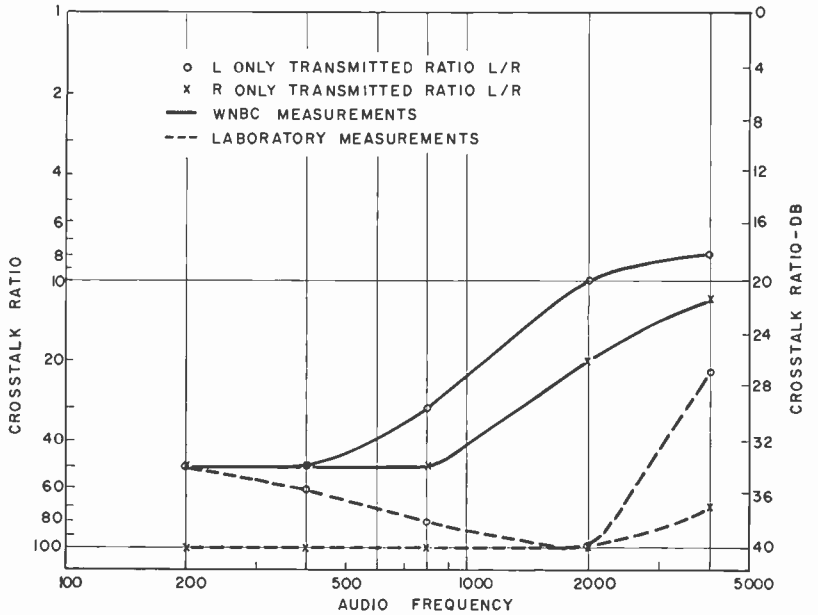


Fig. 18—Stereo crosstalk measurements on WNBC transmitter and on a low-power laboratory transmitter.

of the spurious output of the FM channel of the transmitter monitor receiver, with the de-emphasis time constant removed to show the nature of the incidental phase modulation. (Full scale deflection, 10 divisions, corresponds to 1000-cycle deviation.) It is apparent that the incidental phase modulation is produced on both the peaks and troughs of the AM envelope. Oscillogram (b) in Figure 20 shows the effect of approaching 100 per cent modulation. At approximately 97 per cent modulation, the output of the monitor receiver shows a sharp

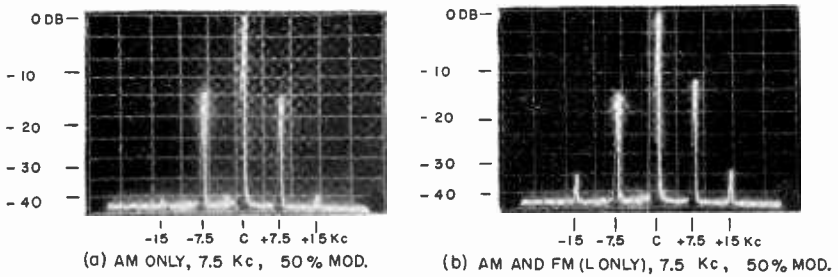


Fig. 19—Spectrum photograph of WNBC signal for 2.5-kilocycle modulation: (a) AM transmission and (b) stereo transmission.

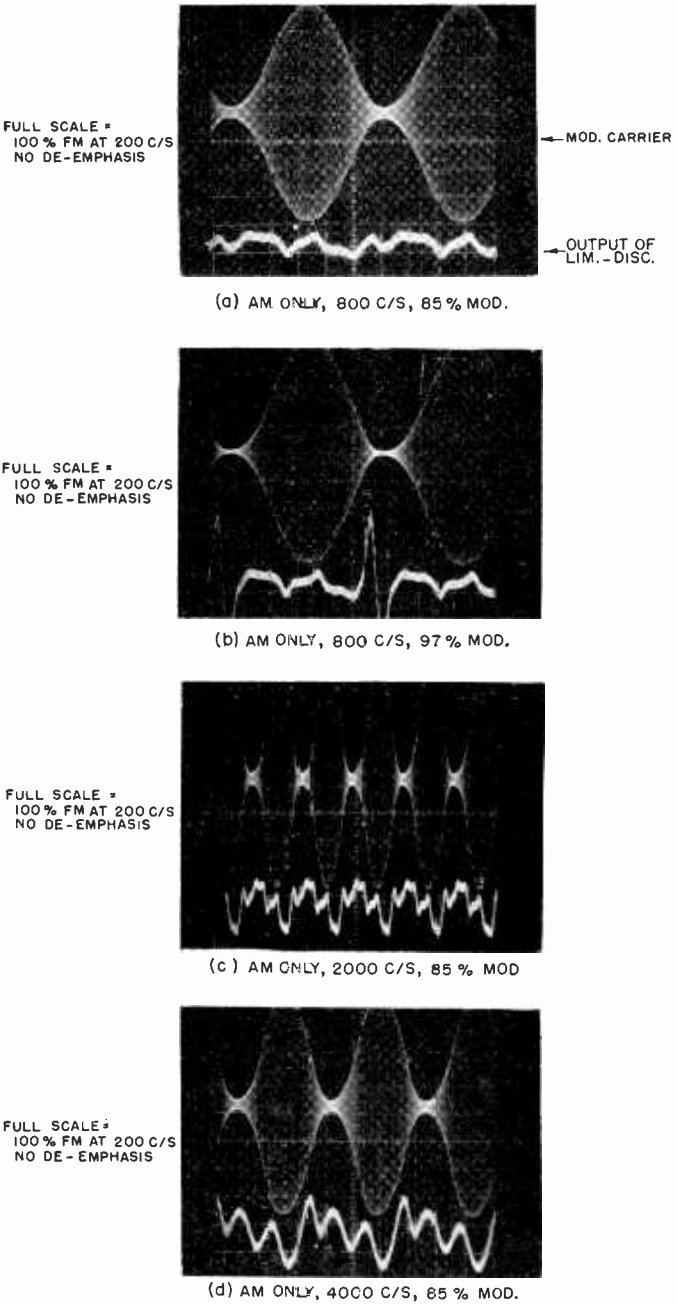


Fig. 20—Effect of modulating percentage and modulating frequency on incidental phase modulation of the WNBC trasmitter.

increase which is normally partially attenuated by the de-emphasis network.

It was not feasible within the off-the-air time available to further reduce the incidental phase modulation; nor was this considered necessary for the purpose of demonstrating compatibility and stereo performance.

Field-Test Receivers

Some general receiver considerations have been described. Additional detail is given here, followed by a description of receivers used during the WNBC field test.

The bandshape affects the performance of a stereo receiver to a greater extent than that of a monophonic receiver. This is because asymmetry in the passband produces distortion and crosstalk between the AM and FM modulations. Maintaining tracking between the r-f and i-f passbands over the complete broadcast band requires closer front-end tolerances than in monophonic practice. By designing the front end for relatively large r-f bandwidth, the effect of tracking errors on the over-all bandshape is reduced.

The output of the common i-f amplifier is fed to the circuits which provide amplitude-insensitive FM detection. The over-all bandwidth of these circuits should be appreciably greater than the common i-f bandwidth in order to reduce distortion, crosstalk and delay. Various types of detectors can be used; for example, a discriminator preceded by a limiter.

The limiter may be of either the static or dynamic type. A form of dynamic limiter used in some of the field-test receivers is shown in Figure 21. Because the diodes are made self-biasing, the output is proportional to the mean level of the i-f signal and no tracking circuitry is required.

Typical operation of a static limiter, together with the tracking means, is described in connection with the field test receiver shown in Figure 22. This receiver uses an open loop to accomplish tracking. Alternatively, the limiter may be controlled in a closed feedback loop by a signal derived from a comparison of the i-f signal before and after limiting.

Under weak signal conditions, the tracking circuits can be designed so as to improve the signal-to-noise ratio at the expense of stereo sepa-

ration. This is accomplished by reducing the relative output of the $(L - R)$ channel.

A number of field-test receivers were made by modifying a commercial radio receiver (RCA Model XF-2). A schematic diagram and photograph are shown in Figures 22 and 23. A similar speaker mounted in another cabinet was used for stereo reproduction; alternatively, two external speakers were used. Several field-test receivers used similar circuitry but were built as tuners and used with standard stereophonic amplifiers and external speakers.

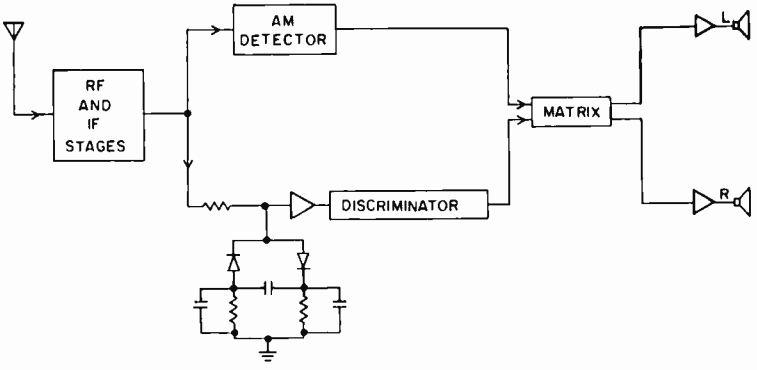


Fig. 21—Double diode dynamic limiter as employed in an AM-FM stereo receiver.

This circuit is one of many ways of designing a receiver. A conventional mixer and a single i-f amplifier stage are used. The output transformer of this i-f stage drives both an AM detector and a grid-biased limiter stage. The limiter feeds a balanced discriminator which detects the FM modulation. The double time constant de-emphasis network in the output of the discriminator is designed in accordance with the signal specifications used during the field tests.

Matrixing is accomplished by bootstrapping the AM detector with the de-emphasized FM signal so as to produce L and $-R$ at the input to the stereo amplifier stages. Volume is controlled at this point by a dual ganged potentiometer. The L and $-R$ signals are amplified in two similar audio amplifiers which drive the loudspeakers. The phase of one of the secondaries of the audio output transformers is reversed so that the L and R signals are applied to the speakers in proper phase.

The $(L + R)$ and the $(L - R)$ channel outputs are tracked by controlling the output of the limiter. This is accomplished by varying

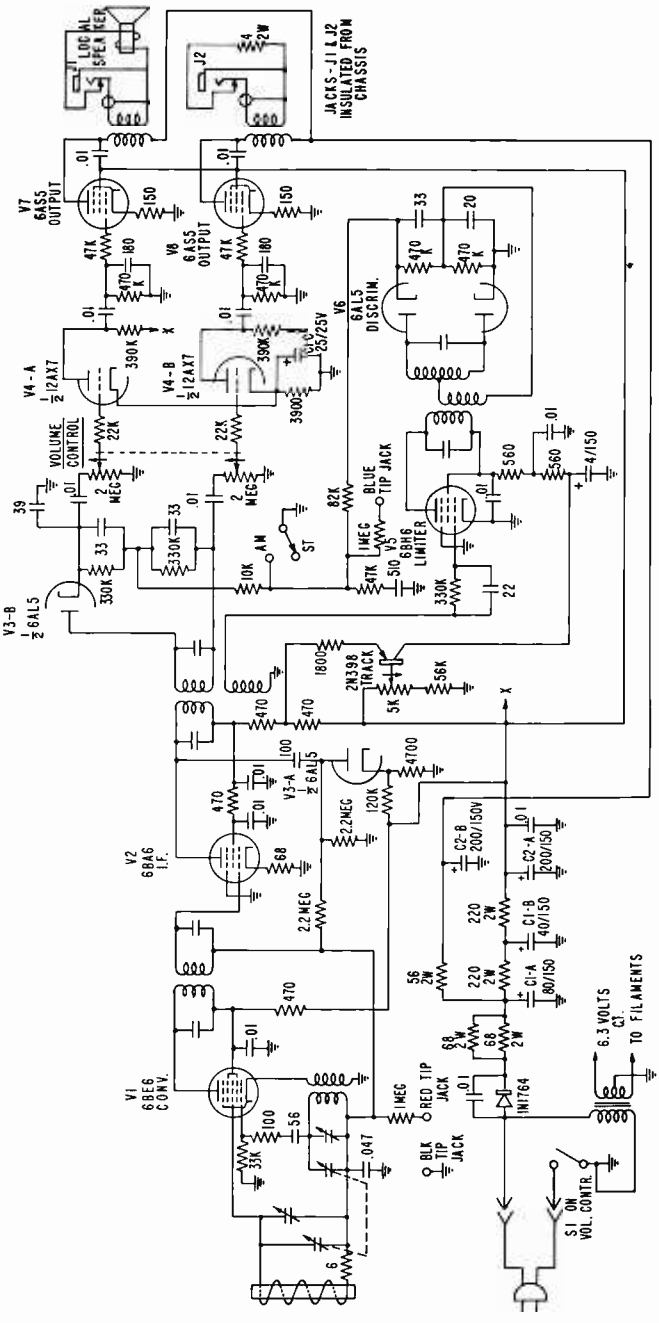


Fig. 22—Schematic diagram of one type of AM-FM stereo receiver employed during field tests.

the limiter screen and plate voltage with a transistor which is controlled by the a-g-c voltage. As the i-f input signal increases, the a-g-c voltage becomes more negative, causing the limiter screen and plate voltage to be increased.

The tuning characteristic of this receiver is essentially the same as that of a conventional monophonic receiver. Tuning is accomplished in the normal manner by tuning to the middle of the passband for minimum distortion and maximum signal-to-noise ratio. This automatically produces correct stereo reproduction. As the tuning is altered

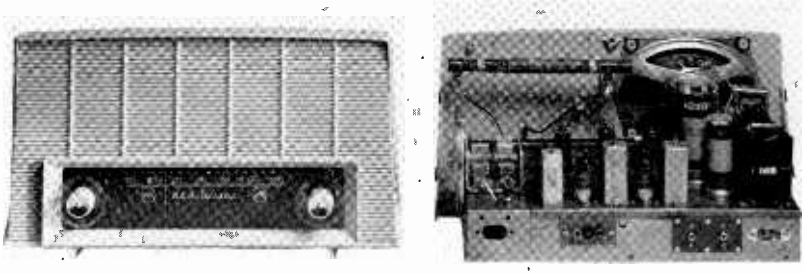


Fig. 23—AM-FM field test receiver (Figure 22).

on either side of center, distortion and noise are introduced as in a monophonic receiver before the stereo reproduction is substantially affected.

The average performance data for three types of receivers used for the field test are given below:

	Radios A-H	Tuners T1-T8	Tuners T9-T11
Sensitivity	310 μv/m	350 μv/m	200 μv/m
Signal strength for 20 db S/N	260 μv/m	420 μv/m	430 μv/m
Bandwidth r-f–i-f	4,600 cycles	10,000 cycles	8,800 cycles
Minimum stereo separation*	20 db	17 db	19 db
Maximum cross talk of FM into AM†	—22 db	—23 db	—24 db
Maximum cross talk of AM into FM†	—20 db	—21 db	—22 db

* The ratio L/R or R/L, measured at the frequency giving the poorest separation.

† Measured at the frequency giving the greatest cross talk.

Here radios A-H are of the type just described; tuners T1-T8 have similar FM detector circuitry, but have wider i-f bandwidths and have 10-kilocycle whistle filters; and tuners T9-T11 have double-diode dynamic limiters (Figure 21) in the FM channel.

Receiver Performance Calculations

Calculations were made on a digital computer to predict the performance obtainable. Transmitted signal spectra and final receiver outputs to the L and R speakers were calculated. A typical receiver bandpass was assumed consisting of two double-tuned transformers ($Q_1 = Q_2 = 50$ at 455 kilocycles; coupling 85 per cent of critical) and

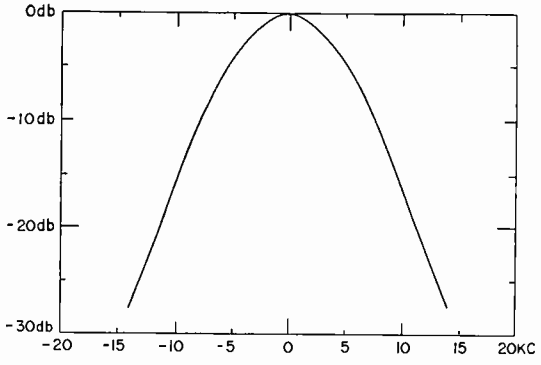


Fig. 24—Bandpass characteristic used for distortion computation for 4-kilocycle receiver.

a single-tuned circuit ($Q = 71$ at 660 kilocycles), giving the 4-kilocycle bandpass of Figure 24. This corresponds to a practical receiver of this bandwidth such as was used during the field test. In addition, calculations were made for receivers having a bandwidth of 1.5, 2.5, and 7 kilocycles to provide data on both compatibility and stereophonic performance for both narrow- and wide-band receivers. The method used to perform the calculations is described in Appendix C.

The calculations are based on a system deviation of 1250 cycles corresponding to 100 per cent modulation and a pre-emphasis of 100 microseconds. The lower energy content at the higher audio frequencies characteristic of program material is taken into account by "de-emphasizing" the L and R signals by a 75-microsecond time constant in accordance with standard FM practice. A delay of 25 microseconds is incorporated in the transmitter $(L + R)$ channel to compensate for an equal delay in the receiver $(L - R)$ channel.

The calculations shown in Figure 25 are for a transmitted signal

of 45 per cent L at 4 kilocycles and 45 per cent R at 5 kilocycles, in order to evaluate the spectrum and intermodulation for high-frequency audio inputs near the limit of the receiver bandpass. It is worth noting that a simple spectrum does not necessarily correspond to low audio distortion. Distortion must be evaluated by passing the spectrum through the r-f–i-f bandpass filter, to produce the AM $(L + R)$, L, and R outputs.

Figure 25a shows the resulting transmitted spectrum. The spectrum components above the first-order sidebands fall off rapidly. The audio outputs calculated from this spectrum are as shown in Figure 25b for the AM channel output corresponding to monophonic reception; and as shown in Figures 25c and 25d for stereo reception.

The high-frequency audio components at 8, 9 and 10 kilocycles would normally be further attenuated by the roll-off in the audio amplifier (assumed flat). The separation in both the L and R channels is greater than 40 decibels.

Comparative data are shown in Figures 25.1, 25.2, and 25.3 for receiver bandwidths of 1.5, 2.5, and 7 kilocycles, respectively. The distortion decreases progressively with increasing receiver bandwidth. The conditions are the same as for the 4-kilocycle receiver in Figure 25.

The next series, Figure 26, gives the receiver outputs for the same R signal as before but with the L signal reduced in frequency from 4 kilocycles to 1 kilocycle. As expected, the distortion components are very small.

A calculation was made under conditions similar to those in Figure 25, for a 4-kilocycle L signal and a 5-kilocycle R signal, but with the single-tuned circuit mistuned by 3 kilocycles, tuned to 663 kilocycles instead of 660 kilocycles. This corresponds to a combination of fairly severe misalignment and mistuning. It will be seen from Figure 27a that the compatibility is good and that an AM receiver is practically unaffected. The stereo outputs shown in Figures 27b and 27c, indicate that the mistuning degrades the stereo separation without significant increase in intermodulation distortion.

Receiver Performance Measurements

In addition to the calculated performance detailed in the preceding section, measurements under similar conditions were made on a field-test receiver having substantially the same r-f–i-f passband. The intermodulation distortion component values are tabulated (Figure 28) for the AM and stereo outputs for a variety of signal inputs corresponding to various combinations of high- and low-frequency

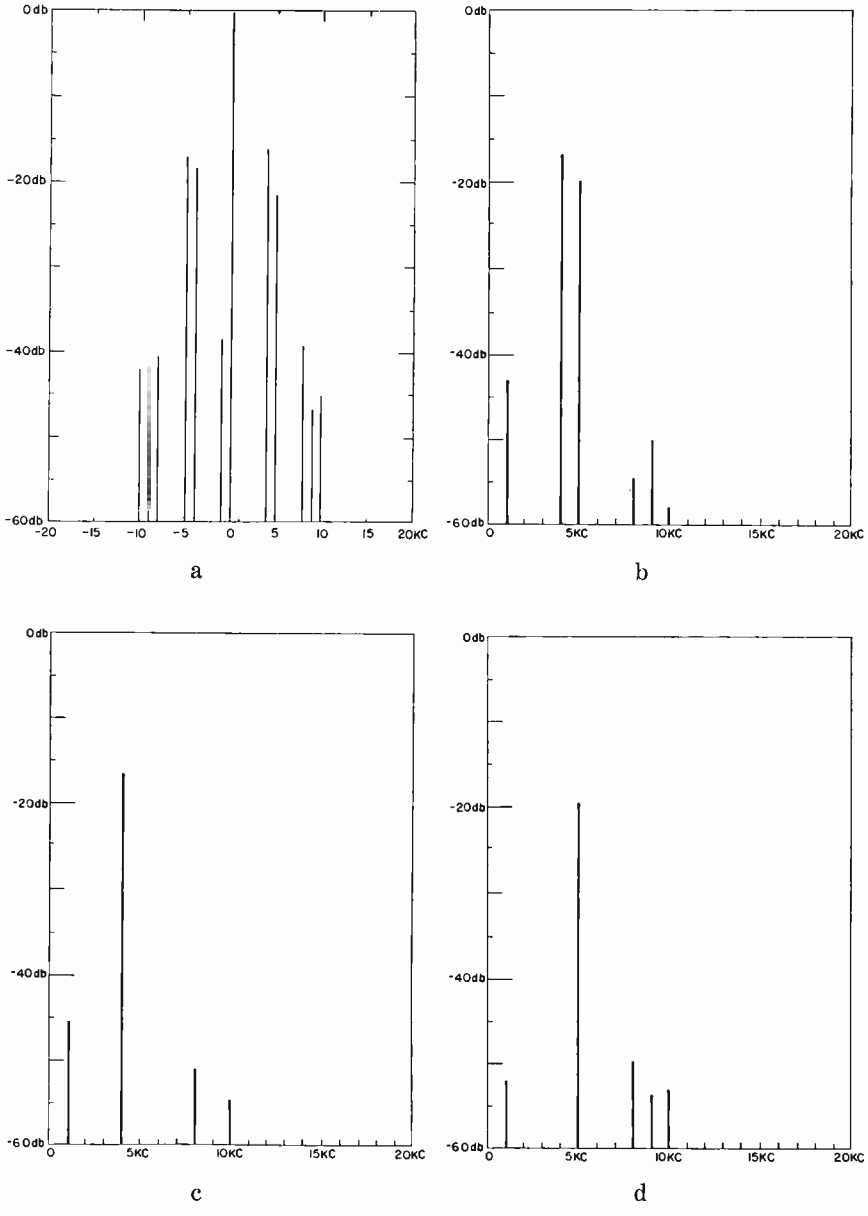


Fig. 25—Performance of AM-FM system on 4-kilocycle receiver. Modulation: $L = 45$ per cent at 4 kilocycles, $R = 45$ per cent at 5 kilocycles, 75 μsec "transmitter de-emphasis" in both L and R. (a) transmitted spectrum, (b) $L + R$ output, (c) L output, (d) R output.

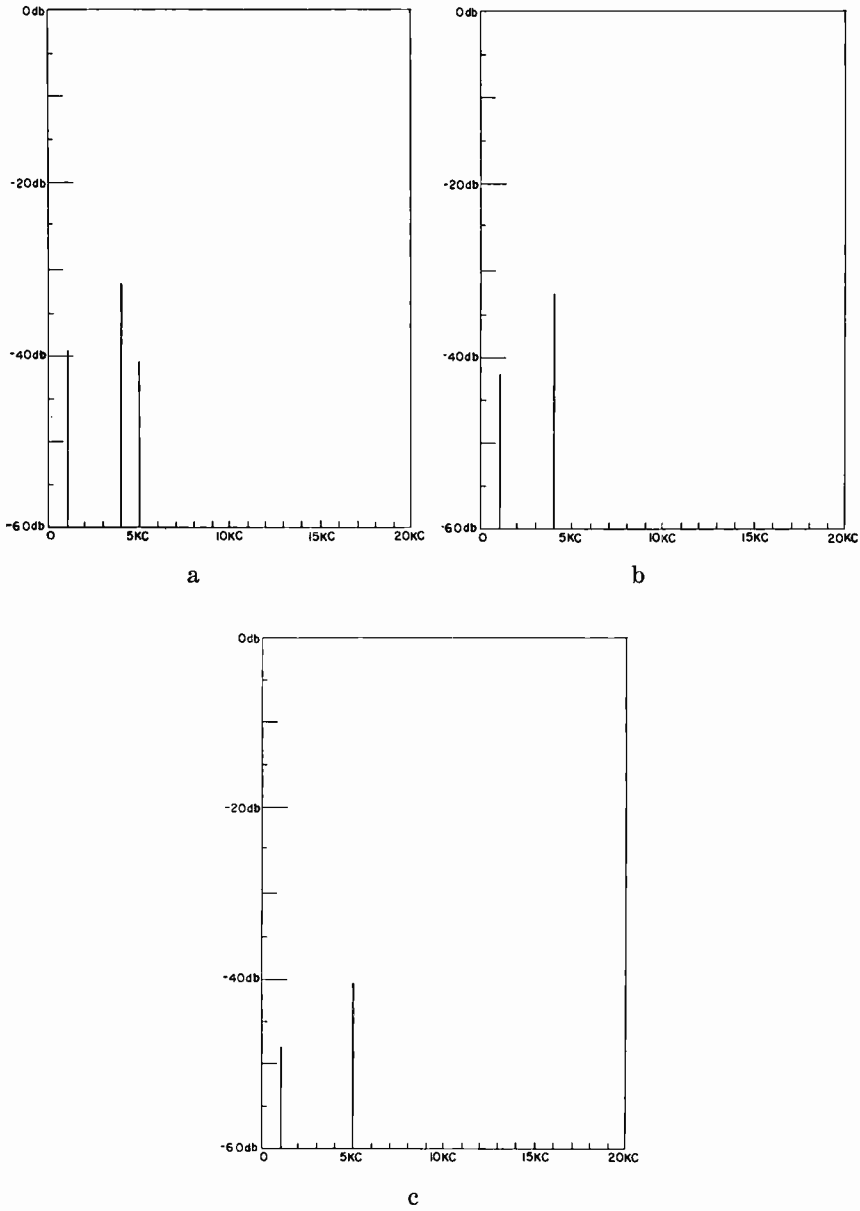


Fig. 25.1—Performance of AM-FM system on 1.5-kilocycle receiver (modulation as in Figure 25); (a) $L + R$ output, (b) L output, (c) R output.

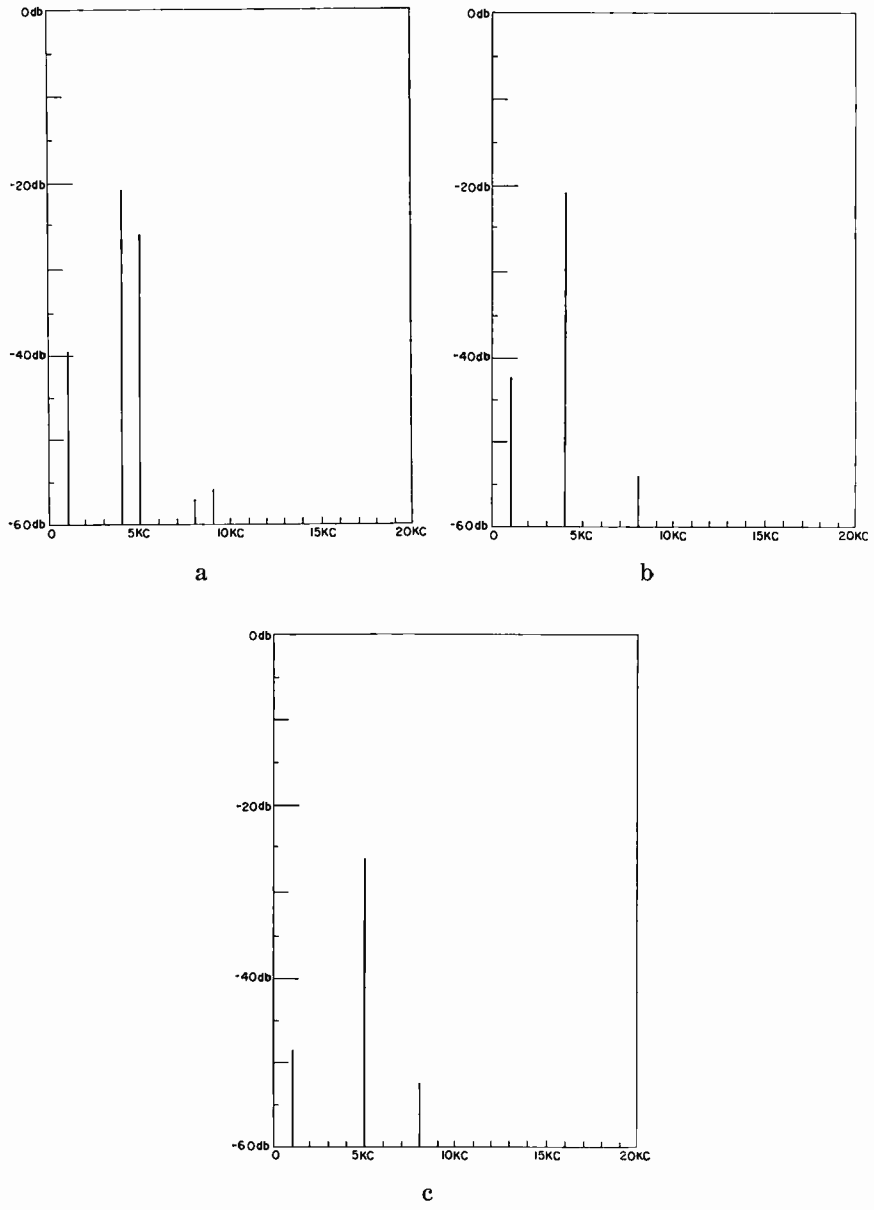


Fig. 25.2—Performance of AM-FM system on 2.5-kilocycle receiver (modulation as in Figure 25); (a) $L + R$ output, (b) L output, (c) R output.

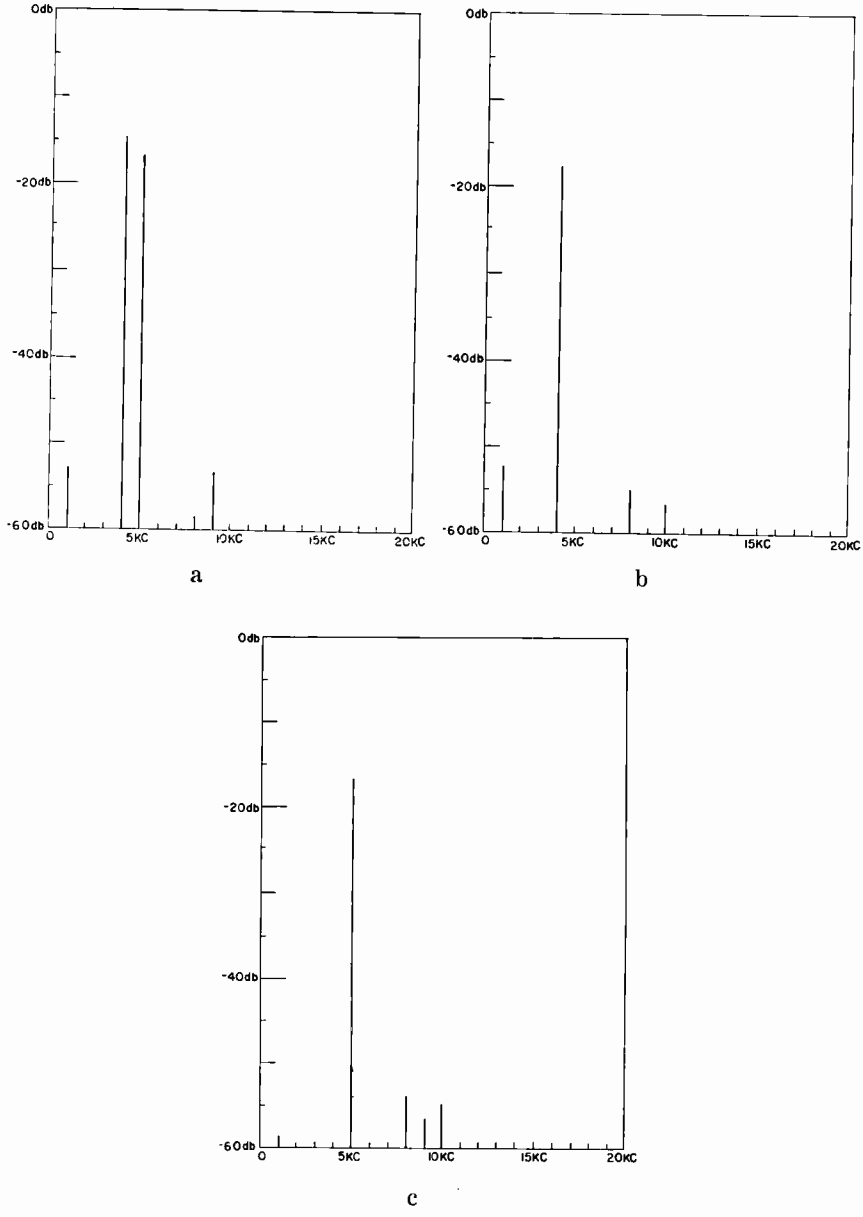


Fig. 25.3—Performance of AM-FM system on 7-kilocycle receiver (modulation as in Figure 25); (a) $L + R$ output, (b) L output, (c) R output.

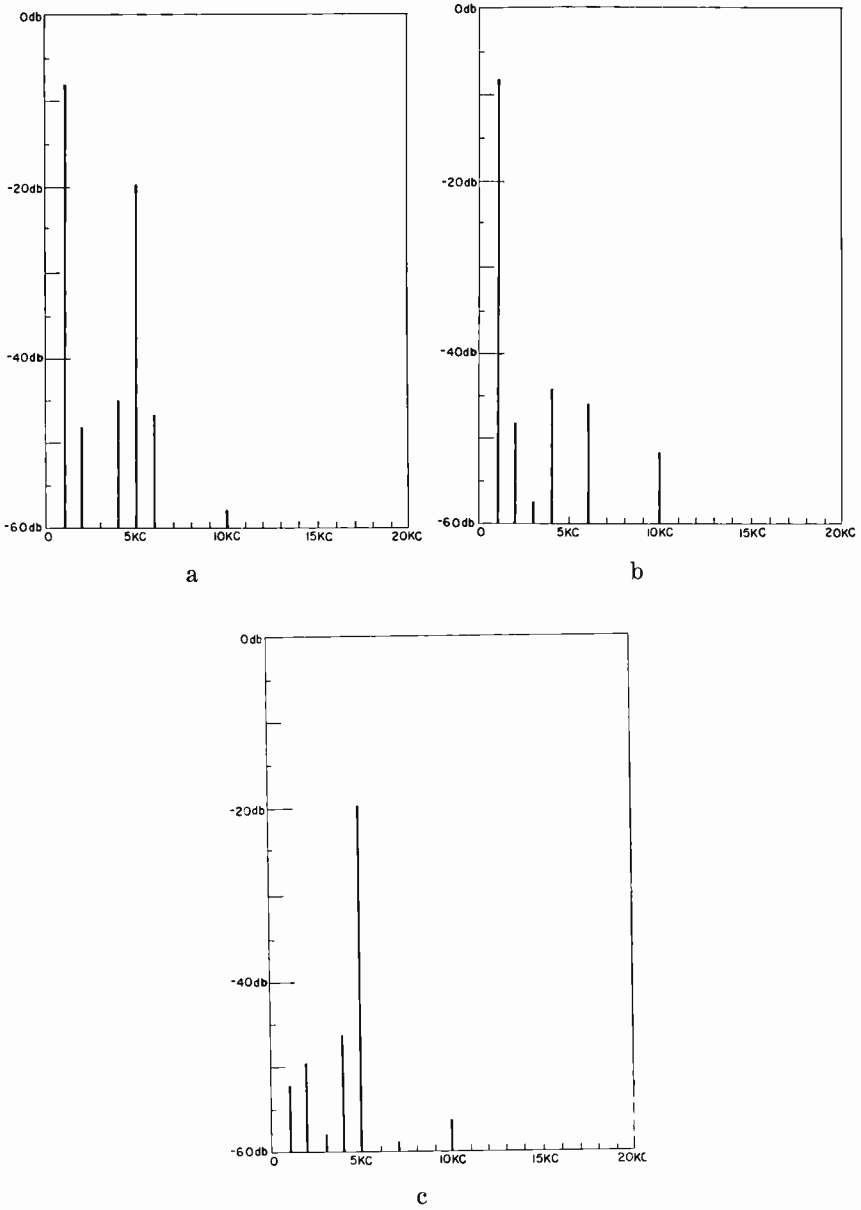


Fig. 26—Performance of AM-FM system on 4-kilocycle receiver. Modulation: $L = 45$ per cent at 1 kilocycle, $R = 45$ per cent at 5 kilocycles, 75 μsec "transmitter de-emphasis" in both L and R. (a) $L + R$ output, (b) L output, (c) R output.

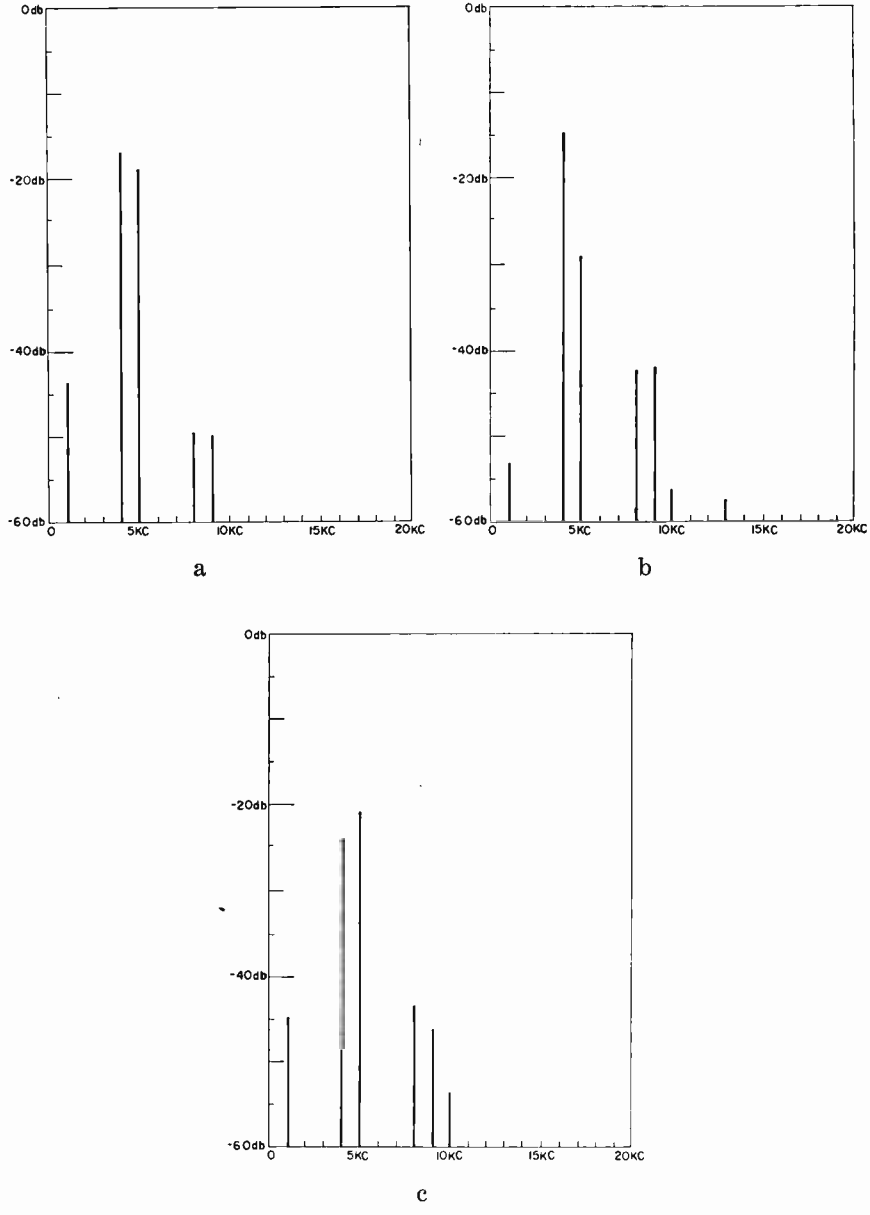


Fig. 27—Performance of a misaligned receiver. Modulation as in Figure 25. (a) $L + R$ output, (b) L output, (c) R output.

tones in the L and R inputs. The calculated values are shown in parentheses. The distortion components are referred to a 100 per cent low-frequency signal in the channel in which they are measured, e.g., 100 per cent AM for the AM distortion, and 100 per cent L for the distortion appearing in the left channel.

Distortion measurements were made on the same receiver to determine the intermodulation distortion as a function of input frequency for a constant difference frequency of 500 cycles. As shown in Figure 29, for low frequencies the difference component is predominantly in the AM channel, while at the higher frequencies distortion arises largely in the FM channel.

Extensive measurements and listening tests were made on a variety of receivers. These observations showed that even for relatively narrow-band receivers, inability to pass the outer sidebands did not introduce significant distortion. The decreasing distortion with increasing bandwidth (characteristic of AM-FM as shown in Figures 25, 25.1, 25.2, 25.3) leads to the desirable result that the distortion is negligible for the wider-bandwidth receivers where fidelity is of greatest importance.

Field Tests over Radio Station WNBC

The field tests were made between March 2 and April 23, 1959 and between December 7, 1959 and January 11, 1960, employing the WNBC 50-kilowatt transmitter at Port Washington, Long Island, New York. The tests were authorized by the FCC.[3]

During the first series of tests, a system deviation of 500 cycles was used with pre-emphasis of 150 microseconds in the $(L-R)$ channel. Based on the results of these tests, the deviation was increased to 1000 cycles and a double time constant pre-emphasis (Figure 15) was used in the second series of tests.

Daily transmissions took place between 2:05 and 2:30 p.m. and 7:05 and 7:30 p.m. The afternoon transmissions were received regularly on field test receivers at RCA Laboratories in Princeton. The evening transmissions were monitored on additional field test receivers at various locations in New Jersey, New York, and Pennsylvania.

A sensitive indication of compatibility during stereophonic transmission was obtained by automatically switching the transmitter at

---

[3] Engineering Report on Experimental Broadcasting of Stereophonic Programs over Radio Station WRCA (660 kc), New York, N. Y., February 8, 1960.

Signal Composition	Measured Frequency	Measured Distortion Present in Channel (Computer Figures in Parentheses)		
		AM	L	R
L = 45% 4 kc	1 kc	.28%	.90%	1.1%
R = 45% 5 kc		(.70)	(.52)	(.25)
L = 45% 1 kc	4 kc	.47	.85	.41
R = 45% 5 kc		(.57)	(.61)	(.48)
	6 kc	.54	.21	.54
		(.45)	(.51)	(.05)
L = 45% 4 kc +45% 5 kc	1 kc	1.9	.72	1.7
R = 0				
L = 56% 1 kc +14% 5 kc	4 kc	.63	.09	.57
R = 0	6 kc	.61	.11	.45

Fig. 28—Intermodulation distortion measurements on a 4-kilocycle stereo receiver. Numbers in parentheses are computed values.

Input Signals		Difference Frequency (db below 100%, 500 cps Modulation)			
45% L	45% R	AM	FM	L	R
200 cps	700 cps	40	49	44	47
1200	1700	42	43	42	44
2200	2700	46	37	41	42
3200	3700	51	35	40	41
4200	4700	54	36	41	42
5200	5700	54	36	42	42
6200	6700	57	35	41	41
7200	7700	56	34	39	40

Fig. 29—Intermodulation distortion measurements as a function of modulating frequencies for a 4-kilocycle stereo receiver.

5-second intervals between stereophonic and monophonic. The $(L - R)$ program signal (as received on a stereo field test receiver) was viewed on an oscilloscope during each 5 seconds of stereo transmission; this $(L - R)$ signal disappeared abruptly as the auto-timer shorted the stereo input to the transmitter. At the moment of switching, observers were able to listen critically for any change in the character of the output of the (monophonic) receiver under test.

Observations were made on representative receivers for a wide variety of reception conditions; the compatibility was excellent for all conditions. In particular, under weak signal reception, no degradation in signal-to-noise ratio was observed as a result of switching the transmitter from monophonic to stereophonic.

When switching between stereophonic and monophonic transmission, no instance was noted of any increase in interference to reception on an adjacent channel. One instance of increased co-channel interference was observed in a test conducted at Princeton using a laboratory signal generator as the desired co-channel signal. A just perceptible increase in co-channel interference was observed when WNBC's transmitter switched from monophonic to stereophonic.

Good stereophonic reproduction was obtained under weak signal conditions. Under such conditions, however, it appeared questionable whether the added audio perspective compensated for the additional noise introduced by the stereo channel. (The set user has the option of either monophonic or stereophonic reproduction through the *mono-stereo* switch.)

Under weak signal conditions or in the presence of an interfering signal, it was observed that the noise and interference were more objectionable when the transmission was 100 per cent amplitude modulated on the negative peaks. For this condition, bursts of noise and interference appeared in the output of the FM channel whenever the desired signal approached zero during the troughs of AM modulation. This effect was overcome by the installation of a negative peak clipper in the transmitter, as previously described, to prevent the troughs of envelope modulation from exceeding 95 per cent.

The field test of the AM-FM system confirmed laboratory measurements and demonstrated that high-quality stereo performance, with substantially the same coverage as in monophonic transmission, can be obtained on simple low-cost receivers.

Acknowledgment

H. F. Olson's interest in extending compatible stereophonic broadcasting to the AM band led to the initiation of this project. A. L.

Limberg contributed to the design of receiving and transmitting equipment, and M. Kanefsky was responsible for the analysis leading to the computer programming for the r-f and audio spectra. The cooperation of George M. Nixon, Stephen White, Lester Looney and John Flynn of the National Broadcasting Company was invaluable in carrying out the field tests over Radio Station WNBC in New York.

APPENDIX A — BROADCAST STANDARDS

Monophonic Standards

The FCC standards for AM broadcast transmitters[4] include the following specifications:

(1) The audio-frequency transmitting characteristics of the equipment from the microphone terminals to the antenna output shall not depart more than 2 decibels from that at 1000 cycles between 100 and 5000 cycles.

(2) The total audio-frequency harmonic distortion shall not exceed 5 per cent when the modulation is from 0 to 84 per cent, and 7.5 per cent when it is from 85 to 95 per cent. Distortion shall be measured with modulating frequencies of 50 to 7500 cycles, up to the 10th harmonic or 16,000 cycles.

(3) Modulation capability shall be at least 85 per cent.

(4) If harmful interference is caused to the reception of other stations, emissions shall be limited as required.

(5) Effective on new transmitters from January 1, 1960: Sidebands at and further than 15 kilocycles from the carrier shall be at least 25 decibels below the carrier. (The requirements are illustrated in Figure 30.)

It is clearly desirable that at least comparable performance be provided during stereophonic transmissions.

Adjacent-Channel Interference

Spectrum calculations for the various possible systems show that out-of-band radiation satisfies the limits of Figure 30. In any case, by using a low-level bandpass filter in the output of the miniature transmitter in Figure 3, the out-of-band components can be reduced

---

[4] Federal Communications Commission Rules and Regulations for Radio Broadcast Services. Sections 3.40(a) and 3.46(c).

to as low a level as desired. Accordingly, out-of-band radiation is not important in evaluating the relative merits of the alternative systems.

Present practice in AM broadcast receivers provides sufficient selectivity so that out-of-band radiation is not observable. On adjacent channels, when the undesired signal level is relatively strong, the most objectionable interference is due to a 10-kilocycle whistle between carriers. For receivers having high 10-kilocycle audio attenuation, interference is due to either radiation from the adjacent station (within its channel), or to cross modulation of the desired carrier. The amount of interference thus depends on the level of sideband energy within the adjacent channel. The additional sideband energy during stereo transmission of program material has been found to be such that it does not contribute significantly to the interference.

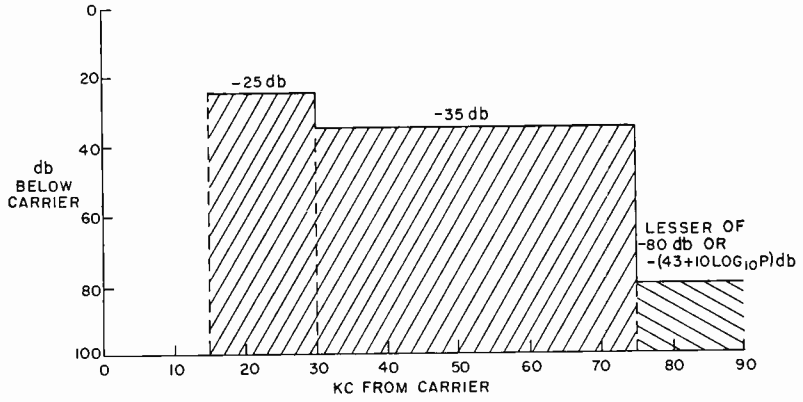


Fig. 30—Specifications for AM transmitter out-of-band radiation.

Appendix B — Audio Considerations

Stereo Audio Requirements

The lack of standard audio practice makes it difficult to generalize with respect to the amplitude and phase requirements for the $(L-R)$ signal. For example, opinion is divided as to the relative merits of the spaced versus coincident microphone techniques. However, the following considerations are generally applicable:

(1) Below about 200 cycles, room acoustics will sometimes distort stereo information. In addition, such low-frequency signals will almost invariably be accompanied by higher-frequency harmonic components which will assist the listener in placing the sound source. Also, other sound sources will usually be present and the "noise" from them will

mask any errors in localization.

(2) From 200 to 1200 cycles, phase differences are important, and the system should maintain the correct phase relationships of the two channels.

(3) Above 1200 cycles, relative L and R amplitude, as well as phase, are important. It is less important to transmit a correct stereo signal for frequencies above some 7500 cycles because voices and instruments other than cymbals and triangles normally have major components below this frequency. In addition, few AM-band receivers will respond to these high frequencies. Loss of high-frequency energy in the $(L - R)$ channel compared with the $(L + R)$ channel is noticed as an apparent spread of the source towards the center.

In additive systems, low-frequency mixing up to a few hundred cycles is necessary where symmetry about the carrier frequency is required to prevent distortion of the extracted carrier phase by the low-frequency components. In multiplicative systems, no low-frequency mixing is required, and full separation is readily attained at the low end of the $(L - R)$ band.

In both additive and multiplicative systems, the contribution of high-frequency $(L - R)$ components to the stereo effect on wide-band receivers may not be commensurate with the accompanying increase in high-frequency noise as well as a possible increase in distortion on all receivers. Extensive audio tests have shown that when the $(L + R)$ bandwidth is 15 kilocycles, negligible improvement results from an increase in the $(L - R)$ bandwidth beyond 10 kilocycles. Even at lower $(L - R)$ frequencies, network transmission makes recovery of the high-frequency $(L - R)$ components in the proper phase difficult; in the receiver, production tolerances in the alignment and tracking of the r-f and i-f passbands, as well as matrixing errors, further restrict the usefulness of the high-frequency $(L - R)$ components. In view of the foregoing, transmission could be restricted to no more than 7500 cycles in the $(L - R)$ channel.

Noise

The thermal noise in an AM system is flat; i.e., it has equal amounts of energy in equal frequency increments. However, the effect on a listener depends on the sensitivity of the ear over the audible frequency range. Standard frequency curves have been specified for noise meters which differ according to the loudness of the noise. Since in an entertainment system the noise is comparatively low, the standard curve corresponding to a 40-decibel equal loudness contour is used.

This curve is directly applicable to AM thermal noise, and is drawn as curve (1) of Figure 31. It will be noticed that much of the effective noise is between 1000 and 5000 cycles. No allowance is made in this curve for the effect of a receiver passband or audio system, which will generally lower the higher frequencies relative to the low frequencies.

The noise triangle characteristic of FM results in less noise output than with AM at frequencies below the deviation of the FM system, and more noise output at higher frequencies. This is illustrated in curve (2) of Figure 31, corrected for the ear response, for FM with 1250 cycles deviation.

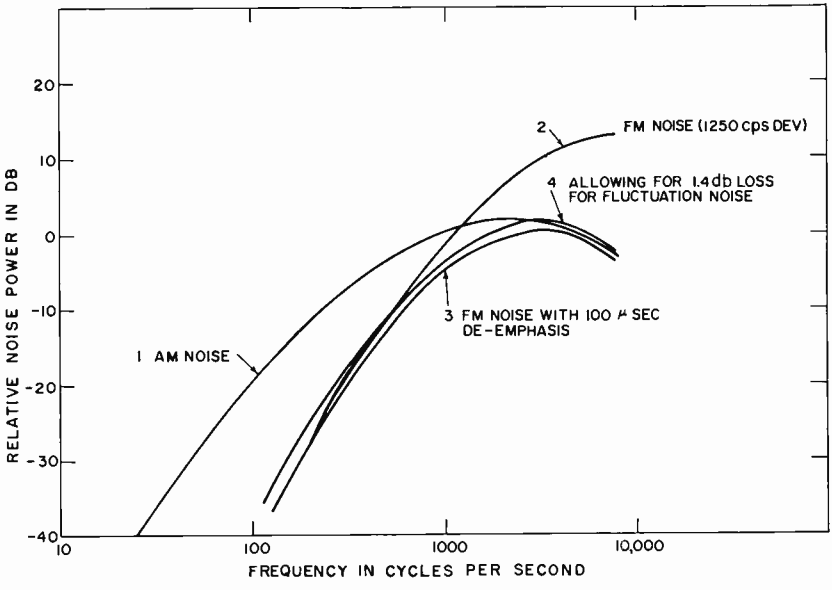


Fig. 31—Audible noise as a function of frequency: (1) AM, (2) FM 1250 cycles deviation, (3) FM with 100 μsec de-emphasis, (4) FM corrected.

To reduce the FM noise, the FM system employs pre-emphasis in transmission and de-emphasis in the receiver.[5, 6] The effect of 100 microseconds de-emphasis is shown in curve (3). For impulse noise, curves (3) and (1) give the best comparison between the AM and FM channels. The noise is now seen to lie well below the AM noise at the lower frequencies, and to approach it at higher frequencies.

---

[5] Murray G. Crosby, "The Service Range of Frequency Modulation," *RCA Review*, Vol. IV, No. 3, January, 1940, p. 349.

[6] H. L. Kirke, "Frequency Modulation: BBC Field Trials," *The BBC Quarterly*, Vol. 1, No. 2, July, 1946, p. 62.

In the quadrature system the noise has the same frequency spectrum in both the $(L+R)$ and $(L-R)$ channels and corresponds in shape to curve (1) of Figure 31.

The noise reaching the listener differs in its importance depending on whether it is derived from the $(L+R)$ or the $(L-R)$ channel. Listening tests show that noise in the $(L-R)$ channel has an unpleasant "inside the head" effect whereas $(L+R)$ channel noise appears to originate from a particular external point and is relatively unobjectionable. The effect of both noises simultaneously and of equal amplitude is to give a stereophonic "curtain of noise." It is therefore desirable to keep the noise in the $(L-R)$ channel low at all times. In very poor signal conditions, it may be desirable to reduce the receiver output from the $(L-R)$ channel. A 3-decibel reduction in the $(L-R)$ channel masks the $(L-R)$ noise to a significant degree but still permits 15 decibels of stereo separation.

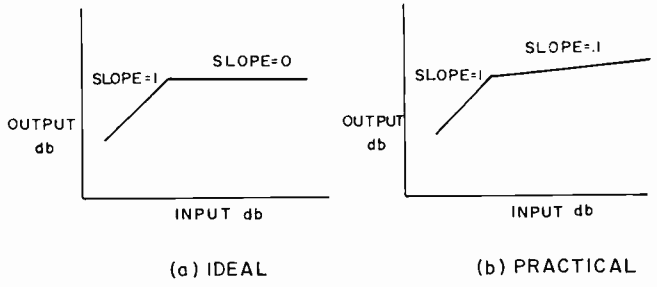


Fig. 32—Gain characteristic of limiting amplifiers.

Limiting Amplifiers

Limiting amplifiers having a characteristic similar to that shown in Figure 32 are conventionally used with monophonic transmitters to enable a high average percentage modulation to be maintained without distortion arising from overmodulation on loud passages. Although these amplifiers are commonly designated as "limiting" amplifiers, the action is one of automatic gain control (rather than clipping) to prevent excursions in the peak signal beyond a prescribed limit. Extensive experience has shown that compression of up to 5 decibels is not discernible.

In stereo transmission, limiting amplifiers afford similar advantages. However, it is necessary to operate both control amplifiers simultaneously; otherwise, there are undesirable shifts of sound image position when one channel is amplified more than the other.

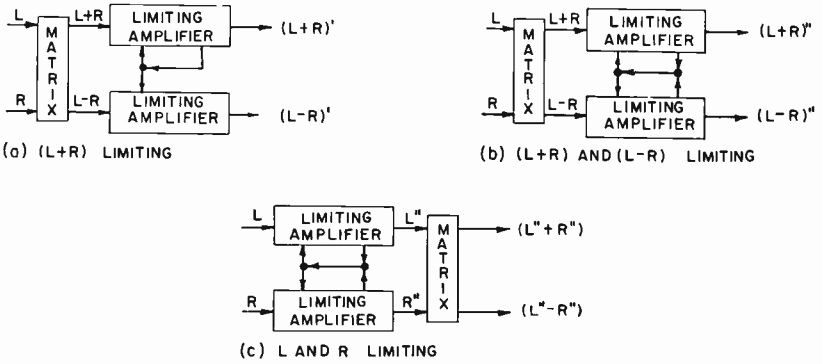


Fig. 33—Three methods of connecting limiting amplifiers.

The compatible channel $(L + R)$ is unaffected by the presence of the stereo program if the connections shown in Figure 33a are used. Here, the L and R signals are matrixed before limiting (compression), and both the $(L + R)$ and $(L - R)$ signals are controlled only by the $(L + R)$ signal. The range of signals possible with this system is shown in Figure 34a. In theory it is possible for a large $(L - R)$ signal to pass through the system if it occurs in the absence of a large $(L + R)$ signal. In practice this does not seem to occur. This system was used during the field tests described in this report.

Two other methods of deriving the control voltages are shown in

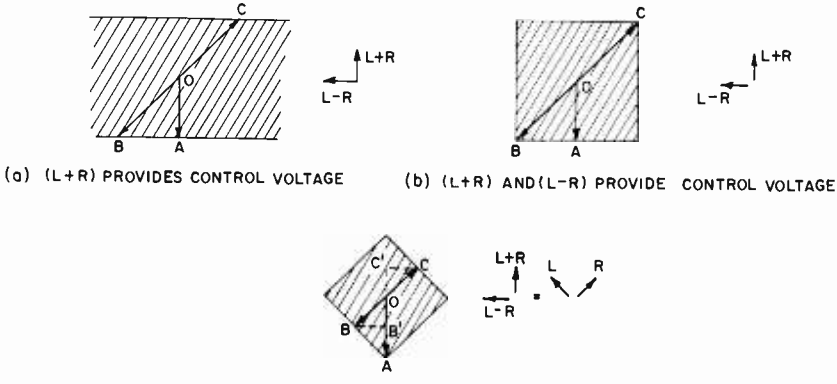


Fig. 34—The gamut of possible modulating signals corresponding to the limiting amplifier connections of Figure 33.

Figures 33b and c, and their respective effects are shown in Figures 34b and c. It will be observed that the area available with L and R limiting (Figure 34c) is only half that with $(L+R)$ and $(L-R)$ limiting (Figure 34b). With L and R limiting, full AM modulation is attainable only in the absence of the $(L-R)$ signal; only at times when the signal is completely monophonic is it possible to amplitude modulate the transmitter fully.

The line BOC in Figure 34 is indicative of how an R-only signal can modulate the carrier for the three control schemes. In the case of L and R limiting (Figure 34c) the maximum attainable $(L+R)$ signal can modulate the AM transmitter by only 50 per cent. Such a signal occurs whenever the program material is derived predominantly from one side—as happens frequently for a solo instrument or singer on one side of the stage.

Accordingly the reduction of modulation resulting from L and R limiting is between 0 decibels (for essentially monophonic programs) and 6 decibels (for programs with wide separation). This has been confirmed experimentally using varied program material. The use of L and R limiting results in a 3-decibel reduction in average AM modulation.[7]

The attack time of the limiting amplifier in the $(L-R)$ channel is not short enough to attenuate high-frequency components resulting from pre-emphasis. The use of a symmetrical clipper to provide this attenuation is effective; subjectively, however, there has been no noticeable effect on reproduction for a wide range of program material.

Appendix C — Computer Calculations

To aid in analyzing the various systems, a computer program was developed for determining the transmitted r-f spectrum, and the outputs from various types of receivers.

This was done as follows: the L and R audio signals were specified as tones—simple or complex. These audio signals were matrixed, the effect of pre-emphasis and time-delay were inserted, and the r-f signal was calculated in terms of the carrier frequency and its sidebands. Subroutines were then used to give the effect on the signal of cascaded single- and double-tuned circuits, with any specified mistuning.

The computer calculated the output from the various detectors—an envelope detector, FM detector, discriminator without a limiter, or synchronous detector, as desired. Correction was made in the computer

[7] D. S. McCoy. "The Application of Limiter Amplifiers to Stereophonic Signals," *Jour. Audio Eng. Soc.*, October, 1960.

for delays introduced in the receiver. A separate subroutine matrixed the outputs to give L and R. The computer calculated the phases of the various components of the signals, as well as the amplitudes. The computer flow chart is shown in Figure 35.

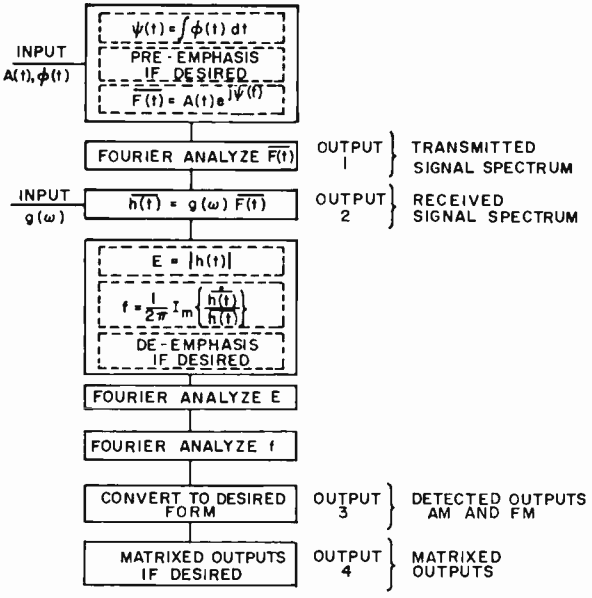


Fig. 35—Computer flow chart for computation of system performance.

The computer results shown in Figure 25 are typical. An L signal of 4 kilocycles at 45 per cent and an R signal of 5 kilocycles at 45 per cent were chosen because they permit both difference frequencies and harmonics to be calculated in detail. The modulations are approaching the maximum for both L and R limiting and $(L+R)$ and $(L-R)$ limiting, which facilitates comparison of different systems. The receiver passband chosen (Figure 24), 3 decibels at $\pm$ 4 kilocycles, is realistic for a good quality receiver. Additional computed spectra and the corresponding $(L+R)$, L and R outputs are shown for the various conditions of interest in other parts of the report.

Appendix D — Quadrature and Independent Sideband Systems

These two systems, with variations, are treated together because their similarities are greater than their differences. Audio processing

alone (i.e., a 90° phase shift) at the transmitter will convert the signal suitable for a quadrature receiver into one suitable for an independent sideband receiver. Both types of signal may be modified in various ways. For example, compatibility may be improved by reducing the $(L-R)$ component or by arranging that the envelope carry $(L+R)$ information only.

Transmitter Considerations

An economic way of designing a transmitter is to use the form shown in Figure 2, in which a frequency-modulated signal is amplitude modulated in the normal manner. The transmitter problems connected with the systems being described resolve themselves into development of the functions $f_1(L,R)$ and $f_2(L,R)$.

The functions f_1 and f_2 will both contain audio frequencies which extend from d-c to several times the highest audio frequency being transmitted. If they did not contain such frequencies, the transmitter output would contain out-of-band components. If the radiated signal is generated with inadequate bandwidth or unequal delays in the f_1 and f_2 "paths," a lack of cancellation results in the generation of out-of-band components. These can easily be of the same order as the out-of-band radiation from an AM-FM transmitter.

The components in f_1 and f_2 may be calculated for particular systems and signals. Suppose that an independent sideband system is being considered.

Let $e_c = E_c \cos \omega_c t,$

$e_L = L \cos \omega_L t,$

$e_R = R \cos \omega_R t.$

$$\text{Then } e_{\text{out}} = E_c \cos \omega_c t \left[1 + \frac{k_m L}{2} \cos \omega_L t + \frac{k_m R}{2} \cos \omega_R t \right] - E_c \sin \omega_c t \left[\frac{k_m L}{2} \sin \omega_L t - \frac{k_m R}{2} \sin \omega_R t \right]$$

The first term shows that the sum information is in phase with the carrier; the second term shows that the difference information is in quadrature with it and in addition there has been a 90° phase shift in the $(L-R)$ audio.

Similar analysis of a quadrature modulation signal derived, for example, from the transmitter of Figure 6, leads to the expression

$$e_{out} = E_c \cos \omega_c t \left[1 + \frac{k_m L}{2} \cos \omega_L t + \frac{k_m R}{2} \cos \omega_R t \right]$$

$$- E_c \sin \omega_c t \left[\frac{k_m L}{2} \cos \omega_L t - \frac{k_m R}{2} \cos \omega_R t \right]$$

As before, the sum information is in phase with the carrier; the difference information is again in quadrature with it, but this time there is no audio phase shift involved. This points up the difference between the quadrature and independent sideband systems.

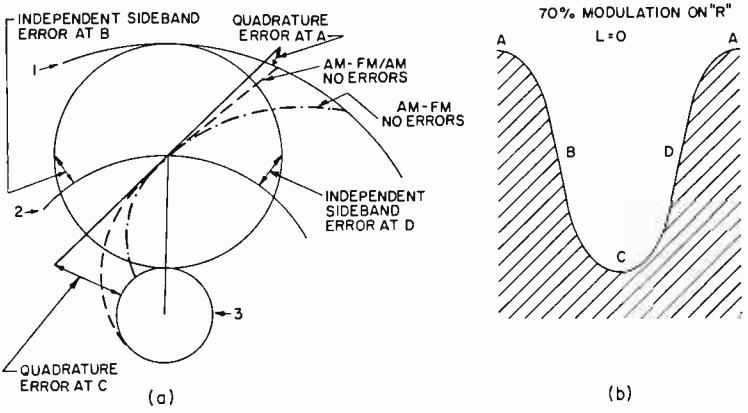


Fig. 36—The origin of compatibility distortion in additive systems for a 70 per cent modulated R signal.

An important consideration here is the limiting which must be used at the transmitter to prevent excessive distortion. This distortion arises as follows. Figure 36 shows 70 per cent modulated signals, derived from an R-only signal, on the four systems. For a compatible envelope, the A, B, and C points should lie on the circles 1, 2, and 3, corresponding to the peak, carrier, and trough values, respectively, of the undistorted envelope. The quadrature system has errors at A and C, but not at B, whereas the single-sideband system has errors at B but not at A and C. These errors are present at low as well as high frequencies. The multiplicative systems do not introduce this type of error.

An $(L - R)$ signal, $(L = 0.5,\ R = -0.5)$ as shown in Figure 37, gives envelope distortion of 16 per cent, which is mostly second harmonic; the envelope should be unmodulated $(L + R = 0)$.

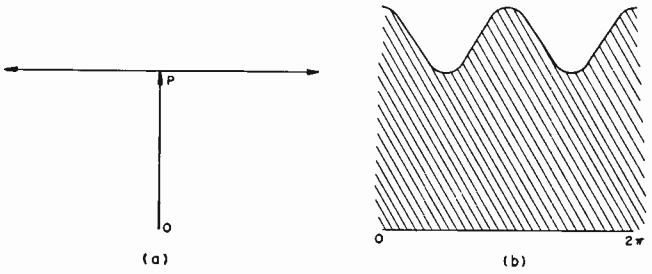


Fig. 37—Envelope (2nd harmonic) distortion in the quadrature system for $(L - R)$ only modulation.

When L and R are at different frequencies, the intermodulation distortion is more significant. Figure 38 shows the locus of the carrier vector for the modulation components $L = 0.5$ (low frequency) and $R = 0.5$ (high frequency). During that part of the modulation cycle when the low-frequency L component is zero, the envelope detector sees a high-frequency R envelope varying between OX and OY. Similarly, when $L = -0.5$, the envelope detector sees a high-frequency R variation between O and OW; and when $L = +0.5$, the variation is between OU and OV. As a result, the high-frequency R component in the output of the envelope detector is modulated 41 per cent by the low-frequency L signal.

Signal Limiting

L and R limiting corresponding to Figure 33c is implicit in quadrature systems where the transmitter is of the form shown in Figure 6f. However, this is not apparent in the alternative transmitter of Figure 4d. In either case, to prevent excessive distortion, L and R

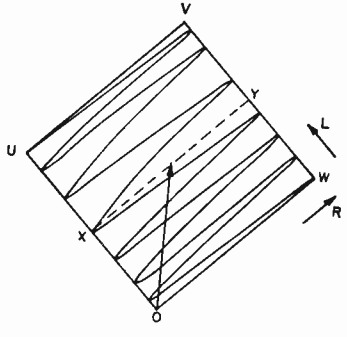


Fig. 38—The origin of intermodulation in the quadrature system.

limiting must be used and this is accompanied by a decrease in loudness of the compatible signal by about 3 decibels, as compared with a monophonic signal.

Receiver

The additive signals are theoretically capable of distortionless reception by a synchronous or exalted carrier receiver. To prevent distortion, the carrier injection must be accurate in phase. To keep the phase accurately locked, it is desirable not to transmit frequencies below approximately 300 cycles in the difference channel. This results in a slight loss of stereo effect and also a loss of low-frequency energy. This latter effect may be compensated at the receiver to provide a flat frequency response.

Synchronous detection has well known advantages, so that there have been many attempts over the past several decades to design and build commercial synchronous detection broadcast receivers suitable for home entertainment use. Unfortunately, such receivers tend to be relatively complex because of the need to incorporate an AFPC loop to recover the carrier, the need to incorporate squelch circuitry, and the more-critical tuning characteristic. These problems are all present in the case of a compatible AM stereo receiver which uses synchronous detection.

In the quadrature modulation system the angular deviation of the carrier is smaller during the peaks of AM modulation than it is during the troughs (Figure 39); for low modulations, the angle is inversely proportional to the instantaneous amplitude. A balanced frequency discriminator which is not preceded by a limiter gives an output directly proportional to the instantaneous amplitude multiplied by the frequency deviation. When such a discriminator (followed by an RC integrating network) is used with a quadrature signal, there is a relatively poor "match" between transmitter and receiver; the resulting distortion is most noticeable for signals in which two or more frequencies occur simultaneously in the $(L - R)$ channel. The absence of the required sum and difference components in the transmitted waveform results in intermodulation components at these frequencies appearing in the output of the receiver.

Reduction of Quadrature Component

The amplitude of the sidebands in quadrature with the carrier may be reduced in order to improve compatibility by lessening distortion. In the quadrature system this is equivalent to a reduction of the angle between the subcarriers (see Figure 6e). A reduction from 90° to 60°, with the retention of L and R limiting, reduces the har-

monic distortion of a maximum amplitude L signal received on an envelope detector from 16 to 7 per cent. The same treatment applied to an independent sideband signal again improves the compatibility. With the same type of signal, the harmonic distortion is reduced from 10 to 6 per cent. It will be noted that this signal no longer has independent sidebands.

Reducing the stereo content in this way improves compatibility. However, it results in a loss of signal-to-noise in the stereo component, which must be added to the loss in both channels already imposed by the need for L and R limiting.

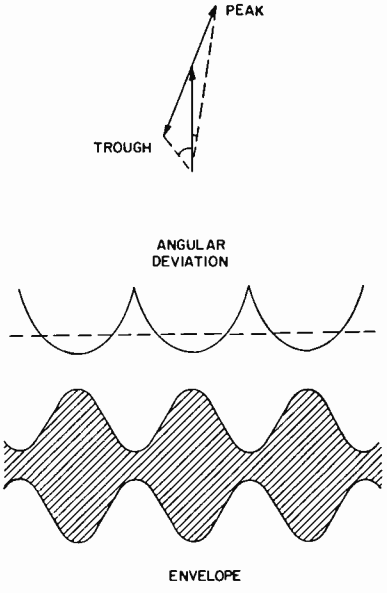


Fig. 39—The angular deviation and envelope which result from quadrature modulation.

Appendix E — Modified Quadrature System

Because of the limited potentiality for improving compatibility by reducing the quadrature component, proposals have been made for relatively complex variations of the additive systems. These result in signals which are multiplicative rather than additive.

If one considers a quadrature signal, then in order to make it compatible, the envelope (Figure 40) must contain only $(L + R)$ information. In the vector diagram of Figure 40a in which OP is the carrier, the envelope amplitude varies between OQ_2 (minimum) and

OQ_1 (maximum). For zero distortion these amplitudes should be OR_2 and OR_1. One method of accomplishing this is by a reduction of the carrier vector OP, the result being that the new locus is $Q'_2PQ'_1$ rather than the straight line Q_2PQ_1.

The carrier vector has now been amplitude modulated with a complex function so that the envelope is simple. Such a signal is suited for reception on a receiver containing an envelope detector (wideband), with a synchronous detector for the stereo channel. It is therefore still relatively complex, and the former theoretical advantage with respect to out-of-band radiation for the additive system has been lost.

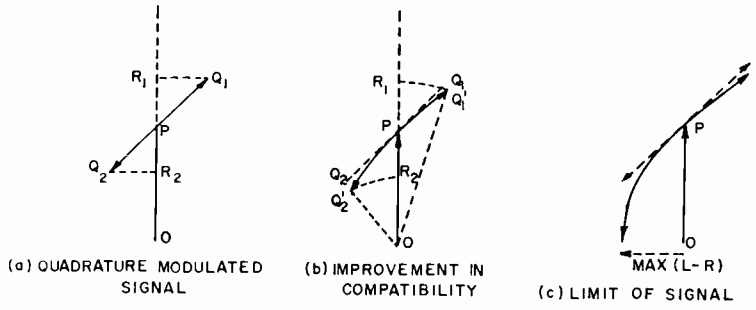


Fig. 40—Vector diagrams for the modified quadrature system. (a) unmodified, (b) modified for improved compatibility, (c) maximum modulation at which correction is possible.

The equation of the waveform applied to the amplitude modulator of the carrier is readily shown to be $\sqrt{(1+L+R)^2-(L-R)^2}$. As the quadrature sideband amplitude is $(L-R)$, the envelope is given by $(1+L+R)$. It will be noted that the carrier amplitude must be real, so that $(1+L+R)^2-(L-R)^2$ must be greater than or equal to 0. This results in a limitation of the amplitude of $(L-R)$ which is identical with L and R limiting and leads to undesirable restriction in modulation capability.

The receiver best suited for this signal consists of an envelope detector to give $(L+R)$ and a synchronous detector for the $(L-R)$ sidebands. The synchronous detector can be replaced, at the expense of some distortion, by a balanced discriminator followed by an integrator (or long-time-constant de-emphasis). A computed spectrum and receiver outputs for this latter case are shown in Figure 41.

It will be observed in Figure 41a that the transmitted spectrum is very complex; it was calculated out to only ± 20 kilocycles. For a

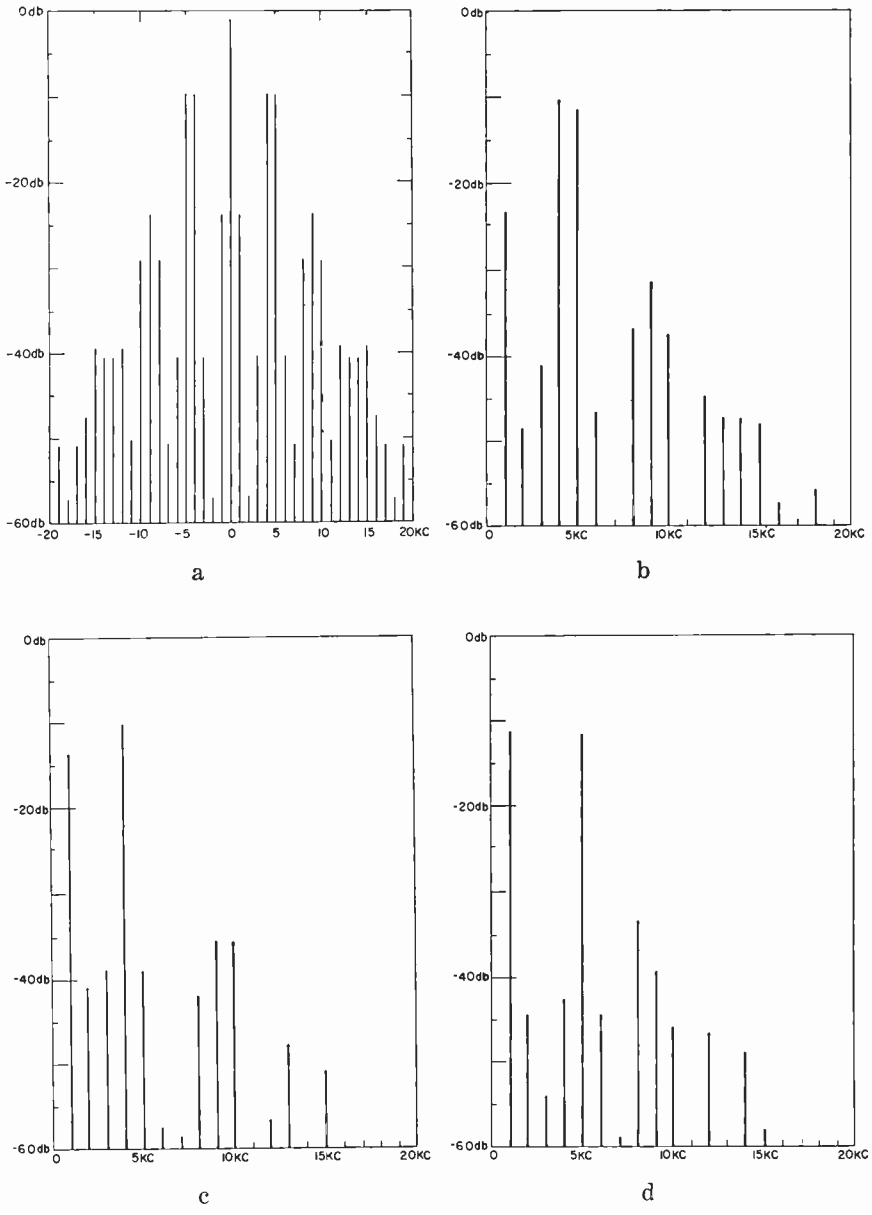


Fig. 41—Performance of modified quadrature system on 4-kilocycle receiver. Modulation: $L = 45$ per cent at 4 kilocycles, $R = 45$ per cent at 5 kilocycles. (a) spectrum, (b) $L + R$ output, (c) L output, (d) R output.

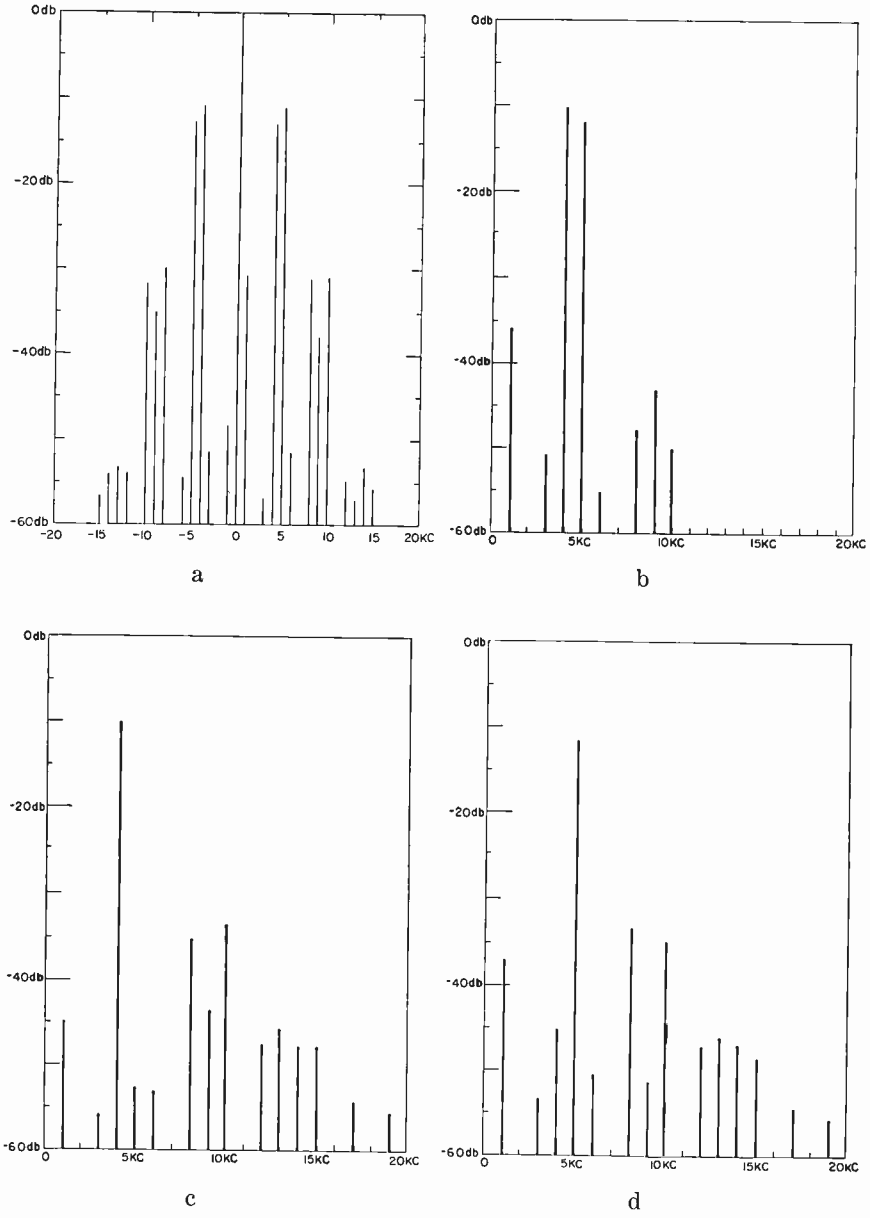


Fig. 42—Performance of AM-FM system with double time constant pre-emphasis under conditions similar to Figure 41. (a) spectrum, (b) $L+R$ output, (c) L output, (d) R output.

wideband receiver the AM output is perfect, but the output from the $(L-R)$ channel contains considerable distortion. The difference frequency (1 kilocycle) term is particularly high.

If the signal is passed through a ± 4-kilocycle simulated receiver before detection, the outputs are as shown in Figures 41b, c, and d. The AM detector output contains distortion components, so that the signal is no longer perfectly compatible. This figure can be compared with Figure 42 for the AM-FM system using a double time constant. (Figures 41 and 42 are computed without restriction of the high-frequency modulation by a 75-microsecond de-emphasis to simulate program material.)

Appendix F — The AM-FM/AM System

The potential advantages of the multiplicative signal generated by frequency modulating the carrier with $(L-R)/(1+L+R)$ and then amplitude modulating the signal with $(L+R)$ made it essential to investigate this system in some detail. This was particularly true because the system appeared inherently insensitive to transmitter incidental phase modulation on the troughs and because the receiver seemed potentially attractive.

Receiver

The basic receiver is shown in Figure 11 as requiring a balanced discriminator which is not preceded by a limiter. In the "on tune" condition, the output from such a detector is the product of the AM and FM modulation; the FM modulation is $(L-R)/(1+L+R)$ and the AM modulation is $(1+L+R)$ so that the output of the discriminator becomes $(L-R)$.

When the null is not at the carrier frequency, the output of the discriminator is readily calculated (Figure 43):

$$E = (1 + m\cos\omega_a t)\ (b + d\cos\omega_f t)$$

$$= d\cos\omega_f t + b\ (1 + m\cos\omega_a t)$$

$$+ \frac{md}{2}\left[\cos\ (\omega_f - \omega_a)\,t + \cos\ (\omega_f + \omega_a)\,t\right]$$

where m is the per cent amplitude modulation, ω_a the $(L+R)$ frequency, ω_f the $(L-R)$ frequency, b the mistuning, and d the FM deviation. The first term is the desired $(L-R)$ carried by FM. The second term is cross talk from $(L+R)$ carried on AM due to receiver mistuning by a frequency b. The third term is cross modulation. It

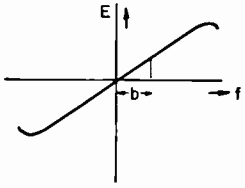


Fig. 43—The output of a balanced discriminator detecting an amplitude and frequency modulated signal with the discriminator null different from the carrier frequency.

can be made to vanish by frequency modulating with an audio signal proportional to $\cos \omega_f t/(1 + m \cos \omega_a t)$. Thus, there is no distortion provided the deviation is inversely modulated with the envelope of the signal, and no cross talk provided the discriminator null is at the carrier frequency.

Because the AM output from the discriminator rises directly with detuning, the cross talk is highly dependent upon receiver tuning. This is in addition to the r-f and i-f slope conversion found in all the systems. As a result, the tuning characteristic of receivers without limiters is unsatisfactory without inclusion of some means for automatically tracking the discriminator center tuning with the carrier frequency. One means for doing this is with a-f-c of the local oscillator. Another is shown in Figure 44; the output of the discriminator is connected to a variable-capacitance diode controlled so as to tend to keep the discriminator null at the carrier frequency.

Transmitter

The transmitter for this system is the basic multiplicative transmitter of Figure 5. As with the AM-FM system, the audio signals are supplied directly to a high-power transmitter, and there is no need to use the miniature transmitter of Figure 3.

The generation of the term $(L - R)/(1 + L + R)$ can be accom-

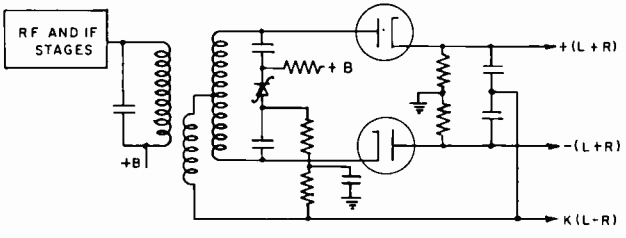


Fig. 44—AFC of discriminator secondary as employed in an AM-FM/AM stereo receiver.

plished by complex feedback means. For experimental convenience, however, the approach shown in Figure 45 was employed. This makes use of the characteristic of a solid-state diode to accomplish the required division without the delay and stability problems inherent in the feedback approach.

A solid-state diode follows, to a reasonable approximation, the equation

$$I = A(e^{\alpha V} - 1),$$

where I is the current, A and α are temperature-dependent constants, and V is the voltage. From this it follows that

$$\delta V = \frac{\delta I}{A\alpha} \frac{1}{(1 + I/A)}.$$

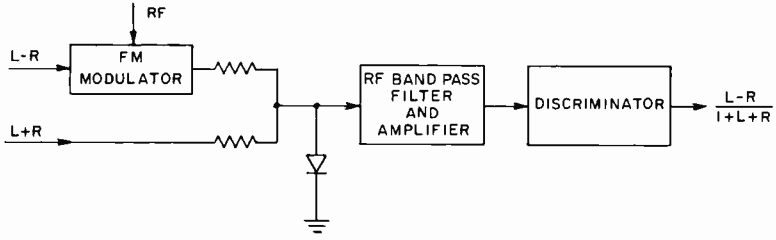


Fig. 45—Method of generating $f_2(L,R) = (L - R)/(1 + L + R)$ for the AM-FM/AM transmitter.

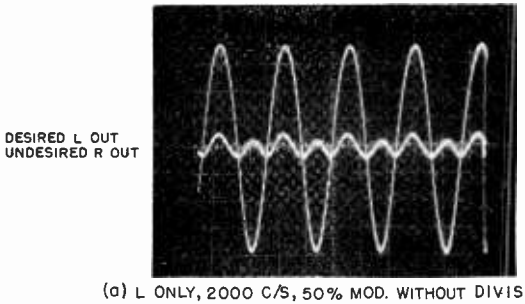
If a current of amplitude $(I/A) = (L + R)$ is added to a small r-f current (δI) and passed through the diode, then the r-f voltage will be proportional to $1/(1 + L + R)$. This r-f voltage is separated from the large audio voltage present, detected, and used in a balanced modulator to modulate $(L - R)$.

The system as developed first frequency modulated an r-f carrier with the $(L - R)$ signal. The r-f signal was then amplitude modulated with $1/(1 + L + R)$, a diode being used as described above. The signal was then fed into a discriminator. The output, proportional to $(L - R)/(1 + L + R)$, was used to frequency modulate the carrier in an AM-FM transmitter.

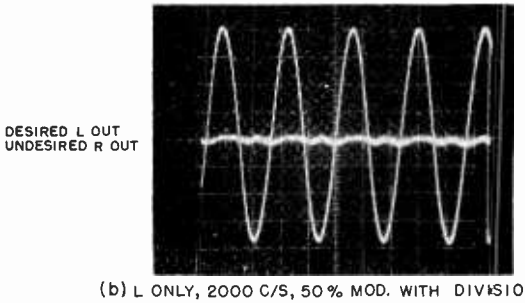
The manner in which the division by $(1 + L + R)$ improves the channel separation and reduces distortion is indicated by the oscillograms in Figure 46 which were taken at the L and R receiver outputs of a complete transmitter–receiver system. Figure 46a shows the

output when the divider circuit is inoperative. The distortion is reduced and the channel separation improved when the divider circuit is adjusted to the proper level, as in Figure 46b.

The fact that the deviation is $(L - R)$ divided by $(1 + L + R)$ leads to the possibility of the deviation becoming infinite near the troughs of envelope modulation. However, if L and R limiting is



(a) L ONLY, 2000 C/S, 50% MOD. WITHOUT DIVISION



(b) L ONLY, 2000 C/S, 50% MOD. WITH DIVISION

Fig. 46—Oscillograms showing effect of $1/(1 + L + R)$ factor on discriminator output. (a) deviation $= L - R$, (b) deviation $= (L - R)/(1 + L + R)$.

used so that neither L nor R individually can cause more than 50 per cent modulation of the carrier, then Figure 47 shows that $(L - R)/(1 + L + R)$ is a constant and that the maximum frequency deviation is the same at the peak and the trough values of envelope modulation down to 100 per cent modulation. Thus, if Q lies on the limit AO or OB, then $x = -(L - R)$ and $y = (1 + L + R)$, so that $(L - R)/(1 + L + R)$ is constant. In practice, it is desirable to limit the values of trough modulation to approximately 90 per cent in order to avoid the highly cusped waveforms required in the division of $(L - R)$ by $(1 + L + R)$ as $(L + R)$ approaches -1.

The average modulation percentage capability can be improved by

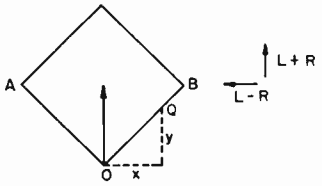


Fig. 47—The peak deviation in the AM-FM/AM system is a constant when L and R limiting is employed.

using an intermediate form of limiting rather than L and R limiting as above. A suitable method is shown in Figure 48. Here L and R are first matrixed to give $(L+R)$ and $(L-R)$. The $(L-R)$ signal is then reduced in amplitude relative to the $(L+R)$ signal, and the two are then rematrixed to give L' and R'. These differ from the original L and R because of the reduced difference component. These signals are now limited, after which they are rematrixed to give sum and difference signals. An attenuator in the sum channel is necessary to match the gains through the system for the two channels.

The elongated diamond in Figure 48b indicates the degree of improvement in modulation gamut. This is obtained at the expense of an increase in deviation near the troughs, a corresponding increase in the criticalness of dividing by $(1+L+R)$, and an increase in distortion because of the more complex spectrum.

Performance Data

The transmitted spectrum and the corresponding outputs were

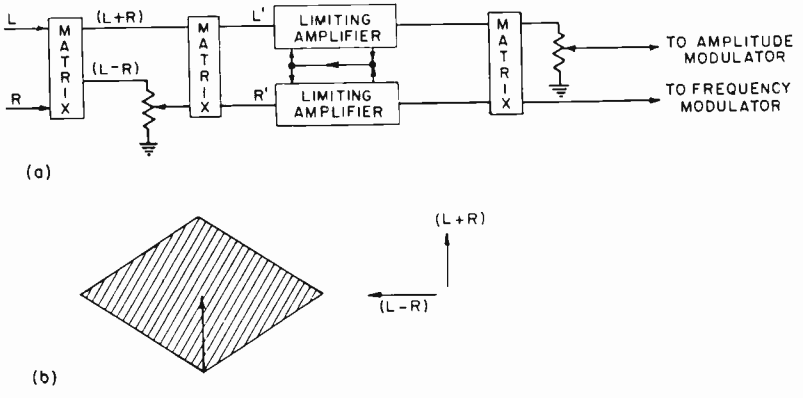


Fig. 48—Method of obtaining limiting intermediate to Figure 34(c) (L and R) and Figure 34(b) ($L+R$ and $L-R$).

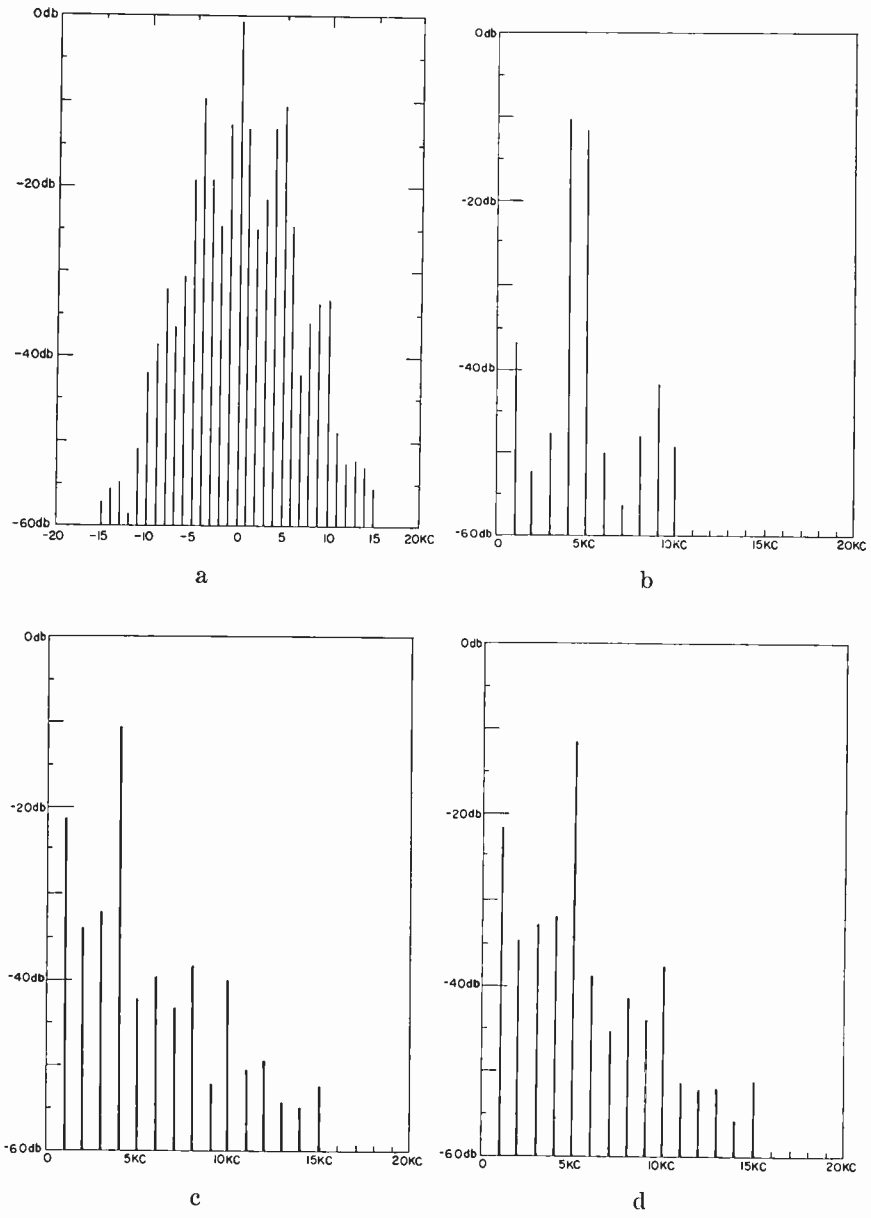


Fig. 49—Performance of AM-FM/AM system on a 4-kilocycle receiver. The conditions are as in Figure 42 to enable direct comparison. (a) Spectrum, (b) $(L + R)$, (c) L output, and (d) R output.

computed, as for the AM-FM system. The calculations were made with a double time constant pre-emphasis, and 25 microseconds delay in the FM channel of the transmitter; the r-f and i-f bandwidth of the receiver was ±4 kilocycles at 3 decibels down, as previously. Figure 49 shows the computed results for the AM-FM/AM system which may be compared directly with Figure 42 for the AM-FM system, and Figure 41 for the modified quadrature system.

The AM-FM/AM system was installed at Radio Station WNBC and field tested on January 11, 1960. The installation was such that direct comparison on program material could be made between the AM-FM/AM system and the AM-FM system. The field test transmission gave stereophonic performance comparable to that of AM-FM. However, the average percentage of AM modulation was lower for AM-FM/AM because of the requirement for L and R limiting as compared with $(L + R)$ limiting for the AM-FM signal.

On the basis of both laboratory and field test experience, it was concluded that the AM-FM system provides significant net advantages over the AM-FM/AM system. While the AM-FM/AM system is less sensitive to incidental phase modulation occurring in the troughs of envelope modulation, the AM-FM system is superior with respect to providing low intermodulation distortion, high average modulation percentage for good signal-to-noise ratio, and simplicity of transmitter adjustment (since there is no complication resulting from an $(L - R) / (1 + L + R)$ division). Furthermore the AM-FM receiver is less complex since a-f-c circuitry (necessary in the AM-FM/AM receiver to decrease distortion and excessive "switching" during tuning) is not required.

VAPOR PRESSURE DATA FOR SOME COMMON GASES*

By

R. E. Honig and H. O. Hook

RCA Laboratories
Princeton, N. J.

Summary—*For 26 gaseous compounds that are frequently encountered in ultra-high-vacuum or low-temperature work, the following data have been selected, evaluated, tabulated, and plotted: vapor pressures; transition, melting, and boiling points; and their associated energies. Vapor pressures have been plotted as log p (mm Hg or Torr) versus log T (° K). This collection contains data available before April 1, 1960.*

Introduction

For studies in ultra-high-vacuum and low-temperature systems, a knowledge of the vapor pressures of various compounds below 400° K is of considerable practical importance. Data for materials that are gaseous at room temperature can be found in the literature and were evaluated and compiled by Kelley[1] in 1935, by Stull[2] in 1947, and by Kubaschewski and Evans[3] in 1958. However, none of these collections present the information in readily accessible form. Kelley's[1] and Stull's[2] tables cover only a limited pressure range (from 10^{-1} to 10^{3} mm Hg), while the equations quoted by Kelley[1] and by Kubaschewski and Evans[3] require a cumbersome numerical evaluation. To bridge this gap, it was decided to prepare a tabulation and graphical representation of the most recent data for the more common gases, the format being patterned after a collection[4] prepared in 1957 for many of the elements.

Data

The vapor pressure data presented in this paper are based on the

* Manuscript received May 6, 1960.

[1] K. K. Kelley, "Contributions to the Data on Theoretical Metallurgy," *U. S. Bureau of Mines Bulletin 383*, 1935.

[2] D. R. Stull, "Vapor Pressure of Pure Substances. Inorganic Compounds," *Ind. and Eng. Chem.*, Vol. 39, p. 540, April, 1947.

[3] O. Kubaschewski and E. L. Evans, *Metallurgical Thermochemistry*, Pergamon Press, Third Edition, New York, 1958.

[4] R. E. Honig, "Vapor Pressure Data for the More Common Elements," *RCA Review*, Vol. 18, p. 195, June, 1957.

three collections mentioned above, as well as on a thorough literature search covering publications and abstracts available before April 1, 1960. Wherever possible, the original publications have been consulted in preference to data presented in the reviews. In most instances, experimental results were expressed in terms of the general equation

$$\log_{10} p = AT^{-1} + B \log_{10} T + CT + DT^2 + E, \tag{1}$$

where

p = pressure, expressed in this paper in millimeters of mercury (mm Hg or Torr)

T = absolute temperature, in °K

A,B,C,D,E = coefficients characteristic of the compound.

A computing machine was employed to determine from Equation (1) the temperatures corresponding to fixed pressures, starting at 10^3 and going in decade steps down to 10^{-13} mm Hg. Newton's method of successive approximations was used until the final approximation matched the equation to five significant figures. The data were checked for accuracy and correctness of transcription by computing the boiling point of each material and comparing the result with the known value.

For some cases, mainly carbon-containing compounds, vapor pressure data were found to be presented in terms of the "Antoine equation," which is of the general form

$$\log_{10} p = a - \frac{b}{T + c}, \tag{2}$$

where

a,b,c = characteristic coefficients.

While this equation appears to be very accurate for pressures in the mm region, it breaks down for low-pressure extrapolations. Therefore, data based on the Antoine equation have not been used in the present collection. For a few materials, particularly He, H_2O, N_2, and CO_2, data were available in tabulated form and could be plotted directly.

Out of the large number of compounds gaseous at room temperature, 26 of the more common gases were selected for which reliable vapor pressure data existed. In order to limit interference between vapor pressure curves, no organic compounds except CH_4 were included. For these 26 compounds, a total of 96 entries were collected,

each of which included the following items: state (solid or liquid); appropriate coefficients for Equation (1) or (2), or a pressure tabulation; the applicable temperature range; transition points, melting points, boiling points, and their associated energies. After the vapor pressure coefficients had been reduced to temperatures for fixed pressures and tabulated, those data deemed most reliable were selected from duplicate entries. The preferences shown are of course subjective and may be arbitrary in some instances, but each case was based on a thorough evaluation of all pertinent parameters.

Data were checked for internal consistency at the transition and melting points where the equations for the two different phases should yield the same pressure. For a number of compounds, solid state data were not available. To indicate that in these cases the data were obtained from the liquid equation, the temperature entry just below the melting point is tabulated in parentheses or shown dashed on the curves.

RESULTS

For the compounds selected, Table I presents literature references, the temperature range of the original data, and the temperature values computed for pressures between 10^{-13} and 10^{3} mm Hg. To simplify the tabulation, temperatures have been quoted to $\pm$ 0.01 between 2 and 10° K; to $\pm$ 0.05 between 10 and 20° K; to $\pm$ 0.1 between 20 and 100° K; to $\pm$ 0.5 between 100 and 200° K; and to $\pm$ 1 between 200 and 400° K, in line with the accuracy to which the curves to be described below can be read. In the case of He, four significant figures have been retained since there exist very accurate temperature tabulations for this element. Since none of the original data extend to pressures below 10^{-3} mm Hg, one may well ask if extrapolations through ten pressure decades down to 10^{-13} mm Hg are permissible. While direct measurements are not available, it was found that equivalent sets of data, obtained analytically from different modifications of Equation (1), usually agree to better than 1°K at 10^{-13} mm Hg. The symbols ⊙ and \ have been inserted for melting point and for transition point, respectively, between the appropriate temperatures in Table I. Table II summarizes temperatures and energies for the transition points, melting points, and boiling points of the selected compound, as well as the references used. The significant figures shown are those of the original articles, but may in many instances give an overoptimistic impression of the accuracy of the data.

Figures 1(a) and 1(b) present the vapor pressure data in graphical

Table I—Vapor Pressure Data for Some Common Gases*

SYMBOL	COMPOUND	REF.	DATA TEMP. RANGE, °K	TEMPERATURES (°K) FOR VAPOR PRESSURES (mm Hg)															
				10^{-13}	10^{-12}	10^{-11}	10^{-10}	10^{-9}	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}
Ar	ARGON	5	82– 88	20.3	21.3	22.5	23.7	25.2	26.8	28.6	30.6	33.1	35.9	39.2	43.2	48.2	54.4	62.5	73.4	89.9
Br_2	BROMINE	1	253–331	102.0	106.5	111.0	116.5	122.0	128.5	135.5	143.5	152.5	163.0	174.5	188.5	204.	224.	248.	282.	339.
CH_4	METHANE	6,7	48–112	24.0	25.3	26.7	28.2	30.0	32.0	34.2	36.9	39.9	43.5	47.7	52.9	59.2	67.3	77.7	91.7	115.0
CO	CARBON MONOXIDE	1	56–133	20.5	21.5	22.6	23.8	25.2	26.7	28.4	30.3	32.5	35.0	38.0	41.5	45.8	51.1	57.9	67.3	84.1
CO_2	CARBON DIOXIDE	7,8	107–196	59.5	62.2	65.2	68.4	72.1	76.1	80.6	85.7	91.5	98.1	106.0	114.5	125.0	137.5	153.5	173.0	198.0
COS	CARBONYL SULFIDE	9	162–224													(124.5)	139.5	159.5	187.0	229.
CS_2	CARBON DISULFIDE	1	194–319													(160.0)	177.5	199.5	228.	269.	329.
Cl_2	CHLORINE	1,10	162–420	66.1	69.1	72.4	76.0	80.0	84.4	89.4	95.1	101.5	109.0	117.5	127.5	140.0	155.0	173.0	201.	245.
F_2	FLUORINE	11	54– 89														(52.2)	59.5	70.5	87.5
H_2	HYDROGEN	12	14– 21	2.67	2.83	3.01	3.21	3.45	3.71	4.03	4.40	4.84	5.38	6.05	6.90	8.03	9.55	11.70	15.10	21.4
HBr	HYDROGEN BROMIDE	1,13	120–205	51.8	54.3	57.1	60.2	63.7	67.6	72.1	77.1	82.9	89.6	97.5	107.0	118.5	132.5	151.0	175.0	209.
HCl	HYDROGEN CHLORIDE	1	132–195	49.7	52.1	54.6	57.5	60.6	64.1	68.1	72.5	77.6	83.4	90.1	98.1	108.5	121.0	137.0	158.5	193.0
HF	HYDROGEN FLUORIDE	14	240–290														(179.0)	207.	245.	301.
H_2O	WATER	1,15	175–380	113.0	118.5	124.0	130.0	137.0	144.5	153.0	162.0	173.0	185.0	198.5	215.	233.	256.	284.	325.	381.
H_2S	HYDROGEN SULFIDE	1,16	153–213	57.1	59.8	62.7	65.9	69.5	73.5	78.0	83.1	89.0	95.7	103.5	113.5	124.5	138.5	156.5	180.5	218.
He	HELIUM	17	0.9–5.2														0.980	1.268	1.738	2.634	4.518
I_2	IODINE	3	298–456	141.5	147.5	154.0	161.5	169.5	178.5	188.5	199.5	212.	226.	243.	262.	285.	312.	345.	389.	471.
Kr	KRYPTON	5,18	63–121	27.9	29.4	30.9	32.7	34.6	36.8	39.3	42.2	45.5	49.4	53.9	59.4	66.3	74.8	85.9	101.0	123.5
N_2	NITROGEN	1,19	54–128	18.1	19.0	20.0	21.1	22.3	23.7	25.2	27.0	29.0	31.4	34.1	37.5	41.7	47.0	54.0	63.4	80.0
N_2O	NITROUS OXIDE	20	103–186	55.8	58.3	61.1	64.2	67.6	71.3	75.5	80.3	85.7	91.9	99.0	107.5	117.5	129.5	144.0	162.5	189.5
NO	NITRIC OXIDE	1	73–180	37.7	39.4	41.3	43.4	45.6	48.1	50.9	54.0	57.6	61.6	66.3	71.7	78.1	85.7	95.0	106.5	123.5
NH_3	AMMONIA	21	145–240	70.9	74.1	77.6	81.5	85.8	90.6	95.9	102.0	108.5	116.5	125.5	136.0	148.0	163.0	181.0	206.	245.
Ne	NEON	22	15– 45	5.50	5.79	6.11	6.47	6.88	7.34	7.87	8.48	9.19	10.05	11.05	12.30	13.85	15.80	18.45	22.1	27.5
O_2	OXYGEN	1,23	57–154	21.8	22.8	24.0	25.2	26.6	28.2	29.9	31.9	34.1	36.7	39.8	43.3	48.1	54.1	62.7	74.5	92.8
SO_2	SULFUR DIOXIDE	1,24	178–263	78.9	82.4	86.3	90.4	95.1	100.0	106.0	112.5	119.5	128.0	137.5	148.5	161.5	177.0	195.5	225.	269.
Xe	XENON	5	110–166	38.5	40.5	42.7	45.1	47.7	50.8	54.2	58.2	62.7	68.1	74.4	82.1	91.5	103.5	118.5	139.5	170.0

* References cited in Tables I and II will be found at the end of the paper.

TABLE II – TRANSITION, MELTING, AND BOILING POINTS AND ENERGIES FOR SOME COMMON GASES

SYMBOL	COMPOUND	REF.	TRANS. P. °K	ΔH_{Tr} kcal/mole	M.P. °K	ΔH_M kcal/mole	B.P. °K	ΔH_V @B.P. kcal/mole
Ar	ARGON	25			83.78	0.281	87.29	1.558
Br_2	BROMINE	25			265.95	2.520	331.4	7.170
CH_4	METHANE	6,26			90.67	0.225	111.67	1.955
CO	CARBON MONOXIDE	27	61.55	0.151	68.09	0.1997	81.61	1.444
CO_2	CARBON DIOXIDE	8,28			216.55	1.99	194.67	6.030
COS	CARBONYL SULFIDE	9			134.31	1.1298	222.87	4.423
CS_2	CARBON DISULFIDE	1,3			162.	1.05	319.35	6.550
Cl_2	CHLORINE	25			172.16	1.531	239.10	4.878
F_2	FLUORINE	25	45.55	0.1739	53.54	0.122	85.02	1.582
H_2	HYDROGEN	25			13.96	0.0280	20.39	0.2158
HBr	HYDROGEN BROMIDE	13,29	89.2 113.2 116.8	0.0634 0.0788 0.0857	186.24	0.575	206.38	4.210
HCl	HYDROGEN CHLORIDE	1,30	98.38	0.284	158.94	0.476	188.11	3.860
HF	HYDROGEN FLUORIDE	3,14,29			190.09	1.094	293.07	1.79
H_2O	WATER	29,31			273.15	1.4357	373.15	9.720
H_2S	HYDROGEN SULFIDE	16	103.52 126.22	0.3682 0.1213	187.61	0.5681	212.77	4.463
He	HELIUM	17,25	2.186				4.2160	0.020
I_2	IODINE	25			386.8	3.770	456.	9.970
Kr	KRYPTON	25			115.9	0.391	119.75	2.158
N_2	NITROGEN	25	35.62	0.055	63.18	0.172	77.36	1.335
N_2O	NITROUS OXIDE	20,32			182.35	1.563	184.70	3.982
NO	NITRIC OXIDE	33			109.49	0.550	121.36	3.293
NH_3	AMMONIA	21			195.36	1.376	239.68	5.614
Ne	NEON	25			24.55	0.0801	27.07	0.422
O_2	OXYGEN	25	23.89 43.80	0.0214 0.1776	54.36	0.1063	90.19	1.630
SO_2	SULFUR DIOXIDE	24			197.64	1.769	263.08	5.966
Xe	XENON	25			161.3	0.549	165.04	3.021

form* for the range 10^{-11} to 10^3 mm Hg. Log p versus log T plots, employed in the previous publication,[4] were chosen again as the most suitable means of presenting these data over wide pressure and temperature ranges. Figure 1(a) covers the range 1 to 40° K and includes He, H_2, and Ne. The remainder of the gases are plotted in Figure 1(b) for temperatures between 10 and 400° K. As in Table I, ⊙ and ⟍ represent melting points and transition points, respectively. Where the melting point falls outside the plotted range, the letters "s" (solid) or "l" (liquid) have been appended to the chemical symbol.

It is often of interest to establish which gases will exert measurable pressures in a vacuum system kept at a fixed temperature. To this end, it is only necessary to place a vertical line on the chart at the desired temperature and to note the pressure levels at which it intersects the individual curves.

* A limited number of "wall-size" reproductions of these charts are available; requests for copies should be directed to RCA Laboratories, Princeton, N. J.

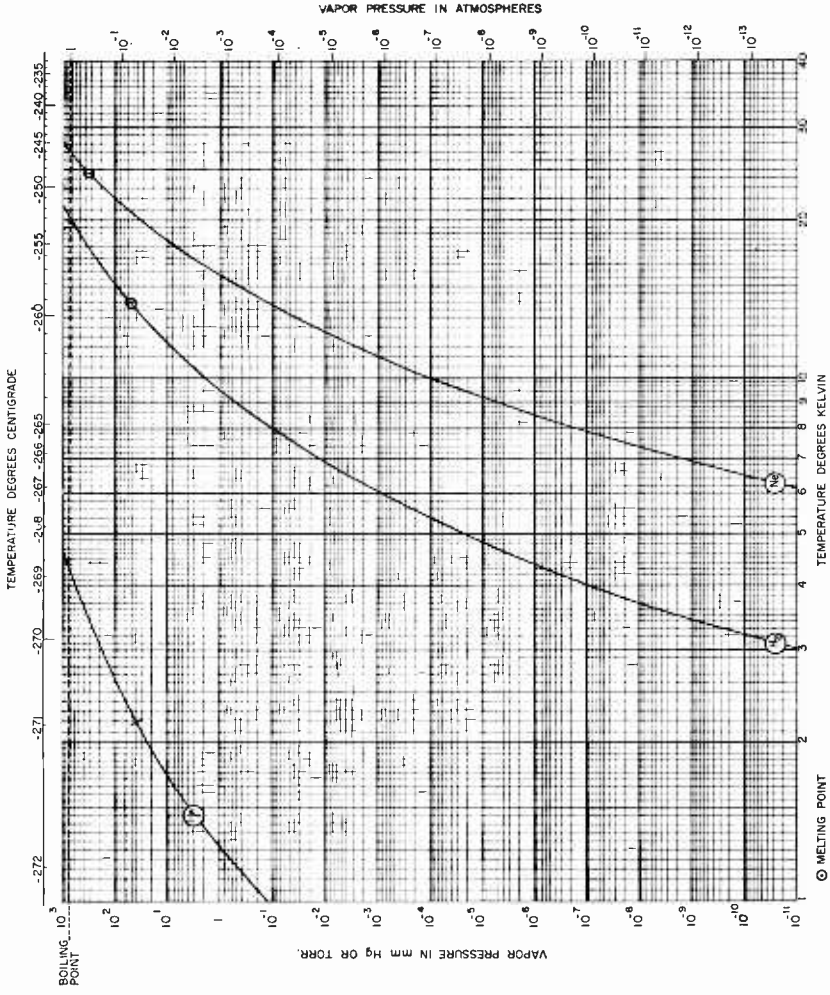


Fig. 1(a)—Vapor pressure curves of common gases.

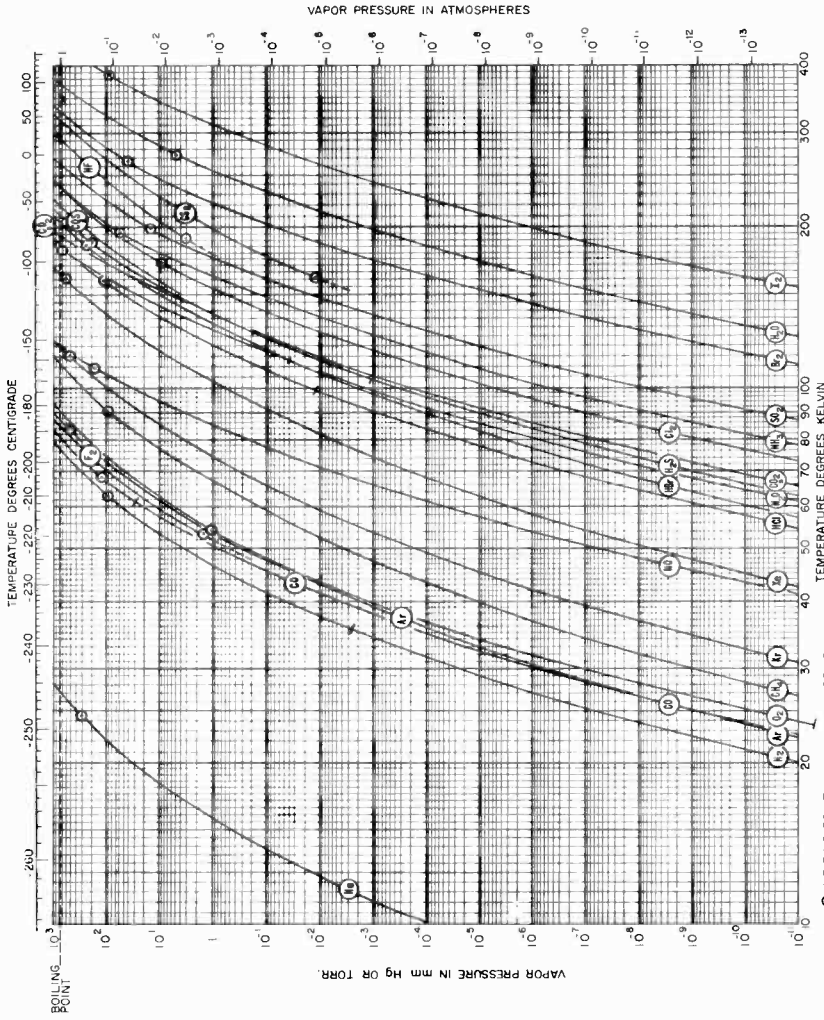


Fig. 1(b)—Vapor pressure curves of common gases.

[5] M. P. Freeman and G. D. Halsey, Jr., "The Solid Solution Krypton-Xenon from 90° to 120° K, The Vapor Pressures of Argon, Krypton and Xenon," *Jour. Phys. Chem.*, Vol. 60, p. 1119, August, 1956.

[6] G. T. Armstrong, F. G. Brickwedde, and R. B. Scott, "Vapor Pressures of the Methanes," *Jour. Res. Nat. Bur. Standards*, Vol. 55, p. 39, July, 1955.

[7] A. W. Tickner and F. P. Lossing, "The Measurement of Low Vapor Pressures by Means of a Mass Spectrometer," *Jour. of Physical and Colloid Chemistry*, Vol. 55, p. 733, May, 1951.

[8] W. F. Giauque and C. J. Egan, "Carbon Dioxide. The Heat Capacity and Vapor Pressure of the Solid. The Heat of Sublimation. Thermodynamic and Spectroscopic Values of the Entropy," *Jour. of Chem. Phys.*, Vol. 5, p. 45, January, 1937.

[9] J. D. Kemp and W. F. Giauque, "Carbonyl Sulfide. The Heat Capacity. Vapor Pressure, and Heats of Fusion and Vaporization. The Third Law of Thermodynamics and Orientation Equilibrium," *Jour. Amer. Chem. Soc.*, Vol. 59, p. 79, January, 1937.

[10] W. F. Giauque and T. M. Powell, "Chlorine. The Heat Capacity Vapor Pressure. Heats of Fusion and Vaporization, and Entropy," *Jour. Amer. Chem. Soc.*, Vol. 61, p. 1970, August, 1939.

[11] J. Hu, D. White, and H. L. Johnston, "Condensed Gas Calorimetry. V. Heat Capacities, Latent Heats and Entropies of Fluorine from 13 to 85°K; Heats of Transition, Fusion, Vaporization and Vapor Pressures of the Liquid," *Jour. Amer. Chem. Soc.*, Vol. 75, p. 5642, November, 1953.

[12] H. W. Woolley, R. B. Scott, and F. G. Brickwedde, "Compilation of Thermal Properties of Hydrogen in Its various Isotopic and Ortho-Para Modifications," *Jour. Res. Nat. Bur. Standards*, Vol. 41, p. 379, November, 1948.

[13] J. R. Bates, J. O. Halford, and L. C. Anderson, "A Comparison of Some Physical Properties of Hydrogen and Deuterium Bromides," *Jour. Chem. Phys.*, Vol. 3, p. 531, September, 1935.

[14] H. Claussen and J. H. Hildebrand, "The Vapor Pressures of Hydrogen and Deuterium Fluorides," *Jour. of Amer. Chem. Soc.*, Vol. 56, p. 1820, August, 1934.

[15] *Handbook of Chemistry and Physics*, 40th Ed., p. 2324, Chem. Rubber Publ. Corp., Cleveland, Ohio, 1959.

[16] W. F. Giauque and R. W. Blue, "Hydrogen Sulfide. The Heat Capacity and Vapor Pressure of Solid and Liquid. The Heat of Vaporization. A Comparison of Thermodynamic and Spectroscopic Values of the Entropy," *Jour. Amer. Chem. Soc.*, Vol. 58, p. 831, May, 1936.

[17] H. Van Dijk and M. Durieux, *Progress in Low Temperature Physics*, Interscience Publishers Co., New York, Vol. II, p. 431, 1957.

[18] B. B. Fisher and W. G. McMillan, "The Sublimation Pressure of Krypton Below 80°K," *Jour. Phys. Chem.*, Vol. 62, p. 494, April, 1958.

[19] H. W. Woolley, "Thermodynamic Properties of Gaseous Nitrogen," *Nat. Adv. Com. Aeron.*, Technical Note 3271, March, 1956.

[20] R. W. Blue and W. F. Giauque, "The Heat Capacity and Vapor Pressure of Solid and Liquid Nitrous Oxide. The Entropy from its Band Spectrum," *Jour. Amer. Soc.*, Vol. 57, p. 991, June, 1935.

[21] R. Overstreet and W. F. Giauque, "Ammonia. The Heat Capacity and Vapor Pressure of Solid and Liquid. Heat of Vaporization. The Entropy Values from Thermal and Spectroscopic Data," *Jour. Amer. Chem. Soc.*, Vol. 59, p. 254, January, 1937.

[22] F. Henning and J. Otto, "Dampfdruckkurven und Tripelpunkte im Temperaturgebiet von 14° bis 90° abs," *Phys. Zeitschr.*, Vol. 37, p. 633, September, 1936.

[23] H. J. Hoge, "Vapor Pressure and Fixed Points of Oxygen and Heat Capacity in the Critical Region," *Jour. Res. Nat. Bur. Standards*, Vol. 44, p. 321, March, 1950.

[24] W. F. Giauque and C. C. Stephenson, "Sulfur Dioxide. The Heat Capacity of Solid and Liquid. Vapor Pressure. Heat of Vaporization. The Entropy Values from Therman and Molecular Data," *Jour. Amer. Chem. Soc.*, Vol. 60, p. 1389, June, 1938.

[25] D. R. Stull and G. C. Sinke, "Thermodynamic Properties of the Elements," *Advances in Chemistry*, Series 18, Am. Chem. Soc., Washington, D. C. (1956).

[26] F. D. Rossini and K. S. Pitzer and others, "Selected Values of Properties of Hydrocarbons," *NBS Circular C461*, p. 132, Issued November, 1947.

[27] J. O. Clayton and W. F. Giauque, "The Heat Capacity and Entropy of Carbon Monoxide, Heat of Vaporization. Vapor Pressures of Solid and Liquid. Free Energy to 5000° K From Spectroscopic Data," *Jour. Amer. Chem. Soc.*, Vol. 54, p. 2610, July, 1932.

[28] C. H. Meyers and M. S. VanDusen, "The Vapor Pressure of Liquid and Solid Carbon Dioxide," *Jour. Res. Nat. Bur. Standards*, Vol. 10, p. 381, March, 1933.

[29] K. K. Kelley, "Contributions to the Data on Theoretical Metallurgy," *U. S. Bureau of Mines Bulletin, 477*, Washington, 1950.

[30] W. Giauque and R. Wiebe, "The Entropy of Hydrogen Chloride. Heat Capacity from 16°K to Boiling Point. Heat of Vaporization. Vapor Pressures of Solid and Liquid," *Jour. Amer. Chem. Soc.*, Vol. 50, p. 101, January, 1928.

[31] N. S. Osborne, H. F. Stimson, and D. C. Ginnings, "Measurements of Heat Capacity and Heat of Vaporization of Water in the Range 0° to 100°C," and "Thermal Properties of Saturated Water and Steam," *Jour. Res. Nat. Bur. Standards*, Vol. 23, p. 197 and 261, respectively, August, 1939.

[32] H. J. Hoge, "Vapor Pressure, Latent Heat of Vaporization, and Triple-Point Temperature of N_2O," *Jour. Res. Nat. Bur. Standards*, Vol. 34, p. 281, March, 1945.

[33] H. L. Johnston and W. F. Giauque, "The Heat Capacity of Nitric Oxide from 14°K to the Boiling Point and the Heat of Vaporization. Vapor Pressures of Solid and Liquid Phases. The Entropy from Spectroscopic Data," *Jour. Amer. Chem. Soc.* Vol. 51, p. 3194, November, 1929.

SPACE-CHARGE EFFECTS IN ULTRA-LOW-NOISE ELECTRON GUNS*†

BY

J. BERGHAMMER

RCA Laboratories
Princeton, N. J.

Summary—The d-c characteristics of a typical electron gun used in low-noise traveling-wave tubes are investigated. It is found that the cathode current exhibits abrupt jumps as the positive potential on the beam-forming electrode is varied. Positive beam-forming potential is known to lead to best noise performance. These d-c current jumps are explicable on the basis of the well-known "6L6 Effect." This space-charge effect, rather than temperature limitation, is the main cause of current saturation in these guns.

INTRODUCTION

VERY LOW NOISE figures have been obtained in traveling-wave tubes by applying a *positive* voltage to the beam-forming electrode surrounding the cathode,[1,2] and by simultaneously increasing the axial magnetic field to prevent interception of electrons. In addition to leading to a lower noise figure, the positive beam-forming voltage has several other direct consequences which are related to the d-c characteristics of the tube. In particular, it has been observed that the cathode current changes abruptly when the beam-forming potential exceeds a certain value. This d-c phenomenon is the subject of the present paper. Experimental evidence is presented which indicates that this jump phenomenon is due to space-charge instability of the type observed in the 6L6.

D-C CHARACTERISTICS OF MODIFIED 6861 TRAVELING-WAVE TUBE

Measurements were made on the modified 6861 traveling-wave tube shown in Figure 1. This tube contains two extra electrodes which facilitate the variation of potential along the beam. A relatively high

* Manuscript received June 30, 1960.

† This work was in part supported by a Signal Corps contract.

[1] M. R. Currie and D. C. Forster, "Low Noise Tunable Preamplifiers for Microwave Receivers," *Proc. I.R.E.*, Vol. 46, p. 570, March, 1958.

[2] M. R. Currie and D. C. Forster, "New Mechanism of Noise Reduction in Electron Beams," *Jour. Appl. Phys.*, Vol. 30, p. 94, January, 1959.

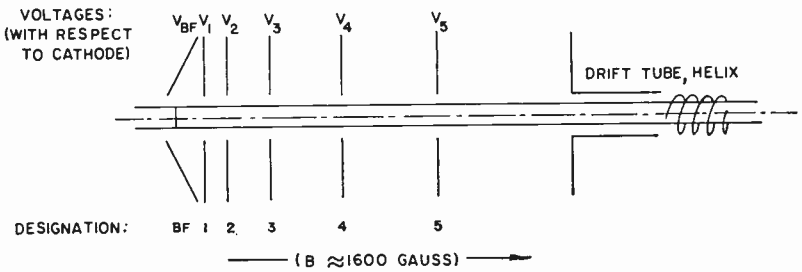


Fig. 1—Schematic view of electron gun used in modified 6861 traveling-wave tube.

magnetic field in the axial direction was used in all measurements to ensure a predominantly laminar flow despite the high transverse electric field produced by high positive beam-forming voltages near the face of the cathode. At moderate positive beam-forming voltages (about 10 volts for a magnetic field of about 1600 gausses) a slight expansion of the beam can be expected, and at higher voltages interception of current would set in. Interception was avoided in all the measurements reported here.

The measured d-c characteristics of the cathode current I_c versus first-anode voltage V_1 with the beam-forming voltage V_{BF} as parameter are shown in Figure 2. All other voltages were kept constant. As can

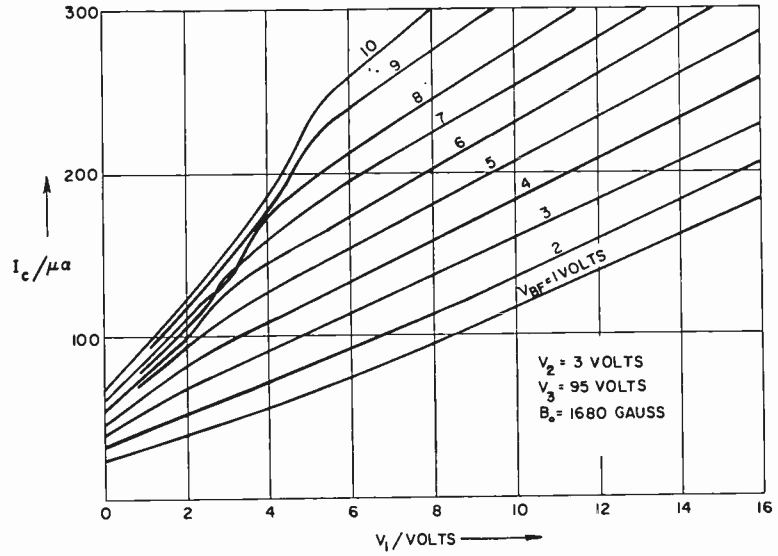


Fig. 2—D-C characteristics of modified 6861 low-noise traveling-wave tube.

be seen, for $V_1 > 5$ volts the curves for constant V_{BF} are almost parallel to each other and have a constant slope of $\partial I_c/\partial V_1 \approx 10$ μa/volt. Furthermore, these lines are approximately equidistant, indicating that $\partial I_c/\partial V_{BF} \approx 20$ μa/volt is constant. In this region, the current control exercised by the beam-forming voltage is about twice as great as that exercised by the first-anode voltage. However, in the region $V_1 < 5$ volts, the curves bunch together as V_1 decreases, as seen in Figure 2 This is referred to as the "cliff region," in analogy with the manner in which the lines of a contour map bunch together in the vicinity of a cliff. To a rough approximation it can be said that $\partial I_c/\partial V_1 \approx 30$ μa/volt, and $\partial I_c/\partial V_{BF} \approx 2$ μa/volt. In the cliff region, then, the beam-forming voltage loses current control to a considerable extent, and the first-anode voltage gains current control. In this region, the first-anode voltage appears to have about 15 times more influence on the current than does the beam-forming voltage.

The existence of a cliff region in the d-c characteristic of ultra-low-noise guns appears to be relatively independent of the geometric configuration of the gun, since very similar curves have been found by Currie and Forster[2] for an annular cathode. These authors attribute the above-mentioned loss of current control of the beam-forming voltage to "some kind of saturation effect," namely *temperature limitation* of the current at the edge of the cathode as the cliff region is approached at constant current.

This explanation is not acceptable for several reasons. For the following argument, it is helpful to construct from Figure 2 a plot of the iso-current curves in a V_1 versus V_{BF} coordinate system, as shown in Figure 3. This plot establishes the linearity properties of the characteristics more clearly. The cliff region now appears to the right of line AA′ of Figure 3 connecting the break-points of slope of the iso-current curves. Temperature limitation in diodes sets in very gradually, whereas in the present structure the saturation effect appears almost suddenly. Figure 3 shows that the larger the V_{BF}, the larger V_1 must be to stay at or to the right of line AA′. This is opposite to what would be required if the knee line AA′ were due to temperature limitation of the cathode edge.

Furthermore, examination of the cliff region in Figure 2 shows that the $V_{BF} =$ constant curves are not parallel. In the most drastic example, the curves for $V_{BF} = 7$ volts and $V_{BF} = 8$ volts cross at $V_1 = 3$ volts. This peculiarity cannot be understood on the basis of temperature-limited emission. It should be emphasized that all of these lines have been repeatedly reproduced.

An explanation of these phenomena is proposed in the following section, wherein space-charge effects of the 6L6 type play a major role.

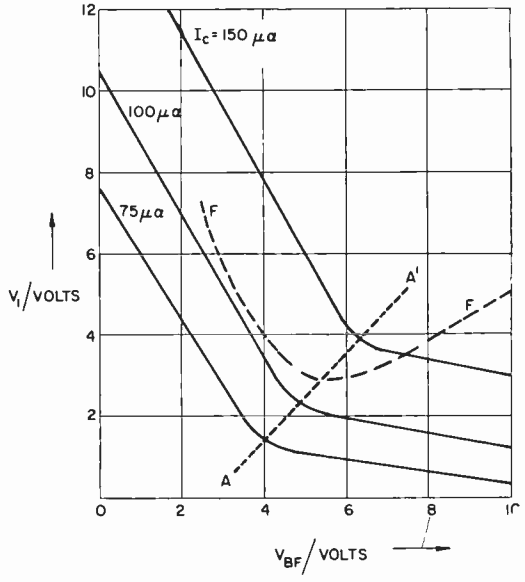


Fig. 3—Iso-current curves exhibiting a saturation-effect at AA'. Dashed line FF is a typical curve for relative noise figure.

Space-Charge Effects

Figure 4 shows the measured cathode current as a function of beam-forming voltage, for three values of V_3 with V_1 and V_2 constant. It is seen that the current varies approximately linearly with V_{BF} up

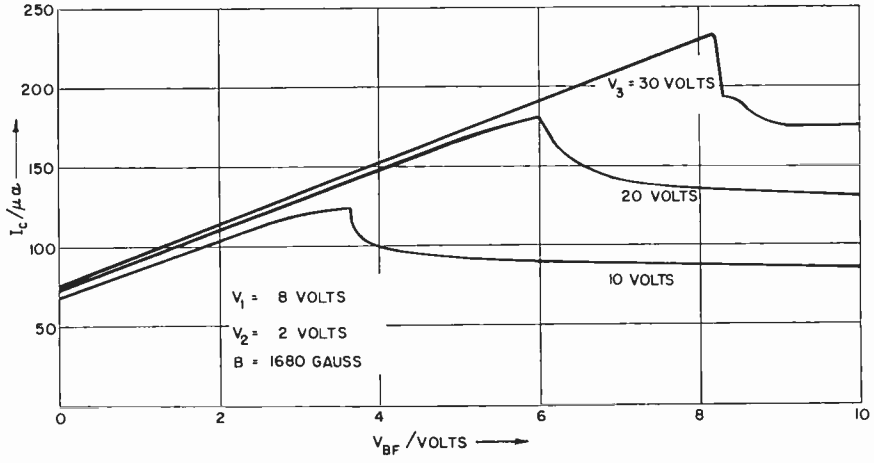


Fig. 4—Observations of abrupt changes of cathode current, I_c, when the beam-forming voltage, V_{BF}, is varied.

to a certain critical point, whereupon the current *drops* abruptly and thereafter tends to decrease slightly. A slope of about 20 μa/volt corresponds to the linear portion of these curves, independently of V_3. The current level of the flat portions of the graphs are seen to depend strongly on V_3. These flat portions correspond, of course, to the cliff region, since the cathode current is essentially constant for increasing V_{RF}.

Similar abrupt changes of current were first observed in beam-power tetrodes, particularly in the 6L6.[3] The explanation given for the 6L6 effect is that a virtual cathode is formed between screen grid and anode when a critical current is reached.[4-9]

Consider two nonintercepting plane-parallel grids at potentials V_a and V_b, separated by a distance d. In the absence of space charge, the potential varies linearly with distance between the two planes, but as a small current of density J is injected through plane a, the potential is depressed. With increasing injected current density a minimum on the potential curve is formed. Near this potential minimum the electrons are slowed down, thus further increasing the charge density, and thereby further lowering the potential. These effects are cumulative and continue until a critical current density, $J_{\max}$, is reached. For the considered one-dimensional case,

$$J_{\max} = \frac{2m\epsilon_0}{9\,ed^2}\,(u_a + u_b)^3, \tag{1}$$

where the velocity of electrons at planes a and b is

---

[3] O. H. Schade, "Beam Power Tubes," *Proc. I.R.E.*, Vol. 26, p. 137, February, 1938.

[4] L. Tonks, "Space Charge as a Cause of Negative Resistance in a Triode and its Bearing on Short Wave Generation," *Phys. Rev.*, Vol. 30, p. 501, October, 1927.

[5] G. Plato, W. Kleen and H. Rothe, "Die Raumladegleichung für Elektronen mit Anfangsgeschwindigkeit," *Zeitschrift für Physik*, Vol. 101, No. 5, p. 509, 1936.

[6] W. Kleen and H. Rothe, "Die Raumladegleichung für Elektronen mit Anfangsgeschwindigkeit," *Zeitschrift für Physik*, Vol. 104, Nos. 11 and 12, p. 711, 1937.

[7] A. V. Haeff, "Space-Charge Effects in Electron Beams," *Proc. I.R.E.*, Vol. 27, p. 586, September, 1939.

[8] B. Salzberg and A. V. Haeff, "Effects of Space-Charge in the Grid-Anode Region of Vacuum Tubes," *RCA Review*, Vol. 2, p. 336, January, 1938.

[9] M. J. O. Strutt and A. vanderZiel, "Über die Elektronenraumladung zwischen ebenen Elektroden, unter Berücksichtigung der Anfangsgeschwindigkeit und Geschwindigkeitsverteilung der Elektronen," *Physica*, Vol. 6, No. 9, p. 62, October, 1939.

$$u_{a.b} = \left(2 \frac{e}{m} V_{a.b} \right)^{1/2}. \tag{2}$$

If the injected current density exceeds J_{max}, the potential distribution collapses. A *virtual cathode* is formed at some mid-plane, causing a certain fraction of the injected current, $(1-p)J$, to return and allowing only a certain fraction of the injected current, pJ, to pass. By integrating Poisson's equation and using the proper boundary conditions, one obtains a relation between the normalized transmitted

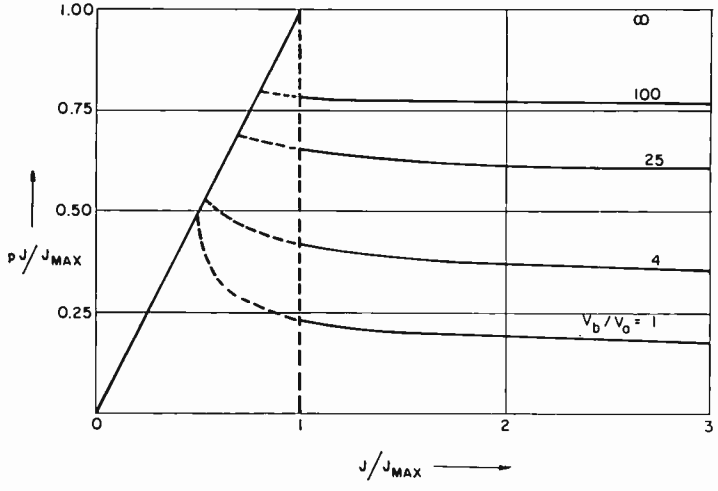


Fig. 5—Normalized transmitted current-density, pJ/J_{max}, versus normalized injected current-density, J/J_{max}, according to one-dimensional model.

current, pJ/J_{max}, and the normalized injected current, J/J_{max}. These results are shown in Figure 5 for various values of V_b/V_a. As long as the injected current is smaller than the maximum current, the full amount of injected current will be transmitted ($p = 1$) independently of the applied voltages V_b/V_a. When J/J_{max} reaches unity, the transmitted current drops suddenly to a value dependent on V_b/V_a. Any further increase of J reduces the transmitted current, pJ, only slightly. The dashed portions of the curves in Figure 5 correspond to a mathematically possible continuation of the curves if the injected current is decreased. This would give rise to a hysteresis-type effect. However, no hysteresis has been observed under the conditions of the present d-c experiments. This is in accord with the theoretical d-c analysis by

Bull,[10] which precludes hysteresis on the basis of the principle of least action. However, experiments by J. M. Hammer and the present author indicate that hysteresis-type behavior can be induced if the injected current is modulated.

The results obtained from this simple model of the 6L6 effect are applicable to the present observations. For the interpretation of Figure 4, the electrodes 2 and 3 are identified with planes *a* and *b*. For the data shown in Figure 4, practically no current is intercepted anywhere in the tube. Thus, the cathode current, I_c, is a true measure of the current transmitted through electrode 3.

It is stipulated that the *injected* current, which is the current emitted by the cathode, is controlled mainly by the beam-forming voltage. Of course, the first-anode voltage also affects the emission; for the conditions of Figure 4, however, V_1 was kept constant and, therefore, any change of current is due solely to a change of beam-forming potential. With this argument, a qualitative correspondence between the observations as presented in Figure 4, and the predictions of the one-dimensional model (Figure 5) is sufficiently well established. Also, it should be mentioned that the slight dependence of current on V_3 in the linear portions of the measured curves (Figure 4) can be attributed to leakage of field through the apertures two and one.

The measurements, then, support the view that the beam-forming voltage does *not* temperature limit the emission, but instead draws so much current from the cathode edge that a virtual cathode is formed which reflects a certain amount of current. Thus, the saturation effect shown by the iso-current curves of Figure 3 is only an apparent effect. The same is true for the loss of current control of the positive beam-forming voltage in the cliff region. The irregularities in the cliff region are most easily explained by the occurrence of a 6L6 effect between any two adjacent electrodes following the cathode. In particular, the overlapping of two $V_{BF} = \text{constant}$ lines of Figure 2 becomes clear. Referring to the previous example, at $V_1 = 3$ volts, a downward current jump occurs as V_{BF} increases from 7 to 8 volts.

The observed maximum transmitted current density usually exceeds the theoretical value of Equation (1) by as much as five times. This discrepancy is attributable to partial ion neutralization of the beam when a virtual cathode is being formed.[11] However, the ratio of trans-

---

[10] C. S. Bull, "Space-Charge Effects in Beam Tetrodes and other Valves," *Jour. Inst. Elec. Eng.*, Vol. 95, pt. 3, p. 17, January, 1948.

[11] K. G. Hernqvist, "Space-Charge and Ion-Trapping Effects in Tetrodes," *Proc. I.R.E.*, Vol. 39, p. 1541, December, 1951.

mitted current densities after and just before the current jump, R_{EXP}, and the corresponding theoretical ratio, R_{TH}, (from Figure 5) are in good agreement for each of the three current jumps, as shown in Table I.

Table I

V_3/V_2	5	10	15
R_{EXP}	0.35	0.56	0.64
R_{TH}	0.43	0.52	0.62

Since the current density is nonuniform across the beam, only the largest current density at the edge of the beam need be counted for determination of R_{EXP}. To eliminate the low-density current flowing independently of V_{BF} in the center portion of the beam, current density is measured from the intercept-points lying near the 70-μa level (Figure 4). The quantitative agreement between experimental and theoretical data demonstrates the validity of the suppositions made to arrive at the results.

Conclusion

The results thus far attained would not evoke much interest—aside from curiosity's sake—unless some correlation between these space-charge effects and the lower noise figures could be pointed out. Although a detailed search for such connections would have to go beyond the scope of this paper, some positive indication in reference to this question can be offered.

To reveal the importance of the above-discussed space-charge effects, a typical relative noise figure curve is shown in Figure 3. This curve descends rapidly as the break-line AA′ is approached and ascends slowly in the portion corresponding to the cliff region. The lowest noise figure usually occurs near the break-line AA′, indicating that the discussed space-charge effects indeed play a significant role in traveling-wave tubes exhibiting ultra-low noise figures.

Acknowledgment

The writer gratefully acknowledges the benefits derived from numerous discussions of this subject with S. Bloom, J. M. Hammer and F. Paschke. Acknowledgment is also due W. A. Conrad who designed the modified 6861 used for the present experiments.

SLOW-WAVE STRUCTURES FOR ELECTROSTATICALLY FOCUSED HIGH-POWER TRAVELING-WAVE TUBES*

By

Erwin F. Belohoubek

RCA Electron Tube Division
Princeton, N.J.

Summary—Two high-power slow-wave structures are described which have provisions for periodic electrostatic focusing of the electron beam. Both slow-wave circuits are of the backward-wave type and electrostatic focusing of the beam is accomplished by the insertion of additional focusing electrodes midway between the drift tubes of the original r-f structure. The incorporation of the electrostatic-focusing arrangement in these structures is achieved without significant deterioration of the propagation characteristics of the original slow-wave structure. Cold test data for X-band designs of both structures are given.

Introduction

THE TREND in many present day traveling-wave-tube applications is toward light weight and high power. Periodic focusing of the electron beam permits a drastic reduction of the weight of the final tube assembly. Of the two basic possibilities, periodic magnetic and periodic electrostatic focusing, the latter is especially advantageous for tubes used in the lower range of the microwave spectrum, where a periodic-magnetic-focusing assembly would represent the larger part of the total weight of the tube. Recent work on periodic electrostatic focusing has shown that solid beams of high perveance can be focused successfully by the use of a prescribed potential distribution between adjacent electrodes.[1] For example, a perveance of 2×10^{-6} ampere/volt$^{3/2}$ can be achieved with a voltage ratio of about 2.5 between alternate electrodes.

Besides a significant reduction in weight, electrostatically focused tubes offer other distinct advantages: (1) Operation over a very wide range of temperatures (the temperature compensation of magnets on

* Manuscript received June 13, 1960.

[1] W. W. Siekanowicz, F. E. Vaccaro, "Periodic Electrostatic Focusing of Laminar Parallel-Flow Electron Beams," *Proc. I.R.E.*, Vol. 47, p. 451, March, 1959.

the contrary is always restricted to a much smaller range); (2) Freedom from ion oscillations and ion bombardment of the cathode; (3) By proper combination of the focusing electrodes with the r-f circuit, electrostatically focused tubes can provide a somewhat higher gain factor than can be obtained from magnetically focused beams of equivalent uniform axial velocity.[2]

Earlier work on periodic electrostatic focusing was confined to bifilar helices.[3-5] However, helices are severely limited in their power-handling capabilities and, therefore, a strong need exists for high-power slow-wave structures which can accommodate electrostatically focused beams.

Most periodic slow-wave structures, except structures of the helix type, have some kind of metal partitions periodically spaced in the axial direction close to the electron beam. There are two principal ways of adapting these structures for periodic electrostatic focusing of the electron beam. One way entails the application of different d-c potentials to alternate partitions and, therefore, requires d-c isolation of adjacent partitions. The other requires the insertion of a second set of focusing electrodes having a potential different from that of the rest of the structure. Because such changes will affect the r-f behavior of the slow-wave structure, a number of new problems arise. The following conditions must be met for successful incorporation of electrostatic focusing:

1. New modes originated by the introduction of the focusing electrodes or by the d-c isolation of the loading plates must be suppressed, shifted, or damped sufficiently to avoid unwanted oscillations.

2. The r-f radiation caused by the d-c isolation of adjacent electrodes must be kept low.

3. The interaction impedance of the working passband must not be lowered appreciably.

---

[2] W. W. Siekanowicz, "A Small-RF-Signal Theory for an Electrostatically Focused Traveling-Wave Tube," to be published in *Proc. I.R.E.*

[3] P. K. Tien, "Focusing of a Long Cylindrical Electron Stream by Means of Periodic Electrostatic Fields," *Jour. App. Phys.*, Vol. 25, p. 1281, October, 1954.

[4] K. K. N. Chang, "Biperiodic Electrostatic Focusing for High Density Electron Beams," *Proc. I.R.E.*, Vol. 45, p. 1522, November, 1957.

[5] D. J. Blattner, F. E. Vaccaro, C. L. Cuccia, and W. C. Johnson, "Medium-Power L- and S-Band Electrostatically Focused Traveling-Wave Tubes," *RCA Review*, Vol. 20, p. 426, September, 1959.

In addition, breakdown and heat dissipation considerations must be taken into account.

This paper describes two slow-wave structures which fulfill the above requirements to a good degree.* Both structures are of the backward-wave type and have focusing electrodes inserted midway between the original cell partitions. This arrangement was chosen because it permits the use of alternating thick and thin focusing electrodes which comes close to the ideal conditions for the focusing of round, solid beams of high perveance.[1]

Crossed-Bar Structure with Focusing Electrodes Individually Supported by Chokes

Slow-wave structures of the backward-wave type are characterized by the fact that group velocity and phase velocity of the fundamental space harmonic have opposite signs. The first structure investigated

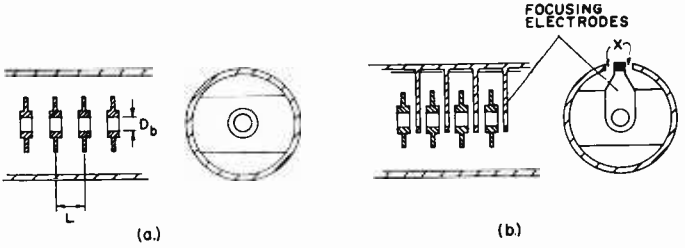


Fig. 1—Round waveguide with periodically spaced bars: (a) basic slow-wave structure, and (b) focusing electrodes inserted.

is shown in Figure 1a. The basic slow-wave circuit consists of a round waveguide with periodically spaced bars, originally proposed by Nalos.[6] Such a structure, operated as a forward-wave amplifier in the first space harmonic, has a cell length nearly twice as large as that of a structure operating in the fundamental space harmonic at the same beam voltage. A space-harmonic structure, therefore, provides enough room for the introduction of a new set of focusing electrodes midway between cell partitions. Figure 1b shows the structure with additional focusing electrodes inserted for electrostatic focusing.

Another advantage of this arrangement is that the focusing period

* Part of the work described in this paper was sponsored by the Naval Research Laboratory under Contract Nonr 2568(00).

[6] M. Chodorow and E. J. Nalos, "The Design of High-Power Traveling-Wave Tubes," *Proc. I.R.E.*, Vol. 44, p. 649, May, 1956.

is less by a factor of 2 than for a structure in which only adjacent partitions are used as focusing electrodes. A reduction of the focusing period in this way results in a larger ratio of beam-hole diameter D_b to focusing period L_f. In addition, alternate focusing electrodes can be made of different thickness. Both factors markedly improve the focusing performance as previous focusing tests[1] have shown.

Figure 2a shows the ω–β diagram of the original slow-wave structure, where β is defined by $\beta = \beta_0 + (2\pi n/L)$, and β_0 is the phase

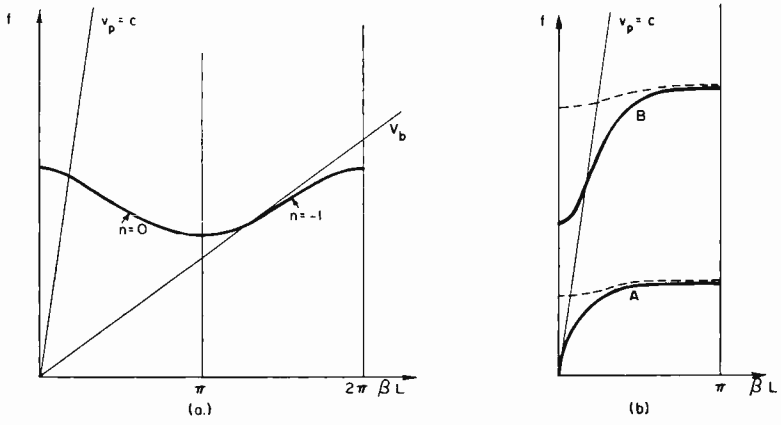


Fig. 2—ω–β diagrams of (a) round waveguide with periodically spaced bars, and (b) round waveguide with focusing electrodes only (for the dashed curve the focusing electrodes are short circuited to waveguide at points X in Figure 1, and for the solid curve they are d-c isolated from waveguide).

shift per unit length for the fundamental space harmonic. In this diagram, the —1 space harmonic is indicated, together with the voltage line V_b for operation as a forward-wave amplifier. All subsequent ω–β diagrams are drawn only for the range $0 < \beta L < \pi$.

New Modes Introduced by Inserting the Focusing Electrodes

The insertion of focusing electrodes into the basic slow-wave structure introduces several new modes. For easier identification of the possible modes in the final structure, the ω–β diagram for a round waveguide containing only the focusing electrodes is shown in Figure 2b. If the focusing electrodes are short circuited to the round waveguide at the points marked X in Figure 1b, the corresponding ω–β diagram is given by the dashed lines in Figure 2b. The lower passband

(curve A) is associated essentially with the $\lambda/4$ resonance of the focusing electrodes, and the next higher passband (curve B) belongs to the $3\lambda/4$ resonance.[7] Both passbands are fairly narrow. If, however, the focusing electrodes are dc-isolated from the round waveguide, the lower passband shows a low-pass characteristic, and the next higher passband also covers a much wider frequency range than before (solid lines in Figure 2b). The upper cutoff frequencies of both passbands remain practically constant.

Thus far, the ω–β diagrams for the basic slow-wave structure and for the focusing electrodes alone in a round waveguide (without the bars present) have been discussed. A superposition of the two diagrams in Figure 2 then should give the ω–β diagram of the composite structure, provided that

1. The passband of the coupled-bar structure (Figure 2a) is not changed when the focusing electrodes are inserted symmetrically between adjacent bars, and
2. The passbands associated with the focusing electrodes are not affected by the horizontal bars.

These conditions can not be fulfilled entirely. For example, some additional capacitive loading normally occurs and shifts the passbands to a slightly lower frequency. However, a good qualitative picture of the possible modes can be obtained this way.

If the field configurations of the modes of both structures were perfectly orthogonal to each other (which situation corresponds to no coupling between those modes), the composite ω–β diagram would correspond to the one shown in Figure 3a. In practice, however, there is some coupling between these modes, and measured curves are of the form shown in Figure 3b. The amount of separation between curves A and B indicates the degree of coupling. The further the curves deviate from the original crossover point, the greater the coupling between the two modes.

The rules concerning the coupling between modes of propagation, as originally described by Pierce,[8] give a very good explanation for the complex-looking ω–β diagrams obtained in structures where several modes of propagation are possible within the same frequency range.

---

[7] For a discussion of qualitative consideration of how to obtain the general shape of the ω–β diagram for various slow-wave structures see E. Belohoubek, "Propagation Characteristics of Slow-Wave Structures Derived from Coupled Resonators," *RCA Review*, Vol. 19, p. 283, June, 1958.

[8] J. R. Pierce, "Coupling of Modes of Propagation," *Jour. of Appl. Phys.*, Vol. 25, p. 179, February, 1954.

Field measurements based on the dielectric bead perturbation method showed clearly that the lower portion of the passband B still has the same field configuration as the original coupled-bar structure, while the upper end of that passband belongs to the $3\lambda/4$ resonance of the focusing electrodes. Passband A shows a similar mode change.

As can be seen from Figure 3b, oscillation problems are to be expected from passbands B and C as well as a strong reduction in the useful bandwidth caused by the mode crossover. Perturbation measurements indicated that the impedance of the competing passbands B and C in the range of interest is only slightly lower than the impedance of the operating passband A.

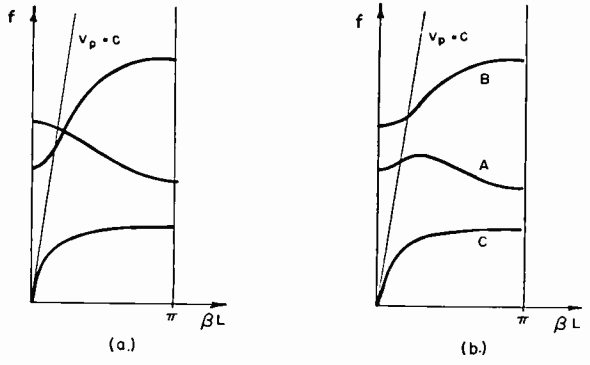


Fig. 3—ω–β diagrams of coupled-bar structure with focusing electrodes d-c isolated from main body of structure: (a) theoretical curves for the assumption of no coupling between modes, and (b) measured curves for the same condition.

Elimination of Unwanted Modes

The structure shown in Figure 1b was discussed in such detail to show the principal problems which arise when a new set of dc-isolated electrodes is inserted. Various schemes which involve damping or perturbing the unwanted modes without affecting the original passband of the coupled-bar structure were tried. These approaches finally led to the structure shown in Figure 4. For a better understanding of the properties of this structure, the structure of Figure 1b, modified so that the focusing electrodes are short circuited to the waveguide at points X, will be further considered.

For this structure, the two competing modes associated with the $\lambda/4$ and $3\lambda/4$ resonance of the focusing electrodes have a fairly narrow bandwidth (Figure 2b). Proper shaping of the electrodes allows one

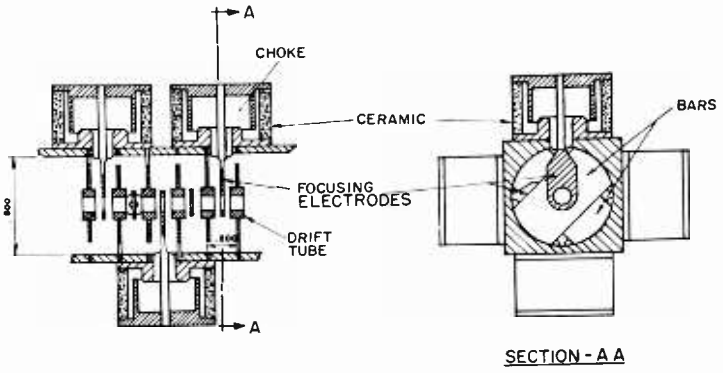


Fig. 4—Crossed-bar structure with focusing electrodes individually supported by choke sections.

passband to be above, and one to be below the working passband of the coupled-bar structure, so that a mode crossover can be avoided. The necessary dc-isolation for the focusing electrodes can be provided by the substitution of separate coaxial choke sections for the short. However, the choke sections are effective only over a certain limited frequency range and their use causes the two passbands associated with the resonances of the focusing electrodes to move slightly toward each other. As shown in Figure 5, a choke below its center frequency has a slightly capacitive input impedance and, therefore, shifts the passband associated with the $\lambda/4$ resonance of the focusing electrodes to a higher frequency. On the other hand, the passband associated with the $3\lambda/4$ resonance of the focusing electrodes is shifted to a lower frequency. A mode crossover, however, can be avoided even for very large bandwidths of the working passband if the choke sections are made with a sufficiently large impedance ratio Z_1/Z_2.

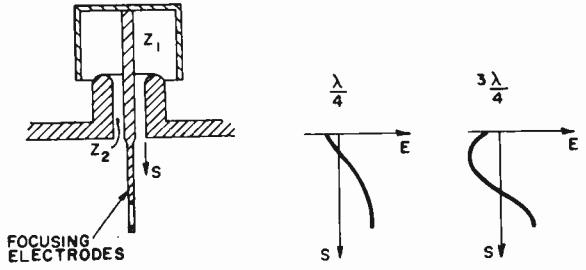


Fig. 5—Choke-supported focusing electrodes with potential distribution for $\lambda/4$ and $3\lambda/4$ resonance.

The final structure is obtained by rotating alternate bars by 90 degrees. This rotation increases the bandwidth of the structure and slightly decreases the stored energy between bars, but does not otherwise affect the propagation characteristic of the structure. The focusing electrodes are then inserted at an angle of 45 degrees with respect to the bars in such a way that they lie on a helical path having a pitch equal to 4 cell lengths, as shown in Figure 4. This arrangement preserves the original periodicity and provides enough room for the placement of choke sections with large characteristic impedance Z_1. In addition, the two passbands which are associated with the resonances of the focusing electrodes, and which still were objectionable from the standpoint of backward-wave oscillations, can no longer be detected on cold test.

A structure of the type shown in Figure 1b, in which all choke sections were aligned axially, already showed strong damping of the passbands associated with the focusing electrodes. This damping was caused by the chokes radiating in these frequency ranges, but the two passbands could still be clearly identified. After the rotation of consecutive chokes by 90 degrees, however, the focusing electrodes were effectively decoupled from each other, and careful measurements did not reveal any energy propagation in the structure associated with the resonances of the focusing electrodes.

The ω–β diagram of a crossed bar structure designed for 5-kilowatt pulsed operation at X-band is shown in Figure 6. At frequencies up to 12,000 megacycles, no other passbands besides the working passband of the coupled-bar structure could be detected. As in any periodic slow-wave structure, other modes will be set up still higher in frequency, but the additional damping caused by strong radiation of the choke sections in these frequency ranges may help appreciably to avoid oscillations.

The introduction of the focusing electrodes reduced the interaction impedance $E_{-1}^2/(2\beta_{-1}^2 P)$ of the working passband by about 30 per cent. This reduction depends on the specific dimensions of the focusing electrodes used, and would be smaller for thinner electrodes. No radiation from the choke sections was found within the frequency range of the working passband. The value of Q for a short-circuited section of the X-band model was about 1600. This figure indicates that there is no significant attenuation from the focusing arrangement.

At X-band the maximum r-f power output of a structure of this type is limited by the voltage breakdown between the focusing disks and the drift tubes. The voltage gradient for the 5-kilowatt design is about 450 volts/mil which comes close to the limit used in standard

tube design. Because the r-f period and, therefore, the spacing between the focusing electrodes varies with the square root of the beam voltage, the voltage gradient becomes prohibitive for very-high-voltage structures at X-band. At lower frequencies, however, the voltage breakdown problem is much less serious and power levels up to a few megawatts at S-band are feasible.

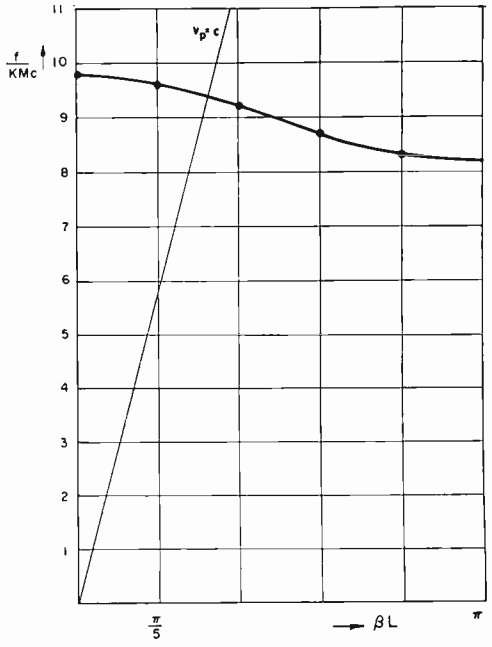


Fig. 6—ω–β diagram of a crossed-bar structure with choke-supported focusing electrodes.

Structure of the Folded-Line Type with Inserted Focusing Electrodes

Two of the principal disadvantages of the structure described above are its somewhat complicated construction and its relatively large over-all dimensions caused by the chokes extending outward all around the basic slow-wave structure. Two versions of a structure which overcomes these difficulties and is, therefore, more suitable for applications at longer wavelengths, are shown in Figures 7 and 8. The basic slow-wave structure, a folded-line-type circuit, also belongs to the group of backward-wave structures and has focusing electrodes inserted midway between the original cell partitions. The field con-

figuration of the fundamental mode of propagation is such that it allows the introduction of the focusing electrodes without an additional choke arrangement. As is the case with a linear antenna inserted into the narrow side of a rectangular waveguide, no energy is coupled out by the supporting rods of the focusing electrodes. In the actual structure, however, fringing fields and slight asymmetries could cause small amounts of energy to be coupled out, lead to irregularities in the propagation characteristic, and increase the attenuation per cell. In addition, the new modes associated with resonances of the focusing electrodes have to be eliminated to avoid spurious oscillations. Two possible methods for minimizing these difficulties are discussed below.

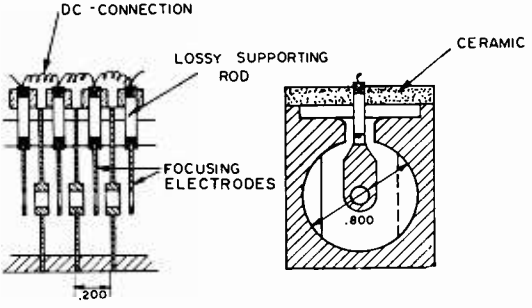


Fig. 7—Folded-line structure with focusing electrodes supported by lossy rods.

Focusing Electrodes Supported by Lossy Rods

Figure 7 shows a folded-line structure in which the focusing electrodes are supported by a lossy material (e.g., carbonized ceramic). The junction between the lossy supporting rod and the focusing electrodes represents a strong reflection, and most of the energy coupled into the coaxial line by fringing fields is reflected there. Furthermore, the small amount of energy which propagates beyond this reflection point is attenuated in the lossy rod. The resistivity of the rods is adjusted to prevent any radiation, but it must also be low enough to supply the necessary d-c voltage to the focusing electrodes at normal amounts of intercepted beam current.

The high attenuation of the rods also effectively damps a narrow passband closely associated with a $\lambda/2$ resonance of the focusing disks. This passband can easily be measured when the lossy supporting rods are replaced by a dielectric. The introduction of loss in the supports, however, made it impossible to detect this mode, and the resulting ω–β diagram is nearly identical to the one of the crossed-bar structure

previously discussed. An X-band model of this type of structure showed that a Q in the order of 1300 can be obtained. This value, according to the equation $\alpha = \pi f/(Qv_g)$ corresponds to an attenuation of less than 0.05 decibel per cell.

The focusing electrodes were made of molybdenum to permit high-temperature operation and to increase the average power-handling capabilities of the structure. A form of silicon carbide† is currently being tested as a material for the supporting rods. This material has good heat-conducting properties, good mechanical strength, and is available in various volume resistivities. The use of a lossy material for the supporting rods also allows the introduction of tapered attenuation by extending the lossy rods somewhat into the round waveguide.

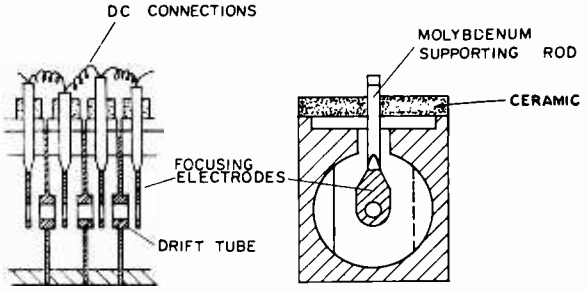


Fig. 8—Folded-line structure with focusing electrodes supported by metal rods of varying lengths.

Focusing Electrodes Supported by Metal Rods of Varying Length

The use of lossy rods as supports for the focusing electrodes produces some difficult technological problems and also limits the heat-dissipating capabilities and the ruggedness of the tube. An alternative solution shown in Figure 8 uses metal supporting rods of varying length. If the lossy rods are replaced by metallic rods, new passbands occur which are associated with $n\lambda/2$ resonances of the focusing electrodes. However, random variations in the length of the supporting rods badly distort the periodicity of the slow-wave circuit formed by the focusing disks and their supporting rods, so that no energy propagation occurs. The d-c connections between the focusing electrodes can then be made either by small wire helices or some attenuating material with low d-c resistivity.

Another problem that arises concerns the manner in which the

† Produced by Semicon Associates Inc.

propagation characteristic of individual cells is affected by the supporting rods of different length, and by the additional d-c connections. Cold-test measurements showed that as long as the focusing electrodes do not have an $n\lambda/2$ resonance within the working passband of the structure, the propagation characteristic is not much affected by changes in the length of the supporting rods, or by d-c connections attached to the rods. In fact, the changes in the propagation characteristic are of the same order or even smaller than the changes which occur between different cells as a result of machining and brazing inaccuracies.

The attenuation per cell was investigated by measurement of the Q of an X-band structure which was short circuited at both ends. The Q values, which ranged from 1200 to 1600 indicated that there is no appreciable radiation loss. These Q values can be obtained only if the focusing electrodes are aligned carefully with respect to the cell partitions. A slight tilting of the focusing elements, for example, leads to coupling out of r-f energy and an appreciable decrease in Q. The ω–β diagram for this tube type is also practically identical to the one of the crossed-bar structure shown in Figure 7. The remarks about voltage breakdown and the reduction in interaction impedance caused by the insertion of the focusing elements apply here as well.

Conclusion

Electrostatic focusing permits the design of extremely lightweight tubes, an advantage especially important at lower wavelengths, where the weight of the focusing magnets even for periodically focused tubes becomes objectionable. Two structures of the backward-wave type, a folded line structure and a crossed-bar structure, were adapted for electrostatic focusing by the insertion of additional focusing electrodes between the original cell partitions. Cold test models of both structures were designed and tested for operation at X-band. Scaling of these models to lower frequencies indicates the feasibility of electrostatically focused megawatt tubes in the frequency range at or below S-band.

The focusing electrode arrangement is compatible with optimum beam focusing requirements and was incorporated into both r-f structures without detrimental effects on the propagation characteristics of the original slow-wave structures.

Acknowledgment

The author wants to express his appreciation to W. R. Beam, F. E. Vaccaro, and W. W. Siekanowicz for many helpful discussions.

DESIGN CONSIDERATIONS FOR GRID-CONTROLLED ELECTRON GUNS FOR PULSED TRAVELING-WAVE TUBES*

By

H. J. Wolkstein

RCA Electron Tube Division,
Harrison, N. J.

Summary—*This paper describes the modification of the convergent-flow electron gun used in traveling-wave tubes to include an additional grid-type aperture for pulsed-beam control. Methods are suggested for the determination of grid parameters, current transmission, and cutoff amplification factor for the modified configuration. Although the use of a grid-type electrode for beam control is not in itself novel, the methods of grid placement and associated design expressly for the convergent-flow gun have not previously been discussed. The method of solution employed should be useful in basic pulsed-electron-gun design.*

Introduction

THE convergent-flow electron gun, which is used as a source of beam current in many diverse applications, was initially suggested by J. R. Pierce in 1940.[1] Boundary conditions for this gun were based on the earlier work of Langmuir and Blodgett,[2] who investigated the inward radial current flow of a spherical diode. The electrode geometry for simple sectors of this convergent-flow diode was empirically determined by Helm, Spangenberg, and Field[3] by means of electrolytic field mapping. The high beam-current density, low cathode loading, and large beam-size compression ratio attainable with the Pierce-type gun caused it to be used extensively. The need for other characteristics in certain areas of usage, however, indicated the need for electrode reshaping. Accordingly, modifications and alternate design methods have been suggested in which the electrode shape is varied to correct the strong divergent effects of the anode

* Manuscript received June 1, 1960.

[1] J. P. Pierce, "Rectillinear Electron Flow in Beams," *Jour. App. Phys.*, Vol. 11, p. 548, August, 1940.

[2] I. Langmuir and K. B. Blodgett, "Currents Limited by Space Charge Between Concentric Spheres," *Phys. Rev.*, Vol. 24, p. 49, July, 1924.

[3] R. Helm, K. Spangenberg, and L. M. Field, "Cathode-Design Procedure for Electron-Beam Tubes," *Elec. Comm.*, Vol. 24, p. 101, March, 1947.

aperture and to extend both the perveance and high-voltage capabilities of convergent-flow electron guns.[4,5] This paper describes further modification of the Pierce-type electron gun by the inclusion of an additional grid-type aperture for pulsed-beam control.

Electron-Beam Pulsing

In conventional beam pulsing of a convergent-flow electron gun, the accelerating anode is driven above cathode potential, as shown in Figure 1. For high-voltage anode operation, the drive potential is of considerable magnitude. The pulse energy required varies as the square of the pulsing voltage and linearly with the pulse repetition rate and shunt input capacitance. This energy transfer often imposes a severe load on the pulse modulator. Further restrictions on modulator size and weight (especially in airborne systems) make it mandatory to seek more efficient beam-pulsing methods.

The problem of beam switching has been alleviated in the past through the use of special beam-pulsing electrodes or nonintercepting-type pulsing members. For example, the beam-forming electrode of the Pierce gun is used widely as a pulse modulator with little structural modification. The use of an intermediate pulsing anode between the main accelerating electrode and the cathode permits low-cutoff operation. These nonintercepting methods and other similar variations are usually restricted, however, to a cutoff amplification factor* of five or less, which is inadequate in many applications.

The need for more efficient beam control made it desirable to examine the use of a grid-type high-amplification-factor pulse electrode. Two methods may be used for determining grid placement relative to the anode and cathode electrodes of the convergent-flow gun. The first method utilizes the classical approach of Langmuir and Compton.[6] The second method, which is preferred, is based on several approximations, but appears to yield very accurate results. In this method, the spherical surfaces are replaced by parallel surfaces. In both methods, the grid-type pulsing electrode operates at relatively

---

[4] M. Muller, "New Points of View in the Design of Electron Guns for Cylindrical Beams of High Space Charge," *Jour. Brit. I.R.E.*, Vol. 56, p. 83, February, 1956.

[5] L. E. S. Mathias, P. G. R. King, "An Experimental Investigation of High Perveance Guns," *SERL Tec. Jour.*, Vol. 6. No. 2, p. 21, August, 1956.

* Cutoff amplification factor is defined as the ratio of accelerating anode voltage to grid-cutoff voltage.

[6] I. Langmuir and K. T. Compton, "Electrical Discharges in Gases. Part II, Fundamental Phenomena in Electrical Discharges," *Reviews of Modern Physics*, Vol. 3, p. 191, April, 1931.

low positive voltage and is positioned to conform with an identical equipotential line between the cathode and anode of the convergent-flow system. However, inherent limitations of the pulsing grid due to thermal considerations impose restrictions on the duty factor.

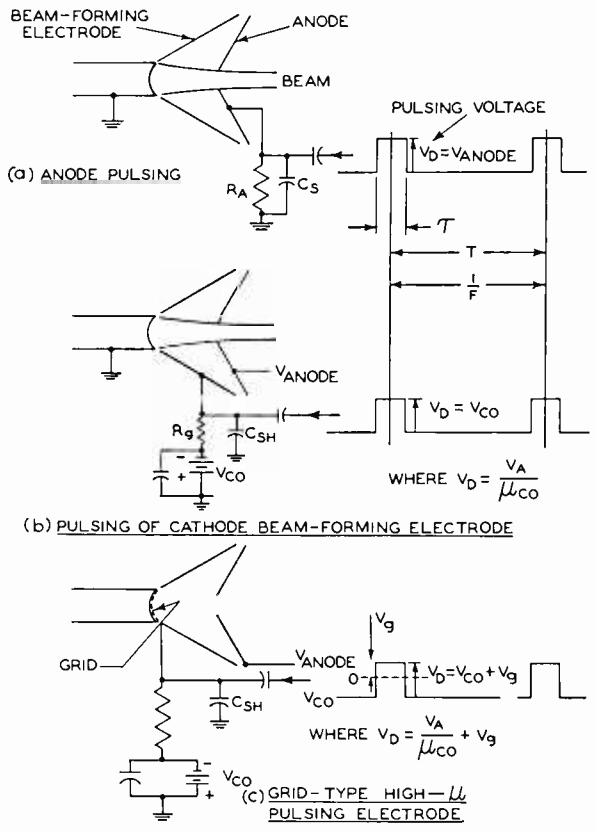


Fig. 1—Illustration of various electron-gun-pulsing methods.

Placement of Pulsing Grid

Langmuir and Compton have shown that the radial space-charge-limited current flow, I, between the surfaces of a spherical cathode and anode is given by

$$I = \frac{16\pi\epsilon\sqrt{2\eta}\,V^{3/2}}{9(-\alpha)^2} \text{ amperes,} \qquad (1)$$

where the Langmuir space-charge-current flow factor $(-\alpha)^2$ is related to the radii of the spherical surfaces with enclosed anode (values for this function are given in Table I).

If the numerical constants for the permittivity of vacuum and for the charge-to-mass ratio of the electron are substituted in Equation

Table I—Tabulation of the Langmuir Space-Charge-Current Flow Factor $(-\alpha)^2$.

R_c/R_r	$(-\alpha)^2$	R_c/R_r	$(-\alpha)^2$
1.0	0.0000	6.5	13.35
1.05	0.0024	7.0	15.35
1.1	0.0096	7.5	17.44
1.15	0.0213	8.0	19.62
1.2	0.0372	8.5	21.89
1.25	0.0571	9.0	24.25
1.3	0.0809	9.5	26.68
1.35	0.1084	10	29.19
1.4	0.1396	12	39.98
1.45	0.1740	14	51.86
1.5	0.2118	16	64.74
1.6	0.2968	18	78.56
1.7	0.394	20	93.24
1.8	0.502	30	178.2
1.9	0.621	40	279.6
2.0	0.750	50	395.3
2.1	0.888	60	523.6
2.2	1.036	70	663.3
2.3	1.193	80	813.7
2.4	1.358	90	974.1
2.5	1.531	100	1144
2.6	1.712	120	1509
2.7	1.901	140	1907
2.8	2.098	160	2333
2.9	2.302	180	2790
3.0	2.512	200	3270
3.2	2.954	250	4582
3.4	3.421	300	6031
3.6	3.913	350	7610
3.8	4.429	400	9303
4.0	4.968	500	13015
4.2	5.528	600	
4.4	6.109	800	
4.6	6.712	1000	
4.8	7.334	1500	
5.0	7.976	2000	
5.2	8.636	5000	
5.4	9.315	10000	
5.6	10.01	30000	
5.8	10.73	100000	
6.0	11.46		

(1), the current flow is given by

$$I = \frac{29.34\ (10^{-6})\ V^{3/2}}{(-\alpha)^2} \text{ amperes.} \tag{2}$$

This basic equation suggests that at a distance R_r from the center of curvature of the diode, the potential variation V_r necessary to sustain radial current flow between the cathode and anode surfaces is given by

$$V_r = V_a \left[\frac{(-\alpha_r)^2}{(-\alpha_a)^2} \right]^{2/3} \tag{3}$$

where V_a is the operating potential of the anode of the basic diode at a distance R_a from the center of curvature of the diode, $(-\alpha_a)^2$ is a function of the ratio of the radii of curvature of the cathode and anode (R_c/R_a), and $(-\alpha_r)^2$ is a function of the ratio of the radii of curvature of the cathode and the equipotential line (R_c/R_r).

The inclusion of an additional spherical electrode, a control grid, between the anode and cathode surfaces with radius R_g will not alter the radial electric field nor the space-charge-limiting conditions, provided its operating voltage conforms to the displaced equipotential line at that radius. This intermediate electrode for pulse beam control can possess any desirable screening fraction without affecting the conditions on the cathode side of the grid configuration. Operation of the control grid at a definite potential V_g normalized by the operating anode potential V_a of the basic diode yields

$$\frac{V_a}{V_g} = \left[\frac{(-\alpha_a)^2}{(-\alpha_g)^2} \right]^{2/3} = \left[\frac{f_1\ (R_c/R_a)}{f_2\ (R_c/R_g)} \right]^{2/3} \tag{4}$$

where $(-\alpha_g)^2$ is a function of the ratio of radii of curvature of the cathode and grid (R_c/R_g). Equation (4) is plotted for various levels of R_c/R_a in Figure 2. The ratio R_c/R_a for the basic Pierce-type gun is adequately described by the desirable perveance and initial convergence half-angle. Values of R_c/R_g slightly larger than unity are obtained for closely spaced grids, as shown by the curve of radial grid parameter $(-\alpha_g)^2$ in Figure 3. Grid-cathode separation d_{gk} is determined by use of the normalizing parameter R_c, the cathode radius of curvature of the basic gun. Figure 4 shows typical values for average

grid–cathode separation based on the desired operating potential, perveance, and tolerable cathode loading.

Approximate Method

An alternate method for determining the grid–cathode separation of a closely spaced spherical system makes use of plane-parallel surface approximations. Space-charge-limited current flow between closely

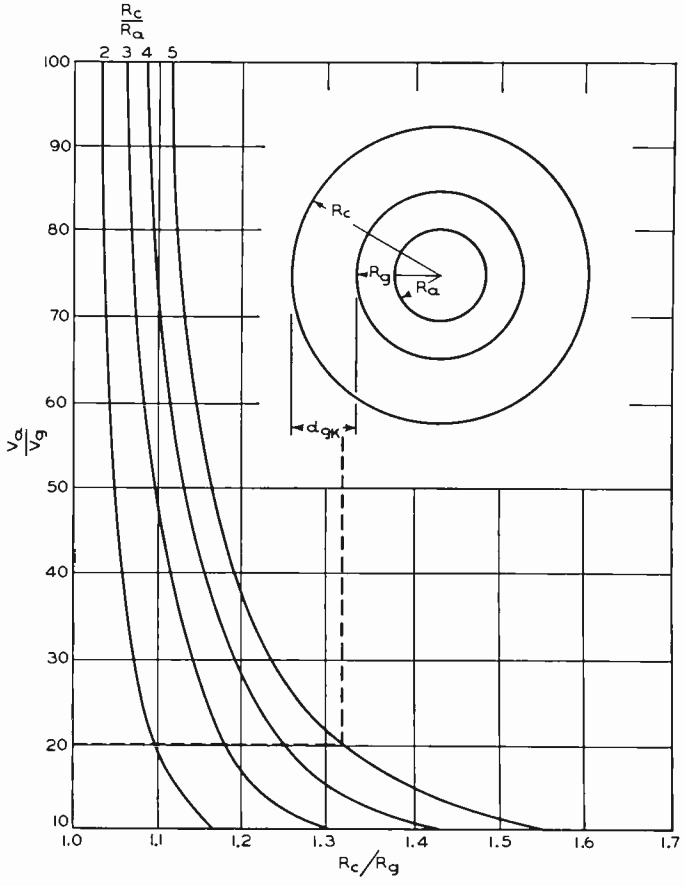


Fig. 2—Curves of R_c/R_g in terms of R_c/R_a and V_a/V_g.

spaced cathode and anode spheres having common centers can be adequately related to the current flow between plane-parallel cathode and anode surfaces having identical spacing, as shown in Figure 5.

The current density for plane-parallel surfaces, J_{11}, is given by

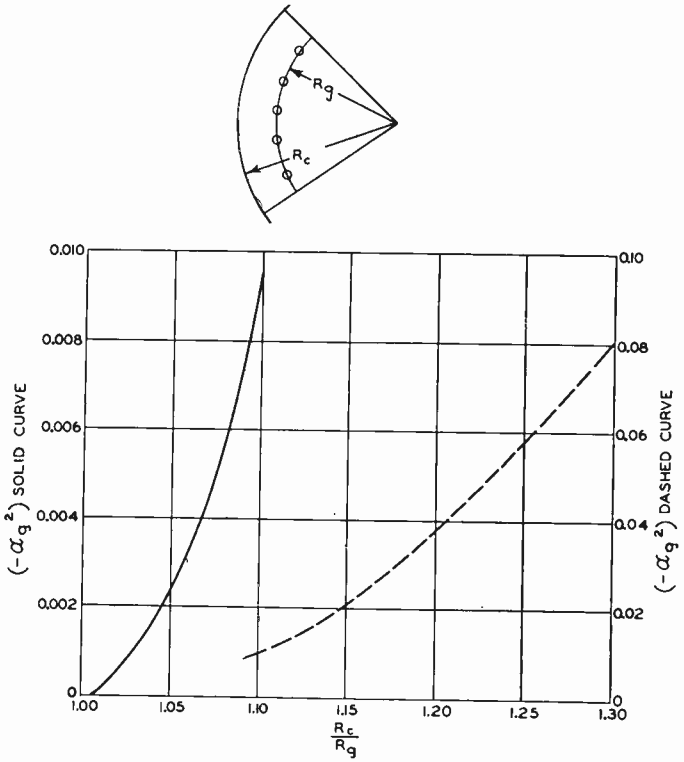


Fig. 3—Values of the space-charge factor $(-\alpha_g)^2$ for small values of R_c/R_g.

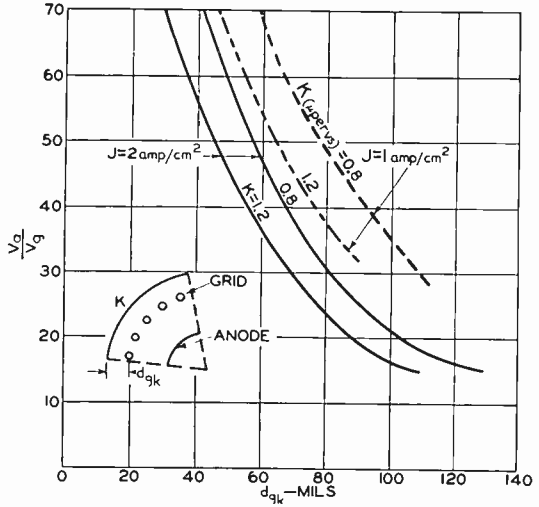


Fig. 4—Grid–cathode separation in terms of perveance, cathode loading, and the ratio of anode potential to grid operating potential.

$$J_{11} = \frac{2.335\ (10^{-6})\ V_g^{3/2}}{d_{gk}^2} \tag{5}$$

where d_{gk} is the distance between closely spaced parallel grid and cathode spheres.

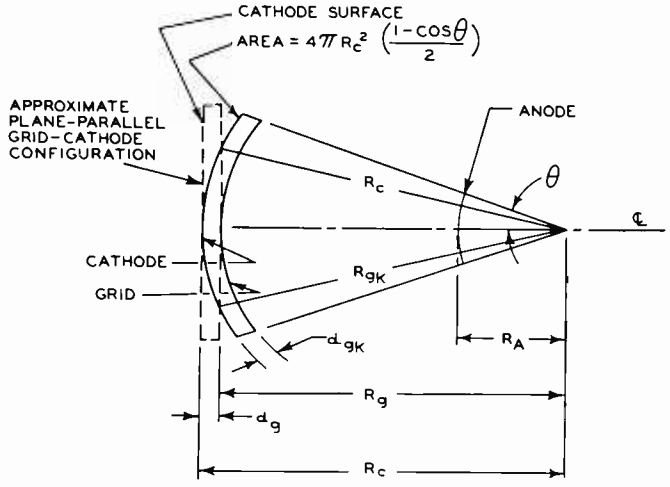


Fig. 5—Approximate plane-parallel grid–cathode configuration.

For a plane-parallel area equal to the area of a spherical surface the current flow, I_{11}, is given by

$$I_{11} = \frac{2.335\ (10^{-6})\ V_g^{3/2}\ 4\pi}{\dfrac{d_{gk}^2}{R_c^2}} \left(\frac{1-\cos\theta}{2}\right) \tag{6}$$

where the quantity $(1-\cos\theta)/2$ represents the half-angle of the conical section when only a portion of the spherical surface is involved.

In addition, for the same conical half-angle, the space-charge-limited current flow, I_0, between spherical cathode and symmetrical anode of any radial spacing is given by

$$I_0 = \frac{29.34\ (10^{-6})\ V_a^{3/2}}{(-\alpha_a)^2} \left(\frac{1-\cos\theta}{2}\right) \tag{7}$$

Placement of the control grid at an equipotential line equivalent to the desired potential of operation V_g does not appreciably alter the space-charge current. Moreover, it can be assumed that the space-charge-limited current flow for the spherical anode–cathode diode is identical to the current drawn by the approximate plane-parallel grid–cathode diode:

$$I_0 = I_{11}. \tag{8}$$

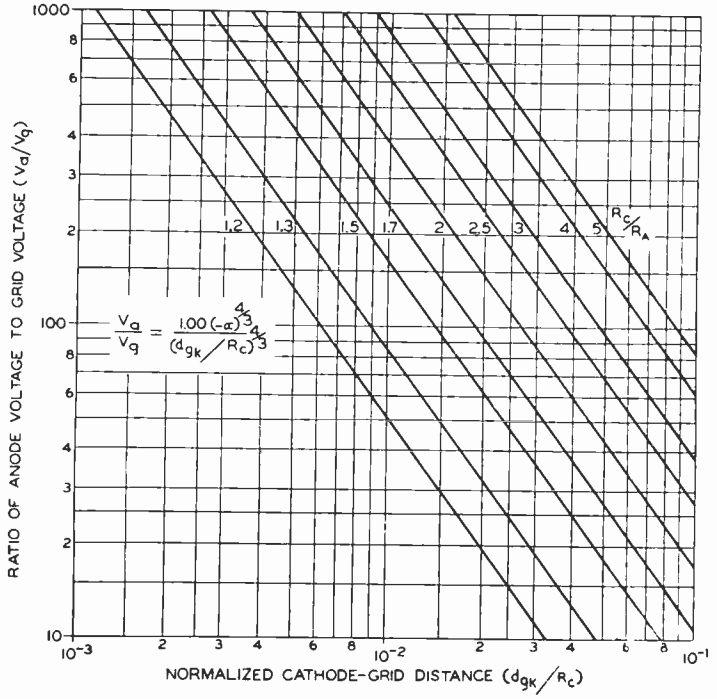


Fig. 6—Normalized approximate cathode–grid distance in terms of R_c/R_a and V_a/V_g.

If the values of space-charge-limited current given in Equations (6) and (7) are equated, the normalized grid–cathode spacing is obtained as follows:

$$\frac{V_a}{V_g} = \frac{(-\alpha_a)^{4/3}}{\left(\frac{d_{gk}}{R_c}\right)^{4/3}}. \tag{9}$$

This equation, which is plotted in Figure 6, is extremely useful for

the determination of grid placement in a convergent-flow Pierce-type gun. Table II lists values for grid placement obtained by exact solution using Langmuir's spherical diode and by the approximate solution given above. These values indicate that the approximate method is reasonably accurate for grid potentials as high as 10 per cent of the anode voltage.

Table II — Comparison of Exact and Approximate Methods for Determination of Grid Placement.

$\frac{V_a}{V_g}$	K (μperv)	J (amp/cm^2)	R_c (inches)	θ	R_c/R_a	Exact Method d_{gk} (inches)	Approximate Method d_{gk} (inches)
30	1.2	2	3.4	15°	1.75	0.18	0.177
60	"	"	"	"	"	0.09	0.105
45	"	"	"	"	"	0.12	0.129
30	1.2	1	3.6	20°	2	0.24	0.241
60	"	"	"	"	"	0.14	0.144
45	"	"	"	"	"	0.18	0.180
70	"	"	"	"	"	0.12	0.126
100	"	"	"	"	"	0.09	0.097
20	0.8	1	3.6	20°	2.2	0.365	0.36

Determination of Cutoff Amplification Factor

After grid–cathode separation is determined, other physical grid parameters must be calculated for the required pulsing performance. Again, approximations are made based on the analogy that exists between plane-parallel and closely spaced spherical surfaces.

If anode-aperture divergent effects are neglected, the electric flux field perpendicular to the spherical cathode, grid, and anode electrodes is comparable to the conditions that exist in a plane-parallel triode. This analogy is particularly valid for closely spaced spherical grid and cathode electrodes acted on by a perpendicular incident, accelerating field introduced by a remote anode. The cutoff amplification factor μ_{co} discussed in this paper, therefore, is based on the use of

such plane-parallel surfaces. The generalized formula for μ_{co}, which is useful for screening fractions of less than 16 per cent, is[7]

$$\mu_{co} = \frac{2\pi d_{yp} L_g - \ln \cosh \pi S}{\ln \coth \pi S}, \tag{10}$$

where L_g is the length of the grid wire per unit area of the grid plane, S is the screening fraction (i.e., the percentage of the area in the grid-wire plane occupied by the laterals), and d_{yp} is the distance between the grid and the anode.

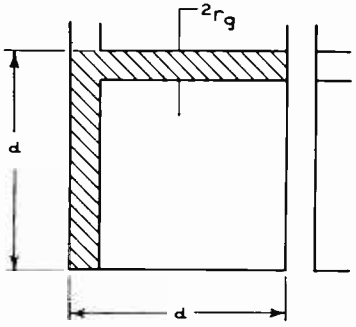


Fig. 7—Square wire mesh showing dimensions used in pulsing-grid approximations.

For a square wire mesh, such as that shown in Figure 7, having a lateral pitch d, the effective length, L_g, of the grid wire per unit area of the grid plane is given by

$$L_g = \frac{d + d - 2r_g}{d^2} = \frac{2d - 2r_g}{d^2}. \tag{11}$$

When the wire size r_g is small compared to the pitch d, L_g can be approximated as follows:

$$L_g \cong \frac{2d}{d^2} = \frac{2}{d}. \tag{12}$$

The screening fraction S is expressed by

[7] K. R. Spangenberg, *Vacuum Tubes*, McGraw-Hill Book Co., Inc., New York, 1948, p. 148.

$$S = \frac{d^2 - (d - 2r_g)^2}{d^2} = \frac{4r_g(d - r_g)}{d^2}. \tag{13}$$

If, again, r_g is very small compared to d, the screening fraction can be approximated by

$$S \cong \frac{4r_g}{d}. \tag{14}$$

The cutoff amplification factor for a square wire mesh, therefore, is given by

$$\mu_{co} = \frac{\frac{4\pi d_{gp}}{d} - \ln \cosh \frac{4\pi r_g}{d}}{\ln \coth \frac{4\pi r_g}{d}}. \tag{15}$$

The cutoff amplification factor as a function of screening fraction, pitch, and grid–anode separation is given in Figure 8. The values of grid–anode distance used are based on the grid-placement approximations made previously. The screening fraction, $4r_g/d$, is a function of the allowable grid-current interception, a percentage of the total beam current intercepted by the projected grid lateral area. The actual grid current experienced in positive grid operation may differ somewhat from the estimated values because the interception-area theory neglects the effects of secondary and primary grid emission and electron-scattering phenomena. In general, however, the use of high ratios of accelerating-anode potential to grid operating potential provide less actual grid-current interception than anticipated due to partial electron scattering and deflection in the grid plane.

It is estimated that Equation (15) is reasonably accurate for grid–cathode spacings in the order of one-third the grid pitch. Larger spacings apparently cause variable grid control and produce remote-cutoff characteristics.

Equation (15) can be extended to accommodate other grid lateral geometries by the use of convenient conversion charts designed by Hsu and Horton.[8] These charts were determined by electrolytic tank measurements and correlate the equivalence of various grid configurations for the same degree of grid control.

[8] H. Hsu and C. E. Horton, "Electrolytic Tank Measurements of Mesh Grid Characteristics, Part 3," *I.R.E. National Convention Record*, p. 114, March 18-21, 1957.

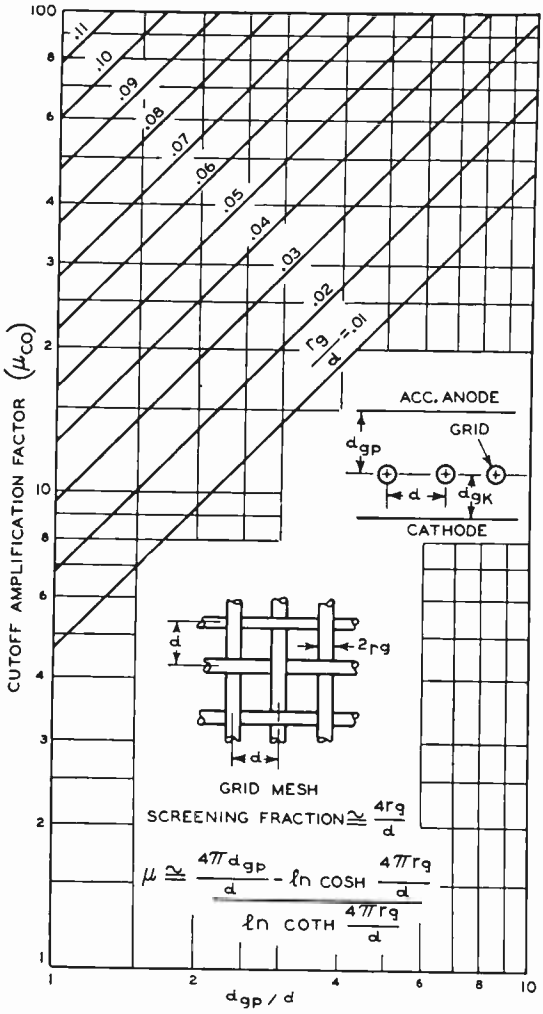


Fig. 8—Physical grid parameters related to screening fraction and cutoff amplification factor.

Pulse Power Requirements

The total power requirements P_p for electron-beam pulsing depend on the power P_c required to charge the total shunt input capacitance at the grid terminals and the actual grid power dissipation P_g.

$$P_p = P_c + P_g. \tag{16}$$

If a constant-current pulsing source is used to charge the shunt input

capacitance C through the total grid-drive potential, the peak charging power P_c is given by

$$P_c = 2.2 \frac{(V_g + V_{co})^2 C}{t} \tag{17}$$

where V_g is the positive grid operating potential, V_{co} is the grid cutoff potential (V_a/μ_{co}), and t is the pulse rise time (between 10- and 90-per-cent points).

If the screening fraction S is considered and electron dispersion is neglected, the grid power dissipation P_g is given in terms of cathode current I_k as follows:

$$P_g = I_k S V_g. \tag{18}$$

Therefore, the total peak pulse power requirements P_p can be approximated by

$$P_p = \frac{2.2\ (V_g + V_{co})^2 C}{t} + I_k S V_g. \tag{19}$$

For very fast rise times and small screening fractions, the grid dissipation term $(I_k S V_g)$ can be neglected in the calculation of the total peak power.

A nomograph of pulse power is shown in Figure 9. In this graph, grid dissipation is neglected and pulse power P_c is given in terms of shunt capacitance C and total grid-drive voltage $(V_d = V_g + V_{co})$.

Care must be used in the selection of the operating parameters of the control grid and the interrelated grid–cathode separation. It is necessary, for example, to insure that allowable grid dissipation (the product of positive grid potential, beam current, screening factor, and maximum duty cycle) is not exceeded for the grid material used. Operation of the grid at elevated temperature can cause excessive grid emission and subsequent loss of grid control. Investigation indicates that titanium metal in the form of grid laterals adequately inhibits grid emission even when exposed to contaminating metallic low-work-function vapors such as barium. Figure 10 compares the performance of titanium with that of tungsten, zirconium, and gold plated molybdenum.[9]

---

[9] G. A. Esperson and J. W. Rogers, "Studies on Grid Emission," *I.R.E. Trans. on Electron Devices*, Vol. ED-3, p. 100, April, 1956.

The basic grid–cathode capacitance C_{gk} of the electron gun, exclusive of all stray wiring, supporting structure, and anode–cathode capacitances, can be approximated from the active cathode emitting area. In the extreme, for spherical electrodes, the grid forms a solid

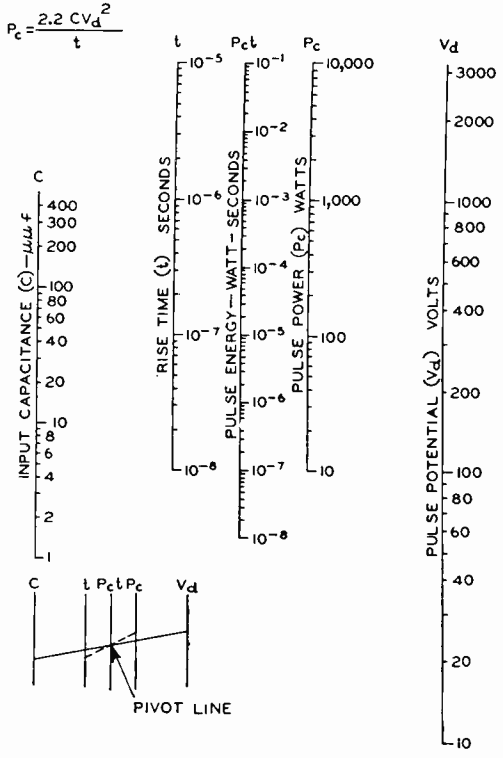


Fig. 9—Nomograph of pulse power related to shunt capacitance, pulsing voltage, and rise time. Enter values of shunt capacitance (C) and pulsing potential (V_d) and draw a straight line connecting these points. The intersection of this line with the index line for pulse energy (P_ct) is the pivot point for the line connecting rise time (t) with the pulse power (P_c).

sphere enclosed by the cathode (the effects of the screening fraction are ignored here). The capacitance C_{gk} of this spherical diode is given by

$$C_{gk} = 4\pi\epsilon \frac{R_c R_g}{R_c - R_g} \text{ (MKS System)}. \tag{20}$$

When R_g approaches R_c, as in the case of a closely spaced grid, this equation becomes

$$C_{gk} = \frac{4\pi\epsilon R_c^2}{R_c - R_g} = \frac{4\pi(8.85)(10^{-12})\,R_c^2}{R_c - R_g}. \tag{21}$$

If a spherical sector of the grid–cathode configuration for the Pierce-

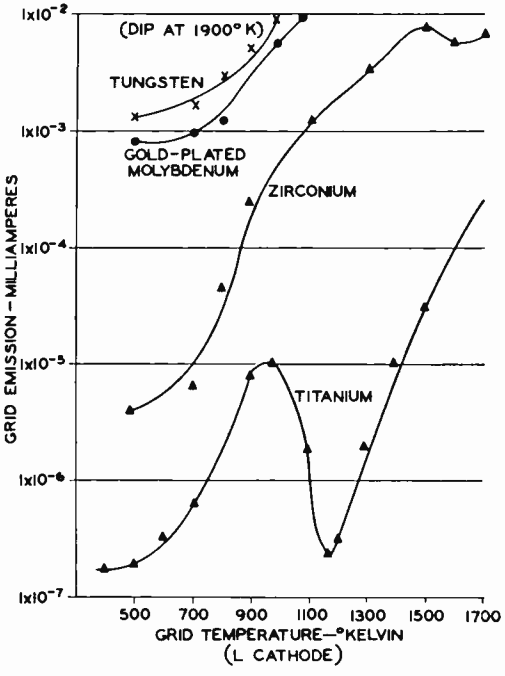


Fig. 10—Grid emission versus temperature for various grid materials.[9]

type gun is considered, and the grid–cathode distance is substituted in Equation (21), the formula for C_{gk} becomes

$$C_{gk} = \frac{4\pi\,(8.85)\,(10^{-12})\,(R_c^2)}{d_{gk}}\left(\frac{1-\cos\theta}{2}\right). \tag{22}$$

This equation, which gives a fair approximation of the grid–cathode capacitance for the active area, can be solved by means of the placement parameters given previously. However, the total grid input capacitance is considerably greater than that predicted by Equation

(22) because of the anode, internal grid-supporting-structure, stray-lead, and basing capacitances. These side capacitance effects are much greater than the small grid–cathode capacitance. A fair approximation of the total input capacitance can be obtained if the actual supporting structure is mechanically approximated and measured and this value is then added to the grid–cathode capacitance derived from Equation (22):

$$C_{in} = C_{gk} + C_{\text{stray}}. \tag{23}$$

Design Examples for Grid-Type Pierce Guns

Use of this method for determining the various grid parameters for convergent-flow electron guns is best illustrated by examples. The first example below presents an academic problem in which the diode-design charts presented by Muller[10] are used. These charts for basic high-perveance guns are shown in Figure 11.

The beam current, minimum beam diameter, minimum beam position, and perveance required for the application are given below. (The notation used is similar to that used in Figure 11):

Beam current $(I_b) = 5$ amperes,
Beam perveance $(K) = 1$ microperv,
$V_a = (I_b/K)^{2/3} = 29{,}300$ volts,
Minimum beam diameter $(D_0) = 0.546$ centimeter,
Minimum beam position $(Z_0) = 2.8$ centimeters.

For tolerable cathode loading (J) of 2 amperes per square centimeter for pulse service, the cathode diameter (D_k) can be approximated as follows:

$$J = 2 \text{ amperes/cm}^2 = \frac{I_b}{\frac{\pi D_k^2}{4}}$$

$$D_k = 1.78 \text{ centimeters.} \tag{24}$$

The value for beam compression $(D_0/D_k = 0.306)$ is then inserted on Muller's chart, and the following ratios are extracted for a microperveance of unity $(K = 1$ microperv):

---

[10] M. Muller, "New Points of View in the Design of Electron Guns for Cylindrical Beams of High Space Charge," *Jour. Brit. I.R.E.*, Vol. 56, p. 83, February, 1956.

$$Z_0/R_c = 1,$$

$$R_a/R_c = 0.49,$$

$$(1 - \cos\theta)/K = 0.052, \text{ where } K \text{ is in micropervs.}$$

Solution of these ratios provides the cathode radius of curvature ($R_c = 2.8$ centimeters), the anode radius of curvature ($R_a = 1.37$ centimeters), and the half-angle ($\theta = 18.50°$).

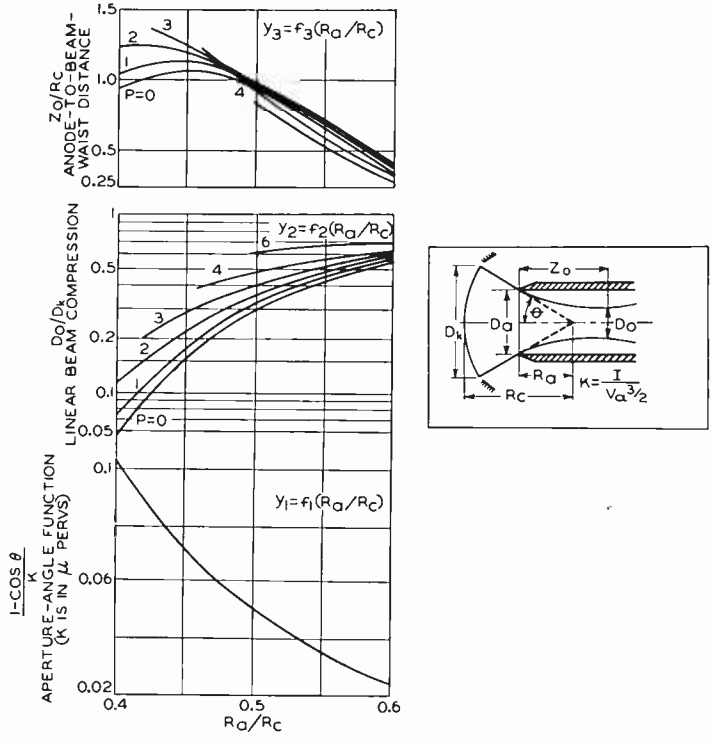


Fig. 11—Design chart for "Muller"-type electron gun.[10]

Grid Parameters

After the basic gun design is determined, it is assumed further that a pulsing grid having the following characteristics is desired:

Cutoff amplification factor (μ_{co}) = 50

Ratio of anode potential to positive grid operating potential (V_a/V_g) = 50:1

Tolerable screening fraction (S) = 10 per cent

The grid–cathode clearance is determined from the grid-placement chart shown in Figure 6 for the values of $V_a/V_g = 50/1$ and $R_c/R_a = 1/0.49$ derived above. The design chart yields the normalized grid–cathode distance ($d_{gk}/R_c = 0.047$); the grid–cathode separation is then found to be 0.132 centimeter.

Other grid parameters such as grid pitch and wire size are determined from the chart shown in Figure 8. For a screening fraction of 10 per cent and a cutoff amplification factor of 50, the following values are extracted:

$$S = 4r_g/d,$$

$$r_g/d = 0.025,$$

$$d_{gp}/d = 6.2,$$

where the grid–plate distance d_{gp} is determined by simple geometry as follows:

$$d_{gp} = (R_c - R_a) - d_{gk},$$

$$d_{gp} = 1.298 \text{ centimeters.} \tag{25}$$

The grid pitch is determined from the ratio of the grid–anode distance to the grid pitch (obtained from Figure 8);

$$d = 6.2\ d_{gp} = 0.209 \text{ centimeter.}$$

The wire size r_g is then calculated from the assumed original screening fraction;

$$\frac{4r_g}{d} = 10 \text{ per cent}; \quad r_g = 0.0105 \text{ centimeter.}$$

The hypothetical gun design derived in this first example is illustrated in Figure 12.

The grid input capacitance for this gun can be estimated by use of Equations (22) and (23). If it is assumed that the stray capacitance effects of the mechanical supporting structure and leads total 15 micromicrofarads, the actual input capacitance is given by

$$C_{in} = C_{gk} + C_{\text{stray}} = 16.8 \text{ micromicrofarads.}$$

The peak input drive power required to charge this capacitance is obtained from the nomograph shown in Figure 9. For a total drive

potential ($V_D = V_{co} + V_g$) equal to 1172 volts and a rise time of 0.03 microsecond, the peak charging power P_c is found to be 1700 watts. The peak grid dissipation P_g can be estimated by means of Equation (18) for a cathode current of 5 amperes, a screening fraction of 10 per cent, and a positive grid drive potential of 586 volts;

$$P_g = I_K S V_g = 293 \text{ watts.}$$

The total peak power P_p is obtained by use of Equation (19):

$$P_p = P_c + P_g = 1993 \text{ watts.}$$

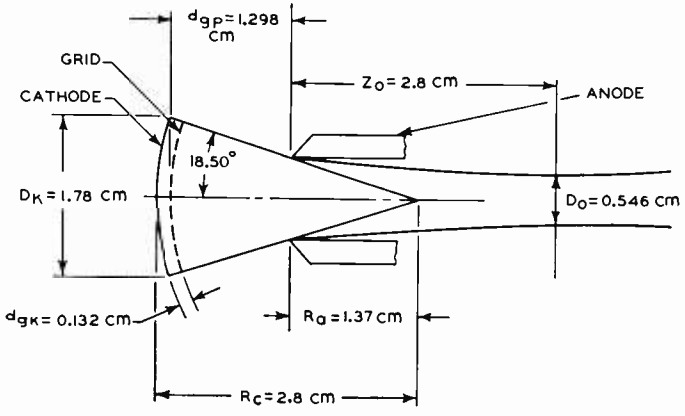


Fig. 12—Description of hypothetical grid-type electron gun.

Practical Examples

The following examples describe the design of control grids for use in modified Pierce-type convergent-flow electron guns which are not of the "Muller design" described in the previous example. The electron guns used in these examples are designed expressly for use in traveling-wave tubes. In each example, both the design objectives and the actual dynamic performance obtained are given.

Example 1

In the design of an S-band traveling-wave tube, the following operating conditions were required:

$$V_a = 4000 \text{ volts}$$

$$V_g = 50 \text{ volts}$$

$$\mu_{co} = 15$$

$$S = 8 \text{ per cent}$$

$$I_k = 250 \text{ milliamperes}$$

$$\text{Duty Cycle} = 10 \text{ per cent}$$

The original convergent-flow gun designed for this tube, without a control grid, had the following electrode geometry:

$$R_c = 0.688 \text{ inch},$$

$$R_a = 0.321 \text{ inch},$$

$$\theta = 22°.$$

This basic gun geometry was preserved, and the grid design parameters were obtained as follows:

(a) The grid–cathode separation was obtained from Figure 6 by use of the ratios

$$R_c/R_a = 2.08 \text{ and } V_a/V_g = 80:1$$

$$d_{gk}/R_c = 3.4 \ (10^{-2}) \ ; \ d_{gk} = 0.024 \text{ inch}.$$

(This distance is "hot" grid–cathode separation. Allowances were made for the anticipated cathode expansion under operating temperature conditions to obtain actual grid–cathode mounting clearance).

(b) The grid–anode clearance was obtained by means of Equation (25):

$$d_{gp} = (R_c - R_a) - d_{gk} = 0.322 \text{ inch}.$$

(c) The grid pitch was found from Figure 8 for the values of screening fraction and cutoff amplification factor given above:

$$d_{gp}/d = 2.25; \quad d = 0.143 \text{ inch}.$$

(d) The grid–wire size was calculated from the screening fraction;

$$S = 4r_g/d = 8 \text{ per cent}; \quad r_g = 2.86 \text{ mils}.$$

Based on this figure, the grid-wire diameter would be 5.7 mils. However, it was desirable to use 5-mil wire which was available (new $r_g = 2.5$ mils) and yet to maintain the same screening fraction. Therefore, the pitch d was changed as follows:

$$S = 4r_g/d = 0.08; \quad d = 0.125 \text{ inch.}$$

(e) The expected peak grid current I_g is given by

$$I_g = SI_k = 0.08(250) = 20 \text{ milliamperes.}$$

(f) The peak grid dissipation P_g is given by

$$P_g = V_g I_g = 50(20 \times 10^{-3}) = 1 \text{ watt.}$$

(g) The input shunt capacitance is approximated by means of Equation (23), including a stray measured capacitance of 7.64 micromicrofarads and a grid–cathode capacitance of 1.91 micromicrofarads calculated from Equation (22).

$$C_{in} = 9.55 \text{ micromicrofarads.}$$

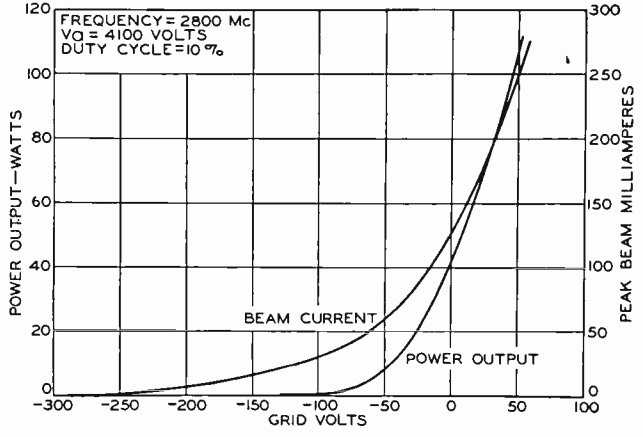


Fig. 13—Grid-control characteristics for 100-watt pulsed S-band tube.

(h) The peak input pulsing power required to charge the shunt capacitance is found from Figure 9 for a total grid drive potential $(V_{co} + V_g)$ equal to 316 volts and a rise time of 0.03 microsecond.

$$P_c = 70 \text{ watts.}$$

The grid control characteristics of the final S-band tube are shown in Figure 13. A cutoff potential of approximately —280 volts is achieved for an accelerating potential of 4200 volts. Table III lists both the design objectives and the actual performance of the grid-type electron gun.

Table III — A Comparison of Design Objectives and Measured Gun Parameters for a 100-Watt Pulsed S-Band Traveling-Wave Tube.

Electrical Parameter	*Design Values*	*Measured Values*
V_a (volts)	4000	4200
V_g (volts)	+50	+50
V_{co} (volts)	—266	—280
μ_{co}	15	15
I_k (ma peak)	250	250
I_g (ma peak)	20	16
P_g (watts)	1.0	0.8
C_{in} ($\mu\mu$f)	9.5	13
P_c (watts)	70	90

Example 2

For a one-kilowatt pulsed S-band traveling-wave tube, the effective geometry and electrical parameters of the basic diode convergent-flow gun were given as follows:

$$V_a = 7750 \text{ volts}$$

$$I_a = 1.2 \text{ amperes}$$

$$R_a = 0.460 \text{ inch}$$

$$R_c = 0.875 \text{ inch}$$

Table IV — A Comparison of Design Objectives and Measured Gun Parameters for a 1-Kilowatt, Pulsed, S-Band Traveling-Wave Tube.

Electrical Parameter	*Design Values*	*Measured Values*	*Interpolated Values**
V_a (volts)	7750	8000	8000
V_g (volts)	+200	100	175
V_{co} (volts)	—150	—160	—160
μ_{co}	51.7	50	50
I_{beam} (amp)	1.2	0.8	1.2
I_g (ma)	120	100	150
P_g (watts)	24	9.5	22.5
C_{in} ($\mu\mu$f)	37	38	38

* Values for grid-drive potentials above 225 volts ($V_g = +95$ volt, $V_{co} = -160$ volt) were interpolated from Figure 15.

Based on these diode parameters, the desirable operating characteristics of the grid-type pulsing electrode for this tube are

$$S = 10 \text{ per cent}$$

$$\mu_{co} = 51.7$$

$$\text{Duty Cycle} = 1 \text{ per cent}$$

$$V_g = +200 \text{ volts}$$

$$V_{co} = -150 \text{ volts}$$

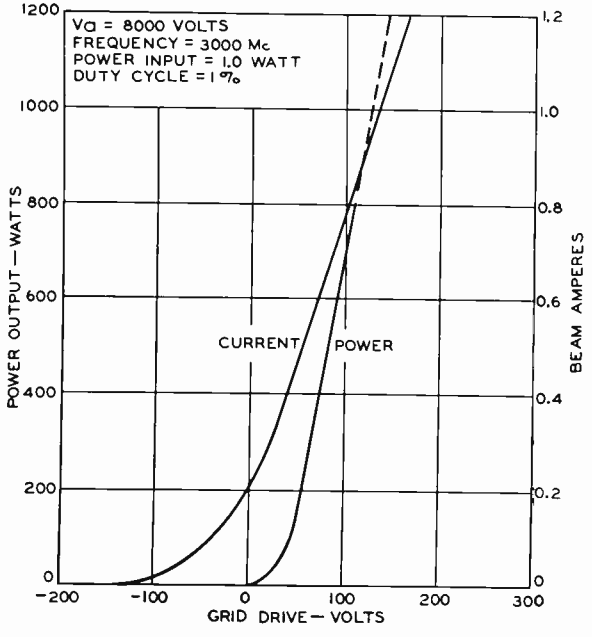


Fig. 14—Grid-control characteristics for 1-kilowatt pulsed S-band tube.

The grid–cathode and grid–anode separation are determined from Figure 6 for the given values of R_c/R_a and V_a/V_g:

$$d_{gk} = 0.046 \text{ inch},$$

$$d_{gp} = (R_c - R_a) - d_{gk} = 0.370 \text{ inch}.$$

The grid pitch is obtained from Figure 8 for the given values of μ_{co} and screening fraction:

$$d_{gp}/d = 6.2, \quad d = 0.060 \text{ inch}.$$

The wire size is obtained from the screening fraction:

$$2r_g = 0.003 \text{ inch.}$$

Results obtained with an electron gun having a pulsing grid with the values determined from the design charts are shown in Figure 14 and Table IV. Actual values of grid drive were limited to +90 volts during testing because of beam-focusing difficulties. However, interpolation of the curves for rated power output and cathode current indicates that fair agreement between design and actual operating parameters has been obtained. The cutoff potential is approximately —160 volts at an anode potential of 7750 volts.

THE MAGNETIC FIELD AND FLUX DISTRIBUTIONS IN A PERIODIC FOCUSING STACK FOR TRAVELING-WAVE TUBES*

By

MAX J. SCHINDLER

RCA Electron Tube Division,
Harrison, N. J.

Summary—*The factors that affect the magnetic field and flux distributions in periodic-focusing stacks for traveling-wave tubes are analyzed. Shortcomings of the present design method are described, and an approach is outlined that is more in harmony with measured results.*

At present, the design of periodic magnetic focusing stacks is based on two solutions of Laplace's field equation for coaxial, oppositely charged cylinders. The first solution, in which each cylinder represents a shim hub, is used for computation of the field along the axis and the flux within the hubs. The second solution, in which the cylinders represent the circumference of the shim discs, is used for determination of the stray field surrounding the stack.

This technique, however, does not consider the contributions to the field made by the disc and the outer hub surface. Furthermore, the influence of neighboring magnets is neglected in determining the operating point of a magnet. As a result, the computed useful center field agrees reasonably well with measured values only in the central part of the stack, where these influences cancel to some degree. This approach, however, cannot predict the demagnetizing fields that exist within the magnet, nor can it account for the irregularity of the axial field at the ends of the stack.

Values of demagnetizing field more than twice those computed conventionally have been measured in the space occupied by the magnets. Computations that take into account the contributions to the demagnetizing field of factors now neglected (i.e., the charges on the disc and the hub surface) yield results that are in very good agreement with measurements made upon actual fields in stack models.

Contributions by neighboring magnets affect both the demagnetizing field and the axial field. The contributions to the demagnetizing field result in a shift of the operating point of the magnet. It has been shown both by theory and experiment that, contrary to present practice, the "regular" demagnetization curve cannot be used in this case; only the "intrinsic" magnetization curve is physically meaningful, and it is therefore used throughout this paper.

The contribution of the alien fields of neighboring magnets to the axial field has been determined experimentally for typical shim geometries. These alien fields may change the original values by as much as 50 per cent. Although the alien fields are equal for all locations in the central part of the stack, they are partially lacking at the ends. This lack accounts for the irregular field distribution on the axis at the end regions of the stack.

Measurements verify the validity of the approach presented in this paper.

* Manuscript received April 22, 1960.

Introduction

SINCE THE DEVELOPMENT of the basic theory of periodic focusing,[1] a number of papers have dealt with design calculations for suitable magnet assemblies.[2-4] All of these papers used the approach to magnetic theory described by Chang.[2] This approach is based upon the solution of Laplace's field equation for coaxial cylinders carrying alternately positive and negative magnetic charges, where each cylinder represents a shim hub. The field on the axis and the flux within the hubs are thus determined. The stray field surrounding the stack is accounted for by a similar computation of the flux outside short cylindrical rings placed at the circumference of the shim discs. The field within the magnet volume is assumed to be uniform, and the influence of neighboring magnets is neglected. In the central portion of the stack, these two influences cancel to some degree, and the computed results agree reasonably well with the measured values in the usable center field.

The field within the magnet volume is far from uniform, however, and field strengths more than twice those calculated with the above method have been measured. The method generally used in permanent magnet design uses the concept of lines of induction and permeance. This method cannot easily be applied to the condition of three-dimensionally varying demagnetizing fields. In the presence of external fields, as is the case in a focusing stack, it is even incorrect. This difficulty can be overcome by use of the "intrinsic induction" or "intensity of magnetization."

Analysis of the magnetic condition in a periodic focusing stack can be simplified by separation of the magnets from the shim structure. This approach is acceptable because of the great difference in permeability (practically unity for the magnets and at least 100 for the shims) that makes the shims act as equipotential surfaces. The shims determine the field configuration, and the magnet merely acts as a source of magnetization as long as the charged surfaces of the magnet are completely covered by the pole pieces. The situation is similar to

---

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

[2] K. K. N. Chang, "Optimum Design of Periodic Magnet Structures for Electron Beam Focusing," *RCA Review*, Vol. 16, p. 65, March, 1955.

[3] F. Sterzer and W. W. Siekanowicz, "The Design of Periodic Permanent Magnets for Focusing of Electron Beams," *RCA Review*, Vol. 18, p. 39, March, 1957.

[4] J. E. Sterrett and H. Heffner, "The Design of Periodic Magnetic Focusing Structures," *IRE Trans. on Electron Devices*, p. 35, January, 1958.

that of an electric network in which the current distribution is not influenced by the kind of current source. In such a network, the current distribution can be determined without the presence of a source.

In a focusing stack, a number of sources are incorporated in the circuit at different locations. Within the range of high permeability of the pole pieces, however, the total fields at each location can be found by superimposition of all of the partial fields at that location. Each of the sources of magnetization contributes to the total field at any location.

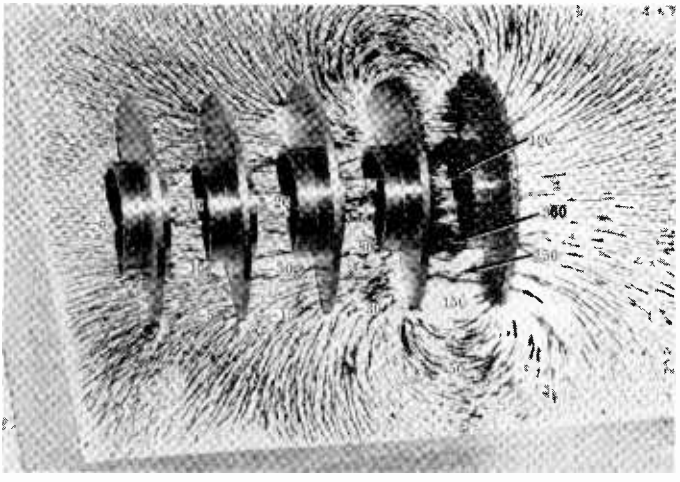


Fig. 1—Iron-filling pattern of a periodic magnet stack model.

The field configuration caused by a magnet at the end of a stack is shown in Figure 1. The numbers indicate the approximate values of the field in oersteds, measured in the plane of the paper. It can be shown experimentally that the field within the magnetized cell (first from right) does not change appreciably if the remaining shims are removed. Therefore, it is possible to consider the condition prevailing within a cell separately from the influence imposed upon neighboring cells.

The Magnetic "Unit Cell"

A pair of shims having opposite magnetic charges may be called a magnetic cell. Because geometrically similar cells exhibit the same magnetic conditions, the dimensions may be divided over the period L to derive a unit cell having dimensionless parameters. The param-

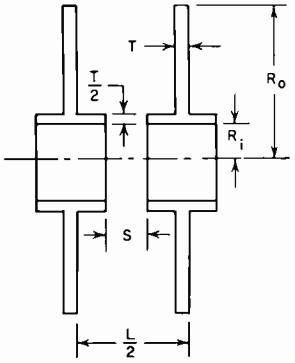


Fig. 2—Dimensions of a magnetic cell.

eters as defined in Figure 2 are

$$\frac{R_i}{L} = r_i, \quad \frac{R_0}{L} = r_0, \quad \frac{S}{L} = s, \quad \frac{T}{L} = t. \tag{1}$$

For ease of computation and understanding, the magnetization "M" is chosen to produce a field of unity if the shim is infinitely thin ($T = 0$) and stray fields are neglected. The fields in the unit cell are denoted by h_i and the fluxes by ϕ_i, as shown in Figure 3. Capital letters are used for actual values (H_i, Φ_i).

If a uniform field is assumed along any line of generation of a

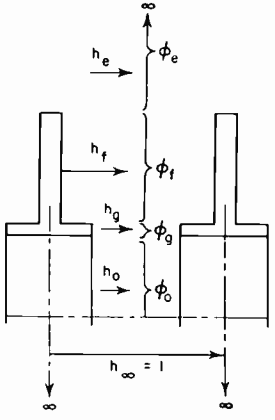


Fig. 3—Fields and fluxes in the unit cell.

cylinder which is coaxial with and limited by the shim, the following equation can be derived from $M = \int Hdl$:

$$M = \frac{L}{2} = h_f \left(\frac{L}{2} - T \right) = h_g S. \tag{2}$$

The solution of Equation (2) for the "flange field" is

$$h_f = \frac{1}{1 - 2t}, \tag{2a}$$

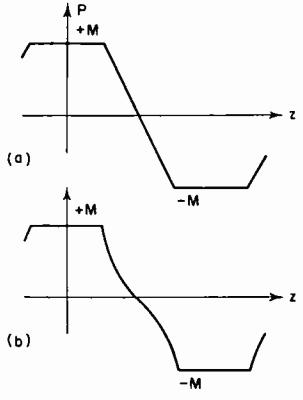


Fig. 4—(a) Idealized potential distribution at $r = r_i$; and (b) true potential distribution at $r = r_i$.

and for the "gap field,"

$$h_g = \frac{1}{s}. \tag{2b}$$

The assumption of a uniform field along a line of generation is known to be incorrect because the magnetic potential does not decrease linearly, as indicated in Figure 4(a), but follows the pattern shown in Figure 4(b). However, it has been shown in the literature[5] that this simplification still gives fairly accurate results, at least for the configuration inside the shim hub. This region is the most important part of the field configuration.

---

[5] S. Bertram, "Calculation of Axially Symmetric Fields," *Jour. Appl. Phys.*, Vol. 13, p. 496, 1942.

The Field within the Inside Diameter of the Pole Piece

Because of the shielding effect of the hubs, the disc charge has little influence on the "hub field" and can, therefore, be neglected. The problem then is to determine the field inside a series of coaxial alternatingly charged cylinders. This problem is solved by use of Laplace's field equation.[6]

The general solution for the axial field component h_z is given by*

$$h_z = \sum_{n=1}^{\infty} (1 - \cos n\pi) \frac{\sin (sn\pi)}{sn\pi} \frac{J_0(i2n\pi r)}{J_0(i2n\pi r_i)} \cos \frac{2n\pi z}{L}. \quad (3)$$

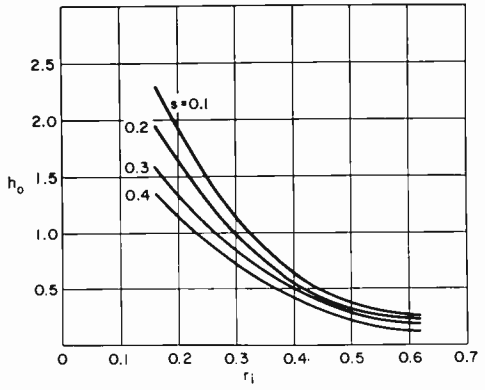


Fig. 5—Maximum axial field.

From this equation, the expression for the field h_r in the gap-center plane can be derived by substitution of $z = 0$;

$$h_r = \sum_{n=1}^{\infty} (1 - \cos n\pi) \frac{\sin (sn\pi)}{sn\pi} \frac{J_0(i2n\pi r)}{J_0(i2n\pi r_i)}, \quad (4)$$

and the expression for the peak field on the axis, shown in Figure 5, can be derived by substitution of $r = 0$;

---

[6] For details, see K. K. N. Chang, "Optimum Design of Periodic Magnet Structures for Focusing Low Voltage Electron Beams," *Jour. Appl. Phys.*, Vol. 24, p. 1247, 1953.

* Bessel functions are designated in agreement with *Tables of Functions with Formulae and Curves*, by E. Jahnke & F. Emde, P. G. Teubner, Leipzig, and G. E. Stechert & Co., New York, 1938.

$$h_0 = \sum_{n=1}^{\infty} (1 - \cos n\pi) \frac{\sin (sn\pi)}{sn\pi} \frac{1}{J_0(i2n\pi r_i)}. \qquad (5)$$

The center gap field has already been computed;

$$h_g = \frac{1}{s}. \qquad (2b)$$

Equation (2b) is very helpful because the series of curves based on Equation (5) converges very slowly for the case of r approaching r_i.

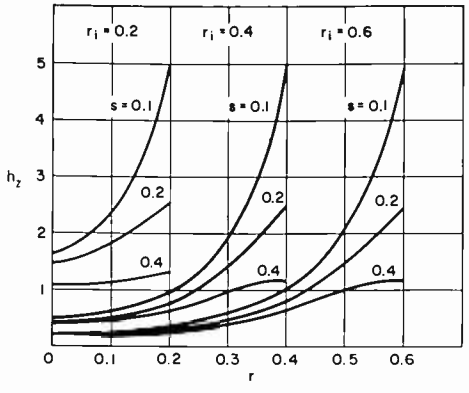


Fig. 6—Field distribution in the gap-center plane.

Figure 6 illustrates the decrease of the axial field from the gap toward the axis for different configurations in a "unit cell." Under actual conditions, the field between the shim flanges is not unity, but is determined in part by the same parameters, r_i and s. Therefore, any curve in Figure 6 shows how the field inside an actual hub decreases under given conditions, but tells little about its relative magnitude compared to any one of the other curves.

If the field configuration inside the hub is known, it is easy to determine the flux in this region. Integration of Equation (4) gives

$$\phi_0 = r_i \sum_{n=1}^{\infty} \frac{1 - \cos n\pi}{n} \frac{\sin (sn\pi)}{sn\pi} \frac{-iJ_1(i2n\pi r_i)}{J_0(i2n\pi r_i)}. \qquad (6)$$

The variation of ϕ_0 with r_i, for two practical values of s, is shown in Figure 7.

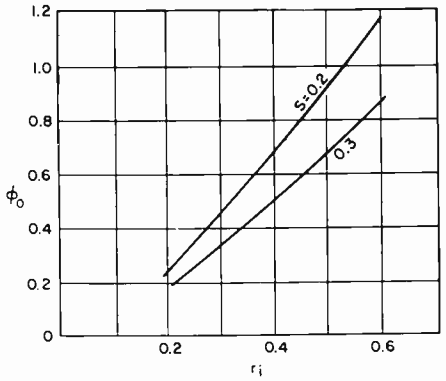


Fig. 7—Unit-cell hub flux.

The Field Outside the Pole Piece Flange

If the circumference of the shim flange is regarded as another set of charged coaxial cylinders, the field and flux in the region beyond the outside diameter may be determined in a manner similar to that used for the field within the hub.

The solution for the field in the gap plane is

$$h_e = \sum_{n=1}^{\infty} (1 - \cos n\pi) \frac{\sin (1-2t)n\pi}{(1-2t)n\pi} \frac{iH_0^{(1)}(i2n\pi r)}{iH_0^{(1)}(i2n\pi r_0)}. \qquad (7)$$

The solution for the external flux is

$$\phi_e = r_0 \sum_{n=1}^{\infty} \frac{1-\cos n\pi}{n} \frac{\sin (1-2t)n\pi}{(1-2t)n\pi} \frac{-H_1^{(1)}(i2n\pi r_0)}{iH_0^{(1)}(i2n\pi r_0)}. \qquad (8)$$

Figure 8 shows the variation of external flux with shim outside diam-

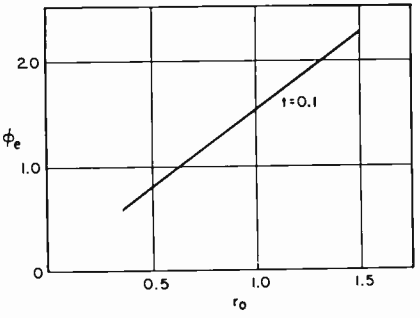


Fig. 8—Unit-cell external flux.

eter for a t value of 0.1. It should be noted that this stray flux is generally larger than the hub flux.

Simplified Shim Field

After the field configuration is determined within the inside diameter as well as beyond the outside diameter of the shim, all that is needed to complete the field picture is the part between the pole pieces themselves. Because this field is very complex, the simplified version used in current literature will be presented first.[2-4]

In the simplified version, two distinct and uniform fields, h_g and h_f, are assumed, as shown in Figure 9. The magnitudes of these fields

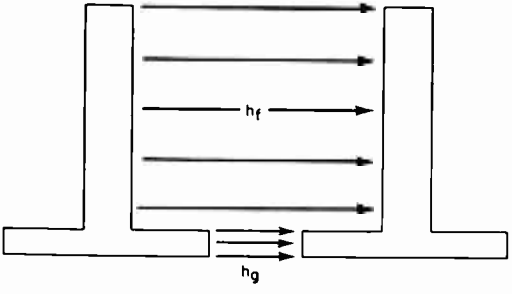


Fig. 9—Simplified gap and flange field.

can be determined by use of the fact that in a magnetic circuit $\int Hdl$ must be the same for the complete length of all field lines.

The fluxes may be computed by multiplying each field strength by its cross-sectional area as follows:

$$\phi_g = \frac{1}{s} 2\pi r_i \frac{t}{2} = \frac{\pi r_i t}{s}. \tag{9}$$

The approximation used for the hub cross section is justified by the fact that it is a thin-walled cylinder.

The expression for the flange flux is given by

$$\phi_f = \frac{1}{1-2t} \pi \left[r_0^2 - \left(r_i + \frac{t}{2} \right)^2 \right], \tag{10}$$

or, if the wall thickness of the hub is neglected, by

$$\phi_f = \pi \frac{r_0^2 - r_i^2}{1-2t}. \tag{10a}$$

Simplified Computation of the Actual Fields

The magnetic potential (or the "charge") that was assumed to exist on the pole pieces is supplied by a permanent magnet. This magnet must obey its hysteresis loop. A coordinate system in which one axis is the field H and the other is the intrinsic induction J is used in this paper. J is defined by

$$J = 4\pi I = B - H, \tag{11}$$

where B is the "regular" or external induction and I is the intrinsic induction in magnetostatic units, as ordinarily used in theoretical physics. The second quadrant of the hysteresis loop, the demagnetization curve, gives the correlation between the intensity of magnetization, usually called the "intrinsic induction," and the demagnetizing field in the magnet. Because the magnetization of the shim structure must follow the "load line," the operating point of a system is determined by the intersection of the load line with the demagnetization curve.

The $J - H$ system has been chosen mainly because the "load line" is basically identical with the "demagnetization factor" N. This demagnetization factor determines the field H produced by a given magnetization J, and is a function of the geometry. The magnetic status is described by

$$H = H_e - NJ, \tag{12}$$

where H_e is an external (applied) field. If no external field exists,

$$N = \frac{-H}{J}. \tag{12a}$$†

In the case of the simplified unit cell, the field is given by

$$h_f = \frac{1}{1 - 2t}, \tag{2a}$$

and the intrinsic induction is equal to the total flux per magnet area;

† Analysis of Equation (12a) indicates why it is basically wrong to work with the "external" induction $B = J + H$. The demagnetization factor, because it is a function of geometry only, cannot be influenced by an external field. On the other hand, a magnet cannot be expected to distinguish between an external field and its own. If B is used instead of J, however, the demagnetizing field would be $H_d = N^*(J + H)$ and would, therefore, be a function of itself: $H_d = N^*(J + H_e + H_d)$.

$$J = \frac{\phi_\Sigma}{a_m}, \tag{13}$$

where a_m is the normalized magnet area, and the total flux is given by

$$\phi_\Sigma = \phi_0 + \phi_a + \phi_f + \phi_e. \tag{14}$$

The demagnetization factor (or load line) is then

$$N = \frac{a_m}{(1-2t)\,\phi_\Sigma}. \tag{15}$$

The preceding equations can be used to determine the operating point of the magnet as well as the demagnetizing field, the latter holding greater interest.

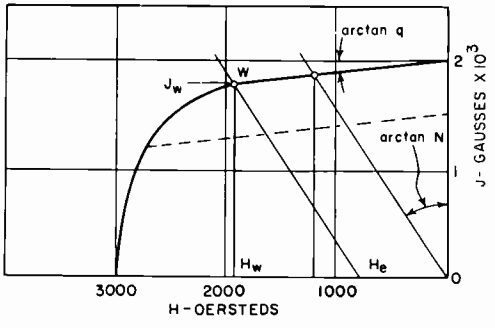


Fig. 10—Operating point of a ceramic magnet.

Figure 10 shows the solution of the problem graphically. An analytical solution is possible if the load line intersects the straight portion of the demagnetization curve. This condition should always be maintained if permanent self-demagnetization of the magnet is to be avoided during assembly. It can be achieved by "stabilizing" the magnet to a "minor loop" as indicated by the broken line.

The external field is caused by neighboring magnets, and is determined by geometry and by the operating induction J_w of these magnets;

$$H_e = pJ_w. \tag{16}$$

Substitution of this relation in Equation (12) gives the equation for the shifted load line;

$$J = \frac{1}{N}(H - pJ_w). \tag{16a}$$

The straight part of the demagnetization curve (shown in Figure 10) is expressed by

$$J = J_R - qH \quad \text{(and specifically } J_w = J_R - qH_w). \tag{16b}$$

The actual operating field H_w is determined by the intersection of these two lines;

$$J_R - qH_w = \frac{1}{N}(H_w - pJ_w) = \frac{1}{N}[H_w - p(J_R - qH_w)]. \tag{16c}$$

The external field, therefore, is given by

$$H_w = J_R \frac{1 + \frac{p}{N}}{\frac{1}{N} + q\left(1 + \frac{p}{N}\right)}, \tag{17}$$

where J_R is the remanence (equal to B_r) and N is given by Equation (15). Because minor loops are almost horizontal for ceramic magnets, it is permissible to set q equal to 0, so that

$$H_w = J_R N\left(1 + \frac{p}{N}\right). \tag{17a}$$

In an isolated cell, no external field is present ($p = 0$). Substitution of $p = 0$ gives the demagnetizing (working) field of an isolated cell;

$$H_{w1} = J_R N. \tag{17b}$$

H_w may then be expressed in terms of the working field of an isolated cell and of a factor P_w that will account for the field of neighboring magnets:

$$H_w = J_R N\, P_w = H_{w1} P_w, \tag{18}$$

where

$$P_w = 1 + \frac{\Sigma p}{N}. \tag{18a}$$

Substitution of Equation (15) for N gives

$$H_w = J_R \frac{a_m}{(1-2t)\,\phi_\Sigma} P_w. \tag{19}$$

The peak axial field in a unit cell h_0 (shown in Figure 5) and the unit-cell flange field h_f have already been computed. The ratio of these two fields must remain the same regardless of the magnitude of the fields. In the actual cell the flange field was found to be H_{w1} instead of h_f, the value for the unit cell. The actual axial field can, therefore, be computed by use of the relation

$$\frac{h_0}{h_f} = \frac{H_0}{H_{w1}}.$$

The actual field on the axis is given by

$$H_0 = H_{w1}\frac{h_0}{h_f} = h_0 J_R \frac{a_m}{\phi_\Sigma}. \tag{20}$$

A factor must again be used to account for the influence of neighboring magnets. Because of the different stray conditions, this factor (P_0) for the axis is not necessarily the same as the one for the flange area (P_w). The final result for the axial peak field is

$$H_p = J_R h_0 \frac{a_m}{\phi_\Sigma} P_0. \tag{21}$$

The formula is valid if the intrinsic demagnetization line is practically horizontal, or if the "external field" is small compared to the internal demagnetizing field, or both. (Because the field is zero on both the J and the B axes, the intrinsic remanence and the regular remanence are identical.)

The value of the factor P_0 in the center of a stack is usually between 1.0 and 1.4, with 1.2 being a reasonable average. Values higher than 1.4 and lower than 1.0 are possible.

The use of the experimental factor P_0 permits the simplified calculation for the peak field to be used despite the inaccuracy involved. All the necessary corrections are put into the factor P_0 until the calculated and the measured values agree best.

The True Demagnetizing Field

As pointed out in the introduction, the field in the magnet region is far from uniform. This nonuniformity is evident in the iron-filing pattern shown in Figure 1. Cross sections of this pattern for the first, second, and third gap plane (H_1, H_2, and H_3), are shown in Figure 11.

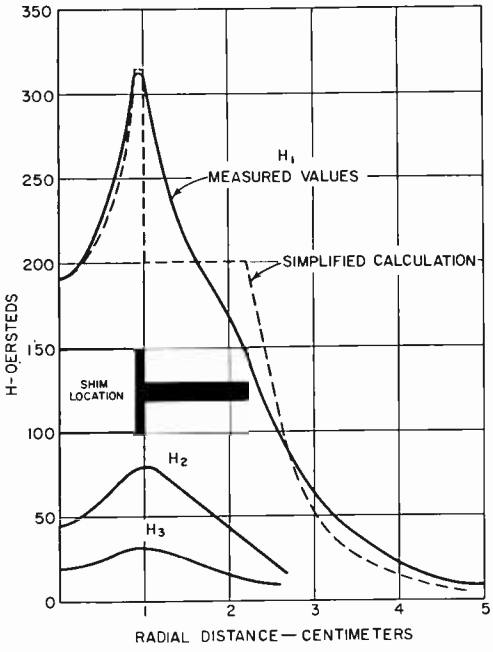


Fig. 11—Field distribution in the stack model of Figure 1.

The broken line indicates the field obtained from Equation (4) and simplified Equations (2a) and (2b). The demagnetizing field is actually much higher than predicted in the region adjoining the hub. This situation is easily understandable, because the hub obviously must cause a "stray field" similar to the one calculated for the shim circumference. Furthermore, the actual field starts to fall off before the edge of the shim is reached. This condition results from the fact that the field lines are no longer straight in the edge region, where an appreciable stray field exists. (The situation is quite similar to that occurring in a plate capacitor.) Finally, the stray field decreases more slowly than would be expected from the calculation. The iron-filing pattern shown in Figure 1 reveals the cause; the stray field is caused not only by the lines originating on the shim circumference, but to a

large extent by a charge on the outer surfaces of the shim flanges. This charge results in an external field at the neighboring cell.

External Hub Field

In the computation of the hub field, it is assumed that the presence of the shim discs does not disturb the charge distribution on the hub surface very much. The results seem to justify this assumption. Equation (7) may then be applied to the configuration shown in Figure 12, neglecting the hub thickness;

$$h_h = \sum_{n=1}^{\infty} (1 - \cos n\pi) \frac{\sin sn\pi}{sn\pi} \frac{iH_0^{(1)}(i2n\pi r)}{iH_0^{(1)}(i2n\pi r_i)}. \tag{22}$$

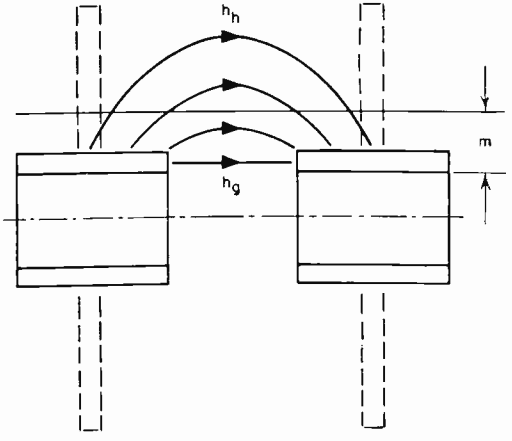


Fig. 12—The external hub field.

The maximum value, at the hub radius is again equal to the gap field;

$$h_g = \frac{1}{s}. \tag{2b}$$

The resulting curves are shown in Figure 13 for several values of r_i and of s. It is noteworthy that for each value of r_i, the hub field decreases in practically the same manner, whereas the gap width s influences the curves appreciably. In Figure 14, the maximum values are reduced to unity and the relative magnitude is plotted over m, the distance from the shim inside diameter (shown in Figure 12). The curves show that for the commonly used gap widths of 0.2 to 0.3, the

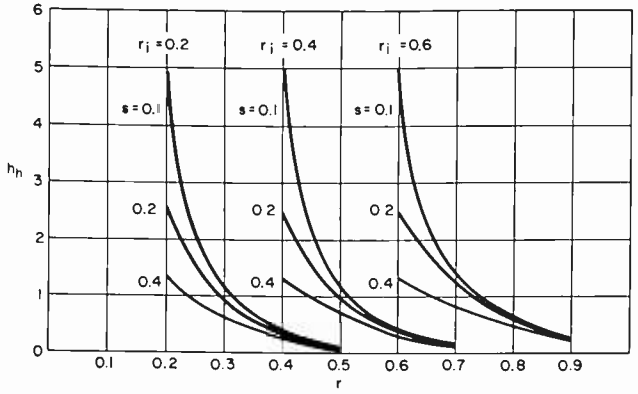


Fig. 13—External hub field versus radial distance.

hub field decreases to half its value at a distance m of 0.1. This knowledge is very useful in practical stack design.

The Disc Field

The flange field was found by the assumption of a homogeneous field between the discs and no lines originating on the disc surfaces (analogous to the ideal disc capacitor). Computation of the field distribution caused by uniformly distributed charges on two parallel discs gives an idea of the deviations to be expected.[7]

This field distribution is shown in Figure 15; r_0' is the ratio of the plate radius to twice the plate distance, equivalent to the period in a focusing stack. If the finite shim thickness is considered, r_0' is defined by

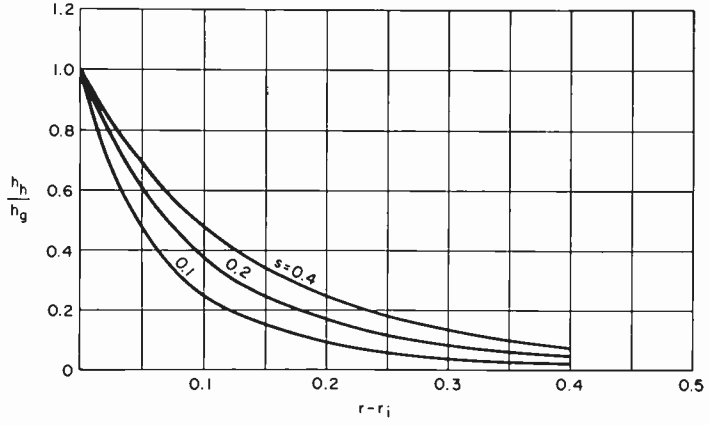


Fig. 14—Relative distribution of the hub field.

[7] M. Schindler, "On the Design of High-Coercivity Permanent Magnets Exposed to External Fields", to be published by the AIEE.

$$r_0' = \frac{R_0}{L - 2T} = \frac{r_0}{1 - 2t}, \tag{23}$$

and usually has a value between 0.7 and 1.2. Figure 15 shows not only that the field falls off in the vicinity of the edges, but also that the maximum field in the center of the disc is a function of relative disc radius, provided the "surface charge" remains constant.

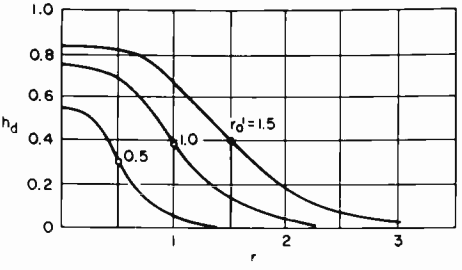


Fig. 15—Field distribution between two disks.

Experimental evidence indicates that the charge distribution is not uniform, as assumed in this computation, but gravitates toward the edges. The true situation, therefore, is somewhat closer to the simplified calculation of the field outside the shim, because the field is more uniform than that shown in Figure 15. The magnitude of the field is not affected significantly, however, and the values of Figure 15 may be used to define a new factor;

$$\kappa = \frac{h_d}{h_\infty}, \tag{24}$$

where h_d is the unit-cell disc field and h_∞ is the "ideal capacitor" field as used in the simplified computation of the actual fields. This factor is identical with the true disc field for $r_0' = \infty$. The value of the factor κ for values of r_0' from 0 to 1.5 is shown in Figure 16.

The Flange Field

In the flange area of the shim, the hub field and the disc field are superimposed. By use of the information given in the preceding paragraphs, this distribution can be calculated and compared to the measured values. As pointed out in the preceding paragraph, the charge concentration at the disc edges improves the uniformity in the flange

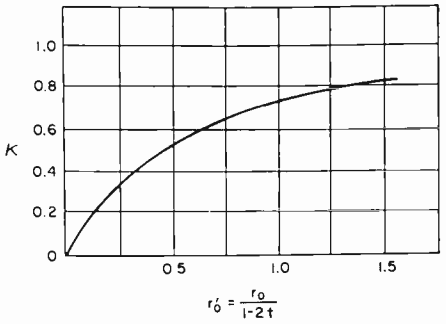


Fig. 16—Center field between two disks (surface charge and distance held constant).

area as compared to the field shown in Figure 15. As a result, the curve may be replaced by a straight line having the proper κ.

It is not necessary to know the absolute value of the hub field. Because the maximum hub field is given by the gap field, h_g, it is only necessary to select the proper curve in Figure 14 and multiply its values by h_g minus h_d. The broken line in Figure 17 shows the result

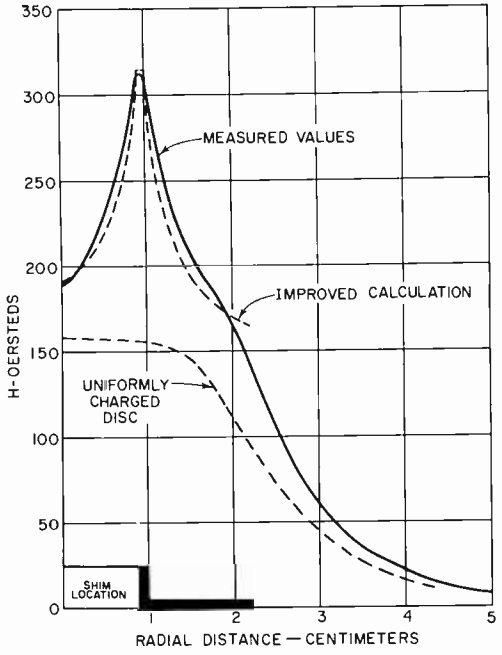


Fig. 17—Comparison between improved calculation and measured field distribution in the stack model of Figure 1.

of this calculation; the measured values are indicated by the solid line. Comparison of this curve with that of Figure 11 indicates the improvement over the simplified calculation. Outside the shim, the field soon approaches the values computed for a homogeneous surface charge (dotted line).

External Demagnetizing Fields

Up to this point, the field in an isolated unit cell has been considered. The iron-filing pattern of Figure 1, however, shows that all the neighboring magnets also contribute to the field in a cell. The contributions of the closest neighbors are highest and, therefore, are used to illustrate the problem.

In a periodic focusing stack, the magnetization of the cells alternates along the length of the axis. The magnet located to the left of the cell under consideration, therefore, introduces a field such as that

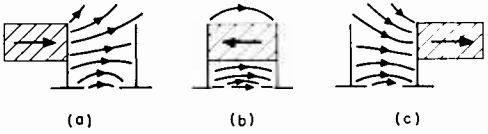


Fig. 18—Contribution of neighboring cells to the local magnetic field.

shown in Figure 18(a). The contribution of the local magnet is shown in Figure 18(b), and that of the right-hand neighbor in Figure 18(c). Because all the fields are in the same direction, the local magnet is exposed to a much higher demagnetizing field than that computed for the isolated cell. It must be expected that the highest external field will be close to the gap, where the contributions of both neighbors are the same. In the flange region, the external field is highest at the edges, and decreases sharply with the distance from the surface on which the neighboring cell's stray lines originate. For this reason, it is necessary to consider the contribution of only one neighbor on the flange, but of both neighbors at the gap.

The stray fields have been determined by measurement of the fields at the different locations in a number of models like the one used in Figure 1. These fields were totaled and divided by the field caused by the local magnet at the same location; the alternating directions of the fields were taken into account. The ratios obtained, which are called P_f for the flange and P_g for the gap, are shown in Figure 19. Both factors appear to be of approximately the same magnitude and,

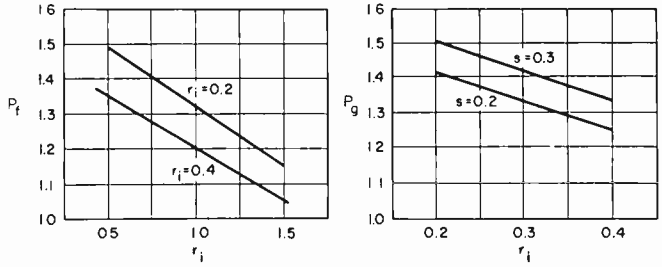


Fig. 19—Stack factors for the flange area (P_f) and the gap area (P_g).

therefore, do not affect the field distribution shown in Figure 17 too much. They may, however, increase the absolute values by as much as 50 per cent.

The Maximum Demagnetizing Field

In the design of a magnet stack, the highest value of the demagnetizing field that can be expected must be determined. There are two critical areas: the area next to the disc surface close to its circumference, and the area in the gap plane close to the hub. The fields are shown in Figure 20. The solid line depicts the gap field, H_g, and the broken line depicts the disk field, H_d. Although the maximum value of the gap field H_M generally limits the operation of the magnet, this field can be reduced by use of an appropriate spacing m between magnet and hub. H_D, on the other hand, limits the possible reduction, usually when the gap spacing m reaches about 0.3.

In the simplified computation of the actual fields, the working field of a stack magnet was approximated by the following expression:

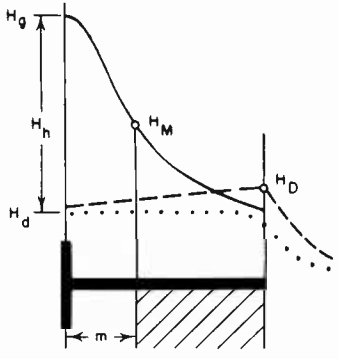


Fig. 20—Potential maximum demagnetizing fields.

$$H_w = J_R \frac{a_m}{(1-2t)\,\phi_\Sigma} P_w, \tag{19}$$

where a_m/ϕ_Σ was the demagnetization factor. It is now known that the disc field is lower by the factor κ, but the same is true essentially for the flux. Therefore, except for the introduction of the proper "stack factor" for the disc area, P_d, the same formula may be used.

$$H_d = J_R \frac{a_m}{(1-2t)\,\phi_\Sigma} P_d. \tag{19a}$$

The equivalent expression for the gap field can be derived from Equation (2b) in the same manner in which Equation (21) was obtained:

$$H_G = P_g H_g = J_R \frac{a_m}{\phi_\Sigma} \frac{P_g}{s}. \tag{25}$$

H_G is the gap field in the assembled stack, while H_g is the gap field in the individual cell. The expression for the maximum hub field in an assembled stack is then given by

$$H_H = H_G - H_d = J_R \frac{a_m}{\phi_\Sigma} \left(\frac{P_g}{s} - \frac{1}{1-2t} \right). \tag{26}$$

At the distance m, the field is found as follows:

$$H_M = H_d + M H_H = J_R \frac{a_m}{\phi_\Sigma} \left(\frac{1-M}{1-2t} + \frac{P_g}{s} \right) \tag{27}$$

where $M = h_h/h_g$ is the function of m and s shown in Figure 14.

The "End Effect"

All the magnets in a stack contribute to the field in any cell. This contribution decreases with increasing distance from the cell concerned. In the middle of the stack, all cells are exposed to practically the same stray fields and, consequently, work under equal conditions. Close to the ends, however, a part of these stray fields is missing, and the previous calculations are not true unless some allowance is made for this end effect. Because the strongest external fields stem from the first neighbors and *increase* the demagnetizing field, there is generally no danger of overstepping the working range of a magnet by removing the cells adjoining it.

There is, however, quite a change in the field on the axis as the cell under consideration comes closer to the end of a stack. In the stack model of Figure 1, an axial field of 190 oersteds was measured in the magnetized cell; in the neighboring cells, values of —45, —20, —10 and —5 oersteds were found. The values are given in the order of increasing distance from the magnetized cell. Although these values were measured with the magnetized cell at the end of a stack, experiments have shown that the values are practically the same for a magnetized cell in the middle of the same stack. The total fields in a given cell are determined by summation of the partial fields contributed by all the neighboring magnets, as shown in Table I. The horizontal-order numbers indicate the location of the cells, and the vertical-order numbers indicate the location of the magnets contributing the partial fields. The total fields in the last line show that the first cell has a relatively low field, while the field of the second cell is quite high. These fluctuations decrease in the succeeding cells and (because all influences extending further than to the fourth neighbor were neglected) a uniform peak field is obtained from cell five on. Although this result agrees well with the familiar pattern, it does not seem to be as bad as the pattern found in many practical cases.

Table I — The Partial Fields in a Cell and their Sum Total

Magnet No.	Cell No.	1	2	3	4	5	6
1		+190	—45	—20	—10	—5	
2		+45	—190	+45	+20	+10	+5
3		—20	—45	+190	—45	—20	—10
4		+10	+20	+45	—190	+45	+20
5		—5	—10	—20	—45	+190	—45
6			+5	+10	+20	+45	—190
7				—5	—10	—20	—45
8					+5	+10	+20
9						—5	—10
10							+5
Total		+220	—265	+245	—255	+250	—250

The end effect can indeed be reduced to a minimum by use of an optimal geometry. In a shim arrangement such as that shown in Figure 21, the field on the axis shows the distribution indicated by the curves. The maximum is located in the gap plane of the magnetized cell, and the negative maximum occurs somewhere between the next gap plane (solid line) and the one thereafter (broken line), depending

on the geometry. The smaller the inside diameter of the structure, the closer the field on the axis must follow the field in the gaps. For wide shim openings, the "penetration" is much higher, and a situation like that indicated by the broken line will prevail. Similar considerations are true for the gap widths, which particularly influence the relative magnitude of the positive and negative maxima.

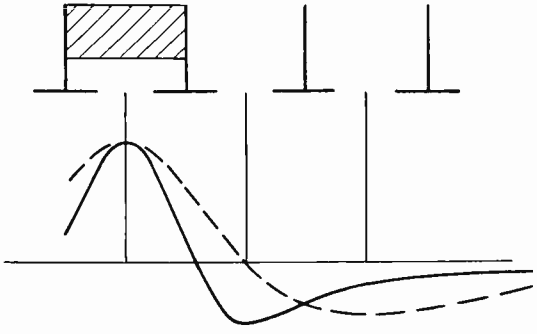


Fig. 21—Axial field caused by a magnetic cell of good (solid line) and poor (broken line) geometry.

CONCLUSIONS

A much better understanding of the situation prevailing in a periodic focusing stack can be gained if the problem is regarded as a static one instead of in terms of the familiar electrodynamic analogy. The distribution of fields and fluxes in space becomes quite obvious in this way, and good approximation by calculation is possible.

In particular, the maximum demagnetizing field in a traveling-wave-tube stack magnet, as well as the peak axial field, can be predicted quite accurately if due attention is paid to the influence of neighboring magnets. This approach was confirmed by measurements on about a dozen stacks of different size and configuration. All results were within 10 per cent of the calculated values. This figure includes any deviations of the magnet materials used (Indox* I and VI) from the typical values published by the manufacturers. A chart can be derived from the formulas given in this paper to permit the design of a focusing stack in little more than a minute.

The analysis furthermore points out that the field on the axis in any location is composed of a number of partial fields stemming from different magnets. This fact accounts for the varying peak fields at the ends of a stack, and knowledge of the factors involved points the way to minimizing the undesirable "end effect."

* Registered trade mark.

EVALUATION AND CONTROL OF DIFFUSED IMPURITY LAYERS IN GERMANIUM*

By

H. S. Veloric and W. J. Greig

RCA Semiconductor and Materials Division,
Somerville, N. J.

Summary—*This paper presents methods for the evaluation of diffused impurity layers in germanium. The calculations reported are for n- and p-type impurities with a complementary-error-function (erfc) distribution in uniformly doped germanium of the opposite conductivity. These methods are then used to evaluate a new solid-phase diffusion process employing a powdered-germanium alloy doped with a known impurity concentration as the diffusant. The surface concentration of a diffused layer is shown to be approximately equal to the concentration in the source powder at high temperature and to decrease with decreasing temperature. Diffusion constants for arsenic and antimony are determined by use of "sheet-resistance" and junction-penetration measurements.*

Because the surface concentration can be predicted within a factor of two, and the penetration within ten per cent, the present experiments show that the powder diffusion method is suitable for the fabrication of carefully controlled impurity layers.

Electron and Hole Mobility in Heavily Doped Germanium

Calculations involving the mobilities of electrons and holes and the surface concentration of diffused layers may be used to evaluate powder diffusion as a technique for the prediction and control of impurity layers in germanium. For an evaluation of the properties of a diffused layer, it is necessary to know the conductivity mobility as a function of impurity concentration.† Prince[1] reports drift mobility values for impurity concentrations up to 2×10^{17} per cubic centimeter for electrons and 5×10^{16} per cubic centimeter for holes. Conwell[2] has calculated the mobility of electrons for impurity concentrations up to 10^{19}, and Trumbore[3] has measured the mobility

* Manuscript received July 18, 1960.

† Conwell defines the conductivity mobility as the mobility which is calculated from the conductivity and the numbers of carriers, the latter quantity being obtained from methods other than Hall measurements.

[1] M. B. Prince, "Drift Mobilities in Semiconductors. I. Germanium," *Phys. Rev.*, Vol. 92, p. 681, November, 1953.

[2] E. M. Conwell, "Properties of Silicon and Germanium," *Proc. I.R.E.*, Vol. 40, p. 1327, November, 1952.

[3] F. A. Trumbore and A. A. Tartaglia, "Resistivities and Hole Mobilities in Very Heavily Doped Germanium," *Jour. Appl. Phys.*, Vol. 29, p. 1511, October, 1958.

of holes up to impurity concentrations of 10^{20} per cubic centimeter. In this section, additional data are presented to extend the range of conductivity mobilities to higher concentrations.

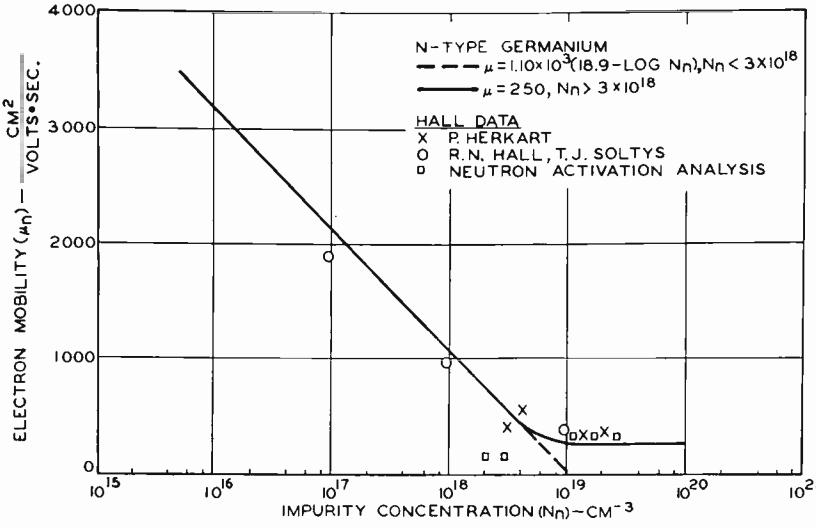


Fig. 1—Electron mobility in n-type germanium as a function of impurity concentration.

Figure 1 is a curve of electron mobility as a function of impurity concentration for n-type germanium. Prince's values of μ are used for concentrations less than 10^{17}. Hall data,[4†] which have been converted to drift mobilities, and values calculated from neutron-activation analysis** are shown in the curve. An extrapolation of Prince's data to higher concentrations is in reasonable agreement with the data. The apparent constant mobility at high impurity concentrations has also been observed in silicon.[5,6] Figure 2 is a curve of hole

[4] R. N. Hall, "Variation of the Distribution Coefficient and Solid Solubility with Temperature," *Jour. Phys. Chem. Solids*, Vol. 3, p. 63, 1957/

† The Hall data were supplied by P. Herkart of the RCA Laboratories, Princeton, New Jersey.

** The neutron activation analysis was performed by G. W. Leddicotte, Oak Ridge National Laboratories, Oak Ridge, Tennessee.

[5] G. Backenstoss, "Conductivity Mobilities of Electrons and Holes in Heavily Doped Silicon," *Phys. Rev.*, Vol. 108, p. 1416, December, 1957.

[6] P. P. Debye and E. M. Conwell, "Electrical Properties of N-Type Germanium," *Phys. Rev.*, Vol. 93, p. 693, February, 1954.

mobility as a function of impurity concentration for p-type germanium which includes both Trumbore's and experimental data. Subsequent calculations are based on Trumbore's data.

In addition to the sources mentioned above, several workers[7-9] have considered the ratio of Hall mobility, μ_H, to the drift mobility, μ_d. All treatments yield a ratio greater than unity for nondegenerate semiconductors, whereas the ratio decreases to unity for the degenerate case. An accepted value[9] for μ_H/μ_d is $3\pi/8$. This value is used to convert all Hall mobility data to drift mobility.

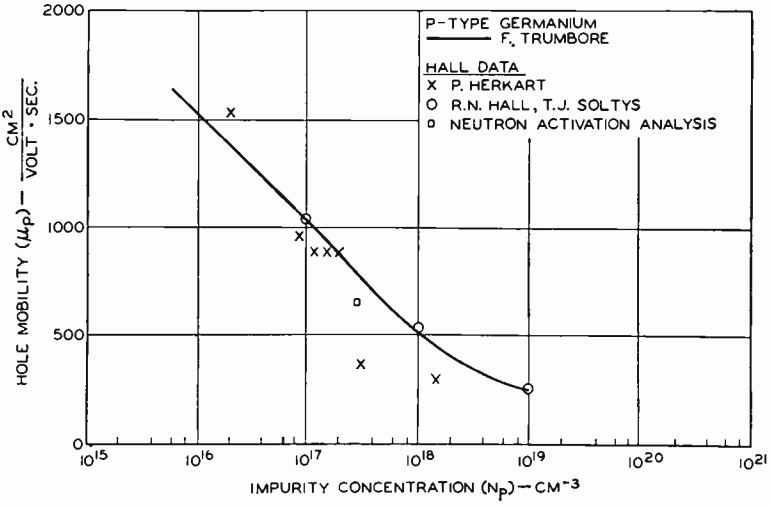


Fig. 2—Hole mobility in p-type germanium as a function of impurity concentration.

Another technique for the calculation of mobilities in the high-concentration range involves determination of the total impurity concentration by neutron-activation analysis. The conductivity mobility, μ, is obtained from the resistivity, ρ, by the relation

$$\mu = \frac{1}{q\rho n} \tag{1}$$

[7] V. A. Johnson and K. Lark-Horovitz, "The Combination of Resistivities in Semiconductors," *Phys. Rev.*, Vol. 82, p. 977, June, 1951.

[8] F. J. Blatt, "Hall and Drift Mobilities; Their Ratio and Temperature Dependence in Semiconductors," *Phys. Rev.*, Vol. 105, p. 1203, February, 1957.

[9] E. M. Conwell, "Properties of Silicon and Germanium: II," *Proc. I.R.E.*, Vol. 46, p. 1281, June, 1958.

where n is the total number of impurities, and q is the electronic charge. Figure 3 shows resistivity as a function of total impurity concentration for n- and p-type germanium. This figure was obtained by means of Equation (1) from the data shown in Figures 1 and 2. Because the impurities used have small ionization energies, they may be assumed to be totally ionized. Therefore, the conductivity mobility is approximately equal to the drift mobility.

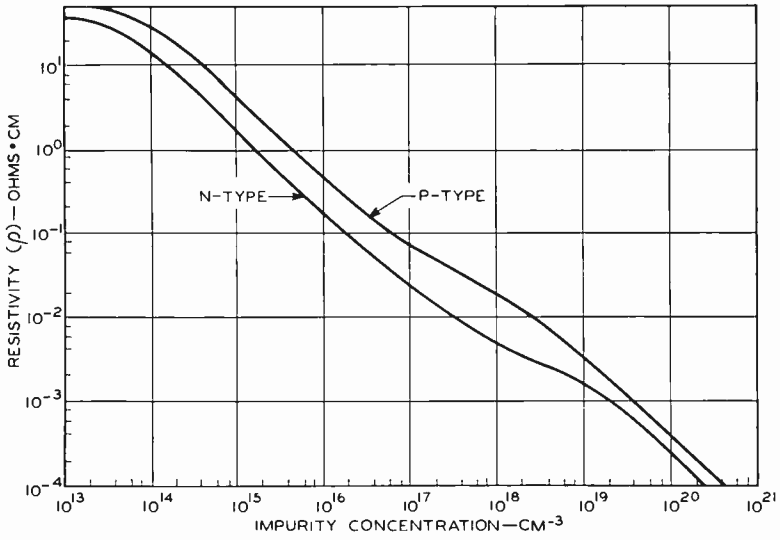


Fig. 3—Resistivity versus impurity density for n- and p-type germanium.

Evaluation of the Properties of Diffused Layers

A diffused layer is characterized by the distribution of the impurities added to the original semiconductor material. To describe a particular impurity distribution, at least two parameters must be known. These parameters may be the diffusion constant–time product, surface concentration, total number of diffused impurities, or concentration at a given depth.

Bakenstoss[10] has shown that the surface concentration of an impurity distribution may be evaluated if the sheet resistivity and layer penetration are known. His calculations are for silicon-diffused layers having several possible impurity distributions. The present investigation extends to include diffused-germanium systems.

[10] G. Bakenstoss, "Evaluation of the Surface Concentration of Diffused Layers in Silicon," *Bell Sys. Tech. Jour.*, Vol. 37, p. 699, May, 1958.

The mathematical form of the impurity distribution is obtained by the solution of the diffusion equation with the appropriate boundary condition. If the diffusant is deposited on the solid and serves as an infinite source, then the resulting concentration $C_{(x)}$ is given by the complementary error function (erfc) as follows:

$$C_{(x)} = C_0 \text{erfc} \left\{ \frac{x}{(4Dt)^{1/2}} \right\}, \tag{2}$$

where $C_{(x)}$ is the concentration at distance x, C_0 is the surface concentration, D is the diffusion constant, and t is the time of diffusion.

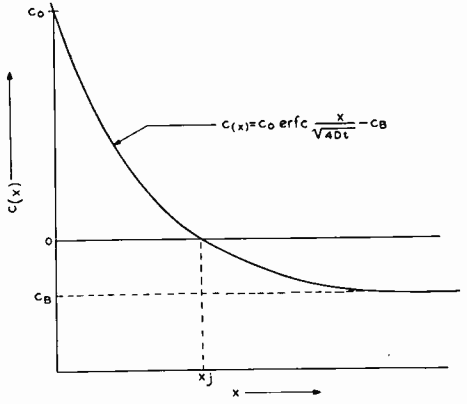


Fig. 4—Impurity distribution across a diffused layer.

Figure 4 schematically illustrates the impurity distribution, with C_B representing the impurity concentration of the original material. At some distance, x_j, the added impurity concentration equals the background and a p–n junction is formed.

The following relation holds[10] for a layer of sheet conductance:

$$\frac{\sigma}{(4Dt)^{1/2}} = qC_0 \int_{Y}^{Y_j} \mu_{(c)} \left(\text{erfc } Y - \frac{C_B}{C_0} \right) dY$$

where

$$Y = \frac{x}{(4Dt)^{1/2}} = \text{erfc}^{-1} \frac{C}{C_0}, \quad Y_j = \frac{x_j}{(4Dt)^{1/2}} = \text{erfc}^{-1} \frac{C_B}{C_0}$$

and where $\mu_{(c)}$ is the mobility as a function of impurity concentration.

The integral can be numerically evaluated* for important values of surface and background concentrations. For ease of computation, the mobility data in Figures 1 and 2 can be described approximately by analytical expressions of the form

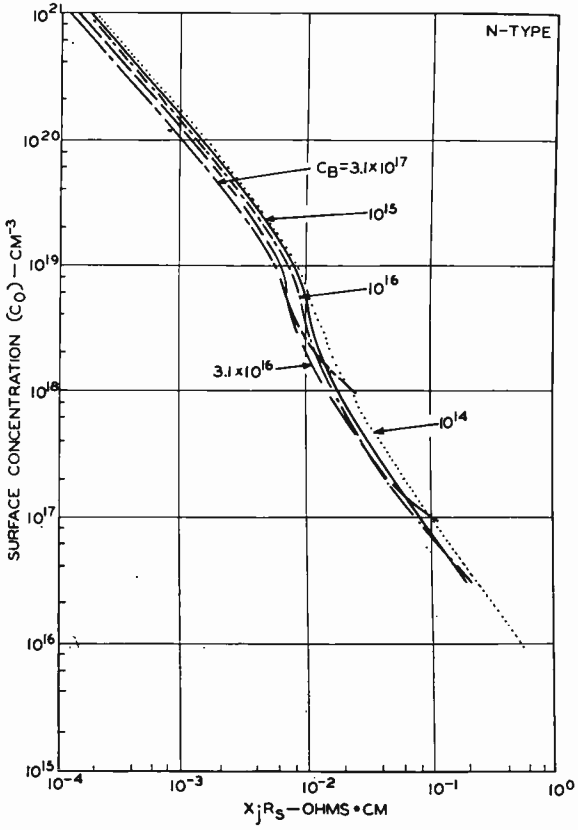


Fig. 5—Surface concentration C_0 versus resistivity for an n-type layer.

$$\mu_n = 1100\,[18.9 - \log\,(C + C_B)\,],$$

for the case $10^{16} \leq C + C_B \leq 5 \times 10^{18}\ \mathrm{cm}^{-3}$

$\mu_n = 250$, for the case of $C + C_B > 5 \times 10^{18}\ \mathrm{cm}^{-3}$

for electrons, and

$\mu_p = 5.4 \times 10^8\, C^{-.33}$, for the case $C + C_B \geq 10^{16}\ \mathrm{cm}^{-3}$

* The integrals were evaluated by W. Hartnet and R. Klopfenstein of the RCA Laboratories, Princeton, New Jersey.

for holes, where μ is expressed in square centimeters per volt-second.

If the integral is evaluated from the junction to the surface, the surface concentration can be calculated from the sheet resistance and Dt product, or from the sheet resistance and junction penetration. In Figure 5, surface concentration C_0 for an n-type layer is shown as a function of sheet resistance R_s and junction penetration x_j, and in Figure 6, it is plotted as a function of sheet conductance σ and Dt.

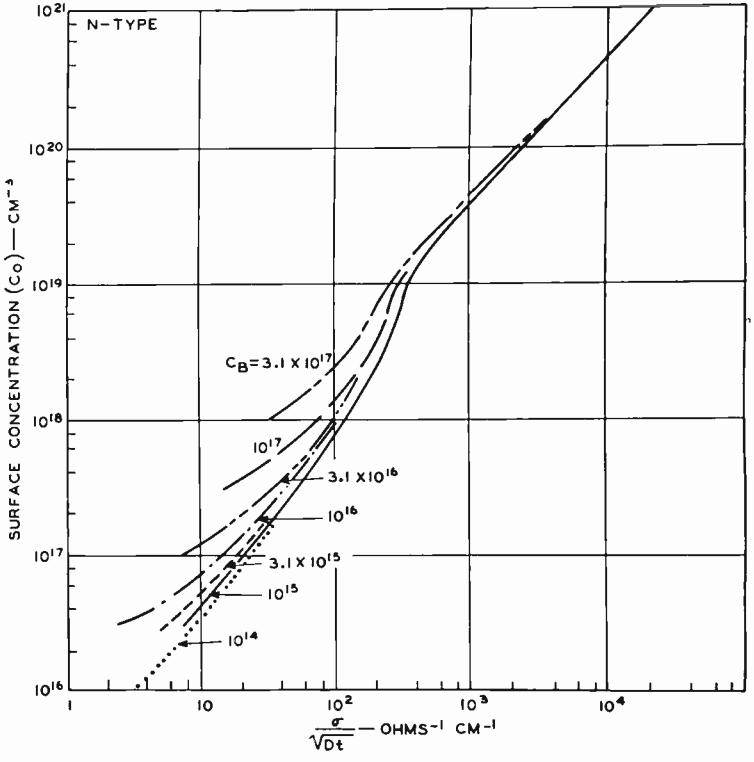


Fig. 6—Surface concentration C_0 as a function of sheet conductance and Dt product for an n-type impurity layer.

The same data are presented in Figures 7 and 8 for a p-type impurity layer. In all cases, the background concentration C_B is the parameter. Although the integral is evaluated for an erfc impurity distribution, the results are similar for exponential and Gaussian distributions.

The conductance of a diffused layer between the junction and any position up to the surface can be calculated from Equation (3). This

quantity is useful in calculations[11] of the intrinsic and extrinsic base resistance and the emitter efficiency of a diffused-base transistor. The calculation of the conductance of a layer is a convenient and sensitive method for measuring the distance between the junction and the surface when part of the layer is removed by an etching or lapping operation.

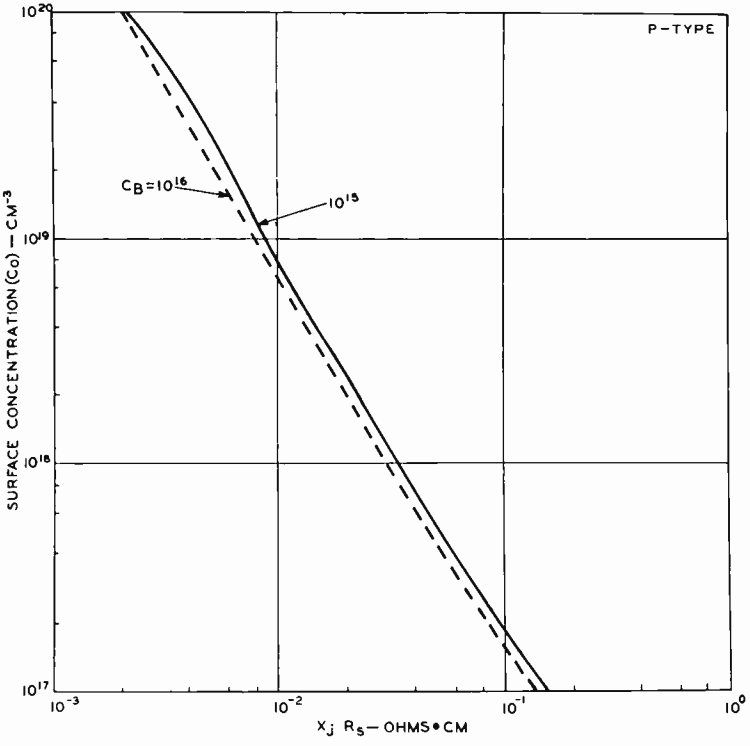


Fig. 7—Surface concentration C_0 versus resistivity for a p-type layer.

Figure 9 shows curves of $\sigma/(4Dt)^{1/2}$ for an n-type impurity layer between Y and Y_j with values of the surface and background concentration as the parameter. Thus, the value of $Y = 0$ represents the total integral between the junction and the surface; the conductance decreases as the layer is removed, becoming 0 at the junction. Figure 10 gives the same information for a p-type impurity layer.

[11] J. M. Early, "Design Theory of Junction Transistors," *Bell Sys. Tech. Jour.*, Vol. 32, p. 1271, November, 1953.

The curves in Figures 9 and 10 show that a large fraction of the layer's conductance is near the surface, where the concentration of impurities is largest. The contribution of the added impurities to the conductance of the layer decreases to a small fraction of the total conductance in the region of the junction.

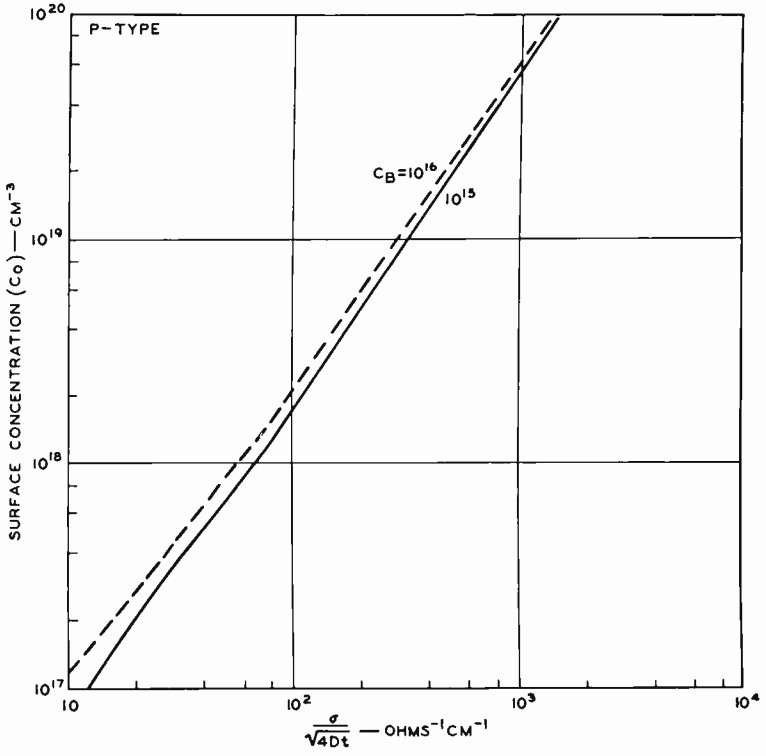


Fig. 8—Surface concentration C_0 as a function of sheet conductance and Dt product for a p-type impurity layer.

Doped Germanium Powder as a Diffusion Source

A solid-phase diffusion process has been developed in which control of the impurity vapor pressure is obtained by use of a powdered germanium alloy doped with a known impurity concentration as the diffusant source. The powder technique offers a means for the prediction and control of both the diffusant surface concentration and the diffused-layer penetration. In addition to having reproducibility, the process is readily adaptable to large-scale production operations requiring a minimum of operator skill and maintenance. Although germanium alloys of arsenic, antimony, indium and aluminum have been

successfully employed, the data reported are primarily for the arsenic–germanium system.

The equipment for the diffusion process consists of a commercially available furnace capable of temperatures up to 1000 degrees centigrade, a furnace tube, a diffusion boat or container, and a thermocouple and controller. The same end results are observed independent of the

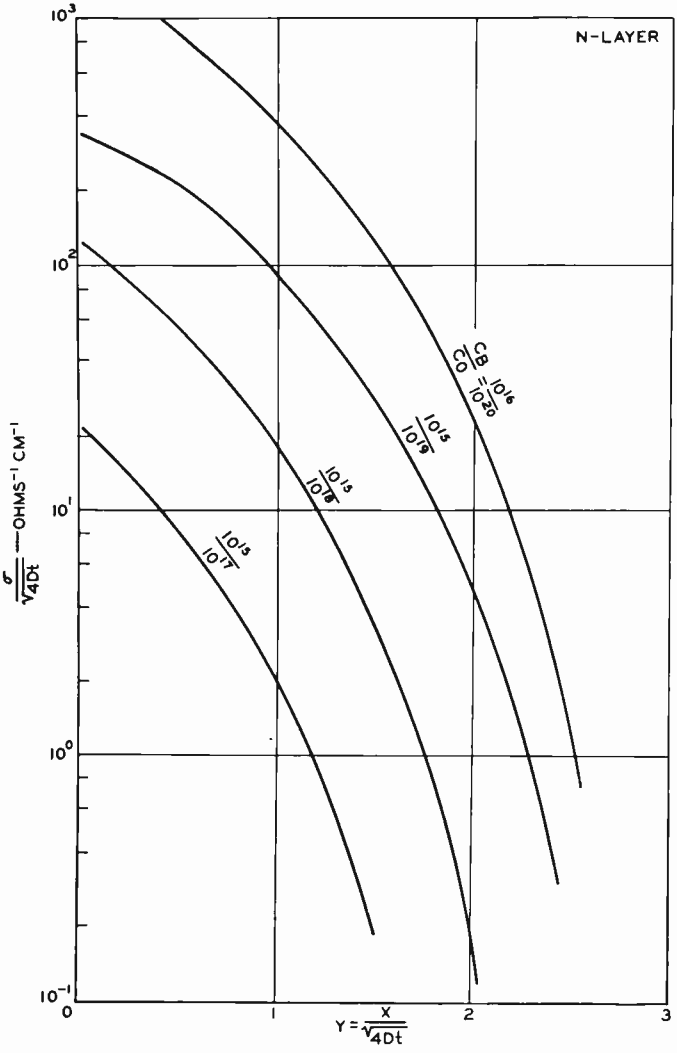


Fig. 9—The conductance of a diffused n-type impurity layer between Y and Y_j as a function of Y for various values of the surface and background concentration.

boat material or design. Materials such as vycor,† quartz, graphite, and molybdenum have been used for containers. Figure 11 shows a typical powder-diffusion furnace setup. In the preparation of the source material, the impurity–germanium alloy is first cleaned in a solution of hydrofluoric and nitric acids. The crystal is then powdered by use of a mortar and pestle. The powdering is followed by a screen-

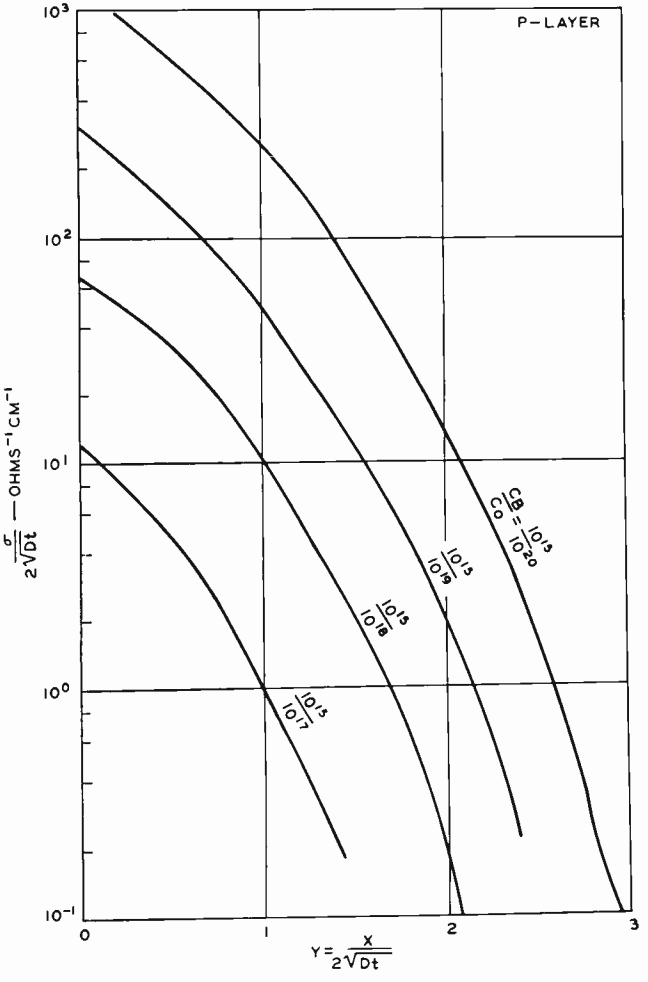


Fig. 10—The conductance of a diffused p-type impurity layer between Y and Y_j as a function of Y for various values of surface and background concentrations.

† Registered trade mark.

ing process, which controls particle size.

The diffusion process consists of physically mixing the source powder and the pellets or wafers, placing the mixture in a boat, and heating at the desired temperature. The diffusion is carried on in a deoxidized atmosphere of dried hydrogen or forming gas.

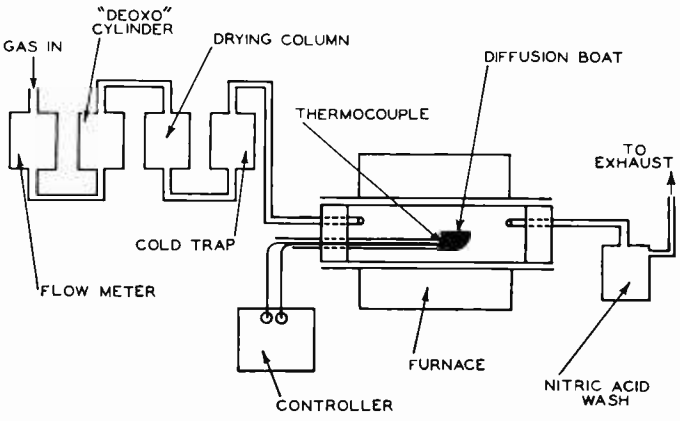


Fig. 11—Diffusion system.

Surface Concentration

The surface concentration is controlled by proper selection of the source powder and the temperature of diffusion. Figure 12 shows the surface concentration as a function of the reciprocal of the absolute temperature for arsenic-doped powders having different impurity concentrations. The values for the surface concentrations were obtained from the data of Figure 6. These values are in good agreement with those obtained through the calculation of the surface concentration from Figure 5 by use of the measured value of $R_s x_j$, the product of the sheet resistance[12] and the observed depth of penetration of the diffused layer. Penetrations were measured by means of angle lapping and staining techniques, using a copper sulfate solution.[13] Junction depths were also checked by interference techniques.[14]

---

[12] F. M. Smits, "Measurement of Sheet Resistivities with the Four-Point Probe," *Bell Sys. Tech. Jour.*, Vol. 37, p. 711, May, 1958.

[13] C. S. Fuller and J. A. Ditzenberger, "Diffusion of Donor and Acceptor Elements in Silicon," *Jour. Appl. Phys.*, Vol. 27, p. 544, May, 1956.

[14] W. L. Bond and F. M. Smits, "The Use of an Interference Microscope for Measurement of Extremely Thin Surface Layers," *Bell Sys. Tech. Jour.*, Vol. 35, p. 1209, September, 1956.

Figure 13 shows curves of surface concentration as a function of impurity concentration in the source material, with temperature as the varied parameter. Figures 12 and 13 indicate the control inherent in the technique. For a given impurity density in the powder, varia-

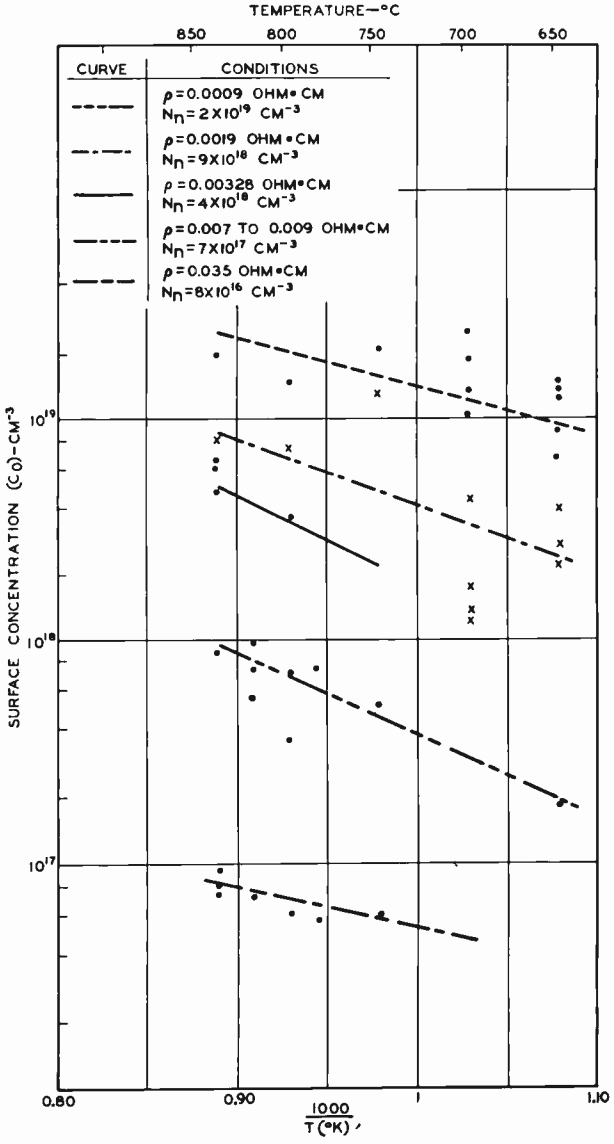


Fig. 12—Diffusant surface concentration versus $1000/T$ with various resistivity source materials as the parameter.

tions in the diffusion temperature of 50 degrees centigrade change the resultant surface concentration by less than a factor of two. For a constant temperature of diffusion, the surface concentration can be controlled entirely by the source material.

In addition to the control, the reproducibility that can be obtained with this diffusion technique is shown in Figure 14, which gives the resultant surface concentrations obtained by the use of a source powder having an impurity concentration of 5×10^{17} per cubic centimeter. In nineteen diffusion runs at a temperature of 825 degrees centigrade, the surface concentrations were between 2 and 3.5×10^{17} per cubic centimeter.

Evaluation of Diffusion Constant

When diffusion constants determined by several workers[15-17] were used, discrepancies were encountered between observed and calculated layer penetrations. In addition, the calculated values of C_0 obtained from the values of $\sigma/(4Dt)^{1/2}$ differed from those obtained from measured values of $R_s x_j$. These experiments indicated smaller diffusion constants than previously reported.

Experiments were made to evaluate the diffusion constant of arsenic in germanium in a temperature range of from 700 to 870 degrees centigrade. Diffusions were made at 870 ± 5 degrees centigrade, 820 ± 5 degrees centigrade, 770 ± 5 degrees centigrade and 700 ± 5 degrees centigrade for periods of 2 hours, 4 hours, 8 hours, and 16 hours at each of the specific temperatures. The same source powder was used for tests made at the same temperature. The background material was 1 ohm-centimeter p-type germanium having an etch-pit density of approximately 8000 per square centimeter. This resistivity was chosen to reduce effects of thermal conversion. The surface concentrations were obtained from the measured sheet resistances and junction depths.

Figure 15 shows the experimental data with a plot of the depth of penetration for an arsenic-impurity layer as a function of the square root of the time of diffusion with temperature as the varied parameter. The values of the diffusion constant were obtained from the slope of

---

[15] Von. W. Bosenberg, "Diffusion von Antimon, Arsen und Indium in festem Germanium," *Zeitschrift fur Maturforschung*, Vol. 10a, p. 285, Heft 4, 1955.

[16] W. C. Dunlap, Jr., "Diffusion of Impurities in Germanium," *Phys. Rev.*, Vol. 94, p. 1531, June, 1954.

[17] C. S. Fuller, "Diffusion of Donor and Acceptor Elements in Germanium," *Phys. Rev.*, Vol. 86, April, 1952.

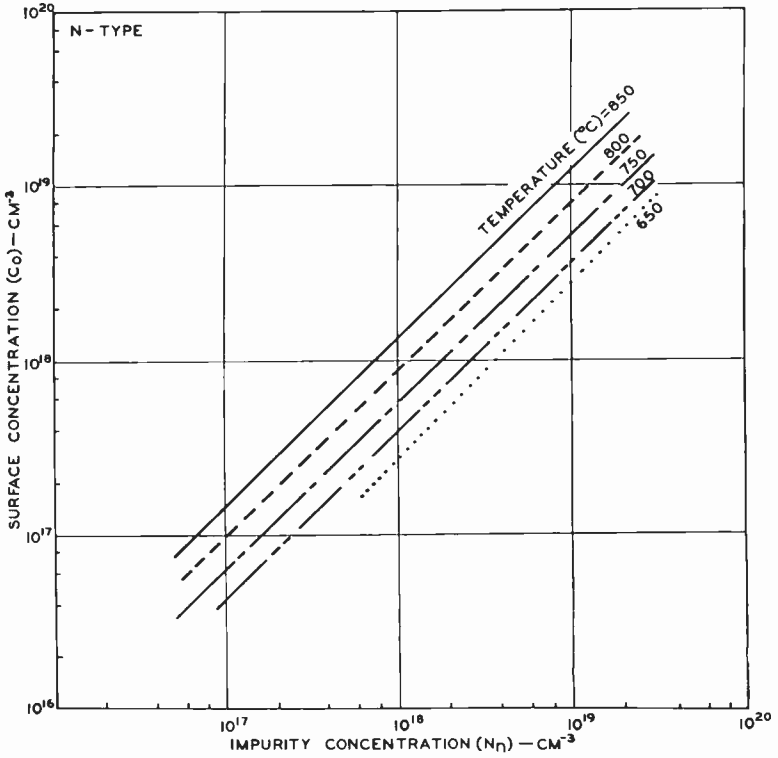


Fig. 13—Surface concentration versus impurity concentration in the source material with temperature as the parameter.

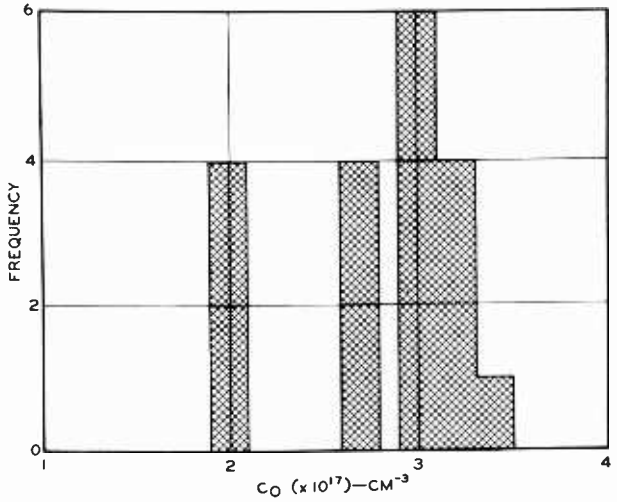


Fig. 14—Surface concentrations obtained from source material of .01 ohm-cm resistivity for 19 diffusion runs.

the lines. Figure 16 shows curves of the sheet conductance $1/R_s$ as a function of diffusion time. The data show that the surface concentration at each diffusion temperature is independent of the diffusion time, which indicates that the boundary requirement for an erfc impurity distribution is satisfied.

Figure 17 shows the values obtained in the present experiment for the diffusion constant of arsenic in germanium as a function of the

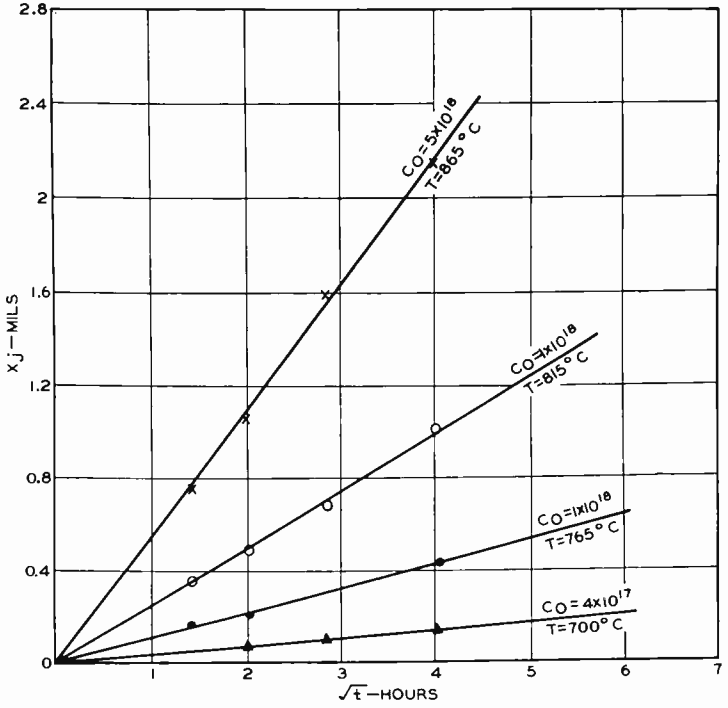


Fig. 15—Arsenic impurity layer penetration versus the square root of the diffusion time with temperature as the parameter.

temperature. They differ from the values of D reported by other workers, which are included in Figure 17, by a factor of from two to ten. The smaller values for the diffusion constant, particularly at the lower temperature, may be attributed to a dependence of D on the surface concentration.[18] The surface concentrations shown in Figure 15 are two and three orders of magnitude smaller than those previously used in experimental determinations of the diffusion constants.[15-17]

[18] F. M. Smits, "Formation of Junction Structure by Solid-State Diffusion," *Proc. I.R.E.*, Vol. 46, p. 1049, June, 1958.

Re-evaluation of the diffusion constant of arsenic in germanium resulted in good agreement with the observed data. When the diffusion constants reported in Figure 17 are used, good agreement is obtained for values of C_0, which are calculated from the observed penetration in the first case, and from the diffusion program in the second case.

This self-consistent check indicates that for the case of donor impurities, the mobility values given in Figure 1 are reasonable. It

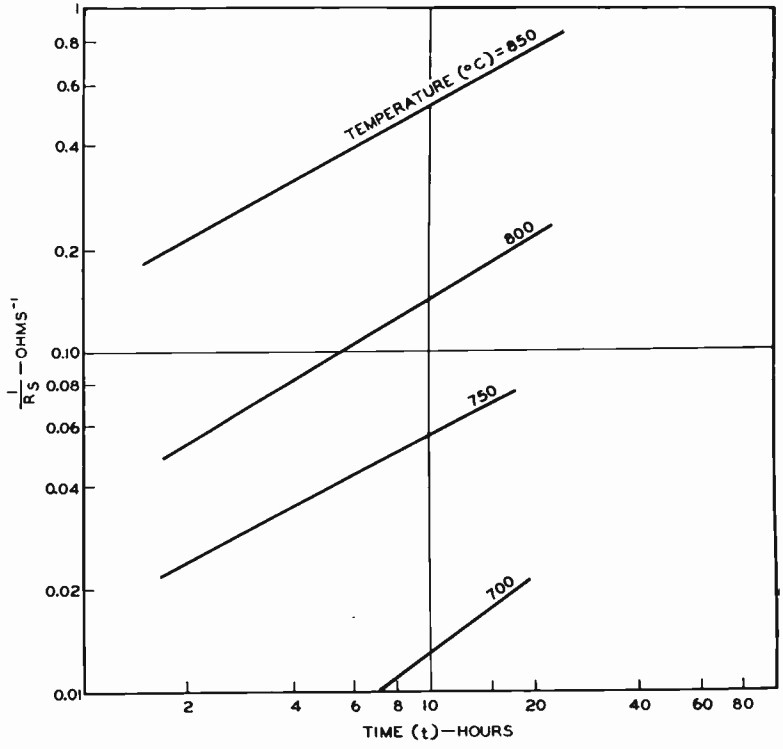


Fig. 16—The reciprocal of the sheet resistance versus the diffusion time with temperature as the parameter.

also suggests a simple technique for checking the diffusion constant of other donor impurities. For example, an antimony–germanium alloy was used as the source powder for diffusions made at 750, 800, and 850 degrees centigrade. The surface concentrations were calculated from the measured values of $R_s x_0$ and from the data shown in Figure 5. The corresponding value of $\sigma/(4Dt)^{1/2}$ for the same C_0 was obtained from Figure 6. The diffusion constant was calculated from $\sigma/(4Dt)^{1/2}$ values for each of the temperatures.

Figure 18 shows the diffusion constant of antimony obtained from the experimental data as a function of the temperature. The antimony data are slightly lower than those reported by other workers, and the diffusion constant for antimony is larger than that for arsenic by a factor of three.

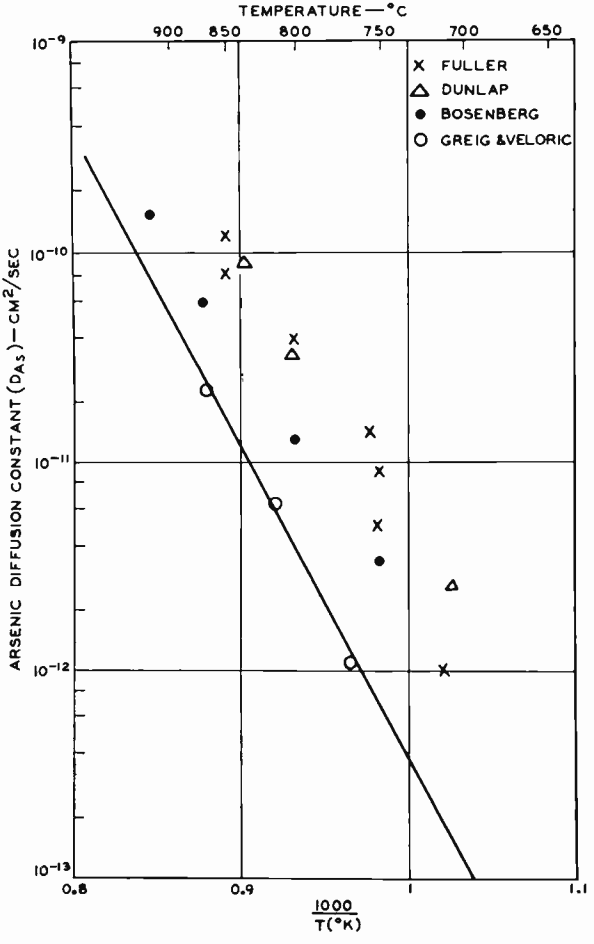


Fig. 17—The diffusion constant for arsenic as a function of temperature.

The diffusion constants as a function of temperature can be approximated by an analytical expression of the form[19]

[19] S. Dushman and I. Langmuir, "The Diffusion Coefficient in Solids and Its Temperature Dependence," *Phys. Rev.*, Vol. 20, p. 113, July, 1922.

$$D = D_0 e^{\frac{-\Delta H}{kT}}$$

where D_0 is the frequency factor, H is the activation energy, k is Boltzmann's constant, and T is the absolute temperature. D_0 and H can be evaluated from Figures 17 and 18.

If k is in electron volts, the equation that fits the data for arsenic is

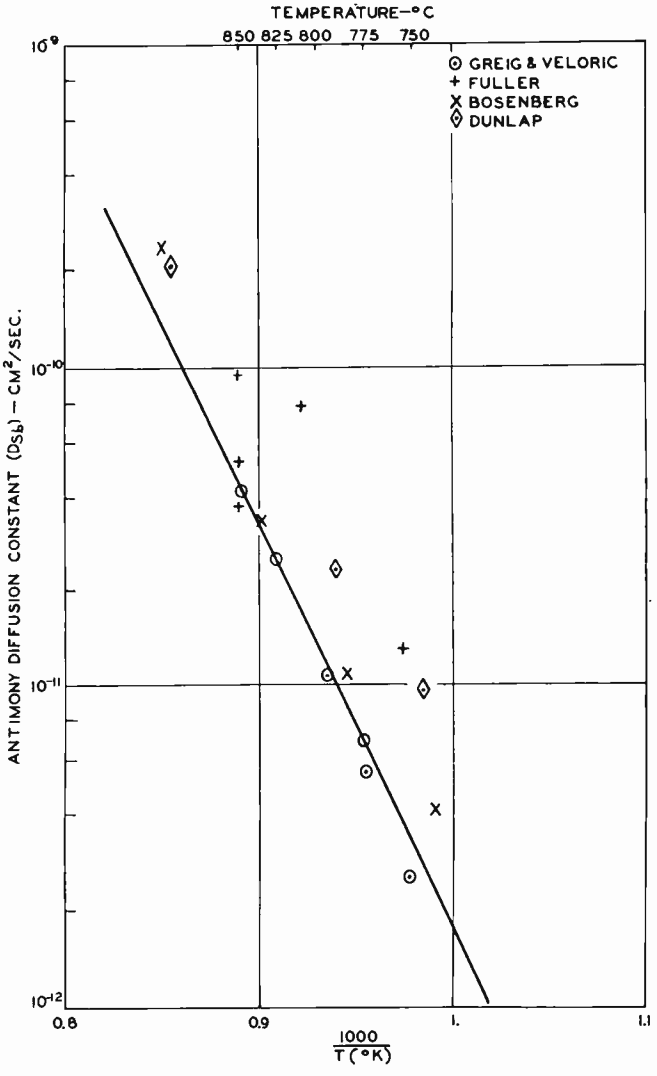


Fig. 18—The diffusion constant for antimony as a function of temperature.

$$D_{As} = 250 \, e^{\frac{-2.9}{kT}}$$

and for antimony

$$D_{Sb} = 0.05 \, e^{\frac{-2.0}{kT}} .$$

Conclusion

Carefully controlled impurity layers in germanium have been fabricated and evaluated. The powder diffusion technique is shown to be a method for obtaining reproducible diffused layers in germanium.

Hall data and neutron-activation analysis indicate that conductivity mobility decreases as a logarithmic function of the total impurity concentration, decreasing at high impurity concentration to a value in the order of 250 cm^2 $volt^{-1}$ sec^{-1} for electrons. The data are less complete for holes but they are in reasonable agreement with those reported by Trumbore.

The conductance of a diffused layer is evaluated as a function of the surface and background concentrations and the diffusion program. Figures are presented from which the surface concentration of a diffused layer can be evaluated from the impurity layer conductance and diffusion program.

Considerable experimental work shows that the surface concentration of added impurities can be estimated from the impurity concentration in the source germanium powder. The diffusion constants for both arsenic and antimony are observed to be smaller by a factor of 2 to 10 from values previously reported.

Acknowledgments

In addition to those previously mentioned, the authors would like to acknowledge numerous discussions with A. Kestenbaum and W. Bosenberg, and the assistance of Miss J. Printon, C. W. Mueller, and A. S. Rose. Much of the experimental work was done by C. Stabinsky and J. Nava.

THE DESIGN OF VARACTOR DIODES*†

By

J. Hilibrand and C. F. Stocker

RCA Laboratories,
Princeton, N.J.

Summary—*The product of diode cutoff frequency and capacitance–voltage sensitivity constitutes a figure of merit for varactor diodes which is inversely proportional to the product of the rise time and the pump power dissipation in a parametric subharmonic oscillator. It also describes the gain–bandwidth product and pump power consumption in a parametric amplifier. This figure of merit can be related in a simple manner to the parameters of the impurity distribution in the diode. Favorable diode designs embody a low-resistance base region supporting a narrow and sparsely doped junction depletion layer.*

Two families of practical diode doping profiles have been considered: those corresponding to alloying to outdiffused surfaces, and those whose impurity distribution can be approximated by a simple power law. Shallow alloy junctions can be improved by outdiffusion of the base wafer until the surface concentration is 30 per cent of the initial concentration.

A still higher figure of merit may be obtained in a structure whose doping approximates a power law, $N = ax^b$, *where* N *is the impurity density,* x *is the distance from the junction,* a *is a constant, and* b *is an arbitrary exponent (between 1 and 2 for optimum figure of merit). For higher exponents, the capacitance–voltage sensitivity decreases faster than the cutoff frequency increases. However, such superlinear structures are shown to be the most favorable for achieving high* Q *values for forward-bias operations. Their low depletion-layer capacitance–voltage sensitivity is compensated for by the high sensitivity of the storage capacitance.*

Introduction

THE USE OF nonlinear energy-storage elements to shift power in the frequency spectrum in amplifier and computer circuits is an important new area in electronics. The variable capacitance of a semiconductor junction diode, which is a lossless storage element well into the microwave range, offers the possibility of microwave-frequency amplification and millimicrosecond switching. This paper is concerned with diode designs aimed at obtaining optimum performance in a parametric subharmonic oscillator circuit[1] such as could be used in an ultra-high-speed computer. This application

* Manuscript received July 18, 1960.

† This work was performed under U. S. Navy BuShips Contracts.

[1] J. Hilibrand and W. R. Beam, "Semiconductor Diodes in Parametric Subharmonic Oscillators," *RCA Review*, Vol. 20, p. 229, June, 1959.

emphasizes a somewhat different set of requirements than those which are of importance for variable-reactance amplifiers or harmonic generators. Within the framework of this set of requirements, an over-all figure of merit for variable-capacitance diodes is established which is directly related to the oscillator rise time and to the power dissipation in the diode. Various impurity distributions are considered and the optimum design is shown to correspond to a superlinearly graded diode (i.e., one in which the impurity density increases more than linearly with the distance from the junction). The design of diodes for forward-bias operation is considered and it is shown that a superlinearly graded diode also offers the possibility of making better use of the forward-bias region.

Various methods of analysis of the operation of the circuits using these elements have been based on the generalized energy transfer properties of nonlinear reactive elements,[2] solutions to the Mathieu equation,[3] and various quasi-linear equivalent circuits.[1,4] In this paper some results obtained from this last approach are used, since this method seems most suitable for device analysis.

A Figure of Merit for Parametric Subharmonic Oscillator Applications

Among the many diode parameters of importance to circuit operation are capacitance–voltage sensitivity, cutoff frequency, avalanche breakdown voltage, and saturation and leakage current. In an ultra-high speed computer, major importance is attached to the first two of these parameters, since a short rise time with low power dissipation is the primary goal. Avalanche breakdown is not a crucial parameter, since the dense packing of diodes necessitates operation with small pump voltages (≈ 2 volts) to reduce power dissipation. Small leakage and saturation currents, which are important for low-noise operation in parametric amplifiers, are not of major consideration in computer circuits since the signal levels are such that noise contributions from the diode current and series resistance are negligible.

---

[2] J. M. Manley and H. E. Rowe, "Some General Properties of Nonlinear Elements, I. General Energy Relations," *Proc. I.R.E.*, Vol. 44, p. 904, July, 1956; H. E. Rowe, "Some General Properties of Nonlinear Elements, II. Small Signal Theory," *Proc. I.R.E.*, Vol. 46, p. 850, May, 1958.

[3] E. Goto, "On the Application of Parametrically Excited Nonlinear Resonators," *Denki Tsuskiu Gakkai-shi*, (The Jour. of Electrical Communication Engineers of Japan), Vol. 38, p. 770, October, 1955.

[4] H. Heffner and G. Wade, "Gain, Band-width, and Noise Characteristics of the Variable-Parameter Amplifier," *Jour. Appl. Phys.*, Vol. 29, p. 1321, September, 1958.

The figure of merit employed in this paper is the product of the capacitance–voltage sensitivity, S_c, and the cutoff frequency, f_{co}. It is defined for a specified applied voltage, V_o, as

$$F = S_c f_{co} = \left| \frac{1}{C} \frac{dC}{dV} \right| \frac{1}{2\pi r C} \tag{1}$$

where C is the diode capacitance and r is the diode series resistance

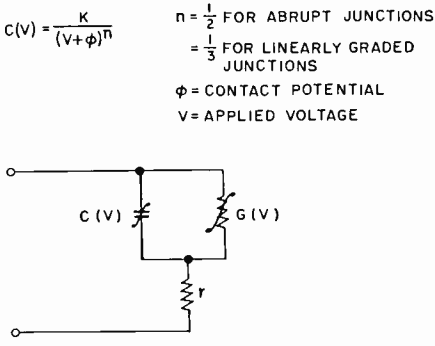


Fig. 1—Nonlinear equivalent circuit for semiconductor diodes.

at voltage V_o (see the diode equivalent circuit in Figure 1). It is demonstrated in Appendix A that

$$F \propto \frac{1}{\tau P}, \tag{2}$$

where τ is the time constant of oscillation build-up, and P is the pump power dissipated in a parametric subharmonic oscillator.

From the usual inverse relation between the rise time in negative-resistance switching and the gain–bandwidth product in negative-resistance amplifiers, it is clear that, for a parametric amplifier, this figure of merit is proportional to the ratio of the gain–bandwidth product to the pump-power consumption. In this sense, the work described in this paper is of broader applicability than the case of the subharmonic oscillator.

This figure of merit may be compared to the figure of merit proposed by Penfield[5] for low-noise-amplifier varactor diodes (which is

[5] P. Penfield, Jr., "The High-Frequency Limit of Varactor Amplifiers," unpublished Microwave Associates, Inc. memorandum, Burlington, Mass., August 27, 1959.

similar to those proposed by others[6] but somewhat simpler in form and more direct conceptually).* Penfield defines a characteristic frequency by the ratio of the available change of diode elastance to the series resistance:

$$f_c = \frac{1}{2\pi} \frac{\Delta\left(\frac{1}{C}\right)}{r}.$$

The figure of merit proposed here can also be written as the ratio of the *rate of change of the diode elastance* with pump voltage to the series resistance:

$$F = \frac{1}{2\pi} \frac{d\left(\frac{1}{C}\right)}{dV} \frac{1}{r}. \tag{1a}$$

F, therefore, is a measure of the cost of subharmonic oscillator rise time (or amplifier gain-bandwidth product) in terms of pump power.

F can also be related in a simple manner to the impurity distribution in the diode. In this paper, a "one-sided" planar diode is analyzed, i.e., it is assumed that the ionized impurity density on one side of the junction is much greater than that on the other side (Figure 2). The more lightly doped side is taken to be n-type so that there is a region of width w in which a positive donor space charge exists. To evaluate the variation of capacitance with voltage, we will imagine that the total voltage across the diode, $V_T = V_0 + \phi$ (the applied voltage plus the contact potential), is increased by dV_T and that there is a resultant increase in depletion layer width by dw. The necessary successive integrations of Poisson's equation and the insertion of suitable boundary conditions can be interpreted graphically as in Figure 2. The electric field goes from zero at $x = w$ to a maximum value at $x = 0$. The effect of added depletion layer width is to increase the field by an increment $(e/\epsilon)N(w)dw$ where e is the electron charge, ϵ is the permittivity of the semiconductor, and $N(x)$ is the net donor impurity density at distance x from the metallurgical junction. The potential is the result of integrating the field, and the incremental voltage is

---

[6] K. E. Mortenson, "Alloyed, Thin-base Diode Capacitors for Parametric Amplification," *Jour. Appl. Phys.*, Vol. 30, pp. 1542-1548, October, 1959; R. C. Knechtli and R. D. Weglein, "Diode Capacitors for Parametric Amplification," *Jour. Appl. Phys.*, Vol. 31, pp. 1134-1135, June, 1960.

* Penfield separates the properties of the diode from those of the diode as pumped.

$$dV_T = \frac{e}{\epsilon} w\, N(w)\, dw. \tag{3}$$

The figure of merit for an arbitrary impurity distribution can be derived;

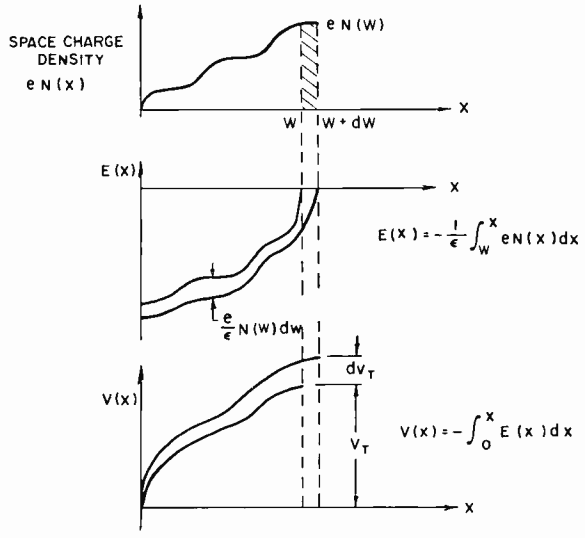


Fig. 2—Space-charge, field, and potential distributions in a one-sided junction diode ($N(x)|_{x<0} \gg N(x)|_{x>0}$)

$$C = \frac{\epsilon A}{w}, \tag{4}$$

$$r = \frac{1}{A}\int_w^t \rho(x)\, dx = \frac{1}{Ae}\int_w^t \frac{dx}{\mu_n N(x)}, \tag{5}$$

$$\frac{1}{C}\frac{dC}{dV_T} = -\frac{1}{w}\frac{dw}{dV_T} = -\frac{\epsilon}{e}\frac{1}{w^2 N(w)}, \tag{6}$$

where A is the junction area, t is the wafer thickness, $\rho(x)$ is the resistivity as a function of distance from the junction, and μ_n is the electron mobility. Planar majority carrier flow is assumed rather than a spreading resistance in obtaining r. The figure of merit, F, which is to be maximized, can be written simply as

$$F = \frac{1}{2\pi\, w\, N(w) \displaystyle\int_{w}^{t} \frac{dx}{\mu_n N(x)}}. \tag{7}$$

The integral is a measure of series resistance, and its reduction calls for heavy doping $N(x)$ in the region $x > w$. The desirability of a low doping $N(w)$ at w, calls for a doping gradient at the edge of the depletion layer. In the following sections, two families of gradients which describe junction fabrication possibilities are considered. F is then evaluated in terms of the respective fabrication parameters.

The Design of Diodes for Reverse-Bias Operation

The families of gradients under consideration are (1) an impurity distribution corresponding to an alloy junction on the surface of an out-diffused wafer, and (2) an impurity distribution given by $N(x) = N_{\max}\,(x/t)^b$, where b is a parameter in the analysis. Both of these families include the abrupt junction as a limiting case.

The Out-diffused, Alloyed Diode Model

The form of Equation (7) indicates clearly that a lightly doped capacitive region on a heavily doped high-mobility substrate is desired. This suggests out-diffusion of impurities[7] from heavily doped n-type germanium crystal followed by shallow alloying. Such a structure was examined to determine the optimum out-diffusion conditions.

Out-diffusion from solids is commonly subject to a rate limitation at the surface so that the governing boundary condition at the surface is[8]

$$KN - D\frac{\partial N}{\partial x} = 0, \tag{8}$$

where N is the impurity density, K is the evaporation rate constant,

---

[7] J. Halpern and R. H. Rediker, "Outdiffusion as a Technique for the Production of Diodes and Transistors," *Proc. I.R.E.*, Vol. 46, p. 1068, June, 1958.

[8] F. M. Smits and R. C. Miller, "Rate Limitation at the Surface for Impurity Diffusion in Semiconductors," *Phys. Rev.*, Vol. 104, p. 1242, December, 1956; R. C. Miller and F. M. Smits, "Diffusion of Antimony Out of Germanium and Some Properties of the Antimony-Germanium System," *Phys. Rev.*, Vol. 107, p. 65, July, 1957.

and D is the diffusion constant for the impurity atoms in the crystal lattice. The impurity distribution obeys the law

$$\frac{N(x)}{N_B} = \operatorname{erf}\left(\frac{x}{2\sqrt{Dt}}\right) + \exp\left(\frac{xK}{D} + \frac{k^2 t}{D}\right) \operatorname{erfc}\left[K\sqrt{\frac{t}{D}} + \frac{x}{2\sqrt{Dt}}\right], \tag{9}$$

where K and D are known constants* and t is the duration of the out-diffusion cycle. The figure of merit was evaluated numerically for various bulk doping levels N_B using the ratio of surface to bulk impurity densities (N_{surface}/N_B) as a running parameter. Through Equation (9), this ratio yields the diffusion cycle variables, time and temperature.

The electron mobility was taken as

$$\mu_n = \begin{cases} 3800 \text{ cm}^2/\text{v-sec} & N < 1.5 \times 10^{15}/\text{cm}^3 \\ 1030\ (18.9 - \log_{10} N) & 1.5 \times 10^{15} < N < 10^{17} \\ 9.2 \times 10^8/N^{1/3} & 10^{17} < N \end{cases} \tag{10}$$

These values are in good agreement with the data of Tyler and Soltys[9] for n-type germanium. A wafer thickness of 2×10^{-3} cm and a total voltage of 1 volt were used. The original, thicker wafer is assumed to be lapped and/or etched down to 2×10^{-3} cm after the out-diffusion cycle, so that it is out-diffused from one surface only.

The resulting values for the figure of merit are shown in Figure 3. It will be noted that F improves by a factor of only 2.5 when the bulk doping density, N_B, is increased by two orders of magnitude. The optimum surface doping level is about 30 per cent of the bulk density. For higher surface concentrations, the doping level in the depletion layer (and hence the depletion-layer capacitance) is excessively high. Lower surface concentrations require more extreme out-diffusion and there is an appreciable decrease of the bulk impurity density so that the series resistance increases. Curves of $N(x)$ for various surface-concentration ratios and surface-evaporation-rate constants appear in Figure 4. The effect of the surface-evaporation-rate constant, K, on the figure of merit is shown in Figure 5. The figure of merit is almost

* The unimolecular evaporation rate constant, K, has been measured by Smits and Miller[8] who indicate that for antimony outdiffused from germanium in vacuum, the ratio K/D is independent of temperature and equal to 2 x 10^3 per centimeter.

[9] W. W. Tyler and T. J. Soltys, private communication.

doubled if the ratio K/D is increased by a factor of five. Such an increase in K/D could come with use of different materials or of different evaporation techniques.

An important upper limit on the figure of merit is established by the onset of avalanche breakdown and/or field emission and tunneling at very high dopings. The curve in Figure 3 labeled $E = 2 \times 10^5$ volts

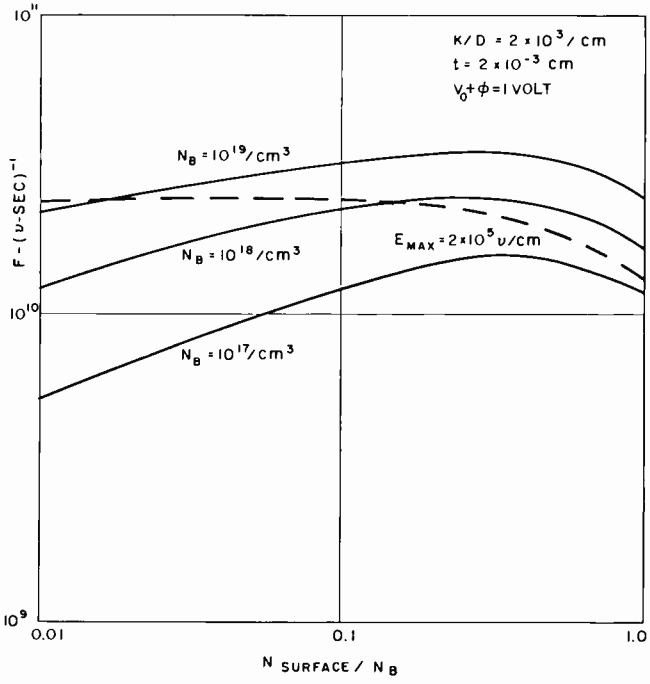


Fig. 3—Effect of degree of out-diffusion on diode figure of merit.

per centimeter indicates approximately the region beyond which improvement in F from increased doping cannot be exploited. Thus, for example, if bulk doping of 10^{18}/cm^3 is used, out-diffusion to less than 10 per cent of the bulk density is required. It should be noted that this is not a rigorously defined boundary; the theoretical basis for considering the two limiting junction-breakdown processes is not yet firmly established. It may be concluded, however, that out-diffusion from germanium where the rate limitation is as shown will yield diodes with a maximum figure of merit of about 2×10^{10} per volt-second, and that to obtain this value of F, the initial bulk doping can be set at any value greater than 10^{18}/cm^3, but out-diffusion to a surface concentration of 1.5×10^{17}/cm^3 is necessary.

The Power-Law Diode Model

A diode model which is more flexible treats the power-law junction where $N(x) = N_{max}(x/t)^b$. The abrupt junction is defined by $b = 0$, the linearly graded junction by $b = 1$, and the cases for $b < 1$ and for $b > 1$ are called sublinear and superlinear doping profiles, respectively.

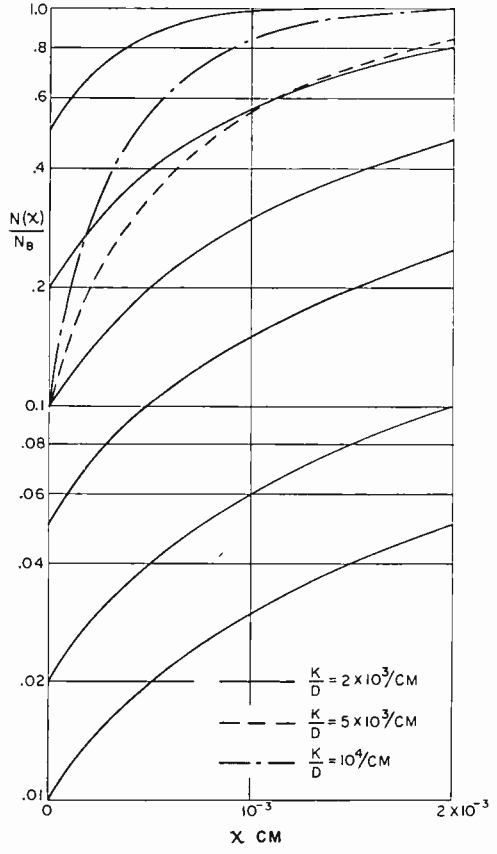


Fig. 4—Impurity profiles for out-diffused wafers.

The examination of power-law doping proceeds from the figure-of-merit formulation in a straightforward manner. The choice of suitable boundary conditions on the doping was intended to represent physical limitations, so that the doping at the base end of the wafer, N_{max}, is chosen as a parameter and may be set by considerations of the solubility limit for impurities. A base thickness, t, of 2×10^{-3} cm is assumed in the computations. The depletion layer width for such a

doping profile is

$$w = \left(\frac{\epsilon V_T}{eN_{\max}} \right)^{1/(b+2)} t^{b/(b+2)}. \tag{11}$$

Numerical integration was used to evaluate the figure of merit, making use of the mobility approximation given in Equation (10).

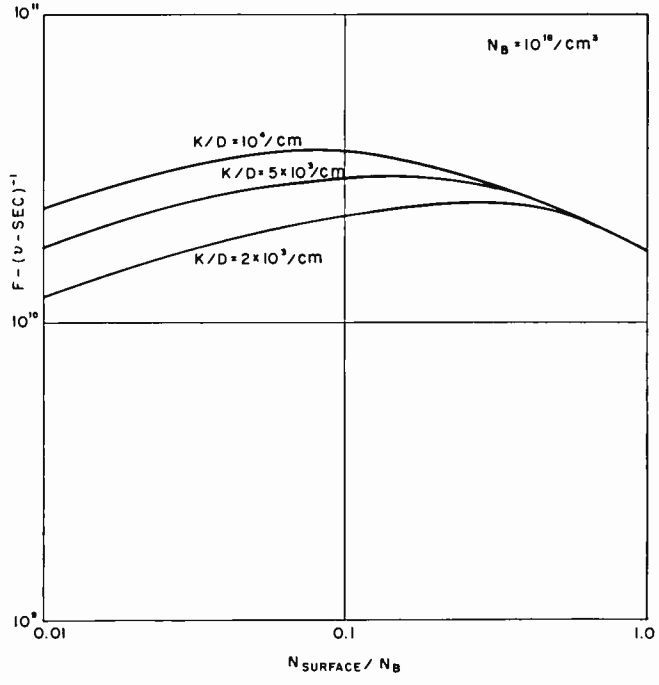


Fig. 5—Effect of surface evaporation rate on figure of merit.

The results are shown in Figure 6 for various doping densities. It should be noted that the optimum designs correspond to $1 < b < 2$. Here again, the requirement that the maximum field be less than 2×10^5 volts per centimeter restricts the range of choice, but the most favorable exponents can still be achieved. The smaller values of b correspond to a high capacitance and poor cutoff frequency, but good capacitance–voltage sensitivity. The larger values of b correspond to poor capacitance–voltage sensitivity but high cutoff frequency. The over-all resistance does not change appreciably along any line of maximum doping level, but the variation of capacitance with b is marked. By the use of a superlinear junction, it is also possible to go

to higher doping levels but still achieve acceptable breakdown voltages. A superlinear junction also has additional properties of interest for forward-bias operation of the varactor diode, as shown in the next section.

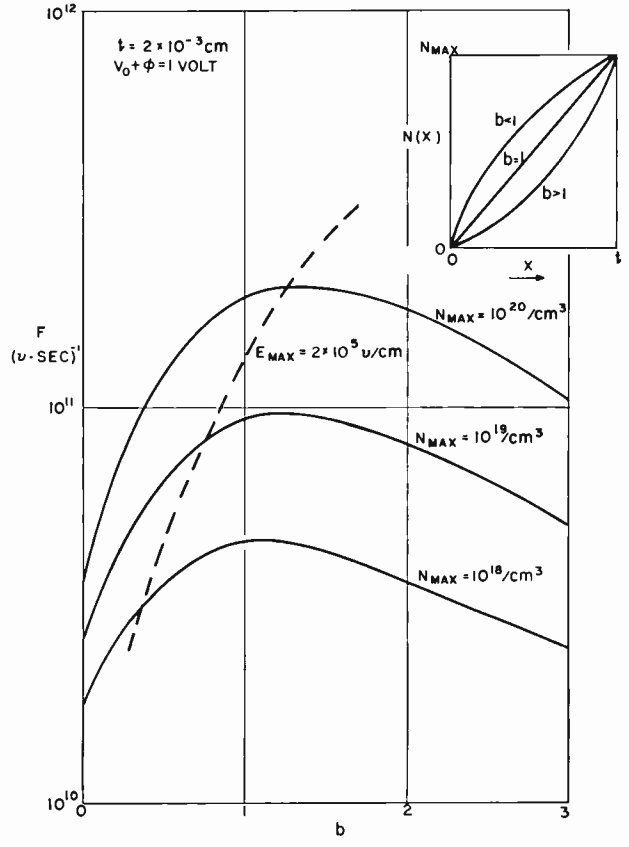


Fig. 6—Effect of exponent in power-law junctions.

Forward-Bias Operation

It is well known[10] that when an increasing forward bias is applied to an abrupt p-n junction diode, the admittance due to the flow of injected minority carriers (the storage or diffusion admittance) soon exceeds the admittance of the electrostatic depletion-layer capacitance. This storage admittance may be represented by equal voltage-sensitive

[10] W. Shockley, "The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors," *Bell Sys. Tech. Jour.*, Vol. 28, p. 435, July, 1949.

susceptance and conductance in parallel at frequencies such that $\omega\tau \gg 1$, where τ is the injected-minority-carrier lifetime. This can be thought of as corresponding to the fact that during the reverse portion of the cycle, half of the injected minority carriers will diffuse back to the junction (the reactive current) while half recombine on the n side of the junction. If a built-in electric field (due to a suitable impurity distribution) could be made to increase the percentage of minority carriers which return, a storage admittance of high quality factor would result and the forward-bias region could be used in various variable reactance applications. A very high capacitance–voltage sensitivity is obtainable since the storage admittance varies exponentially with the applied voltage. ($S_c = e/(kT)$ versus $S_c = 1/2(V_c + \phi)$ for the depletion-layer capacitance of an abrupt-junction diode). An expression for the storage admittance of a linearly graded junction has been obtained by Bakanowski[11] and, indeed, the quality factor is greater than unity. In Appendix B, the p-n junction with doping following a general power-law is examined and an expression for the storage admittance obtained. The admittance is expressed in terms of Bessel functions of the second kind:

$$y_s = \frac{e\mu_p p_{n0}}{w}\left[\frac{\sqrt{1+j\omega\tau_p}}{L_p}\, w \,\frac{K_{(b-1)/2}\left(\dfrac{\sqrt{1+j\omega\tau_p}}{L_p}\, w\right)}{K_{(b+1)/2}\left(\dfrac{\sqrt{1+j\omega\tau_p}}{L_p}\, w\right)}\right] \tag{12}$$

where p_{n0} is the minority-carrier density at the edge of the depletion layer and $L_p = \sqrt{D_p\tau_p}$ is the minority carrier diffusion length. The form of this expression simplifies at high frequencies (where $\omega\tau_p \gg 1$) to

$$y_s = \frac{e\mu_p p_{n0}}{w}\left[\frac{w}{\sqrt{\dfrac{D_p}{\omega}}}\,\frac{K_{(b-1)/2}\left(\dfrac{\sqrt{j}\, w}{\sqrt{\dfrac{D_p}{\omega}}}\right)}{K_{(b+1)/2}\left(\dfrac{\sqrt{j}\, w}{\sqrt{\dfrac{D_p}{\omega}}}\right)}\right] \tag{13}$$

[11] A. E. Bakanowski, *Bell Telephone Laboratories Crystal Rectifier Report*, 9th Interim Technical Report, p. 4, October, 1956.

The magnitude of the argument of the Bessel Functions, $\omega/\sqrt{D/\omega}$, describes whether the carrier flow is dominated by the built-in electrical field at the edge of the depletion layer or by diffusion processes. This argument can be written

$$\frac{w}{\sqrt{\frac{D_p}{\omega}}} = b \frac{\sqrt{\frac{D_p}{\omega}}}{\frac{D_p}{\omega}\frac{b}{w}} = b \frac{l_D}{l_E}. \tag{14}$$

In this expression b is the power-law exponent, l_D is the average length which a carrier will diffuse in a field-free region during one cycle at frequency ω, and l_E is the distance a carrier will drift in one cycle in a field equal to that at the edge of the depletion layer. The argument is of the order of 0.1 to 1.0 for germanium diodes at microwave frequencies.

For small values of the ratio, simplified expressions for y_s can be obtained.

$$y_s = \begin{cases} \dfrac{e\mu_p p_{n0}}{w}(1 + j\omega\tau_p)\left(\dfrac{w}{L_p}\right)^2 \dfrac{1}{b-1}, & b > 1 \\ \dfrac{e\mu_p p_{n0}}{w}(1 + j\omega\tau_p)\left(\dfrac{w}{L_p}\right)^2 \ln \dfrac{\frac{L_p}{w}}{\sqrt{1 + j\omega\tau_p}}, & b = 1 \\ \dfrac{e\mu_p p_{n0}}{w}\left(\dfrac{\sqrt{1 + j\omega\tau_p}}{L_p} w\right)\left(\dfrac{\sqrt{1 + j\omega\tau_p}}{2}\dfrac{w}{L_p}\right)^b \dfrac{\left(\frac{1-b}{2}\right)!}{\left(\frac{1+b}{2}\right)!}\left(\dfrac{1+b}{1-b}\right), & b < 1. \end{cases} \tag{15}$$

These expressions indicate that high-quality-factor storage admittances can, indeed, be achieved through the use of superlinear junc-

tions. At very high frequencies, where $w/\sqrt{D_p/\omega} \gg 1$, Equation (13) yields a storage-admittance quality factor of unity, just as for ar abrupt-junction diode. The intermediate region is examined in Figure 7. It is clear that it is advantageous to use superlinear junctions with large built-in fields, and that it is possible to maintain appreciable Q up to large values of the parameter $w/\sqrt{D_p/\omega}$.

The usefulness of this highly voltage-sensitive, high-quality-factor capacitance is limited by the restricted voltage range in which one

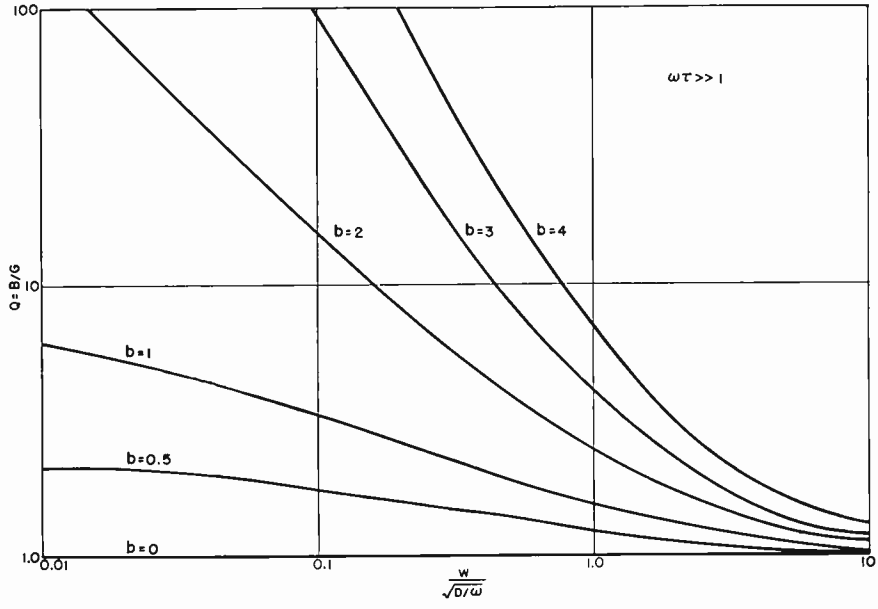


Fig. 7—Quality factor of diode storage admittance.

can use it. For small forward biases the storage capacitance is less than the transition capacitance, and for large bias voltages the losses in the diode series resistance limit the voltage excursion. However, a very approximate analysis of the Fourier coefficients of the capacitance indicates that the contribution of the storage capacitance can be of the same order as that of the transition capacitance despite the small voltage swing involved. Since both the optimization of diodes for reverse-bias applications and the improvement in capacitance–voltage variation at high quality factor for forward bias indicate a superlinear junction design, the development of techniques for producing such impurity distributions takes on considerable importance.

Conclusion

It has been shown that to achieve minimum rise time and power dissipation in a parametric subharmonic oscillator, diodes with high cutoff frequencies and high capacitance–voltage sensitivities are needed. Favorable diode designs embody a low-resistance base region supporting a narrow but sparsely doped junction depletion layer. Design parameters have been optimized for two families of junctions: those formed by alloying to an outdiffused surface, and those whose doping distribution approximates a power law. It is pointed out that the latter are optimized for reverse-bias operation when the doping is between a linear and a quadratic function of distance from the junction. Furthermore, superlinear distributions are most favorable for forward-bias operation.

Appendix A—The Figure of Merit, F, in Terms of Subharmonic Oscillator Circuit Parameters

The results of prior analyses[1] can be applied to interpret F in terms of circuit response and power consumption.

The pump power absorbed in the diode has been shown to be

$$P = \frac{2\omega C_0}{Q} V_p^2 = \frac{2\omega^2 C_0}{\omega_{c0}} V_p^2,$$

where C_0 and Q are the diode capacitance and quality factor at the bias point, ω is the radian oscillation frequency, ω_{c0} is the radian frequency at which Q goes to unity, and V_p is the pump voltage amplitude. The rise time of the subharmonic oscillations can be written approximately as

$$\tau \cong \frac{4}{n\omega},$$

where n is the exponent in the expression for $C(V)$ in Figure 1 and is a measure of the transition capacitance–voltage sensitivity. The figure of merit can be written

$$F = \left| \frac{1}{C_0} \frac{dC}{dV} \right| \frac{1}{2\pi r C_0} = \frac{n}{V_0 + \phi} \frac{\omega_{c0}}{2\pi} = \frac{4}{\pi} \left(\frac{V_p}{V_0 + \phi} \right)^2 (\omega C_0)(V_0 + \phi) \frac{1}{\tau P}$$

In this expression $V_p \cong V_0 + \phi$ and ωC_0 is determined by the desired

impedance level for the circuit so that F is proportional to the inverse of the product of the rise time and the power dissipation. F, then, is a suitable figure of merit for variable-capacitance diodes used in parametric subharmonic oscillator circuits for ultra-high-speed computers.

It must be noted, however, that these results are, strictly speaking, applicable only when the voltage excursion is small and only the transition-capacitance variation is of importance. If the storage capacitance is also to be used, development of a suitable figure of merit is more complicated.

Appendix B—The Storage Admittance of an x^b Junction Diode

The storage admittance of a junction diode with power-law doping is examined on the basis of straightforward considerations of carrier flow. Points where elaboration might be called for are indicated but the elaboration is not carried out here. A one-sided junction is considered and carrier flow in the less heavily doped n-type region is examined.

The doping distribution is $N_D - N_A = N(x) = ax^b$ for $x > 0$. The built-in field will be evaluated for equilibrium conditions but approximate space-charge neutrality is assumed to hold when carrier flow is considered. The built-in field is

$$E = -\frac{kT}{e}\frac{1}{n}\frac{dn}{dx} = -\frac{kT}{e}\frac{b}{x}.$$

The hole continuity equation is written

$$\begin{aligned}\frac{\partial p}{\partial t} &= -\frac{\partial J_p}{\partial x} - \frac{p - p_n}{\tau_p} \\ &= -\frac{\partial}{\partial x}\left[-D_p\frac{\partial p}{\partial x} + p\mu_p E\right] - \frac{p - p_n}{\tau_p} \\ &= D_p\frac{\partial^2 p}{\partial x^2} + b\frac{D_p}{x}\frac{\partial p}{\partial x} - bD_p\frac{p}{x^2} - \frac{p - p_n}{\tau_p}.\end{aligned}$$

If a small applied voltage $v(t) = V_0 + V_1 e^{j\omega t}$ is assumed, the hole density is $p(t) = p_{n0} + p_1 e^{j\omega t}$. Substituting and examining only the small-signal component

$$j\omega p_1 = D_p \frac{d^2 p_1}{dx^2} + \frac{bD_p}{x}\frac{dp_1}{dx} - \frac{b}{x^2} D_p p_1 - \frac{p_1}{\tau_p},$$

or

$$0 = x^2 \frac{d^2 p_1}{dx^2} + bx \frac{dp_1}{dx} - \left(b + \frac{1 + j\omega\tau_p}{{L_p}^2} x^2 \right) p_1.$$

This is a Bessel Equation[10] whose solution is

$$p_1 = x^{\frac{1-b}{2}} \left[C_1 I_{(b+1)/2}\left(\frac{\sqrt{1 + j\omega\tau_p}}{L_p} x \right) + C_2 I_{-(b+1)/2}\left(\frac{\sqrt{1 + j\omega\tau_p}}{L_p} x \right) \right].$$

In order to match the boundary condition on p_1 for $x \to \infty$, it is necessary that $C_1 = -C_2$ and at $x = 0$, $p_1 = p_{n0} eV_1/(kT)$ and $p_{n0} = p_n e^{eV_0/kT}$. The final expression for p_1 is

$$p_1 = p_{n0} \frac{eV_1}{kT} \left(\frac{x}{w} \right)^{\frac{1-b}{2}} \left[\frac{K_{\frac{b+1}{2}}\left(\frac{\sqrt{1 + j\omega\tau_p}}{L_p} x \right)}{K_{\frac{b+1}{2}}\left(\frac{\sqrt{1 + j\omega\tau_p}}{L_p} w \right)} \right].$$

The current density can be evaluated at $x = w$

$$J_{p_1}\Big|_{x=w} = -eD_p \frac{dp_1}{dt} - e\mu_p \frac{kT}{e}\frac{b}{x} p_1$$

$$= eD_p p_{n0} \frac{eV_1}{kT} \left[\frac{\sqrt{1 + j\omega\tau_p}}{L_p} \frac{K_{\frac{b-1}{2}}\left(\frac{\sqrt{1 + j\omega\tau_p}}{L_p} w \right)}{K_{\frac{b+1}{2}}\left(\frac{\sqrt{1 + j\omega\tau_p}}{L_p} w \right)} \right].$$

[12] F. B. Hildebrand, *Advanced Calculus for Engineers*, p. 152, Prentice-Hall, Inc., 1949.

The unit area admittance is simply

$$y_s = \frac{J_{p_1}}{V_1} = \frac{e\mu_p p_{n0}}{w}\left(\frac{\sqrt{1+j\omega\tau_p}}{L_p}w\right)\left[\frac{K_{\frac{b-1}{2}}\left(\frac{\sqrt{1+j\omega\tau_p}}{L_p}w\right)}{K_{\frac{b+1}{2}}\left(\frac{\sqrt{1+j\omega\tau_p}}{L_p}w\right)}\right]$$

RCA TECHNICAL PAPERS†

Second Quarter, 1960

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

"An Automatic Sensitivity Control for Monochrome Film Cameras," S. L. Bendell and K. Sadashige, *Jour. S.M.P.T.E.* (April) 1960

"Dry-Circuit Evaluation of Mechanical Connections," J. W. Kaufman, H. R. Sutton, A. V. Balchaitis, and W. R. Matthias, *Electrical Manufacturing* (April) 1960

"Errors Associated with Thermoelectric Power Measurements Using Small Temperature Differences," R. A. Horne, *Rev. Sci. Instr.* (April) (Notes) 1960

"Forced Periodic Changes of Kinetic Energy of Gas Molecules as a Means of Vacuum Measurement," H. Schwarz, *Rev. Sci. Instr.* (April) 1960

"Method of Sealing Stainless Steel to Glass," J. E. Benbenek and R. E. Honig, *Rev. Sci. Instr.* (April) (Notes) 1960

"New System Design for Detecting Interference to Missiles," H. Kilberg, *Electronic Industries* (April) 1960

"Properties of p-Type InSb in Pulsed High Electric Fields," M. C. Steele and M. Glicksman, *Phys. Rev.* (April 15) 1960

"Receiving Type Tubes," M. B. Knight, *Electronics* (April 29) 1960

"Becquerel Photovoltaic Effect in Binary Compounds," R. Williams, *Jour. Chem. Phys.* (May) 1960

"The Changing Shape of Electronics," W. W. Watts, *Signal* (May) .. 1960

"Comments on 'The Design of Ridged Waveguide,'" R. M. Kurzrok, *Trans. I.R.E. PGMTT* (May) (Correspondence) 1960

"Compact 20-Watt Hi-Fi Amplifier," H. Wittlinger, *Electronics World* (May) 1960

"The Diode-Loaded Helix as a Microwave Amplifier," G. Conrad, K. K. N. Chang, and R. Hughes, *Proc. I.R.E.* (May) (Correspondence) 1960

"The Empreton, A Mercury Pool Arc Tube Allowing Operation at Repetition Rates in the Kilocycle Range," K. G. Hernqvist and F. H. Norman, *Communication and Electronics* (May) 1960

"A Hybrid Mobile Two-Way Radio," R. A. Beers, Jr., *Trans. I.R.E. PGVC* (May) 1960

"An Ivestigation of the Discharge Characteristics of Groups VI-VIII Oxides in an Alkaline Electrolyte," C. K. Morehouse and R. Glicksman, *Jour. Electrochem. Soc.* (May) 1960

"Low-Noise Tunnel-Diode Down Converter Having Conversion Gain," K. K. N. Chang, G. H. Heilmeier, and H. J. Prager, *Proc. I.R.E.* (May) 1960

"Moscow Impressions," W. J. Poch, *Jour. S.M.P.T.E.* (May) 1960

"Photoconduction in Ternary V-VI-VII Compounds," R. Nitsche and W. J. Merz, *Jour. of the Physics and Chemistry of Solids* (May) 1960

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

"The Temperature-Dependence of Electron Spin-Lattice Relaxation Times in Ruby," R. E. Michel, *Jour. of the Physics and Chemistry of Solids* (May) (Letters to the Editor) 1960
"The New Trend in Minified Communications Equipment," J. W. Knoll, *Trans. I.R.E. PGVC* (May) 1960
"A Note Concerning the Precise Measurements of Dipole Antenna Impedance," S. Krevsky, *Trans. I.R.E. PGAP* (May) (Communications) 1960
"Theory of a Negative-Resistance Transmission Line Amplifier with Distributed Noise Generators," K. K. N. Chang, *Jour. Appl. Physics* (May) 1960
"Time Compensation for Speed of Talking in Speech Recognition Machines," H. F. Olson and H. Belar, *Trans. I.R.E. PGA* (May-June) 1960
"Optical Absorption by Degenerate Germanium," J. I. Pankove, *Phys. Rev. Letters* (May 1) 1960
"Diffusion of Cadmium and Zinc in Gallium Arsenide," B. Goldstein, *Phys. Rev.* (May 15) 1960
"Determine Plate Dissipation from Current and Voltage Waveforms," K. W. Angel, *Electronic Design* (May 25) 1960
"Negative-Resistance Amplifier Design," J. B. Schultz and H. B. Yin, *Electronics* (May 27) 1960
"Absolute Spectral Response Characteristics of Photosensitive Devices," R. W. Engstrom, *RCA Review* (June) 1960
"Automatic Control for TV Stations," F. R. McNicol, *Broadcast News* (June) 1960
"Bias in Ferroelectric-Colemanite," E. Fatuzzo, *Jour. Appl. Physics* (June) 1960
"Determination of the Impurity Distribution in Junction Diodes from Capacitance-Voltage Measurements," J. Hillibrand and R. D. Gold, *RCA Review* (June) 1960
"An Electrostatically Focused Traveling-Wave Tube for Wide-Band Amplification in L- and S-Band," C. L. Cuccia and W. C. Johnson, *Proc. National Electronics Conference*, 1959 (June) .. 1960
"Elimination of Cross-Coupling Errors in Rate Gyro Data," P. Mosner, *Trans. I.R.E. PGI* (June) 1960
"Findings of TASO Panel I on Television Transmitting Equipment," J. E. Young (Coauthor) *Proc. I.R.E.* (June) 1960
"Further Improvements in Transistorized Automobile Receivers Using Drift Transistors," R. A. Santilli and C. F. Wheatley, *Semiconductor Products* (June) 1960
"High-Voltage Electron Extraction from an Arc-Discharge Plasma," K. G. Hernqvist, *RCA Review* (June) 1960
"A Low-Wattage Planar Cathode," T. N. Chin, *RCA Review* (June).. 1960
"Microminiature Multichannel Pulse-Position-Modulation System Incorporating Transistor-Magnetic-Core Circuitry," H. Kihn, R. J. Klensch, and A. H. Simon, *RCA Review* (June) 1960
"A New Miniature Beam-Deflection Tube," M. B. Knight, *RCA Review* (June) 1960
"Observations on Angle Diversity," H. Staras, *Proc. I.R.E.* (June) (Correspondence) 1960
"Parametric Amplification, Power Control, and Frequency Multiplication at Microwave Frequencies Using Cyclotron-Frequency Devices," C. L. Cuccia, *RCA Review* (June) 1960
"Photoemission in the Photovoltaic Effect in Cadmium Sulfide Crystals," R. Williams and R. H. Bube, *Jour. Appl. Phys.* (June) 1960
"Picture Quality—Procedures for Evaluating Subjective Effects of Interference," G. L. Fredendall and W. L. Behrend, *Proc. I.R.E.* (June) 1960

"Properties of p-Type GaAs Prepared by Copper Diffusion," F. D. Rosi, D. Meyerhofer, and R. V. Jensen, *Jour. Appl. Phys.* (June) 1960

"Semiconductor Materials for Thermoelectric Power Generation up to 700° C," F. D. Rosi, J. P. Dismukes, and E. F. Hockings, *Elec. Eng.* (June) 1960

"Sine-Squared Pulses in Television System Analysis," R. Kennedy, *RCA Review* (June) 1960

"Sintered Cadmium Sulfide Photoconductive Cells," C. P. Hadley and E. Fischer, *Proc. National Electronics Conference*, 1959 (June) 1960

"Specification of the Linear Feedback System Sensitivity Function," W. M. Mazer, *Trans. I.R.E. PGAC* (June) 1960

"The Stratoscope I Television System," L. E. Flory, G. W. Gray, J. M. Morgan, and W. S. Pike, *RCA Review* (June) 1960

"Telemetry Receiving Antennas at Cape Canaveral," H. A. Roloff, *Trans. I.R.E. PGI* (June) 1960

"Transient Response of a Transmission Line Containing an Arbitrary Number of Small Capacitive Discontinuities," C. Polk, *Trans. I.R.E. PGCT* (June) 1960

"TV Directional Antennas for Taso Experiments," *Broadcast News* (June) 1960

" A Vertical Aperture Equalizer for Television," W. G. Gibson and A. C. Schroeder, *Jour. S.M.P.T.E.* (June) 1960

"Theory of Superconducting Contacts," R. H. Parmenter, *Phys. Rev.* (June 1) 1960

"Television System for Stratoscope I, L. E. Flory, G. W. Gray, J. M. Morgan, and W. S. Pike, *Electronics* (June 17) 1960

"Power-Output Nomograms," L. J. Striednig, *Electronic Design* (June 22) 1960

"Converting Oscilloscopes for Fast Rise Time Sampling," J. J. Amodei, *Electronics* (June 24) 1960

"Antennas for Television Broadcast," H. E. Gihring, *NAB Engineering Handbook*, 5th Edition, Sec. 2, Part 7, McGraw-Hill Book Co., Inc., N. Y. 1960

"Determination of Repetition Frequencies of Intermixed Pulse Trains," R. J. Kern, *I.R.E. International Convention Record*, Part 8, *Aeronautical and Navigational Electronics; Radio Frequency Interference; Vehicular Communications* 1960

"Directional Antennas for Television Broadcasting," G. H. Brown, *I.R.E. International Convention Record*, Part 7, *Audio; Broadcast and Television Receivers; Broadcasting* 1960

"The Effectiveness of Ultrasonic Degreasing as Measured by Radiotracer Techniques," E. L. Romero and H. A. Stern, *I.R.E. International Convention Record*, Part 6, *Component Parts; Industrial Electronics; Production Techniques; Reliability and Quality Control; Ultransonics Engineering* 1960

"How to Edit Your Own Papers," E. M. McElwee, *I.R.E. International Record*, Part 10, *Engineering Management; Engineering Writing and Speech; Human Factors in Electronics* 1960

"The Negative Effective Mass Effect and Quantum Considerations in its Interpretation," G. C. Dousmanis, *Quantum Electronics*, Columbia Univ. Press, N. Y. 1960

"Operation of a Chromium Doped Titania Maser at X and K Band," H. J. Gerritsen and H. R. Lewis, *Quantum Electronics*, Columbia Univ. Press, N. Y. 1960

"Radio Relaying by Reflection from the Sun," D. J. Blattner, *I.R.E. International Convention Record*, Part 5, *Communications Systems; Space Electronics and Telemetry* 1960

"Remote Control of TV Microwave Equipment," J. B. Bullock, *I.R.E. International Convention Record*, Part 7, *Audio; Broadcast and Television Receivers; Broadcasting* 1960

"The Transistor-A New Friend for the Broadcaster," R. N. Hurst and J. W. Wentworth, *NAB Engineering Handbook*, 5th Edition, Sec. 8, Part 1, McGraw-Hill Book Co., Inc., N. Y. 1960

"A Special Effects Amplifier for Noncomposite or Composite, Monochrome or Color Television Signals, R. C. Kennedy, *I.R.E. International Convention Record*, Part 7, *Audio; Broadcast and Television Receivers; Broadcasting* 1960

"A Tunnel Diode Tenth Microsecond Memory," M. M. Kaufman, *I.R.E. International Convention Record*, Part 2, *Circuit Theory; Electronic Computers* 1960

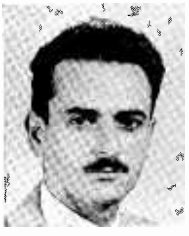
AUTHORS

JACK AVINS received the A.B. degree in Physics from Columbia University in 1932 and the M.S. degree in Electrical Engineering from the Polytechnic Institute of Brooklyn in 1949. Prior to 1941 he was employed by Service Instruments, Inc. where he designed radio test equipment. From 1941 to 1945 he served in the Signal Corps, initially as a radar instructor in the Military College of Science, England, and in 1944 as Chief of the Radar Division, Signal Corps Publications Agency, Fort Monmouth. In 1946 he joined the Industry Service Laboratory, RCA Laboratories Division, New York, where he engaged in radio and television receiver design. In 1956 he transferred to RCA International as Manager of the RCA Industry Service Laboratory, Laboratories RCA, Ltd., Zurich, Switzerland. In 1958, he became Manager of the Research Applications Laboratory, RCA Laboratories, Princeton.

Mr. Avins is a member of Phi Beta Kappa and a Fellow of the Institute of Radio Engineers. He has been active on the IRE Receivers Committee and was Chairman of that committee from 1951 to 1955.

JOHANNES BERGHAMMER received the degree of Diplom-Ingenieur in 1954 and the Dr. Techn. in 1955 from the Technical University, Vienna. From 1954 to 1956 he was a research assistant at the Institute of High-Frequency Techniques in Vienna. Since May 1957, Dr. Berghammer has been a Member of the Technical Staff at RCA Laboratories in Princeton, N. J., and has been engaged in research on electron guns, microwave amplifiers, and electron beam noise.

ERWIN BELOHOUBEK received the degree of Diplom-Ingenieur in 1953 and the Ph.D. in Electrical Engineering in 1955 from the Technical University in Vienna, Austria. From 1953 to 1955 he worked as Research Assistant at the Institute for High-Frequency Techniques of the same university. He joined the RCA Electron Tube Division in Harrison, N. J. in 1956 and transferred to Princeton in 1957, where he is engaged in work on traveling-wave tubes and magnetrons.



LARRY A. FREEDMAN received the B.S. degree in Electrical Engineering from Drexel Institute of Technology in 1948. From 1948 to 1950 he was a research assistant in the Electrical Engineering Department of Rutgers University. He received the M.S. degree from Rutgers University in 1950. In 1950 he joined the RCA Laboratories in Princeton, N. J. where he has been working in the field of broadcast and television receivers. Mr. Freedman is a member of the Institute of Radio Engineers, Phi Kappa Phi, Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.

WILLIAM J. GREIG graduated Fordham University in 1953 with a B.S. in Physics. He joined RCA as a specialized trainee following graduation and was assigned to the Semiconductor Department of the Electron Tube Division, Harrison, N. J. After serving with the U.S. Air Force from 1954 to 1956, he returned to RCA and is presently assigned to the Materials Service Laboratory. He has been concerned primarily with diffusion process developments for both germanium and silicon, in particular the germanium drift-field transistors and both high-frequency and power silicon types. Mr. Greig has taken graduate courses in physics at Stevens Institute and in statistics at Rutgers University. He is a member of the Institute of Radio Engineers.

J. HILIBRAND (see *RCA Review*, Vol. XXI, No. 2, June 1960, p. 294.)

FRANCIS R. HOLT received the B. Sc. degree in Mathematics and Physics in 1941 and the B. Sc. Special Physics degree in 1942, from London University. From 1942 to 1945 he worked in radar and counter-measures circuitry at R.A.E and T.R.E., England. He was also engaged in technical intelligence work with the Royal Air Force. From 1946 to 1956 he was with the British Atomic Energy Department, working mainly on alpha counters, transistors, and technical policy. From 1956 to 1958 he was with Philco Corporation, Leeds & Northrup, and the University of Pennsylvania. He joined the RCA Laboratories in October 1958. He is a Fellow of the Physical Society, an Associate of the Institute of Physics and an Associate Member of the Institution of Electrical Engineers, all of London, England.

RICHARD E. HONIG received the B.S. degree in Electrical Engineering from Robert College, Istanbul, Turkey, in 1938; the M.S. degree in 1939, and the Ph.D. degree in Physics in 1944, both from the Massachusetts Institute of Technology. In 1940-1941, he was instructor of Physics at Bluffton College, Ohio. He taught and did research work at M.I.T. as Research Assistant from 1941 to 1944, and as Research Associate from 1944 to 1946. He was at the Socony Vacuum Research Laboratories, Paulsboro, N. J., from 1946 to 1950, working mainly in the field of mass spectrometry. Since 1950, he has been at RCA Laboratories, Princeton, N. J., engaged in fundamental research in solid-state physics. In 1955/6, on a year's leave of absence from RCA Laboratories, he studied vaporization phenomena at the University of Brussels, Belgium. Dr. Honig is a member of Sigma Xi and the American Physical Society.

HARVEY O. HOOK received the B.A. degree with Chemistry major from Elon College in 1947, the B.E.E. degree from North Carolina State College in 1949, and the M.S.E.E. degree from North Carolina State College in 1950. Since 1950 he has been with RCA Laboratories Division. Mr. Hook is a Senior Member of the Institute of Radio Engineers and an Associate Member of Sigma Xi.

JOHN H. O'CONNELL received the B.S. degree in Electrical Engineering in 1954 from Fournier Institute of Technology, Lemont, Illinois. In the same year, he joined the technical staff of RCA Laboratories. He has been engaged in work on transistor circuit applications to radio and television and developing stereo broadcasting systems. Mr. O'Connell is a member of the Institute of Radio Engineers and the Audio Engineering Society.

JOSEPH O. PREISIG was graduated as a diploma engineer from Technical University, Berlin, Germany. Mr. Preisig worked for Telefunken Laboratories in Berlin from 1937 to 1945. From 1946 to 1948, he was associated with the U.S. Army Signal Corps and in 1949 he was an assistant to Professor Leithauser at the Technical University in Berlin. Mr. Preisig came to the United States in 1950 where he has been associated with the RCA Laboratories since 1951.

ROLAND N. RHODES received the B.E.E. degree from the City College of New York in 1944 and entered the U.S. Navy upon graduation. In 1948, he joined the research staff of RCA Laboratories at Princeton, N. J., where he has worked on various problems associated primarily with the field of home entertainment instruments such as radio and black-and-white, color, and transistor television. He is a member of IRE Subcommittee 17.10 on automatic frequency control.

M. J. SCHINDLER received the M.S. degree in Electronics in 1951 from the Institute of Technology of Vienna, Austria. He remained there in the Department of Metal Physics until 1954, under sponsorship of the Austrian Ministry of Energy, and in 1953 received a doctorate. From 1954 to 1957, he was with the Tungsram-Watt tube plant in Vienna, first as a development engineer and later as Head of Quality Control, working on aspects of receiving tube manufacture. In 1957-58 he was at the Wright Air Development Center, Dayton, Ohio, investigating the a-c magnetization process. In September 1958, he joined the Chemical and Physical Laboratory of the RCA Microwave Division at Harrison, where he has worked on a number of physical problems related to the design of traveling-wave tubes, as well as thin magnetic films and a variety of technological investigations.

CARL F. STOCKER received the B.S. degree in Physics from Loyola University, Chicago in 1954. Before completing his graduate studies, he was employed at the Hawthorne plant of the Western Electric Company. In 1956, he received master's degrees in both Physics and Chemistry from the University of Illinois, where he held a teaching assistantship. In 1956, he joined RCA Laboratories, where he carried out research on phosphors. More recently he has been concerned with the design, fabrication, and testing of semiconductor devices, in particular of variable-capacitance and tunnel diodes. Mr. Stocker is a member of the Electrochemical Society.

HAROLD S. VELORIC received a B.A. degree from the University of Pensylvania in 1951, a Ph.D. degree in 1954 from the University of Delaware. From 1951 to 1954, he was a research fellow concerned with the synthesis and analysis of boron and silicon compounds. He joined Bell Telephone Laboratories in 1954, and the RCA Semiconductor Division, Sommerville, N. J., in 1958. He has since been associated with the development of several classes of silicon diodes, including power rectifiers and voltage reference and computer diodes. He is presently concerned with the analysis and development of high-frequency germanium and silicon transistors. Dr. Veloric is a member of the American Chemical Society and the Electro-chemical Society.

HERBERT J. WOLKSTEIN received the B.S. degree in Electrical Engineering from Newark College of Engineering and is presently working toward an MS degree in the same field. Before joining RCA, he was employed at the Research Laboratories of National Union Electric Corporation as Project Engineer on the design of special-purpose beam-deflection and computer tubes. He joined the Microwave Tube Engineering activity of the RCA Electron Tube Division in 1955 where he has been engaged in the development of traveling-wave tubes. He is currently Engineering leader of the traveling-wave-tube group and has written a number of articles pertaining to this work. Mr. Wolkstein is a member of the Institute of Radio Engineers.