

RCA REVIEW

a technical journal

**RADIO AND ELECTRONICS
RESEARCH • ENGINEERING**

VOLUME XX

JUNE 1959

NO. 2

RADIO CORPORATION OF AMERICA

DAVID SARNOFF, *Chairman of the Board*

FRANK M. FOLSOM, *Chairman of the Executive Committee*

JOHN L. BURNS, *President*

E. W. ENGSTROM, *Senior Executive Vice-President*

DOUGLAS H. EWING, *Vice-President, Research and Engineering*

JOHN Q. CANNON, *Secretary*

ERNEST B. GORIN, *Vice-President and Treasurer*

RCA LABORATORIES

J. HILLIER, *Vice-President*

RCA REVIEW

C. C. FOSTER, *Manager*

R. F. CIAFONE, *Administrator*

PRINTED IN U.S.A.

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

RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH • ENGINEERING

Published quarterly by

RCA LABORATORIES

in cooperation with all subsidiaries and divisions of
RADIO CORPORATION OF AMERICA

VOLUME XX	JUNE 1959	NUMBER 2
-----------	-----------	----------

CONTENTS

	PAGE
Thermionic Emitters	191
L. S. NERGAARD	
Analysis of a Four-Terminal Parametric Amplifier	205
K. K. N. CHANG	
Surface-Immune Transistor Structure	222
H. NELSON	
Semiconductor Diodes in Parametric Subharmonic Oscillators	229
J. HILIBRAND AND W. R. BEAM	
The Propagation of Perturbations along Magnetically Focused Electron Beams	254
F. PASCHKE	
Stabilization of Transistor Gain over Wide Temperature Ranges	284
R. A. SCHMELTZER	
Quality-Control Determinations of the Screen Persistence of Color Picture Tubes	293
J. M. FORMAN AND G. P. KIRKPATRICK	
Microwave Propagation over Rough Surfaces	308
M. P. BACHYNSKI	
The Effect of Several Variables on Phosphor-Dot Size in Color Picture Tubes	336
N. R. GOLDSTEIN	
Reduction of Co-Channel Interference by Precise Frequency Control of Television Picture Carriers	349
W. L. BEHREND	
RCA TECHNICAL PAPERS	365
AUTHORS	368

© 1959 by Radio Corporation of America
All rights reserved

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

Chairman

R. S. HOLMES
RCA Laboratories

A. A. BARCO
RCA Laboratories

M. C. BATSEL
Defense Electronic Products

G. L. BEERS
Radio Corporation of America

G. H. BROWN
Radio Corporation of America

I. F. BYRNES
Industrial Electronic Products

D. D. COLE
RCA Victor Television Division

O. E. DUNLAP, JR.
Radio Corporation of America

E. W. ENGSTROM
Radio Corporation of America

D. H. EWING
Radio Corporation of America

A. N. GOLDSMITH
Consulting Engineer, RCA

A. L. HAMMERSCHMIDT
National Broadcasting Company, Inc.

E. W. HEROLD
RCA Laboratories

J. HILLIER
RCA Laboratories

D. D. HOLMES
RCA Laboratories

C. B. JOLLIFFE
Defense Electronic Products

E. A. LAPORT
Radio Corporation of America

C. W. LATIMER
RCA Communications, Inc.

H. W. LEVERENZ
RCA Laboratories

G. F. MAEDEL
RCA Institutes, Inc.

H. F. OLSON
RCA Laboratories

R. W. PETER
RCA Laboratories

D. S. RAU
RCA Communications, Inc.

D. F. SCHMIT
Radio Corporation of America

G. R. SHAW
Electron Tube Division

L. A. SHOTLIFF
RCA International Division

S. STERNBERG
Astro-Electronic Products Division

I. WOLFF
RCA Laboratories

Secretary

C. C. FOSTER
RCA Laboratories

REPUBLICATION AND TRANSLATION

Original papers published herein may be referenced or abstracted without further authorization provided proper notation concerning authors and source is included. All rights of republication, including translation into foreign languages, are reserved by RCA Review. Requests for republication and translation privileges should be addressed to *The Manager*.

THERMIONIC EMITTERS*

BY

L. S. NERGAARD

RCA Laboratories,
Princeton, N. J.

Summary—A vacuum tube is considered as a chemical system undergoing continuous reaction. It is proposed that the cathode reacts with everything in the tube via the gas phase. If the tube contains reducing agents in the form of getters or electrodes of reducing metals such as titanium, these reducing agents help maintain cathode activity via the gas phase. Thus getters not only prevent poisoning but contribute to keeping the cathode in the reduced state necessary for emission.

INTRODUCTION

THE relation between the performance of a thermionic emitter and the ambient gases which surround it has long been recognized. The classic 1913 paper of Langmuir¹ is ample evidence of early concern with the problem. However, as vacua improved there seems to have been a tendency to regard the ambient gas situation as in hand. It is true that tubes were, and are, checked for gas current, but such measurements are made more as a measure of what resistance can be tolerated in the grid circuit without the tube running away than as a measure of the probable performance of the cathode in tubes. Of course, it has been recognized that certain gases poison a cathode. The poisoning effects of oxygen were beautifully demonstrated by Becker in the case of the oxide cathode.^{2,3} It has also been recognized that gases released from electrodes and the envelope of a tube under electron bombardment poison a cathode. At one stage, poisoning was regarded as responsible for most of the peculiarities of behavior of the oxide cathode. For example, during the war pulsed emissions of 50-100 amperes per square centimeter were achieved with oxide cathodes under short-pulse conditions. As pulse lengths were increased, the emission current was observed to decay. The decay was widely

* Manuscript received February 13, 1959.

¹ I. Langmuir, "The Effect of Space Charge and Residual Gases on Thermionic Currents in High Vacuum," *Phys. Rev.*, Vol. 2, p. 11, December, 1913.

² J. O. Becker, "Phenomena in Oxide-Coated Filaments," *Phys. Rev.*, Vol. 34, p. 1323, November, 1929.

³ J. A. Becker and R. W. Sears, "Phenomena in Oxide-Coated Filaments II," *Phys. Rev.*, Vol. 38, p. 2193, December 15, 1931.

attributed to poisoning and there were some who felt that if poisoning could be eliminated, the oxide cathode could deliver 50 amperes per square centimeter continuously. There were also skeptics who wondered how a cathode reactivated itself in microseconds under pulse conditions when it required a major effort to reactivate a cathode poisoned under d-c conditions.

At about this time, solid-state physics caught the attention of those working on the oxide cathode and there followed an intensive study of the oxide cathode as an extrinsic semiconductor. The energy-level structure of BaO was determined to a degree adequate for a reasonably detailed picture of the performance of the oxide cathode, although the picture cannot yet be described as complete, or at least not wholly satisfactory. Many of the phenomena previously attributed to poisoning found their explanation in the internal behavior of the semiconducting cathode.⁴ Thus emphasis on poisoning and other effects of ambients declined and almost dropped out of sight. Solid-state physics ruled supreme. Of course, there were a few unresolved problems; it was not certain that the F-centers studied so thoroughly by Sproull and his students were in fact responsible for the emission of the oxide cathode.⁵⁻⁷

To resolve this unanswered question, Matheson, Nergaard, and Plumlee undertook to activate an oxide cathode by depositing alkaline-earth metals on an oxide cathode using a mass-spectrometer as a controllable, pure source of the metals.⁸ In the course of these experiments, the entire mass spectrum was scanned to determine the nature and extent of the background. It was soon found that the background depended strongly on the operating conditions of the cathode which was to be the object of study. This again focused attention on the role of ambients, and ambients became the subject of a detailed subsequent investigation.⁹

⁴ For a review of this work up to 1956, see L. S. Nergaard, "Electron and Ion Motion in Oxide Cathodes," *Halbleiter Probleme III*, F. Vieweg and Sohn, Braunschweig, Germany.

⁵ W. C. Dash, "Optical Absorption and Photoconduction in the Visible and Near Infra Red in Single Crystals of BaO," *Phys. Rev.*, Vol. 92, p. 68, October 1, 1953.

⁶ R. L. Sproull, R. S. Bever, and H. Liebowitz, "Oxygen Vacancies in Barium Oxide," *Phys. Rev.*, Vol. 92, p. 77, October 1, 1953.

⁷ Cornelis Timmer, "The Density of Color Centers in Barium Oxide as a Function of the Vapor Pressure of Barium," *Jour. Appl. Phys.*, Vol. 28, p. 495, April, 1957.

⁸ R. M. Matheson, L. S. Nergaard and R. H. Plumlee, "Activation of an Oxide Cathode by Deposition of Alkaline Earth Metals via a Mass Spectrometer," *RCA Review*, Vol. 18, p. 388, September, 1957.

⁹ R. H. Plumlee, "Electrolytic Transport Phenomena in the Oxide Cathode," *RCA Review*, Vol. 17, p. 190, June, 1956.

The specific results of this investigation are important and they will be used now and again to illustrate particular points in the subsequent discussion. More important than the specific results, however, is the general outcome of the investigation, namely, the realization that a vacuum tube as a whole constitutes a chemical system, all components of which react with one another so that the performance of each electrode is tempered by the presence of the others.

It is the purpose of this paper to set forth the view that a vacuum tube is an entity and not a collection of independent electrodes housed in a common envelope, that a tube is a complex chemical system, all components of which are in continuous reaction with one another, and to point out some of the consequences of this view as regards the performance of thermionic emitters.

A THERMIONIC EMITTER AS A REDUCING AGENT

If a vacuum tube is to be considered as a chemical system undergoing reaction, it is necessary to have some measure of the reactivity of the reactants. The chemical activity is such a measure. Before defining the chemical activity, it is necessary to define reducing and oxidizing agents. Pauling has provided very simple definitions. He states that, "An atom, molecule or ion which takes up electrons is called an oxidizing agent and one which liberates electrons is called a reducing agent" and "Oxidation is the removal of electrons from an atom or group of atoms. Reduction is the addition of electrons to an atom or group of atoms."¹⁰ Thus it is apparent that a thermionic emitter is a reducing agent and the emission of electrons is an oxidation process.¹¹

To come to grips with the definition of chemical activity which measures the "strength" of a reducing or oxidizing agent, consider a reaction between an oxidizing agent, A, and a reducing agent, B, in which a molecules of A and b molecules of B react to give c molecules of C and d molecules of D. Then the reaction is written



The mass action law states that when the reaction has reached equilibrium

¹⁰ Linus Pauling, *General Chemistry*, W. H. Freeman and Co., San Francisco, 1947.

¹¹ R. H. Plumlee, "The Electron Donor Centers in the Oxide Cathode," *RCA Review*, Vol. 17, p. 231, June, 1956.

$$\frac{[A]^a [B]^b}{[C]^c [D]^d} = K,$$

where the bracketed quantities are the chemical activities and K is the equilibrium constant. The chemical activities are usually defined with respect to the gas phase, i.e.,

$$a_A \equiv [A] = \frac{n_A}{N_c},$$

where n_A is the concentration in the gas phase and N_c is the "density of states" now in the gas phase;

$$N_c = 2 \left(\frac{2\pi m k T}{h^2} \right)^{3/2}.$$

Thus the activities are proportional to the partial pressures. The equilibrium constant is given by

$$K = e^{\Delta H/kT},$$

where ΔH is the heat of reaction.

In the thermionic emitter, the reactant of interest is the electron. Then density of electrons in the vapor phase is

$$n = N_c e^{-e\phi/kT},$$

where ϕ is the Fermi energy measured with respect to the vacuum level, i.e., the work function. Hence, the chemical activity of the electrons is

$$a_e = \frac{n}{N_c} = e^{-e\phi/kT}.$$

The thermionic emission of the cathode, then, is proportional to the activity of the electrons. This is obvious when the emission current is written out explicitly;

$$\begin{aligned} j &= \frac{1}{4} n e \bar{v} \\ &= \frac{1}{4} N_c e \bar{v} a_e = AT^2 a_e, \end{aligned}$$

in which

$$\begin{aligned} \bar{v} &= \text{the mean thermal velocity,*} \\ A &= \text{the familiar emission constant} \\ &= 120 \text{ amp cm}^{-2} (\text{°K})^{-2}. \end{aligned}$$

It is now apparent that anything that reduces the work function and increases the emission increases the chemical activity in proportion. It follows that the more active a cathode, the more easily it is poisoned (oxidized) by oxidizing agents.

CATHODE EVAPORATION

There is a general feeling that a cathode should evaporate electrons copiously and should evaporate no other of its constituents. There is a certain logic in this because evaporation dissipates the cathode and some cathode constituents evaporate, condense on other electrodes and become the source of poisoning reactants under electron bombardment. However, the picture is not all black because, as will be pointed out subsequently, the evaporation of at least one constituent of the oxide cathode enhances the emission.

Whatever the consequences of evaporation, it exists and deserves a brief scrutiny. To this end, consider a constituent $A(g)$ of the gaseous ambient with a concentration n_A and a constituent $B(s)$ of the solid with a concentration $N(s)$ which react to form a solid product of concentration $n(s)$. The reactant $B(s)$ may be an atom, molecule or just an interstitial site in the solid which is "consumed" when $AB(s)$ is formed. In any event, the reaction is



and the equilibrium equation is

$$[A(g)] [B(s)] = K [AB(s)],$$

or

$$n(g) = n(s) \frac{N_c}{N(s)} e^{-\Delta H/kT},$$

where

$$\Delta H = \text{the heat of sublimation of A.}$$

The equilibrium equation says that any constituent of a cathode also

* The mean thermal velocity *normal to the emitting surface* is $\frac{1}{4} \bar{v}$.

appears in the gas phase. Conversely, it says that to maintain any constituent in the solid there must be an appropriate partial pressure of the constituent in the gas phase. If the heat of sublimation is high, the partial pressure required may be very low indeed, but it remains finite as long as the heat of sublimation is finite.

This consideration has brought about an embarrassing situation. For about 30 years, the activity of the oxide cathode has been ascribed to excess barium. This thesis received a rude shock when Timmer⁷ measured the excess barium content of barium oxide as a function of the partial pressure of Ba over the BaO. He found that a partial pressure of Ba in excess of 10^{-6} mm Hg was required to support the concentration of excess Ba required for an active cathode. Hence, if excess Ba were responsible for the activity of an oxide cathode, the cathode should become inoperable in a tube pumped to 10^{-7} mm of Hg or better. This dilemma is not finally resolved, but a possible way out will be touched upon later. It must be remembered that an excess of one constituent of the host crystal is not the only way to raise the Fermi level and achieve a reduced state; arsenic will raise the Fermi level in germanium.

So far only equilibria have been considered. Under equilibrium conditions as much of a constituent returns to the host via thermal deposition as evaporates so the net rate of evaporation is zero. The equilibrium picture is unrealistic — a vacuum tube is not a system in thermodynamic equilibrium; the system is not all at one temperature nor have all reactions run to completion. However, the fact that a tube will maintain its characteristics over thousands of hours is ample evidence that the system is in a state of dynamic equilibrium. The fact that the system is not in thermodynamic equilibrium in no way vitiates the notion that all electrodes communicate with one another via the gas phase, it merely alters the balance between reactants and products.

A striking example of the interaction between elements of a tube has been noted by R. M. Matheson.¹² The data pertain to a twin triode which had been operated in a power supply voltage regulator for some years. In the course of this use it had developed a high interface resistance, of the order of 400 ohms. To determine if there was any interaction between triode units, triode Unit I was operated at 1.5 milliamperes until the cathode resistance became stable, with zero plate current in Unit II. Then, plate voltage was applied to the second unit for a period of 8 minutes. During this interval, the plate current in Unit II rose from 16 to 22 milliamperes, indicating that

¹² Private communication.

it was being activated by drawing current. During the interval, the contact potential and cathode resistance of Unit I were monitored in a cathode resistance bridge. The results are shown in Figure 1. It will be noted that the cathode resistance of Unit I increased about 15 per cent and that the contact potential dropped about 0.2 volt, both indicating poisoning. When the plate voltage on Unit II was turned off, the contact potential and cathode resistance of Unit I drifted towards their initial values. Because this particular tube was a tube discarded for failure, the behavior is extreme. However, the difference between a good tube and a bad tube is only one of degree, so

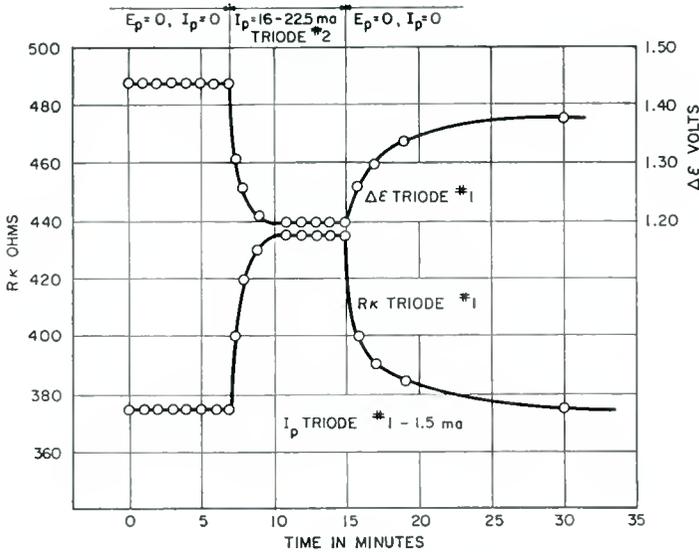


Fig. 1—The effect on the contact potential and cathode resistance of one unit of a twin triode which had failed when current was drawn in the second unit (after Matheson).

the bad tube just makes the interaction between electrodes more apparent.

Before proceeding to the means, obvious and less obvious, used to maintain an emitter in a reduced state in the face of inevitable oxidizing ambients, it is of interest to examine the dynamic equilibrium of a particular component of the ambient over an oxide cathode when energy is supplied to the system from an external source. A case in point is the behavior of the partial pressure of CO_2 as measured by Plumlee, who used a mass spectrometer to measure the constitution and partial pressures of the ambient gases in a diode.⁹

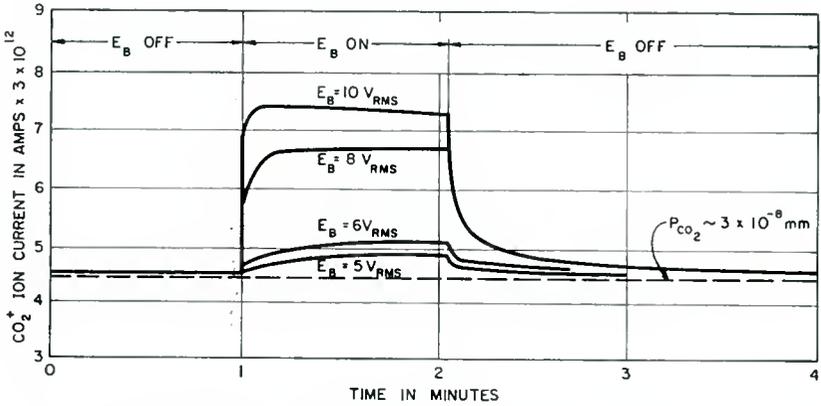


Fig. 2—The evolution of carbon dioxide from an oxide-coated cathode in a diode as a function of time, with anode voltage as a parameter (after Plumlee).

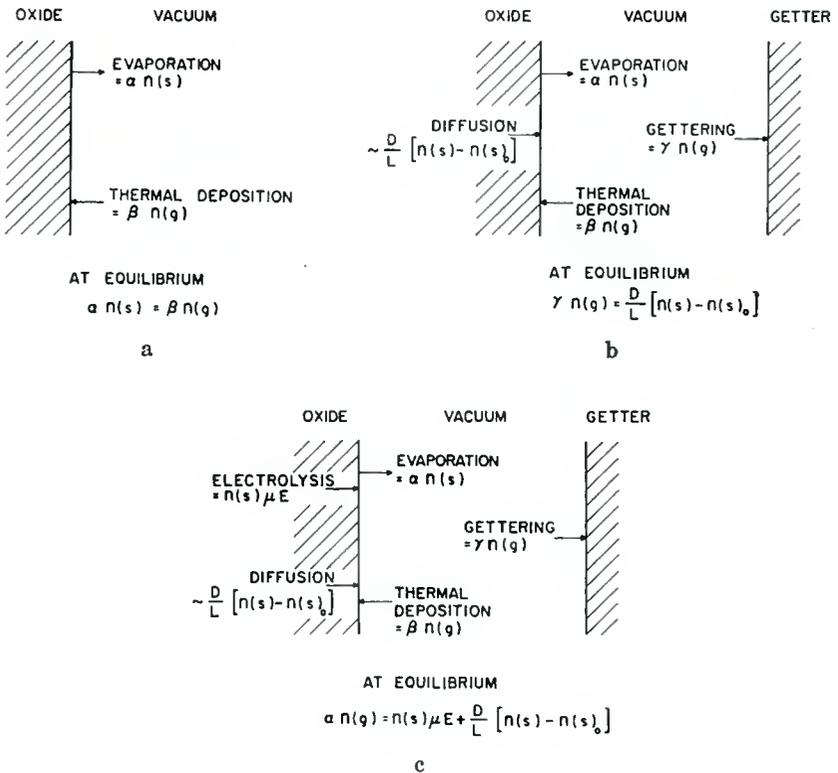


Fig. 3—Cathode-vacuum interface.

Figure 2 shows the behavior of the CO_2 peak when anode voltage was applied to the diode. The partial pressure rose to a new equilibrium value and returned to the original value when the voltage was removed. It will be noted that the equilibrium value reached with voltage *on* increased as the voltage was increased but in every case returned to the initial value when anode voltage was removed. It appears that adding electrical energy to the system changes the balance between the solid and gas phases of CO_2 .

The result may be described in physical terms as follows: Consider the interface between the cathode and vacuum as shown in Figure 3a. CO_2 evaporates into the vacuum in proportion to its concentration at the surface. CO_2 is restored to the surface by thermal deposition, in proportion to its concentration in the gas phase. If no CO_2 is removed from the vapor phase by pumping or gettering, the rate of deposition equals the rate of evaporation. However, if CO_2 is removed from the gas phase by pumping or gettering as in Figure 3b, there is a net rate of evaporation and the surface concentration drops. As the surface concentration drops, diffusion from the bulk material starts and replenishes the surface and when equilibrium is reached, the diffusion current just balances the rate of removal by pumping or gettering. However, the vapor-phase density is lowered by the pumping or gettering. If, in addition, a field is applied, as in Figure 3c, the surface concentration is increased by electrolysis, the vapor phase density increases and so does the thermal deposition and removal by gettering. When equilibrium is reached so that the sum of the diffusion and electrolytic currents just equals the rate of removal by gettering or pumping, the concentration in the vapor phase is higher than in the case of Figure 3b. A detailed computation based on this model shows that the partial pressure should rise roughly in proportion to the voltage for low voltages and then should tend to saturate.

It is noted in passing that Plumlee found that O_2 and H_2 behaved in the manner of CO_2 but that H_2O showed the inverse behavior. This combination of behavior is discussed later.

REDUCING POTENTIAL

Because the ambient gases in a vacuum tube are in part oxidizing, some provision must be made to keep a cathode in the reduced state. This is done by providing a source of reducing potential. As noted earlier, the system is not in a state of thermodynamic equilibrium, but in a state of dynamic equilibrium. Hence the hot cathode tends to evaporate reducing constituents to cooler electrodes where they are less effective in maintaining cathode activity. To keep the cathode active,

therefore, reducing potential must be supplied to the cathode continuously. Thus all useful cathodes are in a sense dispenser cathodes.¹³

In the case of pure metallic cathodes such as tungsten filaments, the metal provides its own reducing potential. It reduces thoria to produce thorium; the thorium diffuses to the surface at a rate which just matches the loss by evaporation at the normal operating temperature and thus maintains the surface in a reduced state.

Hull's dispenser cathode and the L-cathode are familiar cathodes in which reducing potential is dispensed from reservoirs to maintain a high degree of activity for extended periods.¹⁴⁻¹⁶

In the case of the ordinary oxide cathode, it is now recognized that "impurities" in the base metal constitute an important source of reducing potential.¹⁷⁻²¹ Many of the impurities in the base metal are reducing agents, including iron, manganese, magnesium, titanium, silicon and carbon. As is well known, all are not undisguised blessings. Silicon is a very strong reducing agent but is responsible for the barium-orthosilicate interface layer which has been the source of so much difficulty. Hence it is now kept at a very low concentration in cathode base metals. Studies of the diffusion rates and reactivities of various reducing additives to nickel are under way and in the course of time base metals will be "tailored" to meet specific requirements of activity and life.

While reducing additives dispensed by the base metal are useful in effecting the initial reduction of the oxide during activation and in

¹³ L. S. Nergaard, "The Physics of the Cathode," *RCA Review*, Vol. 18, p. 486, December, 1957.

¹⁴ A. W. Hull, "The Dispenser Cathode," *Phys. Rev.*, Vol. 56, p. 86, July, 1939.

¹⁵ E. S. Rittner, R. H. Ablert and W. C. Rutledge, "Studies of the Mechanism of Operation of the L-Cathode I," *Jour. Appl. Phys.*, Vol. 28, p. 156, February, 1957.

¹⁶ W. C. Rutledge and E. S. Rittner, "Studies of the Mechanism of Operation of the L-Cathode II," *Jour. Appl. Phys.*, Vol. 28, p. 167, February, 1957.

¹⁷ H. E. Kern and R. T. Lynch, "Initial Emission and Life of a Planar-Type Diode as Related to the Effective Reducing Agent Content of the Cathode Nickel," *Phys. Rev.*, Vol. 82, p. 574, May, 1951.

¹⁸ E. S. Rittner, "A Theoretical Study of the Chemistry of the Oxide Cathode," *Philips Research Reports*, Vol. 8, p. 184, June, 1953.

¹⁹ R. W. Peterson, "Activation Diffusion and Thermionic Emission in Oxide Coated Cathodes," *Phys. Rev.*, Vol. 99, p. 1651, September, 1955.

²⁰ J. J. Lander, H. E. Kern and A. L. Black, "The Solubility and Diffusion Coefficient of Carbon in Nickel," *Phys. Rev.*, Vol. 85, p. 389, January, 1952.

²¹ R. W. Peterson, D. E. Anderson, and W. G. Shepherd, "Influences of the Cathode Base on the Chemical Activation of Oxide Cathodes," *Jour. Appl. Phys.*, Vol. 28, p. 22, January, 1957.

maintaining activity during life, they are not necessary, either to activate or to maintain activity. This was strikingly demonstrated by Moore and Allison who achieved active cathodes supported on magnesium oxide.²² The mechanism by which this can occur was first noted by J. A. Becker, later by Plumlee and Smith and was studied in some detail by Plumlee.^{2,9,23} In the discussion of the field-dependent evaporation of CO_2 it was noted that O_2 also displays a field-dependent

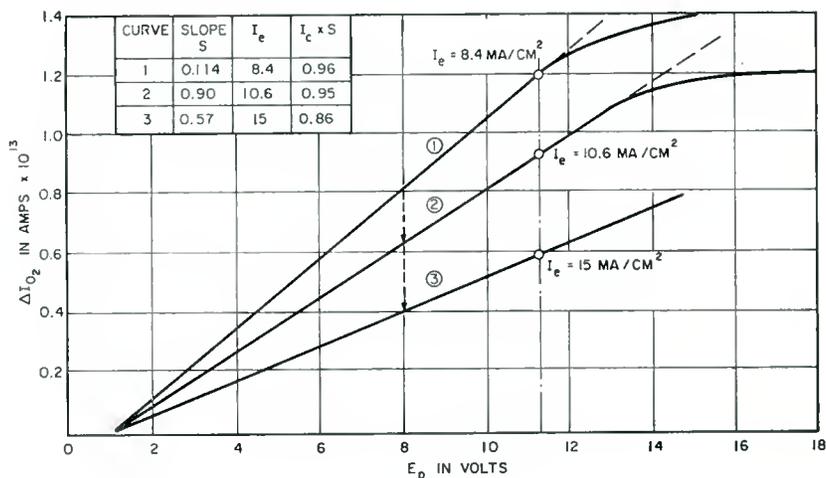


Fig. 4—The evolution of oxygen from an oxide-coated cathode in a diode as a function of anode voltage, with the state of activity of the cathode as a parameter (after Plumlee). The inset shows the slope of the curves (s) in amperes volt⁻¹ $\times 10^{13}$, the activity in terms of the anode current (I_e) in milliamperes at 11.3 volts anode voltage and the product of s and I_e . A constant product indicates a rate of evolution inversely proportional to the activity.

evaporation rate. This is shown in Figure 4. With the cathode in a relatively inactive state, the dependence of oxygen evolution on anode voltage is as shown by the upper curve. The transient behavior is similar to the CO_2 case shown in Figure 2. If the anode is held at a fixed voltage, say 8 volts, the cathode activates and at the same time the rate of O_2 evolution drops to a lower rate as indicated by the dotted line. The dependence on voltage is then as shown by the middle curve. Further activation by operation at a fixed voltage results in

²² G. E. Moore and H. W. Allison, "Emission of Oxide Cathodes Supported on a Ceramic," *Jour. Appl. Phys.*, Vol. 27, p. 1316, November, 1956.

²³ R. H. Plumlee and L. P. Smith, "Mass Spectrometric Study of Solids I. Preliminary Study of Sublimation Characteristics of Oxide Cathode Materials," *Jour. Appl. Phys.*, Vol. 21, p. 811, August, 1950.

curves of lower and lower slope until at high activities, the change in O_2 pressure may become imperceptible. For the curves of Figure 4, the rate of evolution of O_2 is inversely proportional (approximately) to the activity of the cathode as may be seen by multiplying the slope of each curve by the corresponding electron current (measured with 11.3 volts on the anode) which serves as a measure of the activity. As noted earlier, the transient behavior is the same as for the case of CO_2 and the explanation is the same. Whereas, the role of CO_2 in an oxide cathode is uncertain, it is clear that the removal of oxygen is a reduction process which activates the cathode. In this case, the source which supplies the anode power is the source of reducing potential.

It was noted earlier in this section that a cathode can evaporate reducing constituents to cooler electrodes "where they are less effective in maintaining cathode activity." It is now clear, not only that they can, but how they do maintain activity. To be specific, consider reduction by elimination of O_2 from the cathode via the gas phase. As was noted in consideration of the CO_2 behavior, pumping or gettering reduces the partial pressure of the ambient over the surface and reduces the net concentration within the cathode. Thus any material which removes O_2 from the gas phase also reduces the concentration of O_2 in the cathode, i.e., it chemically reduces the oxide and enhances the emission. Thus the concept of a getter serving to prevent poisoning is incomplete. It does more than prevent poisoning, it produces and maintains cathode activity.

It is now clear why getters have been so useful and why titanium, with its high reducing power, is finding favor as a material for grids and anodes.

CAN AN OXIDE CATHODE EXIST IN A VACUUM?

The interesting question of whether or not an oxide cathode can exist in a vacuum was posed by Plumlee as a result of the following considerations, most of which have been discussed above:⁹

1. Timmer's measurements of the Ba partial pressure required to maintain excess Ba in a cathode and its implications with respect to cathode activity.

2. The voltage dependence of the evaporation of O_2 .

3. The voltage dependence of the evaporation of H_2 .

4. The voltage dependence of the evaporation of H_2O . This has not been discussed and needs little discussion at this point. The behavior is as shown in Figure 5. It will be noted that the oxide "soaks

up" water with field applied and releases it when the field is reversed. The usual reaction of water with BaO is



However, items 2, 3, and 4 above, together with other cogent considerations, suggested to Plumlee that at very low water concentrations another reaction is possible:

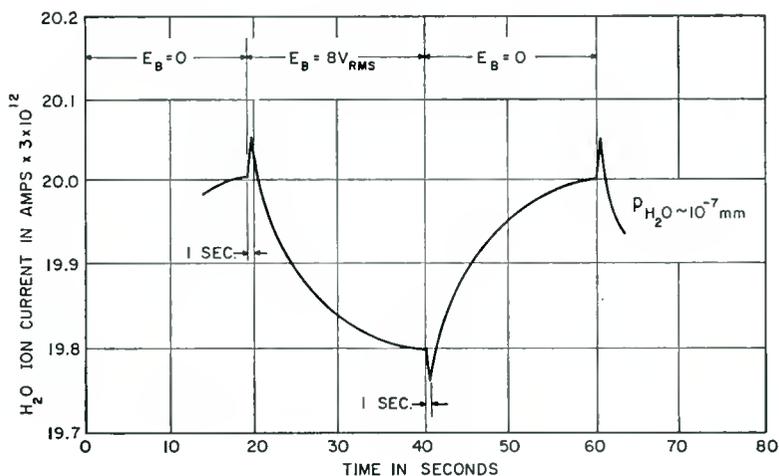


Fig. 5—The rate of evolution of oxygen from an oxide-coated cathode in a diode as a function of time when anode voltage is applied and removed (after Plumlee).

According to this reaction, when water is absorbed, hydrogen and oxygen are released and a new kind of hydroxyl ion forms.¹¹ This hydroxyl ion has an extra electron to charge balance the product and Plumlee suggested that the extra electron is in fact responsible for the activity of the cathode. Note that the reaction makes the behaviors of H₂, O₂ and H₂O as observed with the mass spectrometer, interdependent. It would have been elegant indeed if it had been possible to establish quantitative relations between the partial pressures of H₂O, H₂ and O₂ to check the reaction. However, to calibrate a spectrometer for such an experiment would have been a formidable task, considering the mass discrimination of the spectrometer and the different ionization cross-sections of the molecules involved.

The presence of the hydroxyl ion might also be verified by preparing an oxide cathode of Ba ¹³⁸ and O¹⁶, activating the cathode, and then looking for the hydrogen ion by a proton resonance measurement. This experiment has yet to be performed. Certainly some definitive experiment should be performed to prove or disprove the existence of this ion and its role in the oxide cathode if it exists.

If the hydroxyl ion is in fact the active electron donor which produces the reduced state required for electron emission, an active cathode cannot exist without a small partial pressure of water (Plumlee calculates the required partial pressure to be about 10^{-11} mm Hg.). Hence, an active oxide cathode cannot exist in a true vacuum if Plumlee's hypothesis is correct.

While the question of whether a cathode can exist in a vacuum may be amusing, it is purely academic. Thermodynamics makes it certain that a gas phase will exist and, in view of the arguments set forth above regarding reducing agents which maintain cathode activity via the gas phase, a perfect vacuum would not be desirable if it were possible.

ANALYSIS OF A FOUR-TERMINAL PARAMETRIC AMPLIFIER*

BY

KERN K. N. CHANG

RCA Laboratories,
Princeton, N. J.

Summary—The operation of conventional two-terminal parametric amplifiers is usually handicapped by the interaction between input and output. The difficulty can be eliminated by constructing a four-terminal parametric amplifier which consists of three cascaded stages, a converter, an amplifier, and a modulator, each involving a nonlinear reactance as the interacting element. The input can thus be separated from the output through the frequency conversion among the stages so that no circulators or isolators are needed. The amplifier can be pumped by a lower-frequency pump. Analysis shows that the bandwidth and noise factor of the amplifier can be optimized by independently adjusting the characteristic of the three cascaded stages. The analytical results have been verified by an experimental four-terminal amplifier which operates at a signal frequency of 214 megacycles with an optimized noise factor of 2.5 decibels.

INTRODUCTION

IN A RECENT paper¹ a ferromagnetic amplifier was proposed in which the pump frequency is lower than that of the microwave signal frequency. The scheme involves the use of two idling circuits and one pump. It thus differs from another amplifier scheme, also using lower-frequency pumping, which contains only one idling circuit and one or two pumps.² The difference between these two schemes lies in the order of nonlinearity with which the amplifier works. Essentially, the former requires a quadratic order in the reactance sample, while the latter uses a cubic order.

The present paper puts forward a four-terminal parametric amplifier which also uses a lower-frequency pump, and consists of three cascaded stages: (1) a converter which converts the incoming high-frequency signal into a low-frequency "idling" signal, (2) an amplifier which amplifies the idling signal, and (3) a modulator to convert back the idling signal to an amplified incoming signal. All three stages are

* Manuscript received January 26, 1959.

¹ C. L. Hogan, R. L. Jepen and P. H. Vartanian, "New Type of Ferromagnetic Amplifier," *Jour. Appl. Phys.*, Vol. 29, p. 422, March, 1958.

² S. Bloom and K. K. N. Chang, "Parametric Amplification Using Lower-Frequency Pumping," *Jour. Appl. Phys.*, Vol. 29, p. 594, March, 1958.

pumped by a common pump, the frequency of which is lower than that of the signal. Although the method employed here has some resemblance to that proposed by Hogan et al,¹ there is one striking difference.

In the Hogan-Jepson-Vartanian amplifier, regeneration is employed in the signal circuit so that the amplifier, like ordinary lumped-circuit parametric amplifiers,³ has common input and output terminals. In such two-terminal amplifier networks, circulators or isolators have to be used for low-noise amplification. The amplifier discussed in the present paper, however, has an internal isolation between the input circuit and the output circuit. The isolation results from the amplifier stage operating at a frequency different from the signal frequency; this makes possible the use of distinct and separate input and output

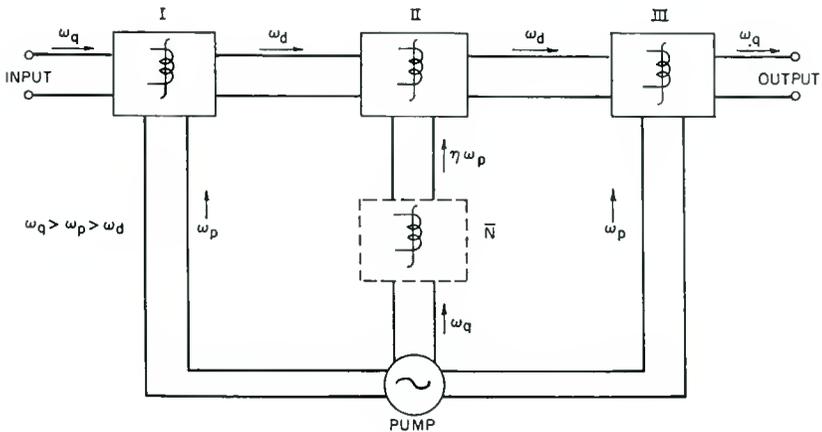


FIG. 1

Fig. 1—Four-terminal parametric amplifier.

terminals. Therefore, no additional circulators or isolators are needed in the operation of the present amplifier.

AMPLIFIER SCHEME

Consider a four-terminal amplifier system (Figure 1) which consists of three cascaded stages (I, II, III). Each stage involves a nonlinear reactance. An input signal at an angular frequency ω_q is applied to the input terminals. The signal frequency is then converted into an idling frequency, ω_d , through nonlinear interaction by a local source at an angular frequency ω_p , which is lower than ω_q . The idling signal is then amplified in stage II by the same local source used as

³M. T. Weiss, "A Solid-State Microwave Amplifier and Oscillator Using Ferrites," *Phys. Rev.*, Vol. 107, p. 317, July, 1957.

the pump. The pump frequency may be multiplied through a multiplier (shown by a dotted box) to yield an effective pump frequency of $n\omega_p$. In either case (pump frequency of ω_p or $n\omega_p$), the amplified idling signal is finally combined with the pump at stage III to form an output at the original signal frequency.

Assume that the nonlinear reactances in all stages are inductive and have a quadratic relation between the flux leakage (ϕ) through these reactances and the current (i), i.e.,

$$\phi = L_0 i - \mathcal{L} i^2, \quad (1)$$

where L_0 is the linear inductance and \mathcal{L} is the nonlinearity coefficient.

The general voltage-current equations for the signal circuit, the pump circuit, and the idling circuit which were derived in a previous paper,⁴ are then, in stage (I),

$$\begin{aligned} V_{d1} &= I_{d1} Z_{d1} - j\omega_d \mathcal{L} I_{p1}^* I_{q1}, \\ V_{q1} &= I_{q1} Z_{q1} - j\omega_q \mathcal{L} I_{p1} I_{d1}, \\ V_{p1} &= I_{p1} Z_{p1} - j\omega_p \mathcal{L} I_{d1}^* I_{q1}, \end{aligned} \quad (2)$$

in stage II,

$$\begin{aligned} V_{d2} &= I_{d2} Z_{d2} - j\omega_d \mathcal{L} I_{s2}^* I_{p2}, \\ V_{s2} &= I_{s2} Z_{s2} - j\omega_s \mathcal{L} I_{p2} I_{d2}^*, \\ V_{p2} &= I_{p2} Z_{p2} - j\omega_p \mathcal{L} I_{d2} I_{s2}, \end{aligned} \quad (3)$$

and in stage III,

$$\begin{aligned} V_{d3} &= I_{d3} Z_{d3} - j\omega_d \mathcal{L} I_{p3}^* I_{q3}, \\ V_{q3} &= I_{q3} Z_{q3} - j\omega_q \mathcal{L} I_{p3} I_{d3}, \\ V_{p3} &= I_{p3} Z_{p3} - j\omega_p \mathcal{L} I_{q3} I_{d3}^*, \end{aligned} \quad (4)$$

where $\omega_s = \omega_p - \omega_d$, and Z 's are the total impedances in the corresponding circuit.

CIRCUIT ANALYSIS

To analyze the amplifier shown in Figure 1, consider the idling circuit which links the three stages. Due to the production of a new voltage and a new impedance in the idling circuit during nonlinear

⁴ S. Bloom and K. K. N. Chang, "Theory of Parametric Amplification Using Nonlinear Reactances," *RCA Review*, Vol. XVIII, p. 578, December, 1957.

interaction in each stage, the equivalent idling circuit is shown as in Figure 2.

In the equivalent circuit, V_{d0}' is the self voltage source, if any, in the idling circuit (for example, the noise source when the noise factor is computed), and Z_{d0}' is the self circuit impedance. V_{d1}' , V_{d2}' , V_{d3}' and Z_{d1}' , Z_{d2}' , Z_{d3}' are, respectively, the voltages and the impedances which are produced by nonlinear interaction in each corresponding stage. These voltages and impedances can be computed as follows.

From Equations (2), I_{q1} and I_{p1}^* can be expressed in terms of I_{d1} ;

$$I_{q1} = \frac{V_{q1}Z_{p1} + j\mathcal{L}\omega_q V_{p1}I_{d1}}{Z_{p1}Z_{q1} + \mathcal{L}^2\omega_p\omega_q I_{d1}^2}, \tag{5}$$

$$I_{p1}^* = \frac{V_{p1}^*Z_{q1}^* - j\mathcal{L}\omega_p V_{q1}^*I_{d1}}{Z_{p1}^*Z_{q1}^* + \mathcal{L}^2\omega_p\omega_q I_{d1}^2}. \tag{6}$$

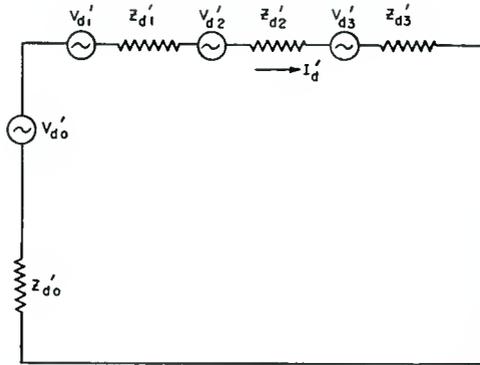


Fig. 2—Equivalent circuit of the amplifier stage.

The product of Equations (5) and (6) in the idling circuit of Equations (2) gives rise to a new voltage V_{d1}' and a new impedance Z_{d1}'

$$V_{d1}' = j\mathcal{L}\omega_d \frac{V_{p1}^*V_{q1}}{Z_{p1}^*Z_{q1}}, \tag{7}$$

$$Z_{d1}' = \mathcal{L}^2\omega_d \frac{-\omega_p Z_{p1}|V_{q1}|^2 + \omega_q Z_{q1}^*|V_{p1}|^2}{|Z_{p1}Z_{q1}|^2}, \tag{8}$$

assuming

$$|Z_{p1}Z_{q1}| \gg \mathcal{L}^2\omega_p\omega_q I_{d1}^2. \tag{9}$$

The assumption (9) is necessary in order for linear amplification, as can be shown in the derivation of Equations (7) and (8). This assumption is quite justified in view of the fact that I_{d1} , being the idling current, is a small quantity.

By the same derivation, the other new voltages and impedances are

$$V_{d2}' = j\mathcal{L}\omega_d \frac{V_{p2}V_{s2}^*}{Z_{p2}Z_{s2}^*}, \quad (10)$$

$$V_{d3}' = j\mathcal{L}\omega_d \frac{V_{p3}^*V_{q3}}{Z_{p3}^*Z_{q3}}, \quad (11)$$

$$Z_{d2}' = \omega_d \mathcal{L}^2 (\omega_p |V_{s2}|^2 Z_{p2}^* - \omega_s |V_{p2}|^2 Z_{s2}) / |Z_{p2}Z_{s2}|^2, \quad (12)$$

$$Z_{d3}' = \omega_d \mathcal{L}^2 (-\omega_p |V_{q3}|^2 Z_{p3} + \omega_q |V_{p3}|^2 Z_{q3}^*) / |Z_{p3}Z_{q3}|^2. \quad (13)$$

Using these idling voltages and impedances, Equations (4) can be rewritten as

$$\begin{aligned} V_d' &= I_d' Z_d', \\ V_{q3} &= I_{q3} Z_{q3} - j\omega_q \mathcal{L} I_{p3} I_d', \\ V_{p3} &= I_{p3} Z_{p3} - j\omega_p \mathcal{L} I_{q3} I_d'^*, \end{aligned} \quad (14)$$

where

$$V_d' = V_{d0}' + V_{d1}' + V_{d2}' + V_{d3}' \quad (15)$$

$$Z_d' = Z_{d0}' + Z_{d1}' + Z_{d2}' + Z_{d3}'. \quad (16)$$

The output signal current, I_{q3} , can be thus computed from Equations (14);

$$I_{q3} = \frac{V_{q3} + j\mathcal{L}\omega_q \frac{V_{p3}V_d'}{Z_{p3}Z_d'}}{Z_{q3}}. \quad (17)$$

In Equations (10), (11), (12), (13), and (15), V_{s2} , V_{q3} , and V_{d0} are equal to zero when amplification on the incoming signal is considered since there are no signal inputs at those corresponding frequencies. These voltages, however, would have nonvanishing values when the noise factor of the amplifier is calculated, as discussed later.

By dropping out V_{d0} , V_{s2} , V_{q3} , the output signal current becomes

$$I_{q3} = \frac{-\mathcal{L}^2 \omega_d \omega_d V_{p1}^* V_{q1} V_{p3}}{Z_{p1}^* Z_{p3} Z_{q1} Z_{q3} Z_{d0} \left[1 + \frac{\mathcal{L}^2 \omega_d \omega_q}{Z_{q1} Z_{d0}} \left| \frac{V_{p1}}{Z_{p1}} \right|^2 - \frac{\mathcal{L}^2 \omega_d \omega_s}{Z_{s2}^* Z_{d0}} \left| \frac{V_{p2}}{Z_{p2}} \right|^2 + \frac{\mathcal{L}^2 \omega_d \omega_q}{Z_{q3} Z_{d0}} \left| \frac{V_{p3}}{Z_{p3}} \right|^2 \right]}. \quad (18)$$

In Equation (18), a distortion term involving $|V_{q1}|^2$ has been ignored since it is relatively small.

At resonance, which is defined as a condition when all reactances are equal to zero, Equation (18) becomes

$$I_{q3} = \frac{-V_{q1}}{R_{q3}} \sqrt{\frac{R_{q3}}{R_{q1}}} \frac{\sqrt{\alpha\gamma}}{1 + \alpha - \beta + \gamma}, \quad (19)$$

where

$$\alpha = \frac{\mathcal{L}^2 \omega_d \omega_q}{R_{q1} R_{d0}} \left| \frac{V_{p1}}{R_{p1}} \right|^2 \quad (20)$$

$$\beta = \frac{\mathcal{L}^2 \omega_d \omega_s}{R_{s2} R_{d0}} \left| \frac{V_{p2}}{R_{p2}} \right|^2 \quad (21)$$

$$\gamma = \frac{\mathcal{L}^2 \omega_d \omega_q}{R_{q3} R_{d0}} \left| \frac{V_{p3}}{R_{p3}} \right|^2 \quad (22)$$

Here R 's are circuit pure resistances. It is interesting to note that the output signal current, Equation (19), is closely analogous to that in a pentode amplifier. The output current is 180° out of phase with respect to the input voltage V_{q1} ; it is a typical characteristic of a four-terminal amplifier. The equivalent amplification factor (μ) of the present amplifier is

$$\mu = \sqrt{\frac{R_{q3}}{R_{q1}}} \frac{\sqrt{\alpha\gamma}}{1 + \alpha - \beta + \gamma}. \quad (23)$$

In the case of vacuum tubes, μ in general depends only upon the geometry of the tube elements. The equivalent μ defined here, however, is determined by the voltages of the pump. As revealed by Equations (23) and (21), the higher the pump voltage of the second stage, the larger the value of μ . A simple equivalent circuit for the amplifier

is shown in Figure 3. The amplifier is characterized by four terminals.

POWER GAIN

The available power gain is

$$G = \frac{\text{Power output}}{\text{Max. Power input}} = \frac{|I_{q3}|^2 R_L}{|V_{q1}|^2 / 4R_g} = \frac{4R_L R_g}{|Z_{p1} Z_{p3} Z_{q1} Z_{q3}|^2} \frac{[\mathcal{L}^2 \omega_d \omega_q V_{p1} V_{p3}]^2}{|Z_d'|^2} \tag{24}$$

On resonance, the gain has its maximum value;

$$G_0 = 4 \frac{R_L R_g}{R_{q3} R_{q1}} \frac{\alpha \gamma}{(1 + \alpha - \beta + \gamma)^2} \tag{25}$$

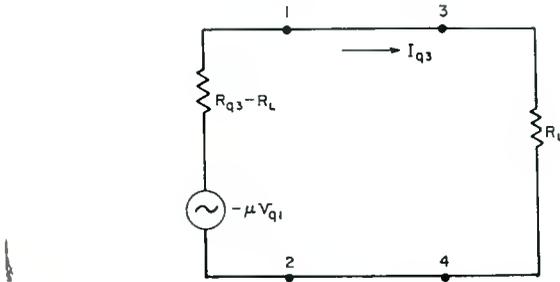


Fig. 3—Equivalent circuit of the four-terminal parametric amplifier.

The necessary condition for achieving an amplifier ($G_0 > 1$) is

$$\beta > 1 + \alpha + \gamma - \sqrt{\frac{4R_L R_g}{R_{q1} R_{q3}} \alpha \gamma} \tag{26}$$

Since β is a parameter proportional to the pumping power according to Equation (21), Expression (26) gives the threshold of pumping power for amplification.

In Equation (25) there are many ways to design the gain function by independently choosing the voltage parameters (α, β, γ). From the point of view of stability, the gain should vary slowly with any of these voltages. One convenient way to achieve a slowly varying gain function is to make α, β, γ related to each other as follows:

$$\frac{\beta}{\alpha} = k_2 = \text{constant}, \quad (27)$$

$$\frac{\gamma}{\alpha} = k_3 = \text{constant}.$$

It follows then that

$$G_0' = \frac{\alpha'^2}{(1 - \alpha')^2}, \quad (28)$$

where

$$G_0' = G_0 \frac{R_{q1} R_{q3}}{4R_l R_p} \frac{(k_2 - k_3 - 1)^2}{k_3},$$

$$\alpha' = (k_2 - k_3 - 1) \alpha.$$

Equation (28) is a normalized gain expression. It is different from the conventional expression which can be written⁴

$$G_0'' = \frac{1}{(1 - \alpha'')^2}. \quad (29)$$

The slope of the gain versus α' in Equation (28) is apparently smaller than that in Equation (29) for α' or α'' less than unity. Therefore, the present gain function varies more slowly than the conventional one.

Another method of achieving a stable gain is to make $1/\alpha$ much smaller than $1 - k_2 + k_3$. By doing this the gain, Equation (25), can be made constant even though the pumping source is subject to a considerable amount of fluctuation, provided that the voltage parameters (α , β , γ) are derived from one single source, and relations (27) are always maintained. To make $1/\alpha$ small is also compatible with the requirement for a low noise factor as will be seen in the computation of the noise factor.

BANDWIDTH

The half-power frequencies which define the bandwidth are the roots of $G(\omega) = G_0/2$, $G(\omega)$ being given by Equation (24), or

$$2 = \left(1 + \frac{X_{q3}^2}{R_{q3}^2}\right) \left(1 + \frac{X_{q1}^2}{R_{q1}^2}\right) \left(\frac{R_d'^2 + X_d'^2}{\bar{R}_d'^2}\right) \quad (30)$$

where X 's and R 's are respectively the reactances and resistances of

the corresponding Z 's. At resonance, Z_d' is a pure resistance which is denoted by \bar{R}_d' . Since the pump is assumed to be maintained at a constant frequency, its self-reactances in all the pump circuits are equal to zero.

Suppose the half-power frequencies are near the signal resonant frequency, ω_q , so that the ratio of the reactance to the resistance can be approximated by

$$\frac{X_{q1}}{R_{q1}} \cong \frac{2(\omega - \omega_q)}{\omega_q} Q_{q1}, \tag{31}$$

$$\frac{X_{q3}}{R_{q3}} \cong \frac{2(\omega - \omega_q)}{\omega_q} Q_{q3}, \tag{32}$$

$$\frac{X_{s2}}{R_{s2}} \cong \frac{-2(\omega - \omega_q)}{\omega_s} Q_{s2}, \tag{33}$$

$$\frac{X_{d0}}{R_{d0}} \cong \frac{2(\omega - \omega_q)}{\omega_d} Q_{d0}. \tag{34}$$

In Equations (31)-(34), Q 's are the circuit Q 's. From Equations (8), (12), (13), and (16), it follows that

$$\begin{aligned} \frac{|Z_d'|^2}{\bar{R}_d'^2} &= \frac{R_d'^2 + X_d'^2}{\bar{R}_d'^2} = \frac{1}{(1 + \alpha - \beta + \gamma)^2} \\ &\quad \left[\left(1 + \frac{\alpha}{1 + \left(\frac{X_{q1}}{R_{q1}}\right)^2} - \frac{\beta}{1 + \left(\frac{X_{s2}}{R_{s2}}\right)^2} + \frac{\gamma}{1 + \left(\frac{X_{q3}}{R_{q3}}\right)^2} \right)^2 \right. \\ &\quad \left. + \left(\frac{X_{d0}}{R_{d0}} \right)^2 \left(1 - \frac{\alpha \left(\frac{Q_{q1}}{Q_{d0}}\right) \left(\frac{\omega_d}{\omega_q}\right)}{1 + \left(\frac{X_{q1}}{R_{q1}}\right)^2} + \frac{\beta \left(\frac{Q_{s2}}{Q_{d0}}\right) \left(\frac{\omega_d}{\omega_s}\right)}{1 + \left(\frac{X_{s2}}{R_{s2}}\right)^2} - \frac{\left(\frac{Q_{q3}}{Q_{d0}}\right) \left(\frac{\omega_d}{\omega_q}\right)}{1 + \left(\frac{X_{q3}}{R_{q3}}\right)^2} \right)^2 \right] \tag{35} \end{aligned}$$

By making use of the gain expression (25), Equation (35) can be written

$$\frac{Z_d'^2}{\bar{R}_d'^2} = G_0 \left(\frac{R_{q1} R_{q3}}{4R_L R_g} \right) \frac{1}{\alpha\gamma} \left[\left(1 + \frac{\alpha}{1 + \left| \frac{X_{q1}}{R_{q1}} \right|^2} - \frac{\beta}{1 + \left| \frac{X_{s2}}{R_{s2}} \right|^2} + \frac{\gamma}{1 + \left| \frac{X_{q3}}{R_{q3}} \right|^2} \right)^2 \right. \\ \left. + \left(\frac{X_{d0}}{R_{d0}} \right)^2 \left(1 - \frac{\alpha \left(\frac{Q_{q1}}{Q_{d0}} \right) \left(\frac{\omega_d}{\omega_q} \right)}{1 + \left| \frac{X_{q1}}{R_{q1}} \right|^2} - \frac{\beta \left(\frac{Q_s^2}{Q_{d0}} \right) \left(\frac{\omega_d}{\omega_s} \right)}{1 + \left| \frac{X_{s2}}{R_{s2}} \right|^2} - \frac{\gamma \left(\frac{Q_{q3}}{Q_{d0}} \right) \left(\frac{\omega_q}{\omega_d} \right)}{1 + \left| \frac{X_{q3}}{R_{q3}} \right|^2} \right)^2 \right]. \quad (36)$$

The percentage of full bandwidth, B , is defined as

$$B = \frac{2(\omega - \omega_q)}{\omega_q}.$$

Equation (30) then becomes

$$2 = (1 + B^2 Q_{q3}^2) (1 + B^2 Q_{q1}^2) \left\{ G_0 \left(\frac{R_{q1} R_{q3}}{4R_L R_g} \right) \frac{1}{\alpha\gamma} \right. \\ \left[\left(1 + \frac{\alpha}{1 + B^2 Q_{q1}^2} - \frac{\beta}{1 + B^2 Q_{s2}^2 \left(\frac{\omega_q}{\omega_s} \right)^2} + \frac{\gamma}{1 + B^2 Q_{q3}^2} \right)^2 \right. \\ \left. + B^2 Q_{d0}^2 \left(\frac{\omega_q}{\omega_d} \right)^2 \left(1 - \frac{\alpha}{1 + B^2 Q_{q1}^2} \frac{Q_{q1}}{Q_{d0}} \frac{\omega_d}{\omega_q} \right. \right. \\ \left. \left. + \frac{\beta}{1 + B^2 Q_{s2}^2 \left(\frac{\omega_q}{\omega_s} \right)^2} \frac{\omega_d \cdot Q_{s2}}{\omega_s Q_{d0}} - \frac{\gamma}{1 + B^2 Q_{q3}^2} \frac{\omega_d}{\omega_q} \frac{Q_{q3}}{Q_{d0}} \right)^2 \right] \left. \right\}. \quad (37)$$

Equation (37) is very involved and difficult to solve for B . However, the solution can be readily attained if the ratios in Equations (31) to (34) are assumed to be small compared with unity. The assumption is purely a trial step. Fortunately, these ratios are known to be small from a rigorous computation for parametric amplifiers² of the same nature. Furthermore, it can be shown that the solution obtained is

consistent with the original assumption. Nevertheless, it should be pointed out that these ratios of reactances to resistances in the present *active* resonant circuit should not be confused with the reactance-to-resistance ratio in a passive resonant circuit. The reactance-to-resistance ratio in a passive resonant circuit for the half-power frequencies is unity.

With the above assumption, Equation (37) becomes

$$2 = (1 + B^2 Q_{q3}^2) (1 + B^2 Q_{q1}^2) \left(1 + B^2 Q_{d0}^2 \frac{\omega_q^2}{\omega_d^2} G_0' \right), \quad (38)$$

where G_0' is the effective gain and is

$$G_0' = G_0 \frac{R_{q1} R_{q3} \left[\frac{1}{\alpha} - \frac{\omega_d}{\omega_q} \frac{Q_{q1}}{Q_{d0}} + \frac{\beta}{\alpha} \frac{\omega_d}{\omega_s} \frac{Q_{s2}}{Q_{d0}} - \frac{\gamma}{\alpha} \frac{\omega_d}{\omega_q} \frac{Q_{q3}}{Q_{d0}} \right]^2}{4 R_L R_g} \frac{\gamma}{\alpha} \quad (39)$$

$$= \frac{\left[\frac{1}{\alpha} - \frac{\omega_d}{\omega_q} \frac{Q_{q1}}{Q_{d0}} + \frac{\beta}{\alpha} \frac{\omega_d}{\omega_s} \frac{Q_{s2}}{Q_{d0}} - \frac{\gamma}{\alpha} \frac{\omega_d}{\omega_q} \frac{Q_{q3}}{Q_{d0}} \right]^2}{\left(\frac{1}{\alpha} + 1 - \frac{\beta}{\alpha} + \frac{\gamma}{\alpha} \right)^2}$$

To be compatible with the original assumption that the bandwidth-Q product is small compared with unity, Equation (38) can be satisfied only when G_0' is large compared with unity.

Accordingly, for large gains and small ratios of reactances to resistances, Equation (38) reduces to

$$B \sqrt{G_0} = \frac{\sqrt{\frac{4R_L R_g}{R_{q1} R_{q3}}} \left(\frac{1}{Q_{d0}} \right) \left(\frac{\omega_d}{\omega_q} \right) \sqrt{\frac{\alpha}{\gamma}}}{\left(\frac{1}{\alpha} - \frac{\omega_d}{\omega_q} \frac{Q_{q1}}{Q_{d0}} + \frac{\beta}{\alpha} \frac{\omega_d}{\omega_q} \frac{Q_{s2}}{Q_{d0}} - \frac{\omega_d}{\omega_q} \frac{Q_{q3}}{Q_{d0}} \frac{\gamma}{\alpha} \right)} \quad (40)$$

In the case of $Q_{q1} = Q_{q3} \ll Q_{d0} \ll Q_{s2}$, Equation (40) becomes

$$B \sqrt{G_0} = \frac{1}{Q_{s2}} \frac{\omega_s}{\omega_q} \frac{\sqrt{\gamma \alpha}}{\beta} \frac{\sqrt{4R_L R_g}}{R_{q1}} \quad (41)$$

This gain-bandwidth expression, being inversely proportional to the idling circuit Q in the amplifier stage, is quite analogous to that for the conventional parametric amplifier⁴ involving only one stage.

NOISE FACTOR

The noise voltages which appear in the three stages are

$$\begin{aligned} |V_{q1}|^2 &= 4k\Delta f(R_g T_0 + R_1 T); & R_g + R_1 &= R_{q1}, \\ |V_{q3}|^2 &= 4k\Delta f R_{q3} T, \\ |V_{s2}|^2 &= 4k\Delta f R_{s2} T, \\ |V_{d0}|^2 &= 4k\Delta f R_{d0} T. \end{aligned} \tag{42}$$

In Equations (42), the internal resistance of the signal generator, R_g , is at the reference temperature, T_0 , and the amplifier resistances R_1 , R_{s2} , R_{d0} and R_{q3} are at the ambient temperature, T . The noise voltages in the pump circuits have been ignored since the pump voltages are assumed to be always much larger than any noise voltage.

These noise voltages, according to Equation (17) will give rise to a noise output which is equal to

$$\begin{aligned} N_0 &= I_{q3}^2 R_L = \frac{|V_{q3}|^2 + (|V_{d0}|^2 + |V_{d1}|^2 + |V_{d2}|^2 + |V_{d3}|^2) \frac{\mathcal{L}^2 \omega_q^2 V_{p3}^2}{R_{p3}^2 \bar{R}_d^2}}{R_{q3}^2} \\ &= \frac{|V_{q3}|^2 + (|V_{d0}|^2 + |V_{d1}|^2 + |V_{d2}|^2 + |V_{d3}|^2) \frac{\mathcal{L}^2 \omega_q^2 V_{p3}^2}{R_{p3}^2}}{R_{q3}^2 / R_L} \\ &= G_0 \frac{(R_{p1} R_{p3} R_{q1} R_{q3})^2}{4R_g R_L (\mathcal{L}^2 \omega_d \omega_q V_{p3} V_{p1})^2}. \end{aligned} \tag{43}$$

Hence, by Equations (7)-(13), the noise factor, F , is

$$F = \frac{N_0}{kT_0 \Delta f} \frac{1}{G_0}$$

$$\begin{aligned}
 &= \frac{|V_{q1}|^2}{4kT_0 R_g \Delta f} \left[\frac{4R_L R_g}{G_0 R_{q3}^2} \frac{|V_{q3}|^2}{|V_{q1}|^2} + 1 + \frac{|V_{s2}|^2}{|V_{p1}|^2} \frac{|V_{p2}|^2}{|V_{q1}|^2} \left(\frac{R_{p1}^2}{R_{s2}^2} \right) \left(\frac{R_{q1}^2}{R_{p2}^2} \right) \right. \\
 &+ \left. \left(\frac{V_{p3}}{V_{p1}} \right)^2 \frac{|V_{q3}|^2}{|V_{q1}|^2} \frac{R_{p1}^2}{R_{q3}^2} \frac{R_{q1}^2}{R_{p3}^2} + \frac{|V_{d0}|^2}{|V_{p1}|^2} \frac{R_{p1}^2 R_{q1}^2}{\mathcal{L}^2 \omega_d^2 |V_{q1}|^2} \right]. \quad (44)
 \end{aligned}$$

By making use of Equations (42) and (20)-(22), the ultimate noise factor is

$$\begin{aligned}
 F = \left(1 + \frac{R_1}{R_g} \frac{T}{T_0} \right) &\left\{ \frac{4R_L R_g}{G_0 R_{q3}} \frac{T}{T_0 R_g + TR_1} + 1 + \frac{TR_{q1}}{T_0 R_g + TR_1} \right. \\
 &\left. \left[\left(\frac{\beta}{\alpha} \right) \left(\frac{\omega_q}{\omega_s} \right) + \left(\frac{\omega_q}{\omega_d} \right) \frac{1}{\alpha} + \left(\frac{\gamma}{\alpha} \right) \right] \right\} \quad (45)
 \end{aligned}$$

It is indicated in Equation (45) that the noise factor of the amplifier is controlled primarily by the three voltage parameters of α , β , γ and the frequency ratios ω_q/ω_s and ω_q/ω_d . The theoretical lowest noise factor is unity; this could be achieved by making

$$\left(\frac{1}{\alpha} \frac{\omega_q}{\omega_d} \right), \left(\frac{1}{\alpha} \right), \left(\frac{\omega_q}{\omega_s} \right), \left(\frac{\beta}{\alpha} \right), \text{ and } \frac{R_1}{R_g} \text{ much smaller than unity.}$$

DISCUSSION

The above results, Equations (25), (40), and (45), show that the gain, the bandwidth, and the noise factor of the amplifier are each dependent upon three voltage parameters (α , β , γ) which are the measure of the pumping power for each stage. It is thus logical to predict that there must be some relations among the gain, the bandwidth, and the noise factor. To find such relations, let

$$\frac{\beta}{\alpha} = x,$$

$$\frac{\gamma}{\alpha} = y,$$

$$\frac{1}{\alpha} = z.$$

Equations (25), (40), and (45) then become, respectively,

$$\frac{2\sqrt{y}}{z+1-x+y} = G_V, \quad (47)$$

$$\frac{2\sqrt{y}}{b[cx-b'+z-c'y]} = M, \quad (48)$$

$$ax+y+bz = F-1. \quad (49)$$

where

$$a = \frac{\omega_q}{\omega_s}, \quad (50)$$

$$b = \frac{\omega_q}{\omega_d}, \quad (51)$$

$$c = \frac{\omega_d}{\omega_s} \frac{Q_{s2}}{Q_{d0}}, \quad (52)$$

$$b' = \frac{\omega_d}{\omega_i} \frac{Q_{q1}}{Q_{d0}}, \quad (53)$$

$$c' = \frac{\omega_d}{\omega_i} \frac{Q_{q3}}{Q_{d0}}, \quad (54)$$

$$G_V = \sqrt{G_0} \sqrt{\frac{R_{q1}R_{q3}}{R_L R_\theta}}, \quad (55)$$

$$M = G_V Q_{d0} B. \quad (56)$$

In Equation (49), the reference temperature has been assumed to be same as the ambient temperature. G_0 is assumed to be large enough so that the noise term involving R_L in Equation (45) can be dropped. R_1 is assumed to be small compared with R_g .

Since a , b , x , y , and z are all positive, x , y , and z must be small to achieve a low noise factor according to Equation (49). If y and z are small, however, x must be in the order of unity for high gains according to Equation (47). These conditions require the use of large values of α as shown by Equation (46). According to Equations (20)-(22),

large values of α mean high pumping powers. This is the price which must be paid for low-noise operation.

If the gain is high (i.e., $x \approx 1$), Equations (47), (48), and (49) can be reduced to the important relation

$$M' = \frac{1}{F + F_0}, \tag{57}$$

where

$$M' = \frac{M}{2\sqrt{y}}, \tag{58}$$

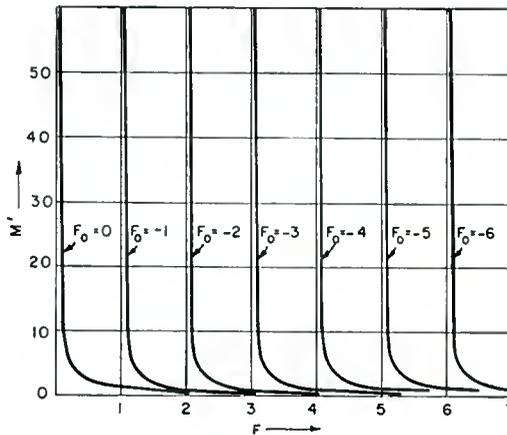


Fig. 4—Bandwidth Parameter M' ($= M/2\sqrt{Y}$) versus noise factor F , F_0 being the limiting noise factor when M' approaches infinity.

$$F_0 = (bc - bb' - 1 - a) - (1 + bc')y. \tag{59}$$

Equation (57) is plotted in Figure 4. Only positive values of M' and F are allowed. Theoretically, with help of eight independent parameters (x, y, z, a, b, c, b', c') it is possible to optimize the bandwidth and noise factor. As an illustration, consider a simple example. Suppose an amplifier with a desired noise factor of 1.5, a full bandwidth of 2 percent, and a voltage gain of 10 has the following constants:

$$Q_{d0} = 5,$$

SURFACE-IMMUNE TRANSISTOR STRUCTURE*

BY

HERBERT NELSON

RCA Laboratories,
Princeton, N. J.

Summary—A new transistor structure is described in which the base region is brought to the surface through the central region of the collector. This structure is found to offer great promise for improving power transistor performance. Silicon transistors of this new geometry are characterized by a high degree of minority carrier conservation and surface immunity. These transistors also exhibit much lower saturation and base resistance than do comparable units of "conventional" structure.

INTRODUCTION

SINCE the advent of the transistor, efforts have been made to achieve a device geometry that would eliminate the influence of the semiconductor surface. Such a device would be unaffected by adverse surface conditions or by surface changes during life, and would have a higher order of reliability and uniformity. In addition, many of the annoying and costly difficulties associated with the present etching and encapsulation procedures would be eliminated.

This objective has now been partially realized with the successful testing of a novel diffused silicon transistor structure whose base region is brought to an external surface through the central region of the collector rather than at the perimeter of the emitter as in conventional types. This effectively eliminates the recombination of injected carriers on the surface between the emitter and the base-lead region; such recombination is particularly troublesome in conventional silicon transistors. In the new structure, the flow of base current is inward from the emitter and collector perimeters to the central base lead. This gives a forward-biasing effect that crowds injection towards the central area of the emitter where no opportunity for surface recombination exists.

Several lots of experimental silicon power transistors of the new geometry have been prepared and tested. Results show that the expected surface immunity is obtained and, furthermore, that the base and saturation resistances are considerably lower than those of comparable conventional units.

* Manuscript received November 4, 1958.

In the following, design considerations and descriptions of the techniques whereby surface immune transistors have been prepared are discussed. Data obtained in tests where surface-immune transistors are compared with comparable conventional units is also presented.

DESIGN CONSIDERATIONS

The new transistor geometry is shown in Figure 1; for purposes of comparison, a transistor of more conventional geometry is shown in Figure 2. In the conventional structure, surface-region B plays an important role in determining the electrical characteristics of the transistor. This is particularly true when surface recombination is excessive and at high current densities when the injection of minority carriers is crowded toward the perimeter of the emitter^{1,2} by the biasing action of the base current. A relatively large fraction of the injected carriers is then lost at surface B, and the current transfer ratio, α_{fe} , comes to depend rather critically upon the recombination rates at this surface. It has been shown, for instance, that in conventional silicon transistors, α_{fe} may be caused to increase by a factor as large as three as a consequence of a surface treatment which leads to a substantial reduction³ of s . In the conventional structure, base resistance also varies with s . When the surface recombination rate of area B is high, minority carrier modulation of the underlying base region is low and vice versa.

The new transistor structure shown in Figure 1 has no counterpart of surface region B of the old geometry. In the new structure the biasing effect due to the base current² leads to a crowding of injection towards the center rather than towards the perimeter of the emitter. Loss of injected carriers in the latter structure, therefore, can occur only to the relatively remote base-lead interface, F, and to bulk recombination. It is consequently relatively independent of surface conditions of the transistor and α_{fe} is unaffected by changes in surface conditions.

In the new transistor structure, the loss of injected carriers depends upon the bulk recombination rate, geometrical dimensions, and the electric field strength in the neighborhood of the base-lead interface. No attempt has been made to quantitatively relate these and

¹R. A. Gundmundsen, "A New Type Transistor Tetrode," *Trans. I.R.E. PGED*, p. 195, April (Abstract), 1957.

²N. H. Fletcher, "Some Aspects of the Design of Power Transistors," *Proc. I.R.E.*, Vol. 43, p. 551, May, 1955.

³A. R. Moore and H. Nelson, "Surface Treatment of Silicon for Low Recombination Velocity," *RCA Review*, Vol. XVII, p. 5, March, 1956.

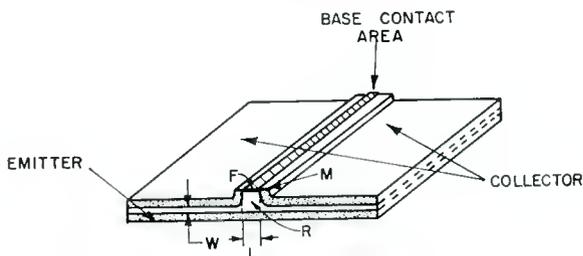


Fig. 1—New transistor structure.

other pertinent parameters to the current transfer ratio. Qualitative considerations as well as experimental results do indicate, however, that in silicon transistors, the new geometry is markedly more conducive to the conservation of minority carriers than is the conventional structure. This is particularly true when the width, L , of the internal base-lead channel is held to a minimum.

The loss of injected carriers to the base-lead interface depends importantly upon the width of the base-lead channel but not as critically as one might expect. Since the distance between the emitter and the base-lead interface is much greater than the base width, w , it follows that the diffusion gradients will drive many more carriers to the collector than to the interface, F . In the internal base-lead channel, minority carrier flow to the collector is favored by the direction of the diffusion gradient. Consequently, the base contact can be made wide enough to be compatible with practicable fabrication procedures without giving rise to substantial minority carriers loss.

A comparison of the two transistor geometries in Figures 1 and 2 suggests further advantages for the new transistor structure with regard to base and saturation resistance. The fact that no injected carriers are lost directly to a base surface adjacent to the emitter, leads to a more effective modulation of the resistivity of the internal base-lead resistance in the surface-immune structure. At high injec-

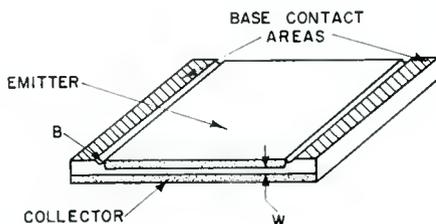


Fig. 2—Conventional transistor structure.

tion levels, channel R of this structure becomes flooded with minority carriers which greatly increase its conductivity.

Low saturation resistance in the new structure is a consequence of its high degree of symmetry.⁴ The loss of injected carriers in this transistor is greater when operated under inverse (emitter and collector interchanged) than under normal conditions. This difference is small since it derives solely from the fact that the base-lead interface collects a slightly greater fraction of the injected carriers under inverse than under normal operating conditions. As a consequence, the condition for low saturation resistance (an inverse current transfer, α_{re} , nearly equal to the normal current transfer ratio, α_{fe}) can be attained in the new transistor structure.

In the conventional geometry, the loss of minority carriers is far greater under inverse than under normal operating conditions. Because of its relatively small area, the inverted emitter is a poor collector for minority carriers injected from the overlap regions of the inverted collector, and practically all of these are lost to the base. As a consequence, $\alpha_{re} \ll \alpha_{fe}$ and high saturation resistance results.

PREPARATION OF SURFACE-IMMUNE SILICON TRANSISTORS

The diffusion-lapping technique⁵ is particularly applicable to the preparation of surface-immune transistors. For the preparation of n-p-n units, 4-6 ohm-centimeter p-type silicon wafers are lapped to a cross section as shown in Figure 3. These wafers are then diffused with phosphorus and afterwards subjected to a second step of lapping to remove the diffused regions at the top of the base-contact ridges. After doping the base-contact areas with boron, the wafers are plated with⁶ "electroless nickel" and then subjected to a final lapping step in which the base-contact areas are isolated from the collector regions, R. Figure 4 shows the cross section of one of the wafers after this final lapping and Figure 1 shows one of the transistor units obtained after the dicing of the wafer. P-n-p transistors similar to these units have also been prepared. In this instance 2-3 ohm-centimeter n-type material has been used for starting material and boron for diffusing the emitter and collector regions.

⁴ J. J. Ebers and J. L. Moll, "Large Signal Behavior of Junction Transistors," *Proc. I.R.E.*, Vol. 42, p. 1761, December, 1954.

⁵ H. Nelson, "The Preparation of Semiconductor Devices by Lapping and Diffusion Techniques," *Proc. I.R.E.*, Vol. 46, p. 1062, June, 1958.

⁶ M. V. Sullivan and J. H. Eigler, "Electroless Nickel Plating for Making Contacts to Silicon," *Jour. Electrochem. Soc.*, Vol. 104, p. 226, April, 1957.

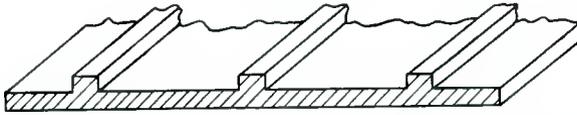


Fig. 3—Silicon wafer lapped for the preparation of surface-immune transistors.

EXPERIMENTAL RESULTS

Silicon power transistors prepared as described in the preceding section were subjected to tests for comparison with comparable conventional silicon units. The comparison was primarily concerned with surface immunity, base resistance, and saturation resistance.

Table I—Effect of Surface Treatment on Current-Transfer Characteristics of Surface-Immune and Conventional Silicon Power Transistors

Unit	α_{fe}	α_{fe}^*	L mils	w mils
Surface-Immune Units				
1	16.0	17.5	4	0.9
2	18.2	20.0	4	0.85
3	12.3	14.0	7	0.8
4	13.5	15.6	7	0.85
Conventional Units ($w = 0.7$ mils)				
5	7.4	30.0		
6	8.4	29.1		

All values of alpha measured at $V_c = 6$ volts, $I_e = 50$ milliamperes.

* After sodium dicromate treatment.

A comparison of conventional and surface-immune n-p-n transistors is shown in Table I. As expected, a large difference in surface immunity is indicated. The current-transfer ratio, α_{fe} , of the con-

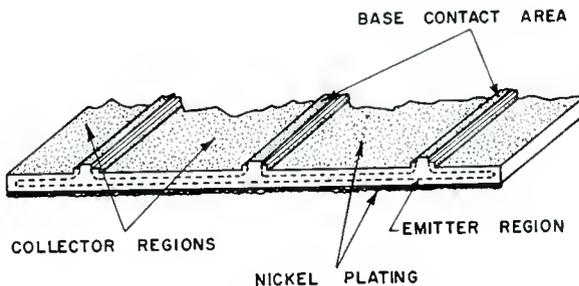


Fig. 4—Silicon wafer after completion of processing.

ventional units varies greatly with surface changes while that of the new geometry units remains substantially unchanged. Data in the same table also show how α_{fe} decreases with an increase in the width, L , of the internal base-lead channel. Transistors 2 and 4, for instance, have equal base widths, but the former with a smaller value of L shows a higher α_{fe} .

Table II—Current-Transfer Characteristics of Surface-Immune Transistors Under Normal and Inverse Operating Conditions

Unit	Normal Operation		Inverted Operation		L mils	w mils
	α_{fe}	α_{fe}^*	α_{re}	α_{re}^*		
7	14.0	16.2	6.4	9.5	4	1.0
8	14.8	16.2	6.4	10.5	7	0.75
9†	17.3	—	<1.0	—	—	—

All values of gain measured at $V_c = 6$ volts, $I_c = 50$ milliamperes.

* After radium dichromate treatment.

† Conventional transistor included for comparison.

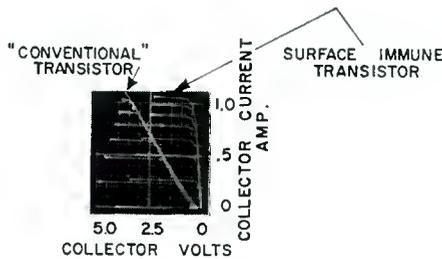


Fig. 5—Common-emitter collector characteristics; $I_b = 20$ milliamperes per step.

In Table II, the normal and inverted operation of two surface-immune transistors are compared. The results indicate a high degree of symmetry for these transistors in that the values of α_{re} are not greatly lower than of α_{fe} . The results also show that α_{re} is more affected by surface treatments than is α_{fe} . This is probably caused by surface region M (see Figure 1) which should have a greater effect upon minority carrier loss in inverted than in normal operation.

Common-emitter collector characteristics of new and old geometry transistors are shown by the two families of curves in Figure 5. The curves show collector current as a function of collector voltage at $I_b = 20, 40, 60$ etc. milliamperes to a collector current maximum of about one ampere. The low saturation resistance of the surface-immune transistor is strikingly evidenced by the early sharp rise of

collector current with voltage. As determined from these curves, the saturation resistances of the surface-immune and the conventional units are 0.4 and 2.5 ohms, respectively.

The transfer characteristics in Figure 6 show how the base-emitter voltage, V_{EBF} , varies with the collector current in the same units. The rate of change of V_{EBF} with collector current at high values of this current is a measure of the rate of change in the voltage drop across the base resistance, r_b . This resistance can be calculated from these curves on the basis of known values of α_{fe} . In this manner, r_b was determined to be approximately 3.5 ohms for the surface-immune unit and 52 ohms for the conventional transistor.

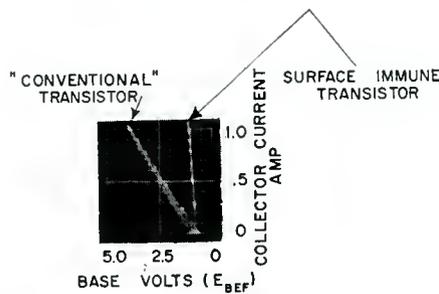


Fig. 6—Common-emitter transfer characteristics; $I_b = 20$ milliamperes per step.

CONCLUSIONS

A new transistor structure in which the base region is brought to an external surface through the central region of the collector is characterized by a high degree of minority carrier conservation and surface immunity. Silicon power transistors of this new geometry exhibit considerably lower saturation and base resistance than do comparable units of conventional structure. The new structural geometry appears to be very promising for the improvement of commercial silicon power transistor performance, and should benefit transistors made of material such as gallium arsenide, where surface recombination tends to be extremely high. The importance of the structure is less for transistors made from germanium where surface recombination can be very low.

SEMICONDUCTOR DIODES IN PARAMETRIC SUBHARMONIC OSCILLATORS*#

BY

J. HILIBRAND[†] AND W. R. BEAM[‡]

Summary—In a parametric subharmonic oscillator, a nonlinear reactance element is driven at frequency $2f$ and gives rise to oscillations in a tank circuit tuned to frequency f . The subharmonic oscillations occur in either of two stable phases. This property makes such oscillators suitable for application in binary computers, where their capability is determined mainly by (1) transient response (rise time), and (2) input power required to establish oscillations.

At present the nonlinear reactance elements most suitable for high-frequency parametric subharmonic oscillators are semiconductor diodes. In this paper the performance of these diodes in the basic oscillator circuit is evaluated; the effects of stray capacitance, spreading resistance, junction conductance, and capacitance-voltage sensitivity are taken into account in the analysis of theoretical oscillator performance. It is found that minimum rise times of 1.3 cycles and 1.9 cycles (of the subharmonic frequency) are theoretically possible using abrupt- or graded-junction diodes, respectively. Minimum quality factors, Q , of 4, for abrupt junctions, or 6 for graded junctions, are required for oscillation.

The subharmonic output voltage is limited, in oscillators designed for minimum rise time, to values much less than the pump voltage, but greater output voltages can be obtained if rise time is sacrificed. In such high-voltage operation the amplitude is limited by available power from the pump source and conversion efficiency is determined by spreading-resistance losses.

INTRODUCTION

ALTHOUGH THE principle of parametric oscillation was described by Faraday as early as 1831, it is only recently that the possibility of improved low-noise microwave amplification has caused tremendous activity in this field and resulted in the development of improved circuits and nonlinear reactance elements for high-frequency operation. An extensive bibliography of recent activity

* Manuscript received April 6, 1959.

This work was performed under U. S. Navy BuShips contracts.

† RCA Laboratories, Princeton, N. J.

‡ Formerly RCA Electron Tube Division, Princeton, N. J.; now at Rensselaer Polytechnic Institute, Troy, N. Y.

¹H. Heffner, "Masers and Parametric Amplifiers," *I.R.E. Wescon Convention Record, Part 3, Electron Devices*, p. 3, 1958. (Contains an extensive bibliography.)

in the parametric amplifier field is given by Heffner.¹ Parametric subharmonic *oscillators* were recognized by von Neumann² and Goto³ to have potentialities in digital computation. In Japan, low-frequency ferrite-core parametric oscillators have been assembled into a complete digital computer.⁴

At present, semiconductor diodes of improved design, operated in the back-biased condition, present an almost lossless voltage-dependent capacitance to an external circuit at microwave frequencies. If r-f voltage at frequency nf is supplied across the diode, the variable capacitance will present a negative resistance to an external circuit tuned to f . This negative resistance will support oscillations in the tuned circuit. Oscillations will begin to build up as soon as the power at nf is applied, beginning from noise or from an injected subharmonic signal. Coupling circuits are used to remove oscillation power and to inject locking signals.

So far as computer applications are concerned, the most important requirements placed on the oscillator are that the rate of build-up of oscillations be fast, and that the pump power (at frequency nf) be small. The former requirement is necessary for high-speed computation, while the latter is important where large numbers of oscillators are required. The mechanics of information handling in such a computer are more dependent on the passive circuitry external to the oscillator than on the active element itself, and will not be treated here.

Although the results derived in this paper apply directly to oscillators incorporating semiconductor diodes, the methods are general and similar to those which have been used to describe parametric amplifiers. Beginning with a voltage-dependent capacitance (the theoretical dependence for junction diodes is used), the development leads to a subharmonic-frequency negative conductance appearing at the diode terminals when pump-frequency voltage is applied to the same terminals. (In this paper, only half-frequency oscillations where $n = 2$ are considered, though it is recognized that even lower-frequency subharmonic generation is possible.) This negative conductance will be a function of applied bias and pump voltages for any given diode. It is useful to determine the optimum voltages to obtain maximum

² J. von Neumann, "Nonlinear Capacitance or Inductance Switching, Amplifying, and Memory Organs," *U. S. Patent No. 2,815,488*, December 3, 1957.

³ E. Goto, "On the Application of Parametrically Excited Nonlinear Resonators," *Denki Tsushin Gakkai-shi*, Vol. 38, p. 770, October, 1955.

⁴ Z. Kiyasu, "The Parametron," *Journal of Institute of Electrical Communication Engineers of Japan*, Vol. 41, p. 377, April, 1958.

negative conductance. Diode losses can be converted to an equivalent positive shunt conductance which partially cancels the negative conductance. When the losses cancel the negative conductance entirely, oscillations will not build up. This leads directly to a minimum diode quality requirement.

Rise time is a function of the diode net negative conductance, shunting capacitance, and the external circuit capacitance. Since transmission-line circuits are a microwave substitute for the lumped-constant circuits used at lower frequencies, rise time is also calculated as a function of transmission-line parameters.

The same simple quasilinear equivalent circuits can be used to explain power flow. All output power comes from conversion of the pump power. Parameters of the pump source must, therefore, determine the maximum output of the oscillator. This analysis becomes simple a-c circuit analysis once the nonlinear relationship between current and voltage in the diode is resolved into two quasilinear circuits at the two frequencies.

Having the rise time and power relationships for these oscillators, the diode designer can understand the importance of each diode parameter and coordinate the diode constants to circuit needs.

DIODE EQUIVALENT CIRCUITS

A semiconductor diode can be represented by the nonlinear equivalent circuit shown in Figure 1. The incremental capacitance,

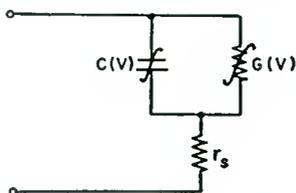


Fig. 1—Nonlinear equivalent circuit for semiconductor diode.

$C(V) = dq/dv$, and the incremental conductance, $G(V) = di/dv$, are functions of the voltage given by semiconductor theory⁵ and shown in Figures 2(a) and 2(b).*

$$C(V) = K(\phi - V)^{-n}, \quad (1)$$

⁵ 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.

* In this paper the reactance and resistance which arise from minority carrier storage effects will be neglected. For a treatment of this problem see J. Hilibrand, et al, "Semiconductor Parametric Diodes in Microwave Computers," to be published in September, 1959, *I.R.E. Transactions PGEC*.

$$G(V) = \frac{qI_0}{kT} \exp \left[\frac{qV}{kT} \right]. \tag{2}$$

The voltage, ϕ , at which $C(V)$ would theoretically go to infinity is called the contact potential, V_0 is the applied bias voltage defined with respect to the contact potential, and K is a constant determined by the diode area, doping, etc. The resistance r_s is nearly independent of diode voltage.

The variable capacitance itself is the element which produces parametric effects. Therefore the circuit operation of the capacitance

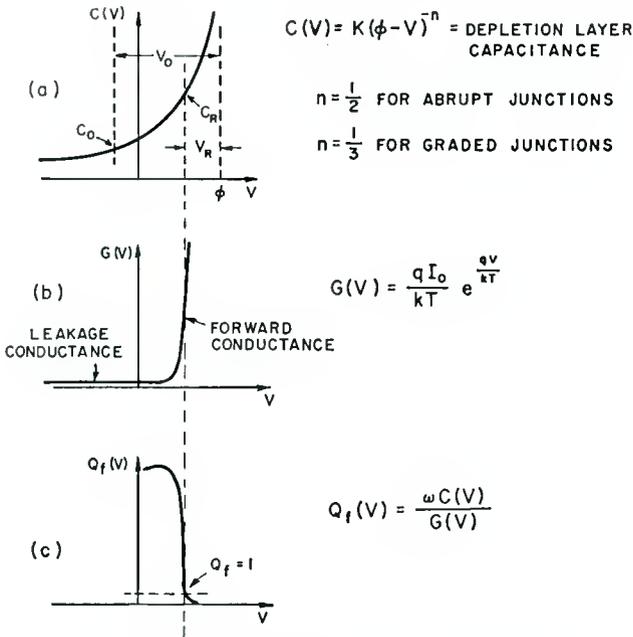


Fig. 2—The voltage-variable diode parameters.

alone will be considered first and the effects of series and shunt resistances will be treated later. Although an approximate method is used, it is sufficiently accurate for most purposes—an exact analytic solution has not been obtained.

The circuit of a typical lumped-constant parametric subharmonic oscillator is shown in Figure 3. A source of power at radian frequency 2ω is applied to a tuned circuit consisting of the diode, various stray capacitances, and an inductance whose value is chosen to tune the circuit to frequency ω . The voltage across the capacitor will consist

of various harmonics of the frequencies ω and 2ω , and cross products between these frequencies. The voltages at frequencies ω and 2ω are the only ones of any appreciable amplitude since the first is the frequency for which the circuit is tuned and the second is the frequency of the pump. The tuned pump and the oscillator circuits are here assumed to have low losses and hence to display appreciable impedances only around the frequencies, ω and 2ω . Therefore, the voltage across the diode can be taken to be of the form

$$v(t) = V_s \sin(\omega t + \theta) + V_p \sin 2\omega t. \tag{3}$$

Here V_s is the amplitude of the subharmonic, V_p is the pump amplitude, and θ is an as yet undetermined phase angle.

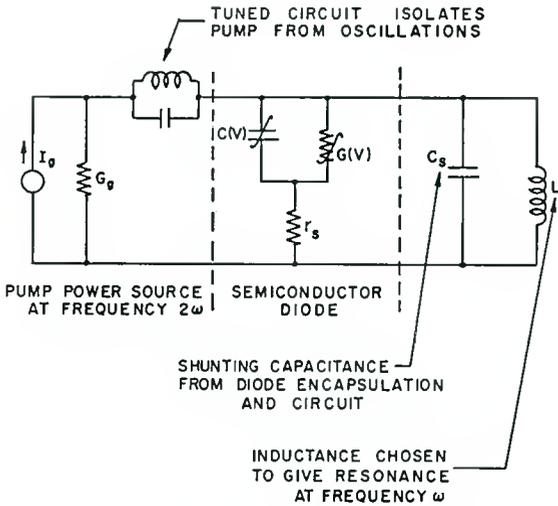


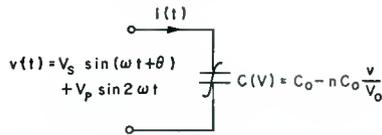
Fig. 3—The lumped-constant parametric subharmonic oscillator.

At this point the actual $C(V)$ characteristic is approximated by a Taylor series expansion around the operating point established by the d-c bias on the diode at $V = \phi - V_0$. This gives the expression

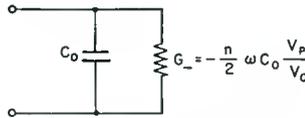
$$\begin{aligned} C(V) &= K(\phi - V + v)^{-n} = K(V_0 + v)^{-n} \\ &= KV_0^{-n} - nKV_0^{-n} \frac{v}{V_0} + \frac{n(n+1)}{2!} KV_0^{-n} \left(\frac{v}{V_0}\right)^2 - \dots \\ &= C_0 - nC_0 \frac{v}{V_0} + \frac{n(n+1)}{2!} C_0 \left(\frac{v}{V_0}\right)^2 - \dots \end{aligned} \tag{4}$$

The first few terms in this series will provide a good approximation to the capacitance when v/V_0 is less than unity. In what follows only the first two terms will be considered, the quiescent value of capacitance C_0 and the linearly varying term $-nC_0v/V_0$. In Appendix I, the effect of higher-order terms is shown to be small normally.

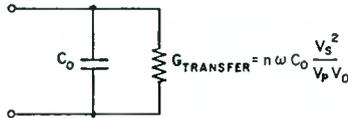
An equivalent circuit for the capacitor at frequency ω can be found by determining the ratio of the current flow at frequency ω to the voltage across the variable capacitor at frequency ω (see Figure 4). The total current is given by $i(t) = C(V)dv/dt$. It is shown in



(a) NON LINEAR EQUIVALENT CIRCUIT



(b) EQUIVALENT CIRCUIT AT FREQUENCY ω



(c) EQUIVALENT CIRCUIT AT FREQUENCY 2ω

Fig. 4—Equivalent circuits for variable capacitance ($\theta = \pm 90^\circ$).

Appendix I that the first-order terms in the expression for the current at frequency ω can be written

$$i_\omega = \omega C_0 V_s \cos(\omega t + \theta) + \frac{n}{2} \omega C_0 \frac{V_p}{V_0} V_s \sin(\omega t - \theta). \quad (5)$$

The first term represents the current due to a pure capacitance of value C_0 (i.e., the diode capacitance at the bias voltage). The second term describes energy conversion from the pump frequency to the frequency of oscillation. The direction and magnitude of the energy

conversion will depend on the phase angle θ . For $\theta = \pm 90^\circ$ this term represents a negative conductance and indicates the transfer of power from the pump into oscillation output. For $\theta = \pm 45^\circ$ or $\pm 135^\circ$ there is no net energy transfer between the two frequencies but there is an energy interchange as represented by a pure positive or negative reactance. For $\theta = 0$ or 180° this term is a positive conductance and indicates that power can be transferred from oscillation frequency to the pump. This condition is satisfied in variable-capacitance-diode harmonic generators.

There are, therefore, two values of θ , $\pm 90^\circ$, for which oscillations will build up at a maximum rate-of-rise. At one of these two phases the subharmonic oscillations will ultimately saturate; two stable oscillation conditions are thus established. The two phases of maximum positive conductance are also useful, for they can be employed to cause oscillations to die out rapidly; thus faster switching from one state to another can be obtained in applications where switching speed is important. At either of the phases where maximum negative conductance is developed, this conductance is

$$G_- = -\frac{n}{2} \omega C_0 \frac{V_p}{V_0}. \quad (6)$$

An equivalent circuit can be established for frequency 2ω in the same way. The current is

$$i_{2\omega} = 2\omega C_0 V_p \cos 2\omega t - \frac{n\omega}{2V_0} C_0 V_s^2 \sin(2\omega t + 2\theta).$$

Here again the first term corresponds to a capacitance, C_0 . The second term relates to the energy transfer process, and for $\theta = \pm 90^\circ$ it will be a positive conductance;

$$G_{\text{transfer}} = \frac{n\omega C_0}{2} \frac{V_s^2}{V_0 V_p}.$$

It will be noted that G_{transfer} is a function of the subharmonic and pump voltage amplitudes. While this is not, then, a linear equivalent circuit, it can, when used with appropriate circumspection, lead to clarification of both equilibrium and dynamic energy transfer situations.

OPTIMIZATION OF THE NEGATIVE CONDUCTANCE

It is desirable to choose diode operating voltages which maximize

the negative conductance when there is considerable stray capacitance in the tank circuit. Since G_- is proportional to both the pump voltage and the slope of the $C-V$ characteristic, it would seem desirable to increase both of these. There is a compromise involved, however, for when the slope is increased (by operating nearer $V = \phi$ in Figure 2(a)), operation extends further into the high forward conductance region. Since the losses in this region must be avoided, it is convenient to limit forward voltage swing to some voltage $V = \phi - V_R^\dagger$. In this case, the pump amplitude must equal $(V_0 - V_R)$, and

$$G_- = -\frac{n\omega KV_0^{-n}}{2} \left(1 - \frac{V_R}{V_0}\right). \quad (6a)$$

Differentiating this expression with respect to V_0 and setting the result equal to zero, the bias for maximum negative conductance can be determined;

$$V_0 = \frac{1+n}{n} V_R.$$

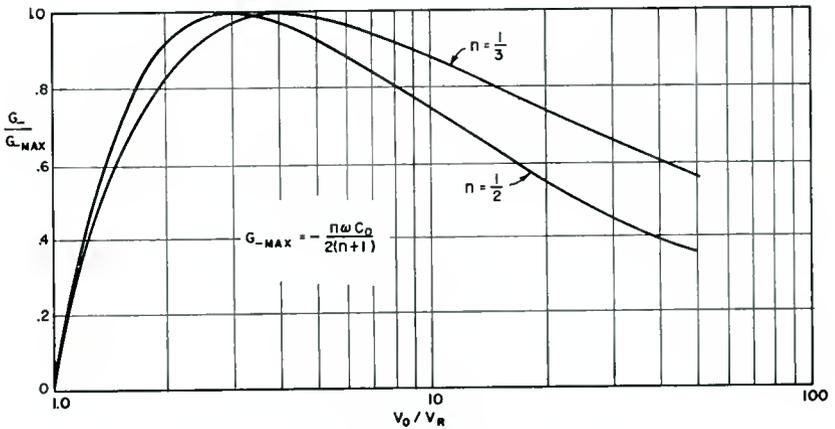


Fig. 5—The effect of bias voltage on the negative conductance.

This bias voltage will serve as a reference value in pump power dissipation considerations. In Figure 5, G_-/G_{-max} is plotted for $n = 1/2$ and $n = 1/3$, the values which are of practical interest. The maxima of these curves are not sharp and, in fact, V_0 can be varied over about a 3:1 range with no more than 10 per cent loss of negative conductance.

[†] The question of choosing a value for V_R is considered in more detail in the next section.

Using this optimum bias voltage, an expression for the maximized negative conductance is obtained;

$$G_- = -\frac{n\omega C_0}{2(n+1)}. \quad (7)$$

DIODE LOSSES

In the diode equivalent circuit of Figure 1 there are two lossy elements, r_s and $G(V)$. At high frequencies and for negative (or slightly positive) biases, the capacitance $C(V)$ practically "shunts out" the conductance $G(V)$; then r_s is the only important lossy element in the circuit. The most common and convenient measure of the loss in a circuit element is the quality factor Q , which for the simple series combination of C and r_s , is defined by

$$Q = \frac{1}{\omega C r_s}.$$

In linear circuits, Q measures the ratio of energy stored to energy lost per cycle; it will not give that measure for large voltage swings if Q is a function of voltage as in the present case. Henceforth Q_s will be defined, *at the bias point*, by

$$Q_s = \frac{1}{\omega C_0 r_s}.$$

If the diode is to be used in a tuned circuit, it is clearly desirable that Q_s be appreciably greater than unity so that the energy circulating between inductance and capacitance will not be dissipated too rapidly. The series resistance r_s and capacitance C can be represented by a shunt equivalent circuit. For reasonably high Q_s , the effective shunt conductance at frequency ω is almost exactly

$$G_{s,\omega} = (\omega C_0)^2 r_s = \frac{\omega C_0}{Q_s}.$$

In the equivalent circuit at frequency 2ω , the corresponding conductance is

$$G_{s,2\omega} = (2\omega C_0)^2 r_s = \frac{4\omega C_0}{Q_s}.$$

The voltage-dependent shunting conductance of a high-quality diode can contribute appreciable power loss only when the diode voltage swings positive. It is necessary that the voltage excursion in the forward conduction direction be limited to prevent excessive energy losses in this conductance. Since these losses vary exponentially with voltage (as illustrated in Figure 2(b)), it is appropriate to choose some reference voltage to mark the limits of the forward voltage excursion. In Figure 2(c) this voltage is chosen at the point where $Q_f = \omega C(V)/G(V) = 1$. (In the forward-bias region, if Q_s is large, Q_f is very nearly the true diode quality factor.) The choice of reference voltage can be made without undue concern about the need to control the forward losses precisely, since a slight change in voltage swing is all that is needed to reduce these parallel conductance losses markedly. Although the over-all Q is around unity at the most positive voltage, the net effect of shunt conductance is very much smaller when averaged over a cycle; in fact, the circuit losses over an entire cycle will be assumed, henceforth, to stem primarily from the series resistance r_s .

The selection of V_R , the voltage of maximum signal excursion relative to contact potential, is thus not critical. The $Q_f = 1$ reference voltage may be used, as in Figure 2(c), or the voltages for which $Q_f = 5$ or $Q_f = 10$ can be chosen. The value of V_R will not differ appreciably in any case. It is, however, desirable to design the diode to make V_R small. Then V_0 will be small, and it follows that G_- will be large, and the required pump power will be small. The total conductance of the diode at oscillation frequency is

$$G_{\text{NET}} = G_- + G_{s,\omega} + \langle G(V) \rangle_{\text{averaged over a cycle}}$$

$$= -\frac{n\omega C_0}{2} \frac{\left[\frac{V_0}{V_R} - 1 \right]}{V_0/V_R} + \frac{\omega C_0}{Q_s} + \langle G(V) \rangle.$$

The shunting conductance term will, in practice, be negligible.

$$G_{\text{NET}} \cong \omega C_0 \left[-\frac{n}{2} \frac{\left[\frac{V_0}{V_R} - 1 \right]}{V_0/V_R} + \frac{1}{Q_s} \right]. \quad (8)$$

Equation (8) specifies the minimum diode quality factor at which

oscillations will be induced. For useable power transfer, it is necessary that G_{NET} be negative so that

$$Q_s \min = \frac{2}{n} \frac{\left[\frac{V_0}{V_R} - 1 \right]}{V_0/V_R} \tag{9}$$

This function is shown in Figure 6. In order to realize subharmonic oscillations, it is necessary for diode quality factor at the bias point to be greater than 4 for abrupt-junction diodes and greater than 6 for graded-junction diodes. Practical limitations on the maximum value of V_0/V_R (stemming from the available pump power) make these lower limits somewhat optimistic—for smaller values of V_0/V_R , better diode quality is needed.

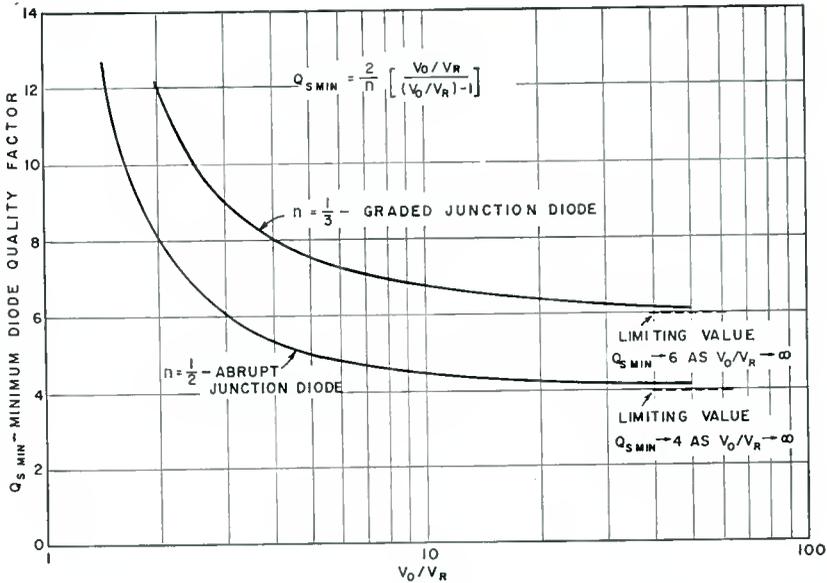


Fig. 6—Minimum diode quality factor for oscillation.

It is clear from the negative conductance relations that if Q_s is very much greater than these minimum values, little improvement will result from further increases in Q_s . Thus, if a Q_s of 20 or more is attainable, the diode designer might better concentrate his efforts on increasing C_0 than on attempting further increase of Q_s . Of course, a higher- Q diode could operate at higher frequencies.

RATE OF OSCILLATION BUILD-UP

From the foregoing, it is possible to calculate the rate of build-up of subharmonic oscillations in a lumped-constant circuit such as that shown in Figure 7. The variation of the envelope of the waveform is described by the function $\exp[-G_{\text{NET}} t/2C_T]$, where G_{NET} is the net negative conductance and C_T is the sum of the diode capacitance and various shunting capacitances, $C_0 + C_s$. The time constant is (following usual practice) defined as $\tau = 2C_T/G_{\text{NET}}$. For switching applications, it is desirable to minimize rise time.

The effects of diode losses will be neglected at this point; they are

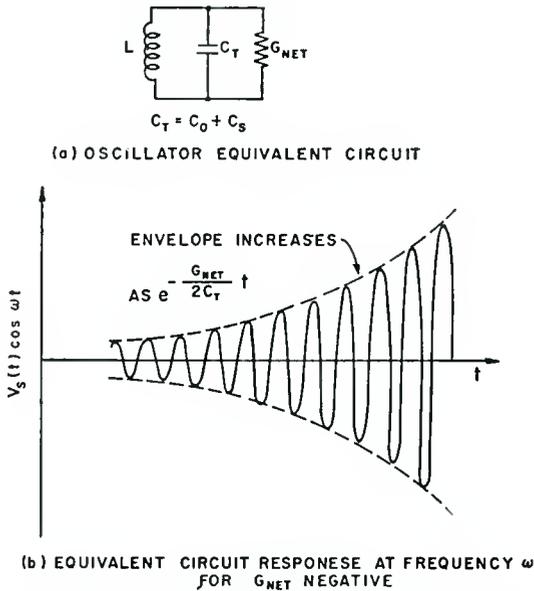


Fig. 7—Rise time in a lumped-constant subharmonic oscillator circuit.

treated later in this section. Combining the value of G_- from Equation (6a) and the value of C_0 from Equation (4),

$$\tau = \frac{2C_T}{G_-} = \frac{4}{n\omega} \frac{V_0/V_R}{\left[\frac{V_0}{V_R} - 1 \right]} \left[1 + \frac{C_s}{C_0} \right]. \quad (10)$$

The number of cycles of the subharmonic frequency required for growth by a factor e is of interest; this is given by

$$\tau f = \frac{2}{\pi n} \frac{V_0/V_R}{\left[\frac{V_0}{V_R} - 1 \right]} \left[1 + \frac{C_s}{C_0} \right], \tag{11}$$

and is shown in Figure 8 for abrupt-junction diodes. The corresponding figure for graded junction diodes is obtained by multiplying the ordinate axis by 3/2. These curves are useful in estimating achievable rise times from diode measurements at the bias point.

For any given diode, the value of C_0 changes with bias point. The dashed line in Figure 8 shows the effect of bias voltage on the rise

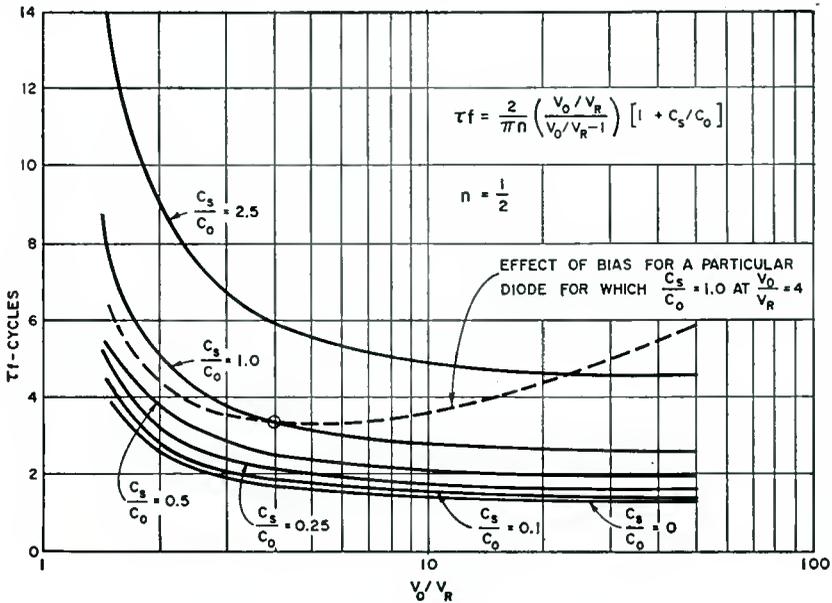
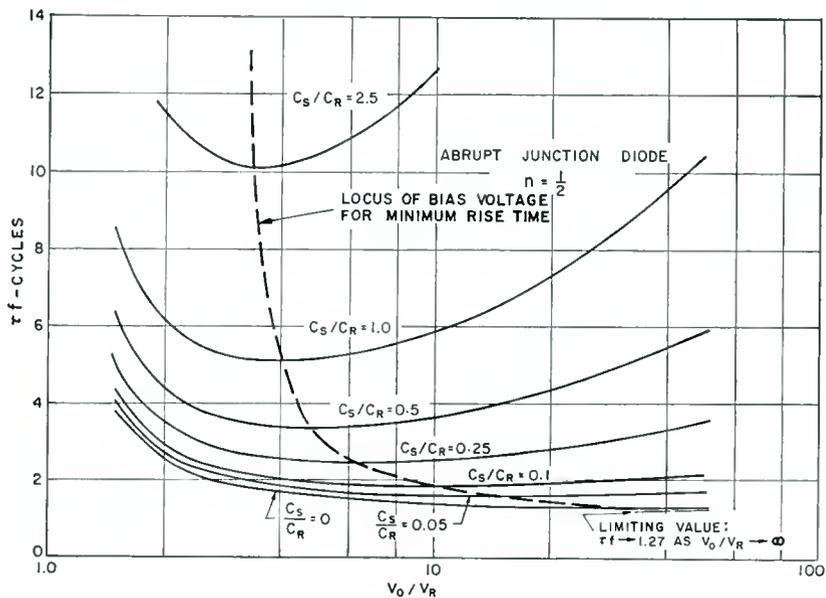
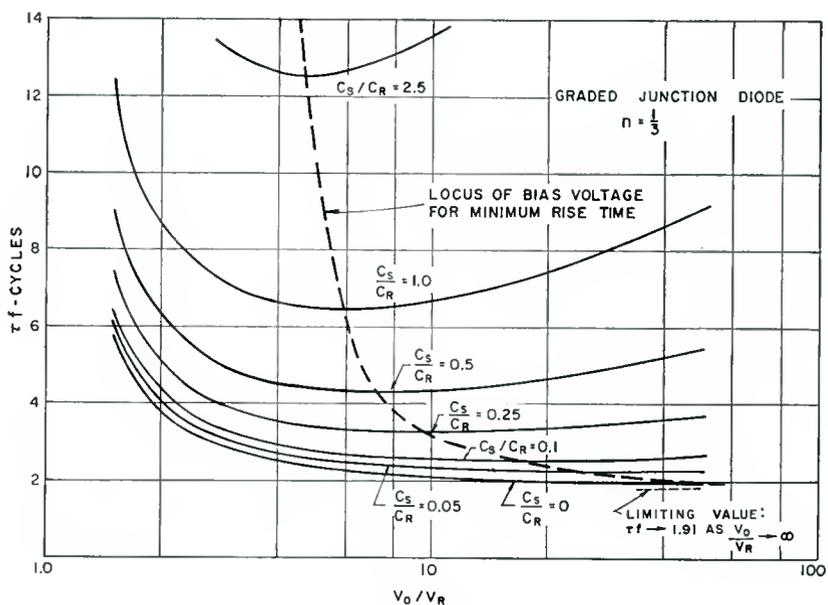


Fig. 8—Effect of bias voltage on rise time.

time achievable with a particular diode. With increased back bias, C_0 decreases and the ratio C_s/C_0 rises. The optimum bias point for a given diode can be determined by taking this variation of C_0 into account. C_R , the depletion layer capacitance at $V = \phi - V_R$ (see Figure 2a), is a useful diode parameter for determining this optimum bias point for a given diode. Then:

$$\tau f = \frac{2}{\pi n} \frac{V_0/V_R}{\left[\frac{V_0}{V_R} - 1 \right]} \left[\left(1 + \frac{C_s}{C_R} \frac{V_0}{V_R} \right)^n \right] \tag{12}$$

Fig. 9—Subharmonic oscillator rise time ($n = \frac{1}{2}$).Fig. 10—Subharmonic oscillator rise time ($n = \frac{1}{3}$).

These rise time curves are plotted with C_s/C_R as the parameter in Figures 9 and 10 for abrupt- and graded-junction diodes, respectively. When the shunting capacitance is much greater than the diode capacitance, the fastest rise time is obtained with the maximum negative conductance, but if C_s is small, a large bias voltage is called for. The limitation in this direction stems from the large pump power required for small decreases of the rise time and, ultimately, from avalanche breakdown of the diode. It will be noted that the minimum rise times (for $V_0/V_R \rightarrow \infty$ and $C_s/C_R = 0$) correspond to 1.27 cycles of the

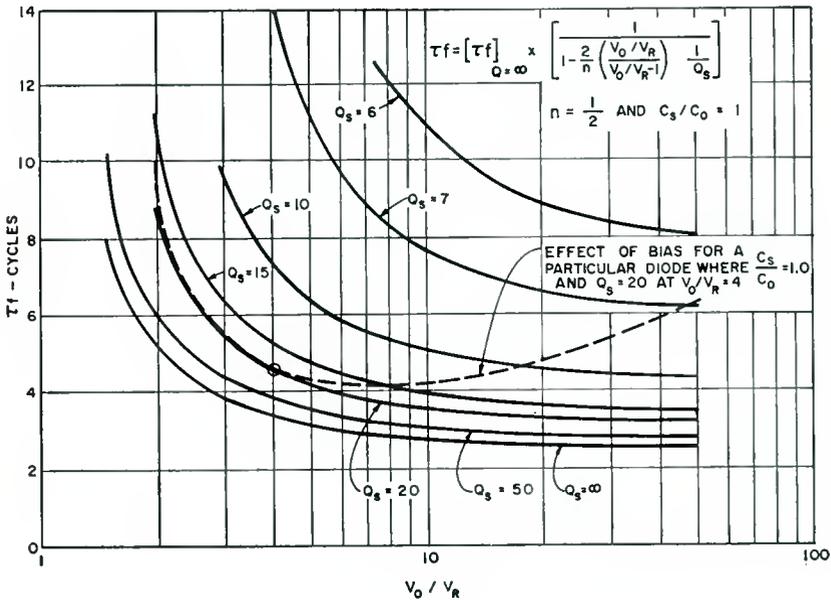


Fig. 11—Effect of diode losses on rise time.

oscillation frequency for the abrupt-junction diode and 1.91 cycles for the graded-junction diode. For practical values of bias ($V_0/V_R = 5$) and stray capacitance ($C_s/C_R = 0.5$), the rise time corresponds to about 3.5 cycles for the abrupt-junction diode and 4.5 cycles for the graded-junction diode.

The effect of diode quality, Q_s , on the rise time is of importance in practical cases. The rise time in the presence of loss is obtained using G_{NET} rather than G_- in Equation (10). The number of cycles corresponding to one time constant is

$$\begin{aligned} \tau f &= \frac{2}{n\pi} \frac{V_0/V_R}{\left[\frac{V_0}{V_R} - 1 \right]} \left[1 + \frac{C_s}{C_0} \right] \left(\frac{1}{1 - \frac{2}{n} \frac{V_0/V_R}{\left[\frac{V_0}{V_R} - 1 \right]} \frac{1}{Q_s}} \right) \\ &= [\tau f]_{Q_s=\infty} \left(\frac{1}{1 - \frac{2}{n} \frac{V_0/V_R}{\left[\frac{V_0}{V_R} - 1 \right]} \frac{1}{Q_s}} \right) \end{aligned} \quad (13)$$

This function is plotted in Figure 11 for an abrupt-junction diode in the case where $C_s/C_0 = 1$. Even comparatively high values of Q_s lead to appreciable deterioration of the response.

The question of determining the operating frequency at which minimum absolute rise time, τ , is achieved using a particular diode is of interest. If the diode quality factor goes to unity at frequency ω_{c0} , Equation (13) can be rewritten in the form

$$\tau = \frac{4}{n\omega} \frac{V_0/V_R}{\left[\frac{V_0}{V_R} - 1 \right]} \left[1 + \frac{C_s}{C_0} \right] \left(\frac{1}{1 - \frac{2}{n} \frac{V_0/V_R}{\left[\frac{V_0}{V_R} - 1 \right]} \frac{\omega}{\omega_{c0}}} \right)$$

It should be noted that this cutoff frequency, ω_{c0} , is defined *at the bias point* and is a function of the bias voltage. The values for ω_{c0} commonly quoted are defined near the avalanche breakdown voltage, V_b , and must be multiplied by a factor $(V_0/V_b)^n$.

The rise time is minimum at the frequency

$$\omega = \omega_{c0} \frac{n \left[\frac{V_0}{V_R} - 1 \right]}{4 \frac{V_0/V_R}{V_R}} \quad (14)$$

This corresponds (for $V_0/V_R = 5$ and $n = 1/2$) to the frequency at which Q_s is 10. Minimum rise time is achieved, then, by using diodes at frequencies near the upper limit of their usefulness and accepting a large number of cycles of the oscillation frequency during the rise time.

Diodes used in microwave subharmonic oscillators are mounted in

a transmission line whose length is chosen to give resonance at the desired subharmonic frequency. The rise time can be determined as a function of the transmission-line parameters.

The behavior of the equivalent circuit of Figure 12 at the complex frequency $s = \sigma + j\omega$ is described by

$$0 = G + C_T s + Y_0 \coth \frac{sl}{v} \quad (15)$$

where G is the net conductance, $C_T = C_s + C_0$ is the total capacitance, Y_0 is the characteristic admittance of the transmission line, and v is the velocity of propagation for the line. If the length of the line is much less than a quarter wavelength, the system equation reduces to the lumped-constant rise-time solution discussed earlier. For greater

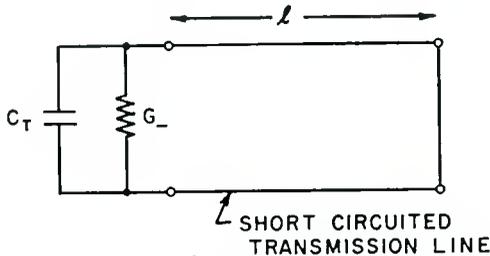


Fig. 12—The microwave parametric subharmonic oscillator circuit.

lengths, the hyperbolic equation has been solved by means of a digital computer for the case where $G = -\omega C_T/6$ (corresponding to an abrupt-junction diode for $V_0/V_R = 3$ and no stray capacitances). These results are shown in Figure 13. Increased line length leads to serious rise-time deterioration.

This result was to be expected since the negative conductance can convert power from the pump to the oscillation frequency at a limited rate. The transmission line has a larger energy storage than the equivalent lumped-circuit reactance for identical terminal voltages. Therefore, the rate of increase must be slower in a long transmission line. The addition of a half-wavelength section of line does not change the resonant frequency, but it will seriously increase rise time. The diode susceptance needed to resonate the indicated length of transmission line is also shown in Figure 13. To obtain the fast rise times associated with short transmission lines a diode (capacitive) admittance which is appreciably larger than the line's characteristic admittance is required.

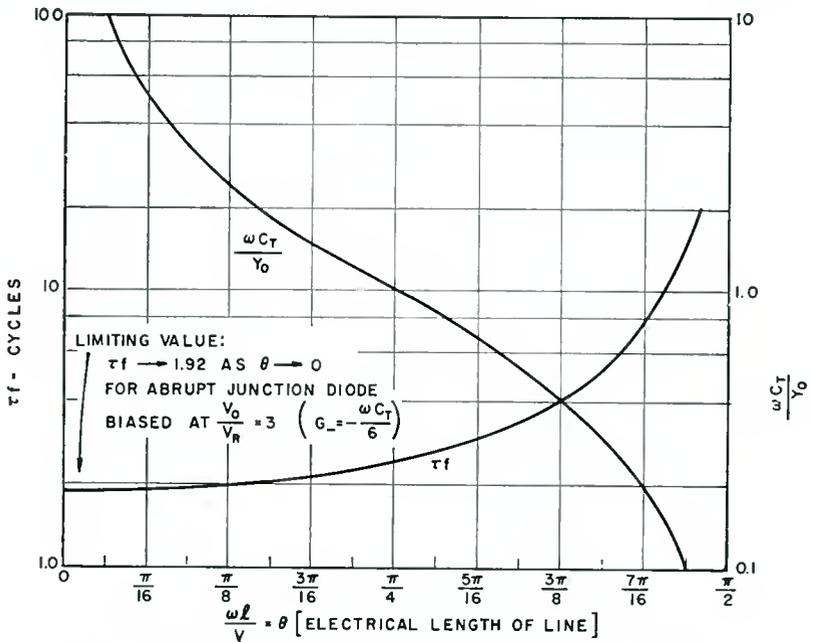


Fig. 13—Rise time in the microwave subharmonic oscillator.

PUMP POWER

Since pump sources at high microwave frequencies have low efficiency and small power output, semiconductor-diode parametric oscillators will not be useful if excessive pump power is required. The pump is required to supply dissipative losses in the parallel conductance and series resistance, as well as the oscillation (output) power (which in practice may be a small part of the total).

If, as was assumed earlier, the losses in $G(V)$ can be neglected at the subharmonic frequency due to the shunting action of $C(V)$, they are a fortiori negligible at the pump frequency. The loss in the series resistor results from the current which is required to produce voltage V_p across the capacitance C_0 (here we assume that $V_p \gg V_s$). This current has a peak value $2\omega C_0 V_p$, hence

$$P_{\text{loss}} = \frac{(2\omega C_0 V_p)^2 r_s}{2} = \frac{2\omega C_0}{Q_s} V_p^2. \quad (16)$$

Since $V_p = V_R/n$ at the bias for maximum negative conductance, this reduces to

$$P_0 = \frac{2\omega C_0}{Q_s} \left(\frac{V_R}{n} \right)^2.$$

Thus V_R should be as small as is practical.

Pump power limitations prescribe the choice of operating bias. Faster rise times can sometimes be achieved by choosing V_0/V_R large, but this also leads to much larger pump power. The pump power goes up in the ratio

$$\frac{P_{\text{loss}}}{P_0} = n^2 \left(\frac{V_0}{V_R} - 1 \right)^2. \tag{18}$$

This function is shown in Figure 14. The rapid increase of losses

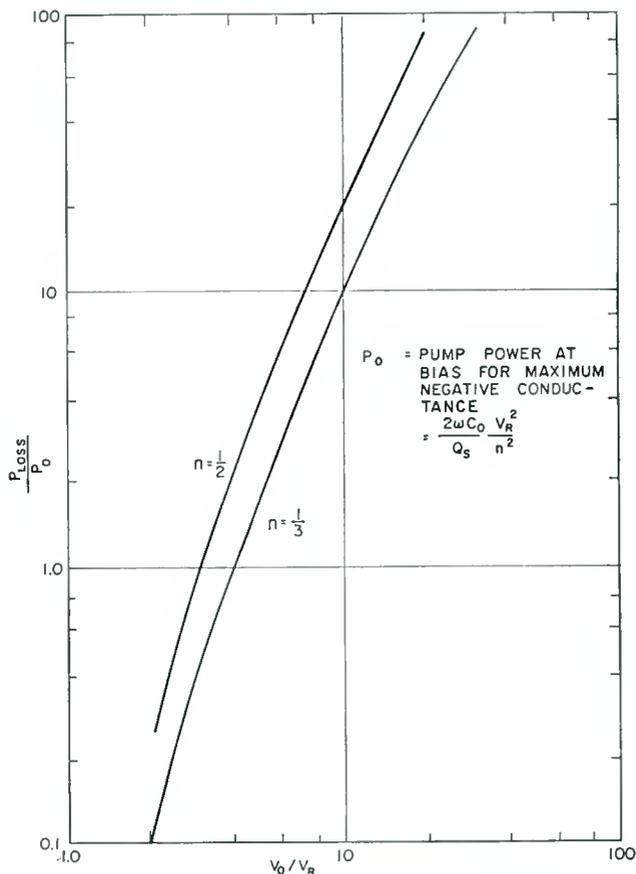


Fig. 14—Pump power required as a function of bias voltage.

precludes bias voltages greater than $10V_R$ in practical subharmonic oscillators.

The conductance G_{transfer} in Figure 4c also absorbs power from the pump. This is the power which is converted from pump to oscillator frequency and is given by

$$P_{\text{conv.}} = \frac{V_p^2}{2} G_{\text{transfer}} = \frac{n\omega C_0}{4} \frac{V_p}{V_0} V_s^2. \quad (19)$$

The ratio of pump-frequency power converted to the subharmonic frequency to pump power lost is of great interest and may be written (assuming bias for maximum negative conductance)

$$\frac{P_{\text{conv.}}}{P_0} = \frac{n}{n+1} \frac{Q_s}{8} \frac{V_s^2}{V_p^2}.$$

The subharmonic oscillator cannot be efficient at small subharmonic voltage levels. If high efficiency is desired, V_s must be appreciable compared to V_p . The limitations on V_s as a function of available power from the pump and on efficiency as a function of diode quality factor are discussed in Appendix II.

ACKNOWLEDGMENT

The authors are indebted to R. J. Schwartz for discussions which led to the concept of a limiting forward voltage and to the subsequent formulation of the optimized value of negative conductance.

APPENDIX I—DERIVATION OF THE EXPRESSION FOR NEGATIVE CONDUCTANCE

The capacitance may be written in the form

$$C = C_0 - nC_0 \frac{v}{V_0}.$$

The voltage across this capacitance may be written, for high-Q diodes, as

$$v(t) = V_s \sin(\omega t + \theta) + V_p \sin 2\omega t.$$

Then the current flowing in the diode is

$$\begin{aligned}
 i_c &= C \frac{dv}{dt} = \left\{ C_0 - \frac{nC_0}{V_0} [V_s \sin(\omega t + \theta) + V_p \sin 2\omega t] \right\} [V_s \omega \cos(\omega t + \theta) + 2V_p \omega \cos 2\omega t] \\
 &= C_0 [V_s \omega \cos(\omega t + \theta) + 2V_p \omega \cos 2\omega t] \\
 &\quad - \frac{nC_0}{V_0} [V_s^2 \omega \sin(\omega t + \theta) \cos(\omega t + \theta) + V_p V_s \omega \sin 2\omega t \cos(\omega t + \theta) \\
 &\quad + 2V_p V_s \omega \sin(\omega t + \theta) \cos 2\omega t + 2V_p^2 \omega \sin 2\omega t \cos 2\omega t] \\
 &= \omega C_0 V_s \cos(\omega t + \theta) + 2\omega C_0 V_p \cos 2\omega t \\
 &\quad - \frac{n\omega C_0}{2V_0} [V_s^2 \sin(2\omega t + 2\theta) + V_p V_s \sin(3\omega t + \theta) + V_p V_s \sin(\omega t - \theta) \\
 &\quad + 2V_p V_s \sin(3\omega t + \theta) - 2V_p V_s \sin(\omega t - \theta) + 2V_p^2 \sin 4\omega t].
 \end{aligned}$$

If only terms of frequency ω are considered,

$$\begin{aligned}
 i_\omega &= \omega C_0 V_s \cos(\omega t + \theta) + \frac{n\omega C_0}{2} \frac{V_p V_s}{V_0} \sin(\omega t - \theta) \\
 &= \omega (KV_0^{-n}) V_s \cos(\omega t + \theta) + \frac{n\omega}{2} (KV_0^{-n}) \frac{V_p V_s}{V_0} \sin(\omega t - \theta).
 \end{aligned}$$

This expression is used in the body of this paper to establish the diode equivalent circuit at frequency ω .

Further terms in the Taylor Series expansion for $C(V)$ can be considered. Including the first four terms, the equivalent shunt capacitance and conductance at frequency ω are

$$\begin{aligned}
 C &= C_0 \left[1 + \frac{n(n+1)}{4} \left(\frac{V_p^2}{V_0^2} + \frac{V_s^2}{2V_0^2} \right) \right], \\
 G_- &= -\frac{n\omega C_0}{2} \frac{V_p}{V_0} \left[1 + \frac{(n+1)(n+2)}{12} \left(\frac{3}{2} \frac{V_p^2}{V_0^2} + \frac{V_s^2}{V_0^2} \right) \right].
 \end{aligned}$$

Since V_s will usually be much less than V_p , the terms including V_s can be neglected. Assume that the bias voltage for maximum negative conductance, $V_0 = V_R(1+n)/n$, is used and that $V_p = V_0 - V_R$. There is then an increase of G_- by a factor $(n+2)/[8(n+1)]$ or about 20 per cent and an increase in C by a factor $n/[4(n+1)]$, or less than 10 per cent. Higher-order terms would be expected to make still

smaller corrections and, as shown later, the negative conductance will be further enhanced. Conclusions in this paper based on the first two terms in the expression for $C(V)$ are a good (perhaps slightly pessimistic) approximation.

Consider now a general term in the expression for the current. The effect of each term will now be examined to determine if any saturation, or reduction, of the negative conductance arises from higher-order terms. The series for $C(V)$ is

$$C(V) = \sum_{j=0}^{\infty} (-1)^j \frac{n(n+1) \cdots (n+j-1)}{j!} C_0 \left(\frac{v}{V_0} \right)^j,$$

so the $(j-1)$ st term for the current is of the form

$$\begin{aligned} i_c^{(j-1)} &= (-1)^{(j-1)} \frac{n(n+1) \cdots (n+j-2)}{(j-1)!} C_0 \left(\frac{v}{V_0} \right)^{j-1} \frac{dv}{dt} \\ &= (-1)^{(j-1)} \frac{n(n+1) \cdots (n+j-2)}{(j-1)!} \frac{C_0}{V_0^{j-1}} \frac{d}{dt} (v^j). \end{aligned}$$

The coefficients preceding the derivatives are constants. Since

$$v(t) = V_s \cos \omega t + V_p \sin 2\omega t = v_s + v_p,$$

the terms in v^j are of the form

$$v_s^j; \quad v_s^{j-1}v_p; \quad v_s^{j-2}v_p^2; \quad \cdots; \quad v_s^2v_p^{j-2}; \quad v_s v_p^{j-1}; \quad v_p^j.$$

It can be shown that *no* terms of frequency ω appear in

$$v_p^j, \quad v_p^{j-2}v_s^2, \quad v_p^{j-4}v_s^4, \quad \dots$$

Furthermore, when j is *odd* all of those terms that contain ω are reactive, whereas when j is *even* all ω terms have *negative* conductance. There are, therefore, no positive conductance terms which could produce a saturation effect as V_s increases. It will be necessary to explain saturation in terms of either (1) shunt conductance increase at high V_s , or (2) limitation of available pump power. A derivation for the latter explanation appears in Appendix II.

APPENDIX II—DIODE OSCILLATOR SATURATION, OUTPUT POWER,
AND EFFICIENCY

The most important application of subharmonic oscillators appears to be in computers. For this application minimum rise time is usually desired and therefore the pump voltage applied should be the maximum possible for negligible forward conduction losses and negligible reverse breakdown losses. As the subharmonic voltage amplitude increases, the excursion into the forward bias region becomes increasingly large and increasingly dissipative. Ultimately, positive shunt conductance can increase to the point at which it equals the net negative conductance, whereupon the subharmonic output saturates.

In certain possible applications the maximum subharmonic output *voltage* (or power) may be desired from the oscillator. The pump voltage and bias point can then be so chosen that forward conduction is not the major limiting factor. The maximum output voltage will then depend on impedance levels and on the available power from the pump source.

Equivalent circuits at the pump and oscillator frequencies are shown in Figure 15. Losses at both frequencies are included and the assumption is made that all reactive elements are tuned out. In this appendix, the assumption that $V_s \gg V_p$ will figure prominently. Although the intermediate cases where $V_s \approx V_p$ can also be treated using the equivalent circuit of Figure 15, analysis with the present assumption clarifies the nature of the power limitations and efficiency for a high-output oscillator of controlled phase.

For large V_s , G_{transfer} in Figure 15a is much larger than $G_{s,2\omega}$ and maximum power transfer from the pump will result when

$$G_{\text{transfer}} = G_g.$$

The pump power output will be equal to the available power from the pump

$$P_{av} = \frac{I_g^2}{8G_g}.$$

(Alternating quantities are represented by their peak amplitudes in this paper.) This power will appear in G_- and be transferred (with some loss) to the load. The maximum power into the subharmonic frequency load is, then, the available power from the pump.

For steady state the total conductance in the oscillation frequency circuit is zero so

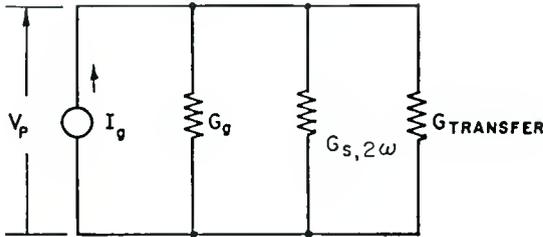
$$G_- + G_{s,\omega} + G_L = 0.$$

The power into the load conductance is

$$P_L = \frac{V_s^2}{2} (-G_- - G_{s,\omega}) = \frac{V_s^2}{2} \left(\frac{n\omega C_0}{2} \frac{V_p}{V_0} - G_{s,\omega} \right).$$

But

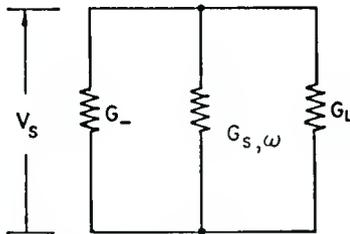
$$V_p = \frac{I_g}{2G_{\text{transfer}}} = I_g \frac{V_0 V_p}{n\omega C_0 V_s^2} = \frac{I_g}{2G_g}.$$



$$G_{S,2\omega} = \frac{4\omega C_0}{Q_s}$$

$$G_{\text{TRANSFER}} = \frac{n\omega C_0}{2} \frac{V_s^2}{V_0 V_p}$$

(a) AT PUMP FREQUENCY, 2ω



$$G_{S,\omega} = \frac{\omega C_0}{Q_s}$$

$$G_- = -\frac{n\omega C_0}{2} \frac{V_p}{V_0}$$

(b) AT OSCILLATION FREQUENCY, ω

Fig. 15—Subharmonic oscillator equivalent circuits.

Therefore

$$P_L = \frac{I_g^2}{8G_g} - \frac{I_g V_0}{2nQ_s} = P_{av} - \frac{I_g V_0}{2nQ_s}.$$

The efficiency is

$$\eta = \frac{P_L}{P_{\text{pump}}} = 1 - \frac{2}{nQ_s} \frac{V_0}{V_p}.$$

Since $V_p \ll V_s$ and $V_s < V_0$ (the voltage cannot be allowed to go too far into forward bias), it is clear that very-high quality factors are required for efficient power conversion. For an abrupt-junction diode where $V_0/V_p = 10$, the quality factor must be $Q_s = 80$ to get 50 per cent efficiency. The conclusion can be drawn, then, that this high-voltage mode of operation is restricted to lower frequencies, where such high quality factors are attainable.

THE PROPAGATION OF PERTURBATIONS ALONG MAGNETICALLY FOCUSED ELECTRON BEAMS*†

BY

F. PASCHKE

RCA Laboratories,
Princeton, N. J.

Summary—The present analysis considers the behavior of beams for magnetic focusing field values between the well known Brillouin field and infinity. By using a model introduced earlier by Zitelli, several discrepancies of previous theories are resolved. The frequency of the radial electron oscillation in an unmodulated scalloping beam is calculated and shown to agree very well with measurements found in the literature. The propagation of high-frequency perturbations in a drift tube is described by two pairs of TM space-charge waves, one with a plasma-frequency-reduction factor smaller than unity (ordinary waves), the other pair with a "reduction" factor larger than unity (extraordinary waves). The gain in traveling-wave-tube amplifiers with thin beams and small space charge is found to be independent of the focusing field strength. For large space charge the gain decreases with increasing magnetic field strength because of the enhanced excitation of passive modes. As a by-product, the maximum current which can pass a drift tube is computed. At the Brillouin field the limiting current rises sharply with the magnetic field. For field strengths larger than about twice the Brillouin field the current increase is insignificant.

INTRODUCTION

THE theoretical evaluation of amplification or generation of signals in klystrons and traveling-wave tubes requires an understanding of the way a perturbation propagates along an ion-free magnetically focused electron beam. For the ideal but impractical case of infinite magnetic focusing field, the theories seem to be satisfactory. However, this is not so for finite magnetic focusing fields, and a number of analyses have been published which have led to contradictory results. One major difficulty arises from the disagreement between the force equations used by different authors. For the well known Brillouin flow with rotational symmetric TM fields, Labus,¹

* A very condensed form of this paper has been presented at the 16th Annual Conference on Electron Tube Research, Quebec, Canada, June, 1958.

† Manuscript received December 30, 1958.

¹ J. Labus, "Einfluss der Lorentzkraft auf die Raumladungswellen im Elektronenstrahl," *Arch. d. elektr. Übertrag.*, Vol. 7, p. 88, February, 1953.

Rigrod and Lewis,² and others³⁻⁶ obtained

$$\ddot{\bar{r}} = \eta \bar{E}_r, \quad (\text{I})$$

where \bar{r} is the radial displacement of an electron from its steady-state path, \bar{E}_r the radial perturbation field, and η the charge-to-mass ratio of an electron. Zitelli,⁷ however, derived

$$\ddot{\bar{r}} + \left(\frac{\omega_c}{2} \right)^2 \bar{r} = \eta \bar{E}_r, \quad (\text{II})$$

where ω_c is the cyclotron frequency. In a later erratum⁸ Zitelli, finding himself in disagreement with other authors, revoked the validity of his Equation (II). There are the following reasons for this discrepancy. In Equation (I), \bar{E}_r is understood to be the perturbation of the radial electric field at *constant radius* of the perturbed beam. In Equation (II), however, \bar{E}_r is the perturbation of the radial electric field seen by an element of the beam *during its radial displacement*. The question arises as to which equation should be combined with the field equations to solve the boundary-value problem. This problem is complicated by the fact that the surface of the beam is rippled and changing with time. To overcome this difficulty one must find a rectilinear equivalent beam, i.e., a beam which is not displaced radially. In References (2) and (3), where Equation (I) is used, the ripples of the beam are claimed to be taken fully into account by the introduction of a surface current. In References (4)-(6), which also use Equation (I), it was realized that the ripples inside of the beam must also be removed. The method employed seems to be rather

² W. W. Rigrod and J. A. Lewis, "Wave Propagation Along a Magnetically-Focused Cylindrical Electron Beam," *Bell Sys. Tech. Jour.*, Vol. 33, p. 399, March, 1954.

³ G. R. Brewer, "Some Effects of Magnetic Field Strength on Space-Charge-Wave Propagation," *Proc. I.R.E.*, Vol. 44, p. 896, July, 1956.

⁴ J. Labus and K. Pöschl, "Raumladungswellen in ionenfreien Elektronenstrahlen," *Arch. d. elektr. Übertrag.*, Vol. 9, p. 39, January, 1955.

⁵ R. Liebscher, "Raumladungswellen bei endlichem Magnetfeld an der Kathode einer zylindrischen Elektronenströmung," *Arch. d. elektr. Übertrag.*, Vol. 11, p. 214, May, 1957.

⁶ J. Labus, "Space-Charge Waves Along Magnetically-Focused Electron Beams," *Proc. I.R.E.*, Vol. 45, p. 854, June, 1957.

⁷ L. T. Zitelli, "Space Charge Effects in Gridless Klystrons," *Stanford University Microwave Laboratory*, Report No. 149, October, 1951.

⁸ L. T. Zitelli, Errata to Reference 7, July, 1952.

inconsistent.⁹ Zitelli's original analysis⁷ implies that the radial electric field at a constant radius of the rectilinear equivalent beam is identical with the field seen by the corresponding element of the rippled beam during its radial displacement. Consequently, in Reference (7) Equation (II) was combined with the field equations. The following facts support this approach:

(1) Equation (I), in contrast to Equation (II), leads to a dispersion equation which is independent of the radial boundary conditions — unusual for a system of finite extent. A pair of so-called "intrinsic" space-charge waves result with propagation constants equal to that of a confined stream of infinite extent ($B = \infty$) and arbitrary radial field distributions, regardless of frequency and geometry. This is unrealistic and can hardly be accepted.

(2) The excitation of an electron beam is completely determined by the three velocity components \bar{v}_r , \bar{v}_ϕ , \bar{v}_z , by the space-charge density $\bar{\rho}$, and by the radial displacement \bar{r} , all quantities being given at a certain axial position $z = z_0$. If the beam travels in a perfect steady state before it enters the modulating field there exists a linear relation between \bar{r} and \bar{v}_ϕ . Consequently, at least four modes are required to meet the boundary conditions at $z = z_0$. Two pairs of waves can be expected to exist because a magnetically focused electron beam inherently has two resonances, the plasma resonance and the cyclotron resonance. There *must* be a pair of modes associated with *each* resonance. The theories which use Equation (I) predict only one pair of waves because the "intrinsic" solutions criticized above lead to inconsistencies and must be abandoned.¹⁰ Also in Reference (7) only two fundamental modes of propagation have been reported. It will be shown in the present paper that a consequent application of Equation (II) resolves this discrepancy.

(3) Theories which are based on Equation (I) predict no wave propagation for the case of the beam completely filling the drift tube. This prediction is not reasonable because for extremely thick beams one expects the propagation constants to be approximately equal to that of a beam of infinite extent, independent of the location of the drift tube. Application of Equation (II) again resolves this discrepancy.

⁹ W. W. Rigrod and J. Labus, "Space-Charge Waves Along Magnetically-Focused Electron Beams," *Proc. I.R.E.*, Vol. 46, p. 358, January, 1958.

¹⁰ A. H. W. Beck, *Space Charge Waves and Slow Electromagnetic Waves*, Pergamon Press, New York, N. Y., 1958, p. 116.

(4) With the aid of either Equation (I) or (II) and the field equations for stationary processes ($\omega = 0$) the scalloping of a beam can be calculated. Experimental investigations¹¹⁻¹³ show very good agreement with the result obtained from Equation (II) (see Figure 8).

In the present paper, the force Equation (II) is generalized so as to include all values of the magnetic field between Brillouin field and infinity, and then combined with the field equations. The resulting differential equation is solved for space-charge-wave propagation in a drift tube and in a slow-wave structure. Formation and maintenance of the beam are also discussed.*

DERIVATION OF THE SPACE-CHARGE-WAVE EQUATION

For the derivation of the space-charge-wave equation the following assumptions are made:

(1) Only first-order perturbations are taken into account.

(2) The electron velocity is very much smaller than the velocity of light.

(3) The thermal velocity spread is neglected.

(4) A rotational symmetrical structure surrounds the solid cylindrical electron beam.

(5) The electric field strength can be derived from a scalar potential. This is tantamount to describing the field distribution by a "slow" TM wave with $\vec{E}_\phi = 0$ and rotational symmetry. Justification for this assumption is demonstrated at the end of this section.

In the Lagrangian description of an electron fluid, each particle is labeled by its position (r_0, ϕ_0, z_0) at a time $t = t_0$. The motion of an individual particle in a homogeneous axial magnetic focusing field B is described by the force equations

$$\ddot{r} - r\dot{\phi}^2 = \eta(E_r + r\dot{\phi}B) \quad (1a)$$

¹¹ J. Berghammer, "Die Welligkeit eines magnetisch geführten Elektronenstrahles," Dipl. Ing. Thesis, Technische Hochschule Vienna, March, 1954.

¹² H. Schnitger, "Die Messung der Welligkeit des Elektronenstrahles einer Wanderfeldröhre," *Arch. d. elektr. Übertrag.*, Vol. 7, p. 415, September, 1953.

¹³ J. D. Lawson, "Some Experiments on a Cylindrical Electron Beam Constrained by a Magnetic Field," *Jour. Electron.*, Vol. 1, p. 43, July, 1955.

* Note added in proof: In the recent publication by R. H. C. Newton, "On Space Charge Waves," *Jour. Electronics and Control*, Vol. 5, p. 510, December, 1958, a wave equation for Brillouin flow has been derived which agrees with the present analysis.

$$r\ddot{\phi} + 2\dot{r}\dot{\phi} = -\eta B\dot{r} \quad (1b)$$

$$\ddot{z} = \eta E_z. \quad (1c)$$

Each superscript dot indicates one differentiation with respect to time at constant r_0 , ϕ_0 , and z_0 . Because of the assumption of a non-relativistic beam, the forces produced by the r-f magnetic field and the d-c magnetic field of the electron current have been neglected. Steady-state solutions are sought for which the radial velocity is zero and the axial velocity constant. From Equations (1)

$$\dot{r} = 0, \quad r = r_0, \quad (2a)$$

$$\ddot{z} = 0, \quad \dot{z} = v_0, \quad (2b)$$

$$\ddot{\phi} = 0, \quad \dot{\phi} = -\Omega_0. \quad (2c)$$

$$r_0\Omega_0^2 + \eta E_{r0} = r_0\omega_c\Omega_0. \quad (3)$$

Here

$$\omega_c = \eta B \quad (4)$$

is the cyclotron frequency. To facilitate a perturbation analysis, the angular velocity Ω_0 and the axial velocity v_0 are assumed to be independent of r_0 . It will be shown in the last section of this paper that these assumptions are justifiable in most practical cases.

The steady-state electric field strength E_{r0} in Equation (3) must be consistent with Poisson's equation:

$$\frac{d E_{r0}}{d r_0} + \frac{E_{r0}}{r_0} = \frac{\rho_0}{\epsilon_0}. \quad (5)$$

With the space-charge density ρ_0 assumed constant across the beam, the proper solution of Equation (5) is

$$\eta E_{r0} = \frac{\omega_p^2}{2} r_0, \quad (6)$$

where ω_p is the plasma frequency defined by

$$\omega_p^2 = \frac{\eta \rho_0}{\epsilon_0}. \quad (7)$$

Equations (6) and (3) combined yield the steady-state condition

$$\Omega_0^2 + \frac{\omega_p^2}{2} = \omega_c \Omega_0. \tag{8}$$

Equation (8) states that at any radius the centrifugal force and the space-charge force are balanced by the magnetic force.

Suppose the steady state is perturbed slightly. Denoting the perturbations by tildes, one can write

$$r = r_0 + \tilde{r}(t, r_0, z_0), \tag{9a}$$

$$\dot{\phi} = -\Omega_0 + \dot{\tilde{\phi}}(t, r_0, z_0), \tag{9b}$$

$$\dot{z} = v_0 + \dot{\tilde{v}}_z(t, r_0, z_0), \tag{9c}$$

$$E_r = E_{r0} + \tilde{E}_r(t, r_0, z_0), \tag{9d}$$

$$E_z = \tilde{E}_z(t, r_0, z_0). \tag{9e}$$

With all second- and higher-order perturbations neglected one finds from the force Equations (1) with the aid of the steady-state solutions, Equations (2), and the steady-state condition, Equation (8)

$$\ddot{\tilde{r}} + k^2 \omega_p^2 \tilde{r} + (2\Omega_0 - \omega_c) C = \eta \tilde{E}_r, \tag{10a}$$

$$r_0 \dot{\tilde{\phi}} = (2\Omega_0 - \omega_c) \tilde{r} + C, \tag{10b}$$

$$\dot{\tilde{v}}_z = \eta \tilde{E}_z. \tag{10c}$$

The factor, k^2 , defined by

$$k^2 = \left(\frac{\omega_c}{\omega_p} \right)^2 - \frac{3}{2}, \tag{11}$$

determines, for a given space-charge density, the radial eigen-frequency of the electrons. k^2 is therefore a measure of the stiffness of the beam, or quality of the focusing system. C is an integration constant. Equations (10) could have been derived also in Eulerian coordinates which will be used henceforth. The superscript dot now indicates the total time derivative. It should be recalled here that \tilde{E}_r is the perturbation of the radial electric field seen by the displaced

stream element at the radius $r + \bar{r}$. The displacement of the stream, \bar{r} , in Eulerian coordinates is a function of time and space whereas the displacement of an individual fluid particle in Lagrangian coordinates (Equation (9a)) is a function of time and initial position. The usual assumption is made that the perturbations vary as

$$\exp(j\omega t - \Gamma z). \quad (12)$$

This assumption is strictly justifiable only for streams with smooth cylindrical surfaces. Thus the rippled beam must be replaced by an equivalent rectilinear beam, i.e., a beam which is not displaced radially. Figure 1a shows a section of a beam element which in the unperturbed stream is bounded by the cylindrical surfaces $r = r_0$, and $r = r_0 - \Delta r_0$. The rectilinear equivalent is depicted in Figure 1b. Equivalence exists if the currents carried by the two elements are identical.

Equalization of the displacement currents leads to the conclusion that *the electric field seen by the rectilinear element is the same as that seen by the rippled element*. The axial convection current carried by the rippled element is

$$\Delta I_z = 2\pi \int_{r_0 - \Delta r_0 + \bar{r}(r_0 - \Delta r_0)}^{r_0 + \bar{r}(r_0)} (i_0 + \tilde{i}_z) r dr. \quad (13)$$

Linearization in both Δr_0 and the r-f quantities leads to

$$\Delta I_z = 2\pi r_0 i_0 \Delta r_0 + 2\pi r_0 \tilde{i}_z(r_0) \Delta r_0 + 2\pi i_0 \left. \frac{\partial(r\bar{r})}{\partial r} \right|_{r=r_0} \Delta r_0. \quad (14)$$

The first term is the d-c current. The second term is due to a density modulation. Both terms appear also in the rectilinear element if the axial convection-current densities in the two elements are identical. The remaining third term in Equation (14) originates from the radial displacement of the particles. In the rectilinear element one can interpret this last term as a superposition of two surface currents,

$$\tilde{I}_{s1} = 2\pi r_0 i_0 \bar{r}(r_0) \quad (15a)$$

at $r = r_0$, and

$$\begin{aligned}
 \tilde{I}_{s2} &= -2\pi (r_0 - \Delta r_0) i_0 \left(\tilde{r}(r_0) - \frac{\partial \tilde{r}}{\partial r} \bigg|_{r=r_0} \Delta r_0 \right) \\
 &= -\tilde{I}_{s1} + 2\pi i_0 \frac{\partial (r\tilde{r})}{\partial r} \bigg|_{r=r_0} \Delta r_0
 \end{aligned} \tag{15b}$$

at $r = r_0 - \Delta r_0$. If all rectilinear stream elements are put together to form the beam one can see that the "surface" currents within the beam cancel and there remains only the surface current at the boundary of the rectilinear beam, $r = b$;

$$\tilde{I}_s = 2\pi b i_0 \tilde{r}(b). \tag{16}$$

Finally, the radial convection current which in the real beam is caused by the *finite* radial displacements must be replaced in the rectilinear model by an equal radial current caused by *infinitesimal* lateral motions (dipole layers). The boundary conditions at the surface of the rectilinear model can now be stated. The axial electric field must be

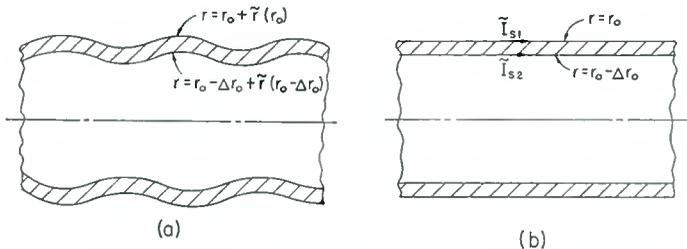


Fig. 1—Annular element of the perturbed beam (a) and its rectilinear equivalent (b). The rippled element is bounded by the time-varying surfaces $r = r_0 + \tilde{r}(r_0)$, and $r = r_0 - \Delta r_0 + \tilde{r}(r_0 - \Delta r_0)$, the rectilinear element by the stationary surfaces $r = r_0$, and $r = r_0 - \Delta r_0$.

continuous while the surface current causes a discontinuity in the tangential magnetic field. Denoting the perturbations inside of the beam with subscripts I and outside of the beam with subscripts II, one can write

$$\tilde{E}_{zII}(b) = \tilde{E}_{zI}(b), \tag{17a}$$

$$\tilde{H}_{\phi II}(b) = \tilde{H}_{\phi I}(b) + i_0 \tilde{r}(b). \tag{17b}$$

Zitelli⁷ combined Equations (10) for Brillouin flow, $k^2 = 1/2$, with

the field equations. This implies that the rippled beam is replaced by the rectilinear equivalent described above. In the force equations of the theories of Labus,¹ Rigrod and Lewis,² and others,³⁻⁶ the radial electric r-f field is not the perturbation field seen by the perturbed stream element at a radius $r + \bar{r}$, as in Equation (10a), but is rather understood as the perturbation field at constant radius of the perturbed beam. In Reference (2) the radial displacement of the stream is claimed to be taken fully into account by the introduction of the surface current from Equation (16). This model, however, is not acceptable, for it leads to the discrepancies outlined earlier. Labus and Pöschl,⁴ and Labus,⁶ realized that the ripples inside of the beam must also be removed. However, their method fails to resolve the discrepancies discussed and furthermore, seems to be rather inconsistent.⁹

From Equations (10) and Relation (12) the Eulerian velocities are, for $C = 0$,

$$\bar{v}_r = \frac{v_0 (j\beta_e - \Gamma)}{v_0^2 (j\beta_e - \Gamma)^2 + k^2 \omega_p^2} \eta \bar{E}_r, \quad (18a)$$

$$\bar{v}_\phi = - \frac{\omega_p^2 / 2\Omega_0}{v_0^2 (j\beta_e - \Gamma)^2 + k^2 \omega_p^2} \eta \bar{E}_r, \quad (18b)$$

$$\bar{v}_z = \frac{1}{v_0 (j\beta_e - \Gamma)} \eta \bar{E}_z, \quad (18c)$$

where $\beta_e = \omega/v_0$. Note that \bar{v}_ϕ is the perturbation of the angular velocity of a stream element during its radial displacement. \bar{v}_ϕ therefore consists of the two terms $r\dot{\phi} - \Omega_0\bar{r}$. The integration "constant" C in Eulerian coordinates can be a quantity which varies as $\exp(j\omega t - j\beta_e z)$. Thus one might expect a so-called kinematic wave to exist which propagates with the drift velocity of the beam. Such a kinematic solution, however, leads to inconsistencies with the field equations and must be abandoned. The analysis is therefore limited to cases where $C = 0$. This is tantamount to the assumption that the beam starts off in a perfect steady state where both \bar{r} and \bar{v}_ϕ are zero. For most applications this is a fair assumption, perhaps noise computations excepted. Because of the proportionality between \bar{r} and $\dot{\phi}$ from Equation (10b), there are the four independent variables to describe the excitation of the beam: space-charge density, $\bar{\rho}$, and the velocity components \bar{v}_r , \bar{v}_ϕ , \bar{v}_z . Thus at least four modes of propagation are

expected to exist in a magnetically focused electron beam. Because of the assumption that the electric field is derivable from a scalar potential,

$$\bar{E}_r = -\frac{1}{\Gamma} \frac{\partial \bar{E}_z}{\partial r}. \tag{19}$$

As a consequence of the continuity equation,

$$-\frac{\partial \bar{\rho}}{\partial t} = \rho_0 \operatorname{div} \vec{v} + v_0 \frac{\partial \bar{\rho}}{\partial z}, \tag{20}$$

the Poisson equation,

$$\operatorname{div} \vec{E} = \frac{\bar{\rho}}{\epsilon_0}, \tag{21}$$

and Equations (18) and Expression (12), it can be shown that \bar{E}_z must satisfy the familiar Bessel-differential equation

$$\frac{\partial^2 \bar{E}_z}{\partial r^2} + \frac{1}{r} \frac{\partial \bar{E}_z}{\partial r} + \gamma^2 \bar{E}_z = 0, \tag{22}$$

where the radial wave number γ is given by

$$\begin{aligned} \gamma^2 &= \Gamma^2 \left[1 + \left(\frac{\beta_p}{j\beta_c - \Gamma} \right)^2 \frac{k^2}{\left(\frac{j\beta_c - \Gamma}{\beta_p} \right)^2 + 1 + k^2} \right] \\ &= \Gamma^2 \frac{\left[\left(\frac{j\beta_c - \Gamma}{\beta_p} \right)^2 + 1 \right] \left[\left(\frac{j\beta_c - \Gamma}{\beta_p} \right)^2 + k^2 \right]}{\left(\frac{j\beta_c - \Gamma}{\beta_p} \right)^2 \left[\left(\frac{j\beta_c - \Gamma}{\beta_p} \right)^2 + 1 + k^2 \right]} \end{aligned} \tag{23}$$

β_p is the plasma-wave number ω_p/v_0 . According to Equation (11), $k^2 = \infty$ for confined flow. Then Equation (23) reduces to the well-known expression

$$\gamma^2 = \Gamma^2 \left[1 + \left(\frac{\beta_p}{j\beta_c - \Gamma} \right)^2 \right]. \tag{24}$$

For a solid beam the only permissible solution of Equation (22) is

$$\tilde{E}_z = A J_0(\gamma r) \exp(j\omega t - \Gamma z) \quad (25)$$

where A is the amplitude of the electric field strength at the axis and J_n the Bessel function of n th order. All other field components can be derived from Equation (25). One finds from Equation (19)

$$\tilde{E}_r = A \frac{\gamma}{\Gamma} J_1(\gamma r) \exp(j\omega t - \Gamma z). \quad (26)$$

The space-charge density is, from Equation (21),

$$\tilde{\rho} = \left(\frac{\beta_p}{j\beta_c - \Gamma} \right)^2 \frac{k^2 \epsilon_0 \Gamma}{\left(\frac{j\beta_c - \Gamma}{\beta_p} \right)^2 + 1 + k^2} A J_0(\gamma r) \exp(j\omega t - \Gamma z). \quad (27)$$

Equations (18), (25), and (26) combined yield for the velocities

$$\bar{v}_r = \frac{j\beta_e - \Gamma}{\beta_p} \frac{\gamma}{\omega_p \Gamma} \eta A J_1(\gamma r) \exp(j\omega t - \Gamma z), \quad (28a)$$

$$\bar{v}_\phi = - \frac{1}{\left(\frac{j\beta_e - \Gamma}{\beta_p} \right)^2 + k^2} \frac{\gamma}{2\Omega_0 \Gamma} \eta A J_1(\gamma r) \exp(j\omega t - \Gamma z), \quad (28b)$$

$$\bar{v}_z = \frac{\beta_p}{j\beta_e - \Gamma} \frac{1}{\omega_p} \eta A J_0(\gamma r) \exp(j\omega t - \Gamma z). \quad (28c)$$

From Equation (28a) and Expression (12), one obtains, for the radial displacement,

$$\bar{r} = \frac{\bar{v}_r}{v_0(j\beta_e - \Gamma)} = \frac{1}{\left(\frac{j\beta_e - \Gamma}{\beta_p} \right)^2 + k^2} \frac{\gamma}{\omega_p^2 \Gamma} \eta A J_1(\gamma r) \exp(j\omega t - \Gamma z). \quad (29)$$

From Equations (2), (27), and (28), the convection-current densities are

$$\bar{i}_r = \rho_0 \bar{v}_r = \frac{j\beta_e - \Gamma}{\beta_p} \frac{\omega_p \epsilon_0 \gamma}{\Gamma} A J_1(\gamma r) \exp(j\omega t - \Gamma z), \quad (30a)$$

$$\bar{i}_\phi = \rho_0 \bar{v}_\phi - \Omega_0 r \bar{p} = - \left[\frac{1}{\left(\frac{j\beta_e - \Gamma}{\beta_p} \right)^2 + k^2} \frac{\omega_p^2 \epsilon_0 \gamma}{2\Omega_0 \Gamma} J_1(\gamma r) \right. \quad (30b)$$

$$\left. + \frac{\left(\frac{\beta_p}{j\beta_e - \Gamma} \right)^2 k^2}{\left(\frac{\beta_p}{j\beta_e - \Gamma} \right)^2 + 1 + k^2} \Omega_0 \epsilon_0 \Gamma r J_0(\gamma r) \right] A \exp(j\omega t - \Gamma z),$$

$$\bar{i}_z = \rho_0 \bar{v}_z + v_0 \bar{p} = \frac{\left(\frac{\beta_p}{j\beta_e - \Gamma} \right)^2 k^2 - j \frac{\omega_p}{\omega} \left(\frac{j\beta_e - \Gamma}{\beta_p} + \frac{\beta_p}{j\beta_e - \Gamma} \right)}{\left(\frac{j\beta_e - \Gamma}{\beta_p} \right)^2 + 1 + k^2} j\omega \epsilon_0 A J_0(\gamma r) \exp(j\omega t - \Gamma z). \quad (30c)$$

Finally, from the first Maxwellian equation,

$$\text{curl } \vec{H} = \vec{i} + j\omega \epsilon_0 \vec{E}, \quad (31)$$

one finds

$$\vec{H}_\phi = \frac{1}{\Gamma} (j\omega \epsilon_0 \vec{E}_r + \bar{i}_r). \quad (32)$$

With the aid of Equations (26) and (30a),

$$\vec{H}_\phi = j\omega \epsilon_0 \frac{\gamma}{\Gamma^2} \left(1 - j \frac{\omega_p}{\omega} \frac{\beta_p}{\left(\frac{j\beta_e - \Gamma}{\beta_p} \right)^2 + k^2} \right) A J_1(\gamma r) \exp(j\omega t - \Gamma z). \quad (33)$$

The second term in the brackets is the ratio of radial convection current to radial displacement current.

It remains to be proven that the electric fields can be derived from a scalar potential. This is tantamount to showing that Equation (19) is a valid approximation, and that \tilde{E}_ϕ is negligible. From Maxwell's first Equation (31) and second equation,

$$\text{curl } \vec{E} = -j\omega\mu_0\vec{H}, \quad (34)$$

and Expression (12), the correct equation relating the radial and axial electric fields is

$$\tilde{E}_r \left(1 + \frac{\omega^2\epsilon_0\mu_0}{\Gamma^2} \right) = -\frac{1}{\Gamma} \frac{\partial \tilde{E}_z}{\partial r} + j \frac{\omega\mu_0}{\Gamma^2} \tilde{i}_r. \quad (35)$$

With

$$\tilde{i}_r = \rho_0 \tilde{v}_r,$$

and Equation (18a)

$$\tilde{E}_r \left(1 + \frac{\omega^2\epsilon_0\mu_0}{\Gamma^2} \right) = -\frac{1}{\Gamma} \frac{\partial \tilde{E}_z}{\partial r} + j \frac{\omega^2\epsilon_0\mu_0}{\Gamma^2} \frac{\omega_p}{\omega} \frac{j\beta_c - \Gamma}{\beta_p} \tilde{E}_r. \quad (36)$$

For slow waves

$$\left| \frac{\omega^2\epsilon_0\mu_0}{\Gamma^2} \right| \ll 1. \quad (37)$$

Equation (19) is a good approximation if

$$\left(\frac{j\beta_c - \Gamma}{\beta_p} \right)^2 + k^2 \neq 0, \quad (38)$$

since in most practical cases

$$\frac{\omega_p}{\omega} \ll 1. \quad (39)$$

Furthermore, if

$$\left(\frac{j\beta_e - \Gamma}{\beta_p}\right)^2 + k^2 \rightarrow 0, \tag{40}$$

then from Equation (23)

$$\gamma^2 = \Gamma^2 \frac{k^2 - 1}{k^2} \left[\left(\frac{j\beta_c - \Gamma}{\beta_p}\right)^2 + k^2 \right] \rightarrow 0, \tag{41}$$

and therefore, from Equation (26),

$$\tilde{E}_r = \frac{\gamma^2 r}{2\Gamma} A \exp(j\omega t - \Gamma z) \rightarrow 0. \tag{42}$$

Thus the last term in Equation (36) is finite and small even at the eigenvalue excluded by Relation (38). To complete the proof the excitation of the angular electric field by the convection-current density from Equation (30b) must be considered. From Equations (31) and (34) one can readily find, under the slow-wave condition, Relation (37), the inhomogeneous Bessel equation

$$\frac{\partial^2 \tilde{E}_\phi}{\partial r^2} + \frac{1}{r} \frac{\partial \tilde{E}_\phi}{\partial r} + \tilde{E}_\phi \left(\Gamma^2 - \frac{1}{r^2} \right) = j\omega \mu_0 \tilde{i}_\phi. \tag{43}$$

With the “driving-current” \tilde{i}_ϕ given by Equation (30b), a particular integral of Equation (43) is

$$\tilde{E}_\phi = (C_1 J_1(\gamma r) + C_2 \Gamma r J_0(\gamma r)) A \exp(j\omega t - \Gamma z), \tag{44}$$

where

$$C_1 = j \frac{2\gamma}{\Gamma} \left(\frac{\omega}{k\Gamma c}\right)^2 \frac{\Omega_0}{\omega} \left(\frac{j\beta_c - \Gamma}{\beta_p}\right)^2 \left[\left(\frac{j\beta_e - \Gamma}{\beta_p}\right)^2 + 1 + k^2 \right] \left[\frac{(\omega_p/2\Omega_0)^2}{\left(\frac{j\beta_e - \Gamma}{\beta_p}\right)^2 + k^2} - 1 \right],$$

and

$$C_2 = j \left(\frac{\omega}{\Gamma c}\right)^2 \frac{\Omega_0}{\omega}, \tag{46}$$

where c is the velocity of light. The minimum value of the focusing parameter k^2 is 0.5, that is, the value for Brillouin flow. In all practical beams $|\Gamma r|$ does not exceed unity by orders of magnitude. Furthermore, in practice,

$$\frac{\Omega_0}{\omega} \ll 1. \quad (47)$$

Thus, under the slow-wave condition, Relation (37), the field strength \vec{E}_ϕ is negligible compared to \vec{E}_z and \vec{E}_r , from Equations (25) and (26). Even at the limit, Expression (40), the first term in the bracket of Equation (44) does not exhibit a pole and is small because of the Relation (41).

It can therefore be concluded, that for slow waves, the electric field strength is in all practical cases derivable from a scalar potential. Thus assumption No. 5, made earlier, is justified.

SPACE-CHARGE-WAVE PROPAGATION IN A DRIFT TUBE

Solutions are sought for which the axial electric field vanishes at the drift-tube wall, $r = a$. Equations (17) must be valid at the steady-state beam surface, $r = b$. These boundary conditions lead to a transcendental dispersion equation which is difficult to solve except for the case where the beam fills the drift tube completely ($b = a$). For the sake of simplicity it is assumed that $b = a$ so that no electrons are intercepted by the drift tube when the beam is modulated. Then, from Equations (25) and (17a),

$$J_0(\gamma b) = 0. \quad (48)$$

From Equation (23)

$$p^2 = \frac{1 + k^2}{2} \left(1 \pm \sqrt{1 - \frac{4k^2}{(1 + k^2)^2 (1 - (\gamma_n b / \Gamma b)^2)}} \right) \quad (49)$$

where p is the plasma-frequency-reduction factor defined by

$$p^2 = - \left(\frac{j\beta_e - \Gamma}{\beta_p} \right)^2. \quad (50)$$

$\gamma_n b$ is a root (eigenvalue) of Equation (48). In References (1)-(6), in contrast to Reference (7) and the present analysis, no space-charge-wave propagation is predicted for Brillouin flow ($k^2 = 1/2$) if $b = a$. The physical unreality of such a result becomes obvious for very large

beam radii where the propagation constant must be approximately equal to that of the infinitely extended beam regardless of the proximity of the drift tube.

Equation (48) has an infinite number of roots (eigenvalues). For the sake of simplicity, only one will be considered in the following discussion. Figure 2 shows typical solutions for various values of the focusing parameter, k^2 . In addition to the familiar pair of waves with $|p| < 1$, there is another pair with a "reduction" factor $|p| > 1$. The two pairs of waves can be interpreted by considering the limits of $|p|$ for large beam radii. For

$$\frac{1}{2} \leq k^2 \leq 1, \quad (51)$$

the effective plasma frequency of the lower branch approaches $k\omega_p$, and for

$$1 \leq k^2 \leq \infty, \quad (52)$$

the effective plasma frequency of the upper branch approaches $k\omega_p$; $k\omega_p$ is, from Equation (10a), the radial eigen-frequency of the electrons. Under the Condition (51) the upper branch approaches the plasma frequency, ω_p , and under the Condition (52) the lower branch approaches the same frequency. It can thus be said that each solution is associated with one of the two possible resonance frequencies of the system, i.e., the radial frequency $k\omega_p$ and the plasma frequency ω_p . Plasma waves of this type have been found by Schumann,¹⁴ and Gould and Trivelpiece.¹⁵ This behavior is typical for systems with two resonances. For example, the dispersion curve of lattice vibrations in a one-dimensional crystal with two kinds of atoms exhibits two branches, the "acoustical" and the "optical" branches, in a manner very similar to that shown in Figure 2. If the beam does not drift ($v_0 = 0$), from Equation (50), $p = \omega/\omega_p$. Then the dispersion of the upper branch is anomalous because the group velocity becomes negative. The dispersion of the lower branch stays normal. In accordance with the plasma theory¹⁴ it is suggested that the waves associated with the lower and upper branches in Figure 2 be called *ordinary* and *extraordinary* waves, respectively. It is note-

¹⁴ W. O. Schumann, "Über Wellenausbreitung im Plasma zwischen zwei unendlich gut leitenden Ebenen in Richtung eines aufgeprägten äusseren Magnetfeldes," *Zeitschr. Angew. Phys.*, Vol. 8, p. 482, October, 1956.

¹⁵ R. W. Gould and A. W. Trivelpiece, "A New Mode of Wave Propagation on Electron Beams." *Symposium on Electronic Waveguides*, Polytech. Inst. Brooklyn, April, 1958.

worthy that Equation (49) is valid also for $b/a \neq 1$ if $\gamma_n b$ is understood as the root (eigenvalue) of the dispersion equation appropriate to the specific value of b/a . The nature of the solutions shown in Figure 2 is therefore independent of b/a .

The degree to which the various modes are excited depends on the way the beam is modulated. For an illustrative example the following assumptions are made:

(1) At $z=0$ the beam is velocity modulated by an ideal double layer. Thus at the input plane only *axial* r-f velocities are present.

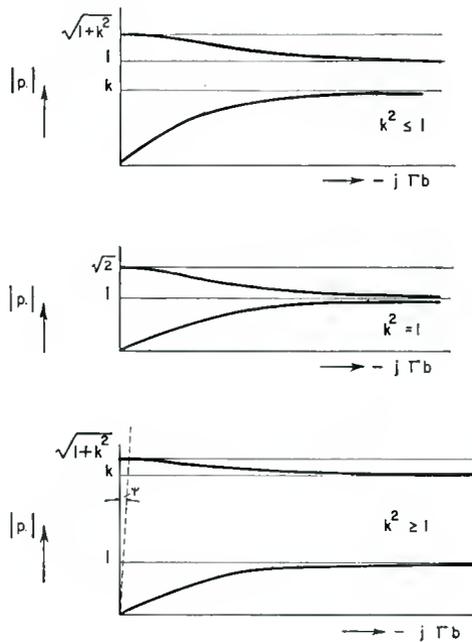


Fig. 2—Typical solutions for the plasma-frequency “reduction” factor $|\rho|$ versus normalized beam radius $-j\Gamma b$. The focusing parameter k is given by Equation (11).

(2) Only modes associated with the first root of Equation (48) (fundamental modes) are taken into account.

(3) The “reduction” factors do not exceed unity by orders of magnitude.

(4) $\omega_p/\omega \ll 1$.

As a consequence of the latter two assumptions

$$\Gamma \doteq j\beta_c. \quad (53)$$

The boundary conditions at $z=0$ are met by a proper superposition of the four fundamental modes. One obtains

$$\bar{v}_r = \frac{v_1 \gamma J_1(\gamma r)}{\beta_c k^2} \frac{p_1^2 p_2^2}{p_2^2 - p_1^2} (\cos \beta_p p_1 z - \cos \beta_p p_2 z) \exp j(\omega t - \beta_c z), \quad (54)$$

$$\bar{v}_\phi = -\frac{v_1 \omega_p \gamma J_1(\gamma r)}{2\Omega_0 \beta_c} \frac{p_1 p_2^2}{p_2^2 - p_1^2} (\sin \beta_p p_1 z - p_1/p_2 \sin \beta_p p_2 z) \exp j(\omega t - \beta_c z), \quad (55)$$

$$\bar{v}_z = v_1 J_0(\gamma r) \frac{p_2^2(p_2^2 - 1)}{k^2(p_2^2 - p_1^2)} \left(\cos \beta_p p_1 z + \frac{p_1^2(1 - p_1^2)}{p_2^2(p_2^2 - 1)} \cos \beta_p p_2 z \right) \exp j \left(\omega t - \beta_c z - \frac{\pi}{2} \right) \quad (56)$$

$$\bar{i}_z = \frac{v_1 \omega_p \epsilon_0 J_0(\gamma r)}{\eta} \frac{p_2^2 - 1}{p_1(p_2^2 - p_1^2)} \left(\sin \beta_p p_1 z + \frac{p_1(1 - p_1^2)}{p_2(p_2^2 - 1)} \sin \beta_p p_2 z \right) \exp j(\omega t - \beta_c z). \quad (57)$$

v_1 is the value of the velocity excitation at the axis of the input plane. p_1 and p_2 , both positive numbers, are the absolute values of the reduction factors of the two pairs of fundamental modes; p_1 and p_2 may be interchanged without changing Equations (54)-(57). As a consequence of Approximation (53) there is, from Equation (49), the following relation between p_1 and p_2 :

$$p_2^2 + p_1^2 = 1 + k^2. \quad (58)$$

Of particular interest, besides the quantities given by Equations (54)-(57), are the ripples of the beam and the total r-f convection current. The ripples from Equation (29), are given by

$$\bar{r} = \frac{v_1 \gamma J_1(\gamma r)}{k^2 \omega_p \beta_c} \frac{p_2^2 p_1}{p_2^2 - p_1^2} (\sin \beta_p p_1 z - p_1/p_2 \sin \beta_p p_2 z) \exp j(\omega t - \beta_c z). \quad (59)$$

The total axial r-f convection current

$$\tilde{I} = 2\pi \int_{r=0}^{b+\tilde{r}(b)} (i_0 + \tilde{i}_z) r dr - i_0 b^2 \pi, \quad i_0 = \rho_0 v_0 \quad (60)$$

consists of two terms,

$$\tilde{I}' = 2\pi \int_0^b \tilde{i}_z r dr, \quad (61)$$

and

$$\tilde{I}'' = 2\pi b i_0 \tilde{r}(b). \quad (62)$$

The first term is due to a density modulation. The second term originates from a radial contraction and expansion process of the beam at constant direct current density i_0 and is identical with the surface current from Equation (16). With Equations (57) and (59) one obtains

$$\tilde{I}' = \frac{v_1 2\pi b \omega_p \epsilon_0 v_0 \gamma J_1(\gamma b)}{\eta \beta_c} \frac{p_1 p_2^2}{p_2^2 - p_1^2} \left(\frac{\sin \beta_p p_1 z}{1 - p_1^2} + \frac{p_1}{p_2(p_2^2 - 1)} \sin \beta_p p_2 z \right) \exp j(\omega t - \beta_c z). \quad (63)$$

$$\tilde{I}'' = \frac{v_1 2\pi b \omega_p \epsilon_0 v_0 \gamma J_1(\gamma b)}{\eta \beta_c} \frac{p_1 p_2^2}{p_2^2 - p_1^2} \left(\frac{\sin \beta_p p_1 z}{k^2} - \frac{p_1}{p_2 k^2} \sin \beta_p p_2 z \right) \exp j(\omega t - \beta_c z). \quad (64)$$

It is now assumed that p_1 is the reduction factor of the lower (ordinary) and p_2 the reduction factor of the upper (extraordinary) branch of the curves in Figure 2. In many practical cases the beam diameter is so small that

$$p_1^2 \ll 1, \quad (65)$$

and consequently, from Equation (58),

$$p_2^2 = 1 + k^2. \quad (66)$$

This approximation can be shown to be independent of b/a . Thus for thin beams the last terms of Equations (63) and (64) cancel upon summation and the r-f convection current is given by the ordinary waves only. Thus the amplitude of the convection current varies

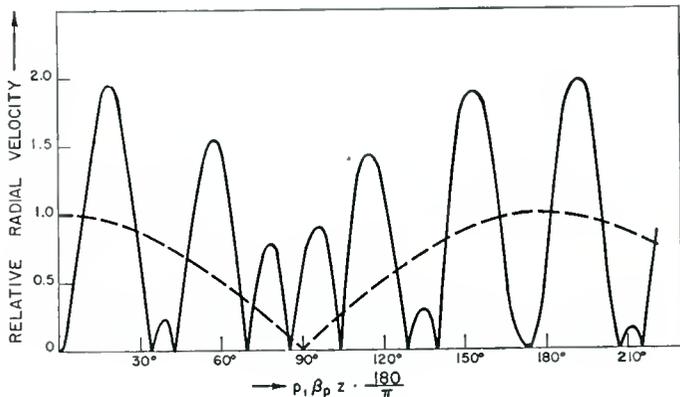


Fig. 3—Relative amplitude of the radial velocity versus distance from the exciting double layer for Brillouin flow, $\beta_c b = 0.55$, and $b/a = 1$. The dashed curve represents the part which is contributed by the ordinary waves. The zeros occur at the same distances as the extrema of the ripples (Figure 4).

sinusoidally along the beam. This is not so for other quantities. Figures 3 to 5 show the functions in the parenthesis of Equations (54), (55) or (59), and (57) for Brillouin flow ($k^2 = 0.5$), $b/a = 1$ and $\beta_c b = 0.55$. The zeros in the amplitude of the radial velocity (Figure 3) naturally coincide with the extrema in the amplitude of the ripples (Figure 4). From Equation (56), the variation of the axial velocity amplitude along the beam is sinusoidal as long as Approximation (65) holds. A comparison between Equations (63) and (64) shows,

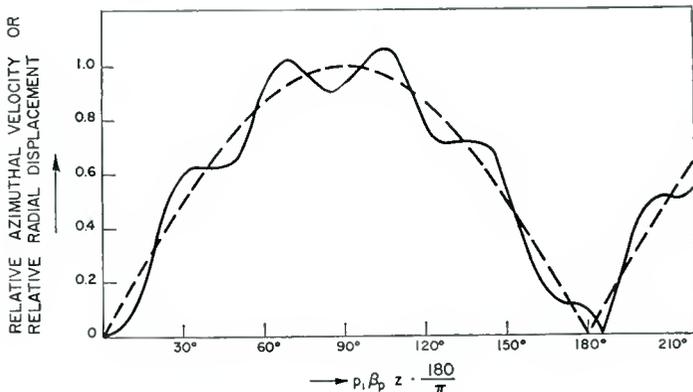


Fig. 4—Relative amplitude of the azimuthal velocity or radial displacement (ripples) versus distance from the exciting double layer for Brillouin flow, $\beta_c b = 0.55$, and $b/a = 1$. The dashed curve represents the part which is contributed by the ordinary waves. The extrema occur at the same distances as the zeros in radial velocity (Figure 3).

again under Condition (65) and for Brillouin flow, that $\frac{2}{3}$ of the total axial r-f convection current originates from the radial contraction and expansion process of the beam at constant unperturbed direct-current density, while only $\frac{1}{3}$ is contributed by the axial density modulation. This changes, of course, as k^2 is increased and, for confined flow, $k^2 = \infty$ and $\tilde{I}'' = 0$. Under Condition (65), one can say that for $k^2 < 1$ the r-f convection current due to the radial contraction-expansion process is predominant, and for $k^2 > 1$ the r-f convection current due to the axial density modulation is predominant. It should be recalled here that in this example modulation by an ideal

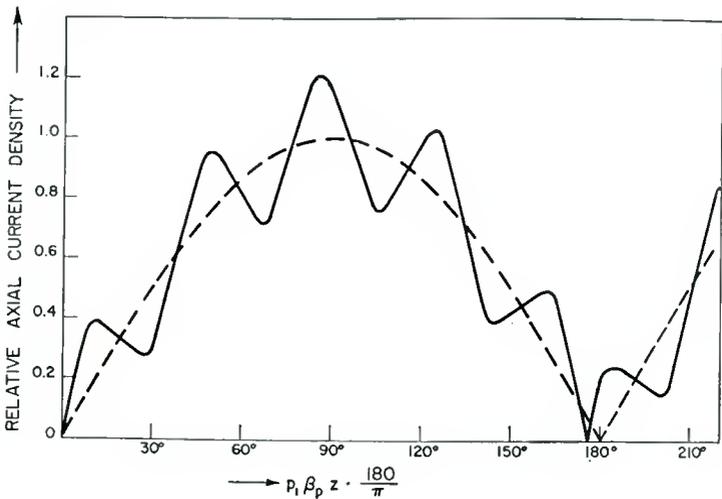


Fig. 5—Relative amplitude of the axial convection-current density versus distance from the exciting double layer for Brillouin flow, $\beta_e b = 0.55$, and $b/a = 1$. The dashed curve represents the part which is contributed by the ordinary waves.

double layer has been assumed. The results may be quite different for another excitation.

SPACE-CHARGE-WAVE PROPAGATION ALONG A SLOW-WAVE STRUCTURE (TRAVELING-WAVE TUBE)

Coupling between beam and delay line is accomplished by the displacement current per unit length \tilde{M} which flows away from the beam;

$$\tilde{M} = 2\pi b j \omega \epsilon_0 \tilde{E}_{r11}(b). \quad (67)$$

With the aid of Equations (17b), (18a), (29), (30a), and (32), one obtains for the radial electric field strength just outside of the beam

$$\tilde{E}_{r11}(b) = \left[\frac{\left(\frac{j\beta_r - \Gamma}{\beta_p} \right)^2 + 1 + k^2}{\left(\frac{j\beta_c - \Gamma}{\beta_p} \right)^2 + k^2} \right] \tilde{E}_{r1}(b). \quad (68)$$

The factor in brackets represents the radial motion of the particles. If the discussion is restricted to *thin beams* one can see from Equations (50) and (68), and from the Approximation (66), that *the extraordinary waves do not couple to the delay line*. To compute the coupling current of the ordinary waves it is assumed that the axial electric field is approximately constant across the beam. Then $|\gamma b| \ll 1$ and the Bessel functions can be approximated by the first terms of the series for small arguments. From Equations (25) and (26)

$$\tilde{E}_{z1} = A \exp(j\omega t - \Gamma z) \quad (69)$$

$$\tilde{E}_{r1} = \frac{\gamma^2}{2\Gamma} r \tilde{E}_{z1}. \quad (70)$$

The boundary condition (17a) and Equations (23), (67), (68), and (70) yield

$$\tilde{M} = j\omega\epsilon_0\Gamma b^2\pi \left[1 + \left(\frac{\beta_p}{j\beta_c - \Gamma} \right)^2 \right] \tilde{E}_{z11}. \quad (71)$$

The so-called "electronic" Equation (71) combined with a "circuit" equation which can be derived, for instance, from Pierce's equivalent network,¹⁶ lead to an algebraic dispersion equation of fourth power in Γ . This dispersion equation seems to be independent of the magnetic focusing field strength because Equation (71) is independent of k^2 . Thus the gain in traveling-wave tubes with thin beams seems to be independent of the magnetic field strength. However, the passive-mode parameter^{16,17} Q can be related to the plasma-frequency-

¹⁶ J. R. Pierce, *Traveling-Wave Tubes*, D. van Nostrand Company, Inc., Princeton, N. J., 1950.

¹⁷ F. Paschke, "Die Wechselseitigkeit der Kopplung in Wanderfeldröhren," *Arch. d. elektr. Übertrag.*, Vol. 11, p. 137, April, 1957.

reduction factor which increases with k^2 . Thus, for appreciable space charge, one can expect a decreasing gain with increasing focusing field strength due to the enhanced excitation of passive modes.

The independence of the coupling current, Equation (71), of k^2 can be explained as follows. Suppose the axial electric field strength is kept constant while the focusing field strength, or k^2 , decreases. Then the bunches become less dense as expressed by the k^2 dependence of γ^2 (Equation (23)) in Equation (70). However, the coupling due to the radial expansion-contraction process of the beam as expressed by the factor in Equation (68) increases so as to fully compensate the reduction in bunch density. This is only true for thin beams. For thick beams both the ordinary and extraordinary waves are coupled to the delay line and the gain depends on the magnetic field strength even for small space charge. A thick-beam analysis, however, leads to a complicated transcendental dispersion equation which, to this author, does not seem to merit further investigation.

It should be noted that Pierce's electronic equation (Equations (2.22) and (2.3) in Reference (16))

$$\tilde{M} = j\omega\epsilon_0\Gamma b^2\pi \left(\frac{\beta_p}{j\beta_e - \Gamma} \right)^2 \tilde{E}_{z11} \quad (72)$$

can be obtained from Equation (71) by reducing the beam to a current thread which means $b^2 \rightarrow 0$ but $b^2\beta_p^2 \neq 0$. The assumption of a current thread is equivalent to neglecting the axial displacement current within the beam compared to the convection current. The propriety of this procedure and the influence of the first term of Equation (71) are discussed in Reference (17).

FORMATION AND MAINTENANCE OF THE BEAM

There is a practical method of forming an electron beam whose steady state is approximated by Equations (2), (3), and (6). This is depicted in Figure 6. The current is emitted from a cathode at zero potential which is threaded by a uniform axial magnetic field, B_c . After proper acceleration in the gun region, the electrons acquire an angular velocity by changing their radial position and by passing through a magnetic shield into a drift space with a homogeneous magnetic field B . The angular velocity is given by Busch's theorem;

$$\Omega_0 = \frac{\omega_c}{2} (1 - h), \quad (73)$$

$$h = \frac{r_c^2 B_c}{r_0^2 B}. \tag{74}$$

Here r_c and r_0 are the radii of an electron at the cathode and at equilibrium, respectively. h is positive if B and B_c are in the same direction and negative if B and B_c are in opposite directions. In conjunction with Equations (2), the angular velocity, Ω_0 , has been assumed to be independent of r_0 . Thus h must be independent of r_0 . It follows from Equation (74) that the rate of convergence must be the same for all electrons, which is a fair assumption. The steady-

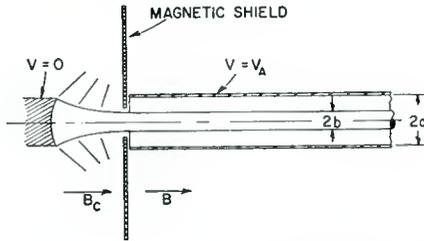


Fig. 6—Formation of the beam. After leaving the cathode the electrons are accelerated and pass through a magnetic shield into a drift tube.

state Condition (8) can be rewritten

$$\omega_p^2 = \frac{\omega_c^2}{2} (1 - h^2), \tag{75}$$

and the focusing parameter k^2 from Equation (11)

$$k^2 = \frac{1}{2} \frac{1 + 3h^2}{1 - h^2}. \tag{76}$$

From Equation (75) one can conclude that only for

$$-1 \leq h \leq 1 \tag{77}$$

is a steady state achieved. The potential distribution outside of the beam is given by the following solution of Laplace's equation

$$r \geq b, \quad 2\eta V_{II} = 2\eta V_A - \omega_p^2 b^2 \ln \frac{a}{b} \left[1 - \frac{\ln \frac{r}{b}}{\ln \frac{a}{b}} \right]. \tag{78}$$

Inside of the beam the potential V_I is obtained from Poisson's equation:

$$r \leq b, \quad 2\eta V_I = 2\eta V_A - \frac{\omega_p^2 b^2}{2} \left(1 + 2 \ln \frac{a}{b} \right) + \frac{\omega_p^2}{2} r^2. \quad (79)$$

The boundary conditions are met by the Equations (78) and (79); that is, at the drift-tube wall, $r = a$, the potential is V_A , and, at $r = b$, both the potential and its gradient are continuous. Equation (79) states that the potential is depressed by space charge. This affects the electron velocity. Assuming zero emission velocity, from the conservation of energy,

$$v_0^2 + r^2 \Omega_0^2 = 2\eta V_I. \quad (80)$$

With the aid of Equations (73) and (75) one obtains from Equations (79) and (80)

$$v_0^2 = 2\eta V_A - \frac{b^2 \omega_c^2}{4} (1 - h^2) \left(1 + 2 \ln \frac{a}{b} \right) + \frac{\omega_c^2}{2} r^2 h (1 - h). \quad (81)$$

Thus the axial velocity depends on the radius. This is due to the r -dependant term of the potential depression which is partly compensated by the conversion of axial velocity into angular velocity described by Equation (80). The compensation is complete only for Brillouin flow ($h = 0$) where the axial velocity is independent of r . For confined flow ($h = 1$) the last two terms in Equation (81) are finite for finite space-charge densities because of Relation (75). For derivation of the space-charge-wave equation the assumption has been made that v_0 does not depend on the radius. The r -f analysis is therefore valid only if, from Equation (81),

$$\frac{\omega_c^2 b^2}{2} h (1 - h) \ll 2\eta V_A - \frac{b^2 \omega_c^2}{4} (1 - h^2) \left(1 + 2 \ln \frac{a}{b} \right), \quad (82)$$

which is true in most practical cases.

In the following, the maximum current which can pass the drift tube is derived. The current density is obtained by multiplying Equation (81) by ρ_0^2 and taking the square root. Substituting with the aid of Equation (75), the cyclotron frequency for the plasma frequency, one obtains

$$i_0 = \frac{\omega_c^2 (1 - h^2) \epsilon_0 \sqrt{2\eta V_A}}{2\eta} \sqrt{1 - \frac{b^2 \omega_c^2}{4} (1 - h^2) - \frac{1 + 2 \ln \frac{a}{b}}{2\eta V_A} + \frac{\omega_c^2 r^2 h (1 - h)}{2 \cdot 2\eta V_A}} \quad (83)$$

Integration over the beam cross section yields

$$I_0 = \frac{2\pi \epsilon_0}{3\eta} \frac{1 + h}{h} (2\eta V_A)^{3/2} \left[\left(\frac{V_0}{V_A} + \frac{2}{1 + 2 \ln \frac{a}{b}} \frac{h}{1 + h} \left(1 - \frac{V_0}{V_A} \right) \right)^{3/2} - \left(\frac{V_0}{V_A} \right)^{3/2} \right] \quad (84)$$

V_0 is the potential at the axis;

$$\frac{V_0}{V_A} = 1 - \frac{\omega_c^2 b^2}{4} (1 - h^2) - \frac{1 + 2 \ln \frac{a}{b}}{2\eta V_A} \quad (85)$$

A focusing system is characterized by the parameter k^2 or, from Equation (76), by h . If this parameter is kept constant and the magnetic field is increased, the effect on the current will be twofold. Firstly, there will be an increase in current because the space-charge density from Equations (75) and (7) increases with the square of the magnetic field. Secondly, there will be a reduction in current because the increased space-charge depression of the potential tends to decrease the axial velocity. Thus there must be an optimum value for the magnetic field or for V_0/V_A from Equation (85). This optimum is readily obtained from Equation (84).

$$\frac{V_{0 \text{ opt}}}{V_A} = \frac{(1 - \sigma)^2}{3(1 - \sigma) + \sigma^2}, \quad (86)$$

where σ is defined by

$$\sigma = \frac{2h}{(1 + h) \left(1 + 2 \ln \frac{a}{b} \right)} \quad (87)$$

From Equations (86) and (84) the maximum current is

$$I_{0 \max} = \frac{4\pi \epsilon_0 (2\eta V_A)^{3/2}}{3\eta \left(1 + 2 \ln \frac{a}{b}\right)} \frac{1}{\sqrt{3(1-\sigma) + \sigma^2}}. \quad (88)$$

The function of σ appearing in Equation (88) exhibits a maximum at $\sigma = 3/2$. This corresponds, however, to a value of h which is prohibited by Relation (77). Therefore the absolute limit of current appears at the maximum permissible value of h . This is, from Relation (77), the value for confined flow $h = 1$, which is to be expected. Pierce¹⁸ has given the limiting current for $a = b$ both for Brillouin flow and confined flow. For Brillouin flow, $h = 0$ and $\sigma = 0$, so that

$$I_{0 \max} = \frac{4\pi \epsilon_0 (2\eta V_A)^{3/2}}{3\eta \sqrt{3}}. \quad (89)$$

This agrees with Pierce's Equation (9.32) in Reference (18). For confined flow, $h = 1$ and $\sigma = 1$, so that

$$I_{0 \max} = \frac{4\pi \epsilon_0 (2\eta V_A)^{3/2}}{3\eta}. \quad (90)$$

This is higher than predicted by Pierce's Equation (9.61) in Reference (18) because Pierce obtained the limiting current for confined flow by assuming constant current density across the beam while in the present analysis the space-charge density is kept constant. In a recent paper¹⁹ a maximum current has been calculated for Brillouin flow ($h = 0$) and varying a/b . This analysis was based on the assumption that the optimum value of the potential at the axis, V_0 , is independent of a/b and equal to one third of the potential at the beam surface, $V_I(b)$. This would be valid if $V_0/V_I(b)$ were not a function of a/b . Equation (79) shows that this assumption is incorrect. From Equations (86) and (87) the optimum value of the potential at the axis for Brillouin flow is $V_{0 \text{ opt}} = V_A/3 \cong V_I(b)/3$. Consequently Equation (88) predicts higher limiting currents than the corresponding equation in Reference (19). Figure 7 shows the maximum perveance versus magnetic focusing field over Brillouin field

¹⁸ J. R. Pierce, *Theory and Design of Electron Beams*, D. van Nostrand Company, Inc., Princeton, N. J., 1949.

¹⁹ H. J. Wolkstein, "Effect of Collector Potential on the Efficiency of Traveling-Wave Tubes," *RCA Review*, Vol. 19, p. 259, June, 1958.

with the ratio a/b as parameter. For small gaps between beam and drift tube the maximum current rises sharply with the magnetic field. For field strengths larger than about twice the Brillouin field the current increase is insignificant. For wide gaps the magnetic field has little influence on the maximum current.

In practice it is very difficult to keep the beam in a smooth cylindrical shape. There are many factors which cause perturbations. For example, the magnetic field does not change abruptly from B_c to B because there are always fringing fields present, or the beam is injected at an angle to the axis or with an improper space-charge

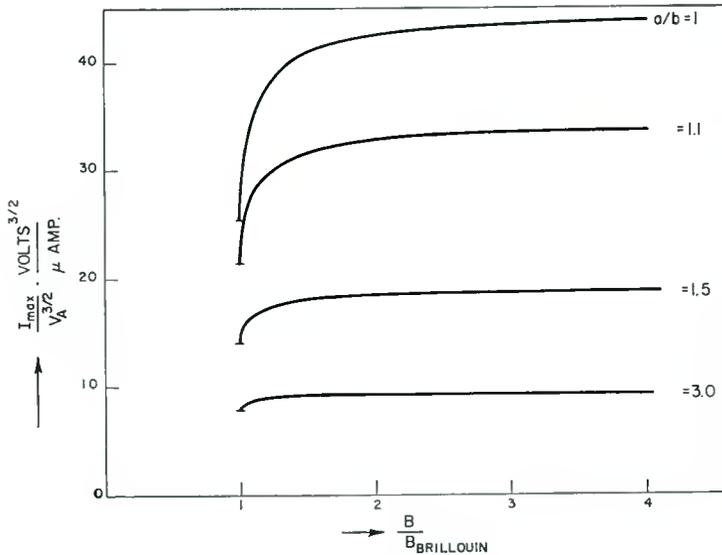


Fig. 7—Maximum direct current which can pass a drift tube versus magnetic focusing field over Brillouin field. V_A is the potential at the drift-tube wall. a/b is the ratio of drift-tube radius to beam radius.

density. Consequently the beam will exhibit ripples. These are described by the space-charge-wave Equation (22) and its solution given by Equations (25)-(30) for $\omega = 0$. In this case, from Equation (50),

$$p^2 = - \left(\frac{\Gamma}{\beta_p} \right)^2. \quad (91)$$

This represents a straight line in the diagrams “ $|p|$ versus $(-j\Gamma b)$ ” discussed in conjunction with Figure 2. The angle ψ between the ordinate and the straight line is given by

$$\tan \psi = \beta_p b. \tag{92}$$

Because

$$\beta_p b \ll 1, \tag{93}$$

this angle is very small so that there is an intersection only with the upper branch of the curves. Also, because of Equation (91) and Relation (93) one can approximate $|p|$ by the value for small $(-j\Gamma b)$ given by Equation (66);

$$|p| = \sqrt{1 + k^2}.$$

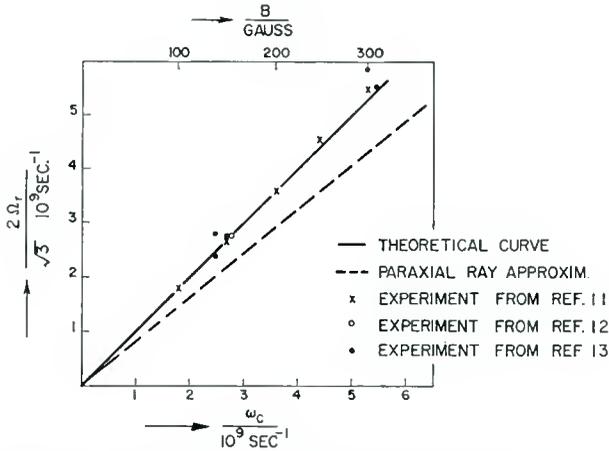


Fig. 8—The frequency Ω_r of the radial electron oscillations in a scalloping Brillouin-flow beam plotted versus the cyclotron frequency, ω_c , or magnetic focusing field strength, B . The theoretical value is given by Equation (98). Previous theories lead to the paraxial-ray approximation.

From Equation (91)

$$\Gamma = \pm j\beta_p \sqrt{1 + k^2}. \tag{94}$$

Thus the ripples vary sinusoidally in space as

$$\sin \left(\frac{\omega_p \sqrt{1 + k^2}}{v_{01}} z + \text{const.} \right). \tag{95}$$

This corresponds to a radial oscillation of the electrons with a frequency

$$\Omega_r = \omega_p \sqrt{1 + k^2}. \tag{96}$$

With the aid of Equations (75) and (76) this can be rewritten

$$\Omega_r = \frac{\omega_c}{2} \sqrt{3 + h^2}. \quad (97)$$

For Brillouin flow, $h = 0$ and

$$\Omega_r = \omega_c \frac{\sqrt{3}}{2}. \quad (98)$$

The excellent agreement of Equation (98) with the experimental results found in the literature¹¹⁻¹³ is depicted in Figure 8. Only two experimental points ($B = 140$ and 300 gauss, both from Reference (13)) deviate substantially from the theoretical values. A demountable vacuum system with relatively high pressure was used in the experiment of Reference (13), whereas low-pressure sealed-off tubes were used in the experiments of References (11) and (12). Thus, the deviation may have been caused by neutralization of the electron space charge by positive ions. Furthermore, the data from References (11) and (12) represent averages over a relatively large number of scalloping wavelengths whereas in Reference (13) the beam was probed over only about one wavelength. This not only makes an evaluation of the radial frequency somewhat inaccurate but also leaves some uncertainty about the laminarity of the stream. In contrast to the good agreement of Equation (98) with the experiments, the corresponding equations in References (1)-(6) lead to the paraxial-ray approximation $\Omega_r = \omega_c/\sqrt{2}$ (Equation 9.44 in Reference (18)) which underestimates the radial frequency by about 20 per cent. As h is increased Ω_r also increases. For strong focusing $h \rightarrow 1$ and the electrons oscillate with the cyclotron frequency

$$\Omega_r = \omega_c. \quad (99)$$

The amplitude and "phase" of the ripples naturally depend upon the way they are excited. This problem can readily be solved with the general Equations (25)-(30).

ACKNOWLEDGMENT

The author is greatly indebted to J. Berghammer for many valuable suggestions. Interesting and helpful discussions were had with K. Pöschl and W. W. Rigrod. The assistance of S. Bloom and L. S. Nergaard in writing the manuscript is gratefully acknowledged.

STABILIZATION OF TRANSISTOR GAIN OVER WIDE TEMPERATURE RANGES*

BY

R. A. SCHMELTZER

RCA Semiconductor and Materials Division,
Somerville, N. J.

Summary—The small-signal gain of a junction-transistor amplifier stage varies with ambient temperature as a result of changes of not only the d-c operating point but also the temperature-dependent a-c parameters of the transistor. The operating point can be stabilized by use of d-c degeneration or feedback methods. The effects of junction-temperature variation on the a-c stability of the amplifier are analyzed in terms of the temperature dependence of the hybrid- π equivalent-circuit parameters of the transistor. It is shown that:

1. When a junction-transistor amplifier stage is driven from a source having an effective a-c resistance approximately equal to $\frac{1}{2}r_{be}$, the rate of change of gain is zero.

2. When the quiescent emitter current of the transistor is maintained constant, the optimum a-c source resistance is a slowly varying function of temperature.

3. By proper choice of a-c source resistance the gain of a junction-transistor amplifier stage can be maintained substantially constant over a wide range of ambient temperatures without the use of temperature-sensitive circuit components, a-c degeneration or feedback, or additional circuit elements.

4. By proper design, and without the use of a-c feedback, it is possible to achieve a gain stability of ± 0.125 decibel over the temperature range between -70° and $+80^{\circ}\text{C}$ in a stage having a gain of about 30 decibels.

INTRODUCTION

THE PERFORMANCE of an alloy-junction-transistor amplifier may be seriously impaired by variations in ambient temperature. Such variations have particularly adverse effects upon the gain parameters of the transistor because they affect not only the position of the operating point, but also the junction temperature of the transistor. The operating point is affected because it is determined by the d-c input conductance and saturation currents of the transistor, both of which are functions of temperature.

The operating point of a small-signal-type transistor amplifier stage can be stabilized against the effects of variations in d-c input conductance by biasing the emitter junction from a source having a

* Manuscript received February 5, 1959.

sufficient amount of d-c resistance. It can be stabilized against the effects of variations in saturation currents by the use of d-c degeneration or inverse feedback. When a junction transistor is properly stabilized, its quiescent emitter current remains substantially constant over a wide range of ambient temperatures. To achieve maximum gain stability, however, it is necessary to minimize the effects of variations in junction temperature.

This paper presents an analysis of the effects of junction temperature on the gain of a germanium p-n-p alloy-junction-transistor amplifier stage. It also describes a method by which these effects can be minimized.

GAIN VARIATION IN D-C STABLE CIRCUITS

The analysis presented below is based upon a d-c stabilized amplifier stage using the common-emitter circuit shown in Figure 1. It is applicable whether the stage is used alone or in cascade with other stages. The transistor is driven from a source having an effective a-c resistance R_s , and operates into a load having an effective a-c resistance R_L . The gain of the stage is directly proportional to its effective transconductance, G_m , generally defined as

$$G_m = I_c/V_s, \quad (1)$$

where I_c is the small-signal collector current, and V_s is the small-signal Thevenin equivalent driving voltage.

The transistor itself can be represented by Giacoletto's hybrid- π small-signal equivalent circuit¹ shown in Figure 2a. If the operating frequency is low enough so that all capacitive effects within the transistor are negligible, and if the effective load resistance, R_L , is much smaller than the output resistance of the transistor, the transistor can be represented by the simplified equivalent circuit shown in Figure 2b. The amplifier stage can then be represented by the circuit shown in Figure 2c.

The effective transconductance of the stage in terms of the parameters shown in Figure 2c is

$$G_m = \frac{1}{\frac{1}{g_m} + [1 + g_{b'e}(R_s + R_c + r_{bb'})] + R_c}, \quad (2)$$

¹ Giacoletto, L. J., "Study of P-N-P Alloy Junction Transistor from D-C Through Medium Frequencies," *RCA Review*, Vol. 15, pp. 506-562, December, 1954.

It is evident that the stability of the effective transconductance, G_m , with temperature depends upon the temperature sensitivity of the parameters $g_{b'e}$, g_m (intrinsic transconductance) and $r_{bb'}$. The quantitative variations with temperature of these parameters in p-n-p germanium alloy-junction transistors have been measured and are in agreement with theory.² For temperatures below about 70° C the intrinsic base-to-emitter conductance $g_{b'e}$ decreases approximately as $T^{-3.2}$, that is,

$$g_{b'e} = g_{b'e_0} \left(\frac{T}{T_0} \right)^{-3.2}, \quad (3)$$

where $g_{b'e_0}$ is the value of $g_{b'e}$ at temperature T_0 ; the intrinsic trans-

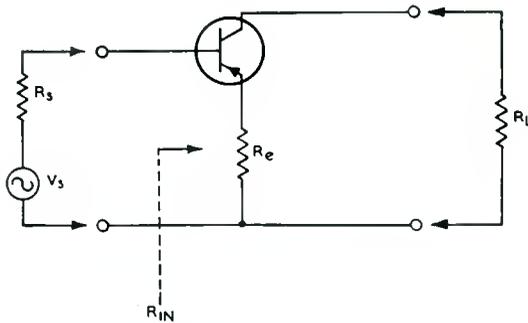


Fig. 1—Basic common-emitter amplifier stage upon which the analysis is based.

conductance, g_m , decreases as T^{-1} , that is,

$$g_m = g_{m_0} \left(\frac{T}{T_0} \right)^{-1}, \quad (4)$$

where g_{m_0} is the value of g_m at temperature T_0 ; and the internal base-lead resistance, $r_{bb'}$, increases as $T^{2.3}$, that is,

$$r_{bb'} = r_{bb'_0} \left(\frac{T}{T_0} \right)^{2.3}, \quad (5)$$

where $r_{bb'_0}$ is the value of $r_{bb'}$ at temperature T_0 . The effective trans-

² Eshelman, C. R., "Variation of Transistor Parameter with Temperature," *Semiconductor Products*, pp. 25-30, January/February, 1958.

conductance of the transistor is most adversely affected by the variations with temperature of $g_{b'e}$ and g_m , and is only slightly affected by the variation of $r_{bb'}$ (see Appendix).

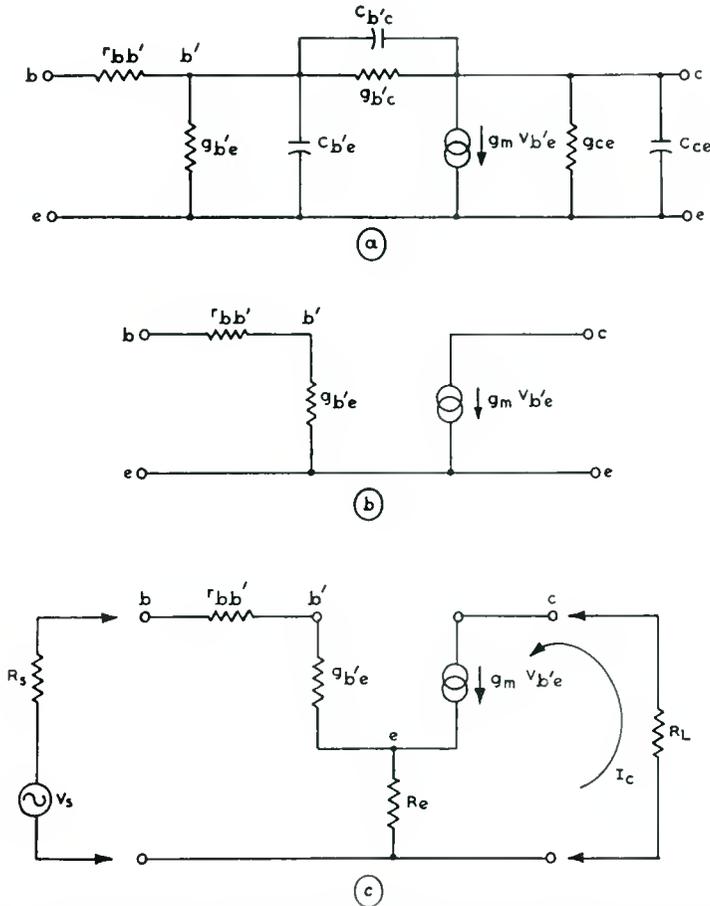


Fig. 2—(a) Giacoletto's exact hybrid- π equivalent circuit of the p-n-p alloy-junction transistor; (b) simplified hybrid- π equivalent circuit of the transistor; (c) equivalent circuit of the amplifier stage shown in Figure 1.

It is convenient to express the temperature-dependent parameters G_m , $g_{b'e}$, and g_m in terms of their logarithmic derivatives with respect to temperature, or temperature coefficients. Thus, the temperature coefficient of G_m is

$$\phi_G = \frac{d}{dT} (\ln G_m) = \frac{1}{G_m} \frac{dG_m}{dT}, \quad (6)$$

the temperature coefficient of $b_{b'e}$ is

$$\phi_b = \frac{1}{g_{b'e}} \frac{dg_{b'e}}{dT} = -\frac{3.2}{T}, \quad (7)$$

and the temperature coefficient of g_m is

$$\phi_g = \frac{1}{g_m} \frac{dg_m}{dT} = -\frac{1}{T}. \quad (8)$$

The values of ϕ_G , ϕ_b , and ϕ_g represent respectively, the fractional change in G_m , $g_{b'e}$, and g_m per degree centigrade at temperature T . It is evident from Equation (2) that when the amplifier is operated at room temperature (300°K) with both R_s and R_e equal to zero, G_m tends to vary as g_m , decreasing at a rate $\phi_G = \phi_g = -1/300 = -0.3$ percent per degree centigrade. It is also evident that when R_s is made sufficiently large, G_m becomes proportional to $g_m/g_{b'e}$, and tends to increase with temperature at a rate $\phi_G = \phi_g - \phi_b = - (1/300) + (3.2/300) = 0.7$ percent per degree centigrade. There is, therefore, an optimum value of source resistance R_s for which the G_m is constant with temperature.

As shown in the Appendix, ϕ_G may be expressed in terms of the temperature coefficients ϕ_b and ϕ_g ;

$$\phi_G \cong \frac{2.2(R_s + R_e) - r_{b'e}}{T(R_s + R_{in})} \quad (9)$$

where R_{in} is the effective input resistance of the amplifier stage. It is evident from Equation (9) that the fractional change in G_m per degree centigrade is directly proportional to the difference between $2.2(R_s + R_e)$ and $r_{b'e}$, and varies inversely as R_{in} . It should be noted that the emitter resistance R_e actually tends to reduce the value of ϕ_G because it increases the value of R_{in} . This effect is predictable, because the emitter degeneration provided by R_e reduces the effects of variations in transistor parameters.

The preceding analysis shows that the temperature coefficient of G_m can be reduced to zero by the use of a source resistance such that

$$R_s \cong \frac{r_{b'e}}{2.2} - R_e. \quad (10)$$

Because $r_{b'e}$ is an increasing function of temperature, ϕ_G can have

zero value at only one temperature, e.g., $T = T_0$. At this temperature, the G_m of the stage is maximum and constant. At temperatures above $T = T_0$, the temperature coefficient of G_m is negative; at temperatures below $T = T_0$ it is positive. Because the optimum value of R_s is a slowly varying function of T , ϕ_G remains quite small over a wide range of temperatures. Furthermore, because the variation in the optimum value of R_s is a second-order effect, the G_m does not change appreciably over the range of temperatures for which Equation (10) is approximately satisfied.

EXPERIMENTAL RESULTS

Figure 3 shows the measured relative gain of a one-stage transistor amplifier as a function of temperature for three values of a-c source resistance R_s , namely, 10,000 ohms, 1,230 ohms, and 30 ohms. The transistor used is a high-gain version of the 2N139 germanium p-n-p alloy-junction type having an $r_{b'e}$ of 2,700 ohms at 0°C. For each value of R_s , the circuit of the amplifier was stabilized to provide a constant emitter current of one milliampere over the temperature range between -70° and +80°C. All emitter resistance was adequately bypassed. The variation in gain between -70° and +80°C with the optimum value of R_s (1,230 ohms) is only 0.25 decibel, whereas with R_s equal to 10,000 ohms it is 4.25 decibels. The decrease in gain observed at temperatures above +70°C is due to the departure of $r_{b'e}$ from its assumed variation, $r_{b'e} \propto T^{3.2}$. Above approximately +80°C, $r_{b'e}$ decreases with temperature, thereby rendering this technique useless at temperatures higher than +80°C. This decrease is caused by the falloff of the diffusion length for minority carriers in the emitter at temperatures above +80°C.^{2,3}

There is a price to be paid for minimizing gain variations with the use of a low source resistance. If $R_s = r_{b'e}/2.2$ is used, the current gain will be decreased by a factor of 3.2 (or 10 decibels). If one were to consider, alternatively, the use of 10 decibels of negative feedback to stabilize the a-c gain, then one would find that the use of the proper source resistance is far more advantageous than the use of negative feedback for operating temperatures below about +65°C, as is easily verified by the curves of Figure 3. But a low d-c source resistance is also required to achieve d-c stability. In the design of d-c stable transistor circuits, the ratio of the d-c emitter resistance to the d-c source resistance should be made as large as possible. In a given application, however, the amount of d-c emitter resistance that can be used is limited by the available power-supply voltage. Hence the maximum value of d-c source resistance is determined by both the d-c

stability requirements and the available power-supply voltage. In many applications, the available power-supply voltage is sufficiently limited that the proper d-c source resistance is comparable to R_g optimum. The additional loss in current gain suffered when using an a-c source resistance of $r_{b'e}/2.2$ is therefore often much less than 10 decibels.

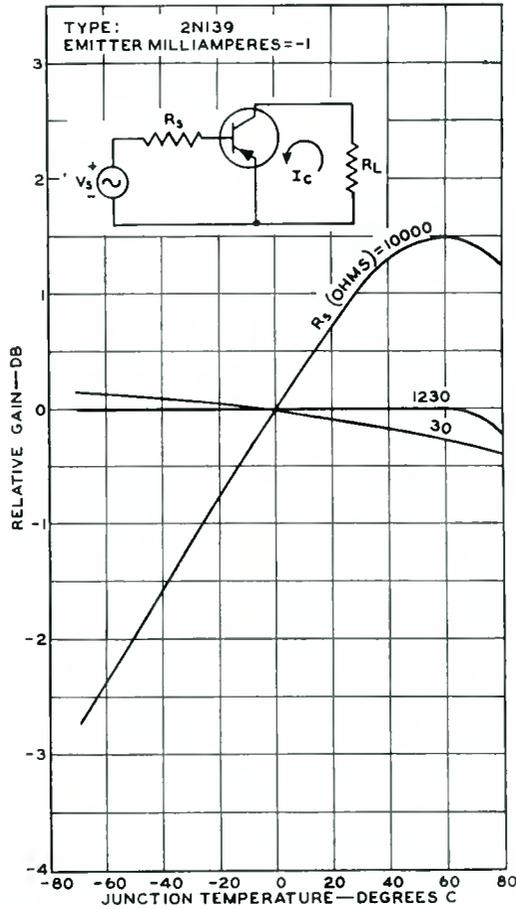


Fig. 3—Relative gain of a d-c stabilized transistor amplifier stage as a function of junction temperature, for three values of a-c source resistance, R_s .

APPENDIX

An expression for the temperature coefficient of the effective transconductance of a transistor amplifier stage in terms of the temperature coefficients of the equivalent-circuit parameters of the transistor is

derived. In the derivation, it is assumed that the operating point of the stage is constant with temperature.

The effective transconductance, G_m , is primarily a function of the variables $g_{b'e}$, g_m , and $r_{bb'}$, each of which is a function of the junction temperature, T . Therefore, from the theory of partial derivatives,

$$\frac{dG_m}{dT} = \frac{\partial G_m}{\partial g_m} \cdot \frac{dg_m}{dT} + \frac{\partial G_m}{\partial g_{b'e}} \cdot \frac{dg_{b'e}}{dT} + \frac{\partial G_m}{\partial r_{bb'}} \cdot \frac{dr_{bb'}}{dT}. \quad (11)$$

From Equation (2) it can be shown that

$$\frac{\partial G_m}{\partial g_m} \cdot \frac{dg_m}{dT} = + \frac{G_m^2}{g_m} \phi_y [1 + g_{b'e}(R_s + R_e + r_{bb'})], \quad (12)$$

$$\frac{\partial G_m}{\partial g_{b'e}} \cdot \frac{dg_{b'e}}{dT} = - \frac{G_m^2}{g_m} \phi_b [g_{b'e}(R_s + R_e + r_{bb'})], \quad (13)$$

$$\frac{\partial G_m}{\partial r_{bb'}} \cdot \frac{dr_{bb'}}{dT} = - \frac{G_m^2}{g_m} \phi_r [g_{b'e}r_{bb'}]. \quad (14)$$

By substitution of Equations (12), (13), and (14), into Equation (11), and rearrangement, we obtain

$$\phi_G = \left[\frac{G_m g_{b'e}}{g_m} \right] [\phi_y] \left[\frac{1}{g_{b'e}} + (R_s + R_e) \left(1 - \frac{\phi_b}{\phi_y} \right) + r_{bb'} \left(1 - \frac{\phi_b}{\phi_g} - \frac{\phi_r}{\phi_y} \right) \right]. \quad (15)$$

Because $G_m g_{b'e} / g_m = 1 / (R_s + R_{in})$ and $1 / g_{b'e} = r_{b'e}$, Equation (15) becomes

$$\phi_G = \frac{\phi_y}{R_s + R_{in}} \left[r_{b'e} + (R_s + R_e) \left(1 - \frac{\phi_b}{\phi_y} \right) + r_{bb'} \left(1 - \frac{\phi_b}{\phi_g} - \frac{\phi_r}{\phi_y} \right) \right]. \quad (16)$$

Equation (16) thus expresses the temperature coefficient of G_m in terms of the temperature coefficients of $g_{b'e}$, g_m , and $r_{bb'}$. The temperature coefficients of $g_{b'e}$ and g_m are given by Equations (7) and (8) respectively. The variation with temperature of $r_{bb'}$ is essentially the same as that of the resistivity of the base material. For resistivities less than 10 ohm-centimeters, $r_{bb'}$ varies approximately as the reciprocal of the mobility of the majority carriers in the base. For germanium

p-n-p type alloy-junction transistors,³ $r_{bb'}$, therefore, varies as $T^{2.3}$. The temperature coefficient of $r_{bb'}$ is

$$\phi_r \approx + \frac{2.3}{T}. \quad (17)$$

Substitution of Equations (7), (8), and (17) into Equation (16) yields

$$\phi_G \cong \frac{2.2(R_s + R_e) - 0.1(r_{bb'}) - r_{b'e}}{T(R_s + R_{in})}. \quad (18)$$

Because the quiescent emitter current, I_e , is usually not more than a few milliamperes, the term $0.1(r_{bb'})$ is usually quite small in comparison with $r_{b'e}$. Therefore, to a good approximation,

$$\phi_G \cong \frac{2.2(R_s + R_e) - r_{b'e}}{T(R_s + R_{in})}. \quad (19)$$

To find the optimum value of R_s , set ϕ_G equal to zero. Then

$$R_s \cong \frac{r_{b'e}}{2.2} - R_e. \quad (20)$$

It is evident that the optimum value of R_s is determined by $r_{b'e}$, and, therefore, depends upon the quiescent emitter current employed.

³ Gartner, W. W., "Temperature Dependence of Junction Transistor Parameters," *Proc. IRE*, Vol. 45, pp. 662-680, May, 1957.

QUALITY-CONTROL DETERMINATIONS OF THE SCREEN PERSISTENCE OF COLOR PICTURE TUBES*

BY

J. M. FORMAN AND G. P. KIRKPATRICK

RCA Electron Tube Division,
Lancaster, Pa.

Summary—Protection against objectionable color trailing in color picture tubes requires quality-control determinations of screen persistence in finished tubes. This paper describes an accurate method for making these determinations in which the control grids of the tube under test are pulsed individually by a continuous square wave, either symmetrical or non-symmetrical, while conventional interlaced television scanning is simultaneously applied. A multiplier phototube behind a narrow slit monitors the varying light output of a small portion of the television raster and produces a signal output as a function of time on a linear calibrated oscilloscope. Persistence measurements are made directly on the displayed waveform.

INTRODUCTION

IN THE production of picture tubes in general, and in that of color picture tubes in particular, constant surveillance of a great variety of physical and chemical properties is required to assure the quality of the final product. For example, some one thousand specifications cover the production of the 21CYP22 color picture tube.

By no means the least exacting of these specifications pertain to the materials, processing, and characteristics of the phosphor screen. It is with the quality-control determination of a specific characteristic of the screen that the present paper is concerned, namely, the persistence of the component phosphors.

While viewing a scene on a color picture tube screen, one is sometimes aware of a greenish-yellow trail on the edge of moving objects, which is a consequence of the phosphor persistences. These persistences cannot be appreciably reduced without changing some of the more desirable characteristics of the phosphors, such as color or luminous efficiency. It is, therefore, necessary to minimize the trailing effect by maintaining, within specified limits, the shortest practical persistences. Table I gives the specified persistences of the color phosphors used in the 21CYP22. Persistence is defined as the ratio of the bright-

* Manuscript received February 5, 1959.

Table I—Some Luminescent Characteristics of Phosphor Components of a P22 Screen

Color	Composition	I.C.I. Coordinates				Persistence*	
		Dry Powder		Screen		Elapsed Time (Milli-seconds)	% Persist- ence
		\bar{x}	\bar{y}	\bar{x}	\bar{y}		
Blue	ZnS:Ag	.146	.052	.151	.073	40	<0.01
Green	Zn ₂ SiO ₄ :Mn	.203	.728	.200	.718	33	8.9
						40	5.1
Red	Zn ₃ (PO ₄) ₂ :Mn	.674	.326	.650	.322	33	12.7
						40	8.2
White	(8,500°K + 27 MPCD)		\bar{x}	\bar{y}		0	100
Greenish-Yellow			.454	.514		33	9.1

* Per cent of initial brightness after designated elapsed time.

ness of the phosphor after a specified decay time to the peak fluorescent brightness of the phosphor expressed in per cent.

PULSED-RASTER, OPTICAL-SLIT, PERSISTENCE-MEASURING EQUIPMENT

A schematic diagram of the equipment developed for the persistence-measuring test is shown in Figure 1. The square-wave generator provides a triggered pulse signal to the color picture tube. The circuit used with the tube is so designed that the electron gun in the tube is turned on during the positive portion of the pulse and off during the negative portion of the pulse. The buildup and decay of light emitted by the phosphor during scanning are observed by a 6217 multiplier phototube, and displayed on an oscilloscope.

In adopting this method, it was realized that a pulsed-raster, optical-slit method was applicable only to exponential decays (see Appendix II). However, this limitation was considered insignificant for the following reasons:

(1) As shown in Figure 2, red and green phosphors follow nearly exponential laws during the important parts of their decay. Although the decay of the blue phosphor follows a t^n law, it is so fast that its contribution to the group persistence is negligible.

(2) The test is designed to indicate screen persistence at a specific time after excitation to determine whether the finished tube meets persistence specifications, and is not intended to provide an academic study of phosphor persistence.

(3) Conventional television scanning which is readily available in existing color test sets can be used.

(4) The possibility of screen burning due to operator setup error

or internal arcing is greatly reduced when the pulsed-raster, optical-slit method is used rather than a method using single horizontal-line scanning or a focused or defocused spot.

(5) The combination of pulsed raster and narrow slit opening permits measurements of phosphor persistences from one millisecond to greater than 50 seconds, and, if minor modifications to the equipment are made, decays in the order of microseconds can be measured. Furthermore, true persistence curves can be obtained by conversion of

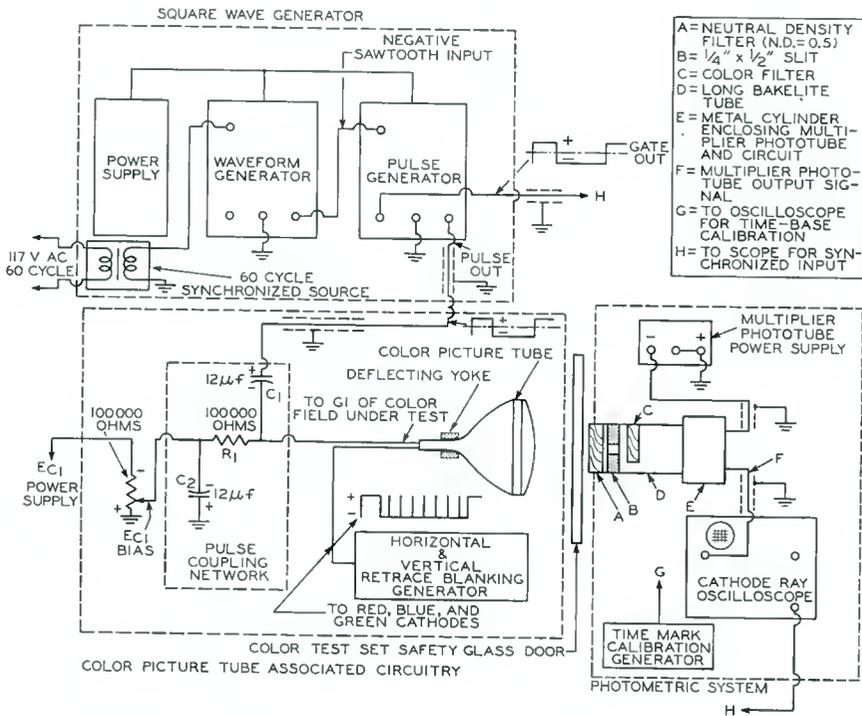


Fig. 1—Schematic diagram of persistence-measuring equipment for color picture tubes.

the equipment for use with a pulsed focus or defocused spot system. Conversion is accomplished simply by removing scanning and a small shunting capacitor across the load resistor of the multiplier phototube.

SQUARE-WAVE GENERATOR

The square-wave generator supplies a nonsymmetrical or symmetrical output pulse through a low-impedance a-c coupling network

to the red, green, or blue control grid (G_1) of the color picture tube under test. It consists of a power supply, a waveform generator, a pulse generator, and a power-line synchronizing source.

The waveform generator supplies a synchronized sawtooth wave for the pulse generator which determines the waveform duration, or total time period, of the output square-wave signal. A 12.6-volt filament transformer operating from the power line is used to supply a

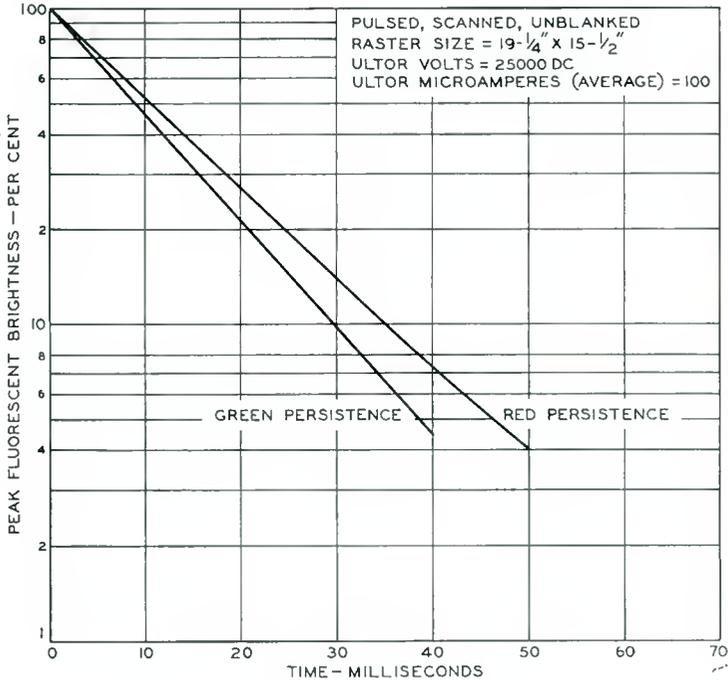


Fig. 2—Persistence of red and green screens.

triggering voltage, synchronized with the power line, to the waveform generator. Operation of the pulse generator in a triggered mode assures locking between the square-wave input to the control grid of the tube under test and the vertical scanning frequency of the deflecting yoke which is also synchronized with the power line. Use of the 60-cycle power line for locking has the advantage that should low-level a-c power-line ground currents or stray light signals be present, these extraneous signals will remain stationary in the oscilloscope display and not roll through the signal being viewed.

The pulsed output from the square-wave generator is also used as

a fixed-level triggering signal for controlling the sweep of the oscilloscope during persistence measurements. This fixed synchronized input also maintains a constant phase relationship between the input signal from the photometric system and the vertical gain adjustment of the oscilloscope.

Measurements of the persistence of phosphors to meet the specifications listed in Table I require a square wave having a period of 120 ± 20 milliseconds and a positive pulse of 50 ± 5 milliseconds. The nearest subharmonics of 60 cycles per second which can be used to supply a square wave having the required period are 8.57 and 7.5 cycles per second (116.67 and 133.33 milliseconds, respectively). Locking of the pulse generator with the raster of the color picture tube is accomplished by varying the pulse frequency by means of the waveform duration control on the waveform generator until the raster is completely stationary. Phasing of the pulse generator pulse output can be adjusted to any point throughout the duration of the sawtooth output of the waveform generator. Nonsymmetrical square-wave pulsing was chosen to permit long-duration persistence readings. However, symmetrical pulsing could have been used just as readily provided the pulse is wide enough to furnish persistence data for the elapsed time desired.

COLOR PICTURE TUBE AND ASSOCIATED CIRCUITRY

An a-c coupling network is used to couple the square-wave generator output to the red, blue, or green control grid of the tube under test. The time constant circuit R_1C_1 in Figure 1 must be such that negligible distortion or only a slight tilt of the square-wave pulse occurs at the picture tube control grid. The use of a 12-microfarad capacitor and a 100,000-ohm resistor assures reasonably good signal application at the control grid for either the 8.57 or 7.5 cycle-per-second low-frequency pulse input. Capacitor C_2 not only by-passes any a-c output of the E_{c1} bias potentiometer voltage source, but also prevents the development of any of the pulse voltage from appearing on the output terminals of the bias power supply and potentiometer system.

A blanking generator is connected to the three paralleled cathodes of the tube. Positive pulses thus applied blank the electron gun beam current during horizontal and vertical retrace, thereby avoiding excitation of the phosphor during retrace. If desired, retrace blanking can be eliminated to reduce circuit complexity, but cleaner persistence waveforms are possible if blanking is used.

The positive portion of the square-wave signal applied to the appropriate control grid of the tube turns on the corresponding electron-gun

beam current. The negative portion of the waveform together with the d-c bias on the control grid completely cuts it off. The d-c control grid bias level must be adjusted so that the effective additive combination of it and the negative portion of the square wave pulse are at cutoff or beyond. Cutoff can conveniently be determined by inspection of the output waveform from the multiplier phototube on the oscilloscope screen. The bias voltage is first reduced so that notches become visible on the output signal. These notches represent normal scanning background superimposed on the pulsed raster. The bias is then increased until the notches definitely disappear. By this procedure, the average ultor current of the tube during pulsing has been adjusted so that no background notches will appear on the oscilloscope during the persistence test. The elimination of normal scanning allows for pulsed raster monitoring only. Thus, the color phosphors cannot be excited during the decay period, and sufficient peak brightness magnitudes can be obtained for proper determination of phosphor persistence.

PHOTOMETRIC SYSTEM

The components of the photometric system are a 6217 multiplier phototube, its bleeder resistors, a slit and color filters, a phototube power supply, a cathode-ray oscilloscope, and a time-mark generator. A general diagram of the photometric system is shown in Figure 1, and a schematic diagram of the 6217 and its circuits is shown in Figure 3.

The 6217 is a 10-stage, head-on multiplier phototube having a spectral response covering the range from about 3,000 to 8,000 angstrom units, and is, therefore, ideal for detecting red, green, or blue light emitted from color picture tubes during persistence testing. The 6217 was selected because it has an unusually low anode dark current; thus, its influence on the anode load resistor is minimized. The use of a low supply voltage of approximately 600 volts d-c also helps reduce dark current. The output current of a multiplier phototube is linear with light flux on the photocathode and nearly exponential with the applied voltage. Linearity of output current with light flux, however, necessitates a voltage divider bleeder current in the order of twenty times the output current. A linearity characteristic of better than one per cent can be achieved by this technique.

A peak-to-peak signal output of 2 volts across the 100,000-ohm load resistor was used to provide sufficient signal gain and to assure multiplier phototube circuit linearity. The 0.001-microfarad load capacitor shunting the 100,000-ohm resistor shorts out any high-frequency horizontal scanning pickup that may appear across the load resistor. This

capacitor can be removed if a defocused spot system without scanning is desired for measuring true persistences of fast decay type phosphors.

As a result of the application of square-wave pulsing to the color picture tube, pulsating light output is produced by the build-up and decay of the color phosphor on the screen. Because the multiplier phototube detects the emitted light through a narrow slit ($\frac{1}{2}$ by $\frac{1}{4}$ inch) at the end of a long piece of bakelite tubing, it optically senses only a small portion of the screen. Red, green, or blue color filters are inserted in the tubing to prevent the effects of stray light or color impurities from influencing the persistence of single-color fields.

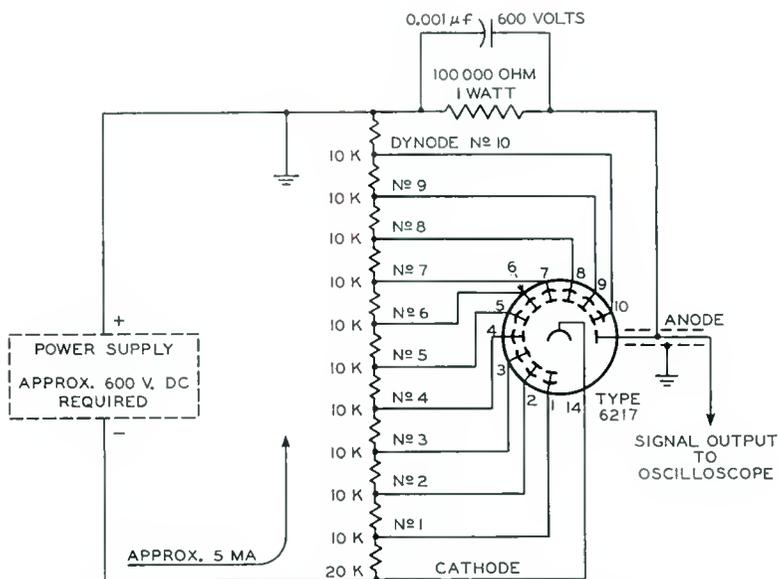


Fig. 3—Schematic diagram of multiplier phototube and associated circuits.

The color filters used are glass mounted Wratten filters #47B (blue), #29 (red), and #61N (green). Because the sensitivity characteristics of the 6217 in the blue and green regions is high relative to the red region, a neutral density (0.5) filter is used with the blue and green color filters to eliminate 6217 overloading. It was necessary to use a vernier-controlled power supply of approximately 600 volts having exceedingly low ripple and noise level because the phototube gain is extremely sensitive to small changes in power-supply voltage.

A precision cathode-ray oscilloscope equipped with stable d-c amplifiers and accurately calibrated vertical amplitude and time base scales is used to monitor the signal output from the multiplier phototube.

Accurately calibrated vertical amplitude scales make it possible to expand the signal by increasing the sensitivity, and thus to expand the scale and permit more accurate persistence measurements. Accurate time scales co-ordinated with a crystal-controlled, temperature-compensated, time-mark generator allow for accurate cross-calibration and adjustment of the horizontal sweep or time base of the oscilloscope. In order to provide persistence measurements accurate to ± 0.5 per cent on this equipment, the oscilloscope trace must be finely focused and not too intense. The use of d-c amplifiers makes it possible to reference the persistence curve to the true zero extinction level of the phosphor.

The effective reduction of 60-cycle hum and other background noise is of primary importance, as difficulty in measurement will otherwise result. Shielding of the high-impedance multiplier phototube load circuit is essential if noise pickup is to be effectively reduced. Extraneous stray fluorescent or other light should be eliminated as the tremendous sensitivity of the multiplier phototube will superimpose these unwanted signals on top of the persistence curve, resulting in decreased accuracy in the determination of the persistence.

WAVEFORM ANALYSIS

The sequence of events which link all portions of the persistence measuring equipment together is given in Figure 4. The particular analysis shown is for the red phosphor, but the approach is the same for the green and blue phosphors.

The nonsymmetrical square-wave pulse of approximately 50 volts peak-to-peak magnitude shown in Figure 4b is superimposed on the d-c grid-bias control of the color picture tube. The leading edge of the positive pulse is synchronized with the 60-cycle power line, point A in Figure 4a, and the vertical-scanning frequency, Q , in Figure 4c. A phase-delay control is incorporated in the pulser to phase the pulse generator output square wave relative to vertical scanning.

Numerous subharmonic frequencies of the square-wave-generator pulse, as shown in Table II, will lock with the 60-cycle power line.

With a positive "on" pulse (t_1 in Figure 4b) of approximately 50 milliseconds, the determination of persistence characteristics specified for the phosphors in the P22 group (given in Table I) will require periods of $T \geq 100$ milliseconds. From the standpoint of longer-persistence measurements, locking frequencies of 8.57 or 7.5 cycles per second were chosen because they produce a cycle having a total duration of 116.67 and 133.33 milliseconds, respectively. The negative portion of the square-wave pulse, t_2 , during which the electron beam

Table II—Subharmonic Frequencies of Square-Wave Generator Pulse

Freq. (<i>f</i>)	Total Time <i>T</i>	
	Millisec.	Sec.
60	16.67	1/60
30	33.33	2/60
20	50	3/60
15	66.67	4/60
12	83.33	5/60
10	100	6/60
8.57	116.67	7/60
7.5	133.33	8/60

* Integral multiples of 60-cycle field time are indicated in this column.

is cut off for these two cases would be 66.67 (116.67 - 50) or 83.33 (133.33 - 50) milliseconds, respectively. In Figure 4b, a locking frequency of 8.57 cycles per second is used.

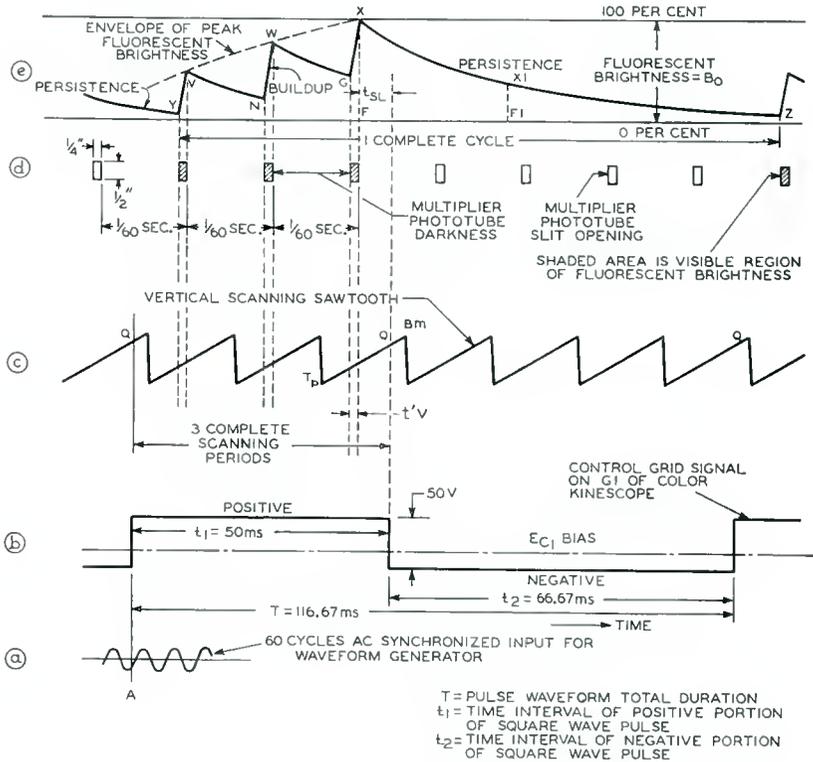


Fig. 4—Analysis of persistence waveforms.

The sawtooth waveform of the current in the deflecting yoke is shown in Figure 4c. Conventional television scanning proceeds horizontally from left to right and vertically from top to bottom. Point T_p is the top or start of vertical scanning, B_m is the finish or bottom of vertical scanning, and Q shows the phase sequence of a given point during vertical scanning with the beginning of the negative portion of the pulse generator output.

The illumination of the narrow slit between the multiplier phototube and the picture tube screen is depicted in its phase relation to the other waveforms in Figure 4d. The long side of the slit is positioned at right angles to the main vertical scanning axis of the color picture tube; the short side coincides with a portion of the vertical deflection indicated by $t'v$ in Figure 4c. The shaded slit areas represent successive scans during which the pulsed raster illuminates the multiplier phototube. Point X in Figure 4e, the point of peak fluorescent brightness, coincides in time with the bottom edge of the narrow slit. The time t_{SL} establishes the location of the bottom edge of the slit with regard to Q , the point at which cutoff of the electron gun first extinguishes the raster. This time depends upon the phase relationship between the negative-going portion of the square-wave pulse and the position of the slit. Because the slit position is fixed, any light picked up by the multiplier phototube will re-occur at 1/60-second intervals.

The solid curve in Figure 4e is the observed build-up and decay curve as displayed on the oscilloscope. The entire buildup and decay cycle, Y to Z , is made up of two principal parts, Y to X and X to Z . Y to X contains both build-up and decay segments, whereas X to Z is a long-decay segment.

In the example given, synchronization of the positive pulse, vertical scan, and slit is such that build-up can occur only three times. Therefore, points V , W , and X represent three peak fluorescent brightness levels during the buildup cycle Y to X . The number of these peaks is determined by the time duration of the positive pulse and the phase position of the square wave pulse relative to the slit. The peak fluorescent brightness envelope indicated by the dotted line in Figure 4e depends on the electron-beam current and pulse duration.

The determination of persistence is made in segment X to Z of the decay or persistence curve. Figure 4e shows that persistence measurements are not limited by the duration of the negative portion of the square-wave pulse, t_2 , because there is an appreciable time interval of darkness from one slit position to the next.

Per cent persistence, by definition, is the ratio of phosphorescent brightness at the desired decay time (for example, 40 milliseconds)

to the peak fluorescent brightness at zero time. As shown in Figure 4e, persistence is the per cent brightness at time F_1 divided by the per cent brightness at time F ; i.e., (X_1F_1/XF) . The peak fluorescent brightness level at point X is considered, therefore, as the 100 per cent persistence, and its time location (F) is thus identified as zero time. The 0 per cent magnitude location is established at the point where the light output is zero. In the example shown in Figure 4e, this condition cannot quite be obtained because the next positive pulse is applied before the zero level can be reached. An easy means for determining this quiescent 0 per cent base line is to use a cathode-ray oscilloscope equipped with d-c amplifiers. With the color picture tube and multiplier phototube at required operating conditions, the multiplier phototube supply voltage is adjusted until the peak fluorescent brightness point X reaches the top line of the scope graph, previously designated as the 100 per cent persistence value. The color picture tube is then cut off by shorting out grid No. 2. The vertical centering of the d-c oscilloscope is adjusted to move the visible, horizontal, single-line scope trace to the bottom horizontal graph line which is now designated as the 0 per cent persistence line. The process is repeated until both these top and bottom graph lines include the peak-persistence point, X , and the zero point, F . The elimination of phototube leakage currents in the load resistor and influence of the stray light is accomplished by this procedure. Persistence measurements from point X , the zero time point, can now be made at 10, 20, 30, 40, etc. milliseconds. The cathode-ray oscilloscope d-c vertical amplifiers will now automatically reference the persistence curve to the true zero extinction level of the phosphor.

A number of oscilloscope waveforms were photographed; these pictures are shown in Figures 5 through 10 to illustrate the considerations which were covered in this waveform analysis section. The electrical signal in Figures 6 and 7 is shown inverted. Auxiliary amplifiers can be used to convert the persistence waveform to its normally expected relationship as shown in Figures 5, 8, 9, and 10.

APPENDIX I—GROUP PERSISTENCE AND COLOR

The luminous group persistence and color of a screen at any time t , following cessation of excitation, depend upon

- (1) The color balance before cessation of excitation.
- (2) The colors of the phosphors.
- (3) The persistences of the different phosphors forming the screen.

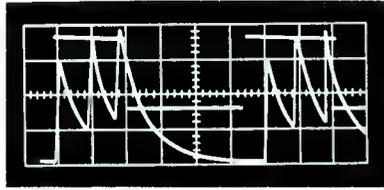


Fig. 5—Persistence curve for red phosphor: (a) Oscilloscope 20-millisecond/centimeter time base; (b) Square wave pulse, 50 milliseconds positive, 67 milliseconds negative; (c) 3 peaks prior to curve decay; (d) Both signals right side up. Auxiliary amplifier technique used to invert persistence waveform; (e) Square-wave pulse shown synchronized with red persistence curve.

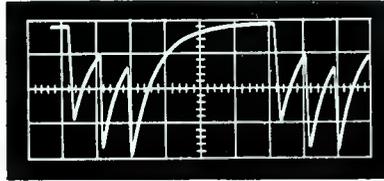


Fig. 6—Same as Figure 5 but persistence waveform upside down: (a) Load circuit generating a negative voltage output with excitation causing waveform inversion; (b) Oscilloscope 20-millisecond/centimeter time base.

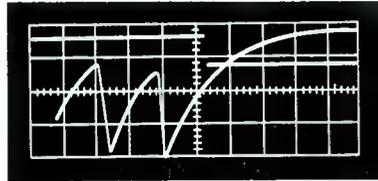


Fig. 7—Same as Figure 5 but emphasizing the axial shift of the negative-going portion of the square-wave pulse with regard to peak fluorescent brightness point: (a) Oscilloscope 10-millisecond/centimeter time base; (b) Axial movement can be accomplished by mechanical movement of the slit in relation to vertical scanning, by pulse phasing, or by adjustment of vertical centering of the raster.

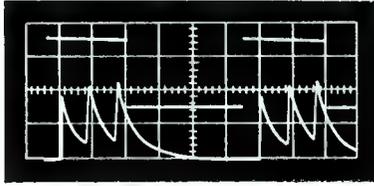


Fig. 8—Persistence curve for green phosphor: (a) Oscilloscope 20-millisecond/centimeter time base; (b) same as Figure 5 but representing the green field persistence curve.

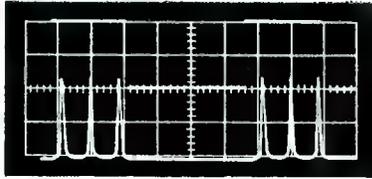


Fig. 9—Persistence curve for blue phosphor: (a) Oscilloscope 20-millisecond/centimeter time base; (b) Same as Figure 5 but representing short persistence of the blue field.

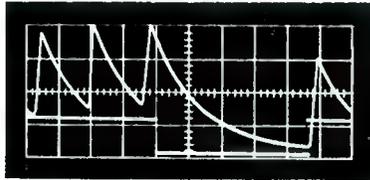


Fig. 10—Red persistence curve: (a) Red persistence curve inverted by auxiliary amplifier technique to maintain normal expected persistence waveform; (b) Four peaks instead of usual 3 due to the increase of positive pulse width from 50 to approximately 70 milliseconds; (c) Oscilloscope 10-millisecond/centimeter time base; (d) Signal generator pulse superimposed on multiplier phototube signal.

Consider a color gamut formed by three phosphors having the I.C.I. coordinates $[x_b, y_b]$, $[x_g, y_g]$, and $[x_r, y_r]$ and balanced for the color x_a, y_a . If, then, the intensities of the components are changed to B, G , and R respectively, the new color balance will have the I.C.I. coordinates x_j, y_j and the relative brightness or persistence Y_j/Y_a . It can be shown that

$$x_j = \frac{B(x_b k_b) + G(x_g) + R(x_r k_r)}{B(k_b) + G + R(k_r)},$$

$$y_j = \frac{B(y_b k_b) + G(y_g) + R(y_r k_r)}{B(k_b) + G + R(k_r)},$$

and

$$Y_j/Y_a = \frac{B(y_b k_b) + G(y_g) + R(y_r k_r)}{y_b k_b + y_g + y_r k_r},$$

in which

$$k_b = \frac{(y_g - y_a)(x_a - x_r) - (x_g - x_a)(y_a - y_r)}{(x_b - x_a)(y_a - y_r) - (y_b - y_a)(x_a - x_r)},$$

and

$$k_r = \frac{(y_g - y_a)(x_a - x_b) - (x_g - x_a)(y_a - y_b)}{(x_b - x_a)(y_a - y_r) - (y_b - y_a)(x_a - x_r)}.$$

Application of these equations to the typical I.C.I. values of the dry powder phosphors (Table I) results in the following equations:

$$x_j = \frac{232B + 203G + 539R}{1591B + 1000G + 799R},$$

$$y_j = \frac{83B + 728G + 261R}{1591B + 1000G + 799R},$$

$$Y_j/Y_a = \frac{83B + 728G + 261R}{1071}.$$

In these equations B, G , and R may be considered, for application to the present problem, the persistences of the blue, green, and red phosphors, at a time t , expressed as fractions of their fluorescent brightnesses. The fluorescent color balance is taken as $x_a = 0.287$ and

$y_a = 0.316$. x_j and y_j are, then, the I.C.I. coordinates for the screen color at the time t , and Y_j/Y_a is the group persistence of the screen expressed as a fraction of the screen fluorescent brightness, Y_a .

APPENDIX II—THE MAGNITUDE OF THE ERROR INTRODUCED IN THE OBSERVED PERSISTENCE BY THE PREVIOUSLY SCANNED LINE

It must be realized that the pulsed-raster, optical-slit method presented in this paper is for quality control monitoring of screen persistences and not for the general empirical derivation of persistence curves. Although the observed curves for the phosphors in the P22 group are identical with those derived by more precise and generally applicable methods, such as defocused or focused spot, or line selection, it will not be true for nonexponential decays. The slit method will give values which, for a number of screen types, will be considerably in error if more than one scanning line fills the optical slit.

Screens using P4, P7, P11, P14, P20 and P23 type phosphors having persistences of approximately 10 per cent at 54 microseconds are particularly subject to these errors. The errors are 37 and 21 per cent for these screens, assuming hyperbolic decay constants of -1 and -2 respectively, and assuming that the previously scanned line was at peak intensity 63 microseconds before the primary scanned line.

The secondary scanned line does not appreciably affect screens having phosphors of very short persistence (error 1 per cent if the screen has 10 per cent brightness at $t \leq 1$ microsecond), such as P16 or P24, or screens having long persistence phosphors (10 per cent at $t \geq 20$ milliseconds), such as P2 and the red and green of the P22 group.

As pointed out in the text, the pulsed-raster, optical-slit equipment can readily be modified for precise academic determinations of persistence curves. The method described in this paper, however, is more satisfactory for quality control purposes.

MICROWAVE PROPAGATION OVER ROUGH SURFACES*

BY

M. P. BACHYNSKI

RCA Victor Company, Ltd.,
Montreal, Canada

Summary—The field of propagation of short radio waves over rough surfaces is surveyed. The results of available experimental measurements are summarized and theories which are readily amenable to calculation are discussed. The experimental findings may serve as an indication of typical values to be expected in practical situations. The effects of both irregular terrain and sea surfaces are considered. The limitations of various theoretical approaches are outlined and aspects requiring further study are indicated. An extensive bibliography is also included.

INTRODUCTION

THE PROBLEM of reflection of electromagnetic space waves above approximately 30 megacycles by irregular surfaces is of considerable interest in radar and communications. The irregularities of the earth's surface and man-made obstructions impose limitations on the coverage of television stations and communication channels, introduce multipath effects which may cause "ghosts" and "crosstalk," degrade the free-space patterns of short-wave antennas (in some instances drastically), and limit the range, accuracy and definition of search radar. Further, these irregularities are constantly changing—seasonally in the case of vegetation and with the prevailing wind systems and lunar cycle for sea surfaces. For this reason, accurate prediction of radar and communication performance is rather difficult.

During the last decade many experiments were performed and numerous attempts were made to develop theories to explain these measurements. On the one hand there has been an accumulation of field experience to guide the engineering of practical systems and on the other, an attempt to relate the pertinent parameters to ready calculation. The greatest difficulty has been the inability to establish the relationship of the important parameters with the practical terrain profiles and electrical properties, primarily because of the vast and random fluctuations of these quantities both in location and time.

* Manuscript received March 31, 1959.

Thus, the many theoretical approaches are limited by assumptions which have to be introduced in order to obtain a solution.

This paper summarizes results which have been obtained by experiment and presents, in condensed form, various theories which permit rapid calculation in order to predict performance. The experimental measurements are an indication of the values which would be encountered in practical situations. The theoretical formulas listed apply when the terrain profile closely approaches the model assumed in their derivation. An extensive bibliography of current literature forms the final section of the paper.

PROPAGATION OVER IRREGULAR TERRAIN (EXPERIMENT)

Smooth-Earth Propagation

Reflection from Smooth Surfaces

In wave propagation, a smooth flat earth introduces a specularly reflected wave whose interference with the direct wave is determined by the phase difference due to the difference in path lengths over which the energy travels. The relative strength of the reflected wave is given by the well-known Fresnel equations for specular reflection from a smooth-plane surface. (In these discussions it is assumed that a "surface wave" does not exist.) Thus it is found that the reflection coefficient for a smooth surface is determined almost exclusively by the dielectric constant of the surface. The conductivity, upon which the absorption in the material of the surface is primarily dependent, exerts only a minor influence. Conversely, when specular reflection occurs, the electrical constants of the material forming the reflecting surface can generally be estimated from the reflection coefficient. For example, measurements at a wavelength of 9 centimeters¹ have yielded values of the dielectric constant in agreement with values determined by other means. However, this method is not satisfactory for determination of conductivity of surface material.

Region of Reflection

In analyzing the influence of the reflected field it is necessary to locate the areas on the earth's surface from which the reflected radiation has a specified phase relative to the direct ray from the primary source to the field point. This is normally done by plotting reflected paths which are an integral number of half-wavelengths longer than the slant range measured directly between transmitter and receiver.

¹L. H. Ford and R. Oliver, "An Experimental Investigation of the Reflection and Absorption of Radiation of 9 cm Wavelength," *Proc. Phys. Soc.*, Vol. 58, p. 265, May, 1946.

A family of ellipses is obtained by the intersection of the ground plane with a series of ellipsoids of revolution whose foci are the transmitting and receiving terminals and whose axis of revolution is the direct ray between the transmitter and receiver. These Fresnel zones are fairly narrow but highly elongated ellipses; the vast area of the first zone suggests that a large, smooth elongated area is necessary in order that a well-developed lobe pattern be formed. Some limited experiments on wire netting² suggest that the region around the point of reflection (as defined by geometric optics) should be flat for an area of the order of half the first Fresnel zone before the reflected wave will be of the same order as that reflected from an infinite reflecting plane. The essential difference between the Fresnel zones of a reflecting region on the earth's surface and those encountered in physical optics is that the optical zones are concentrically symmetrical about the direction of propagation. Hence the amplitude of excitation varies slowly from zone to zone, so that the contributions from successive zones almost completely cancel, leaving the contribution from approximately half the area of the first zone as the net result. The zones of a reflecting region, on the other hand, are inclined at large angles to the direction of propagation so that due to their great eccentricity, especially at small grazing angles, the amplitude of the incident wave will vary markedly over the zone. Thus it is not clear whether the effect of the first zone should predominate as in the former case. This is a problem which has not yet been extensively treated.

Effect of Antenna Position

For line-of-sight transmission paths which are near grazing, the reflection coefficient of a smooth terrain may be taken as -1 for both polarizations when divergence due to earth curvature is neglected. Under such conditions the field strength from the direct and specularly reflected rays results in a lobe structure given by

$$\left| \frac{E}{E_0} \right| = 2 \sin \left\{ \frac{kh_1h_2}{d} \right\},$$

where h_1, h_2 are the transmitter, receiver heights above ground,

d is the transmitter-receiver separation,

$k = 2\pi/\lambda$, where λ is the wavelength.

²J. S. Hey, J. J. Parsons, and F. Jackson, "Reflection of Centimeter e.m. Waves over Ground and Diffraction Effects With Wire Netting Screens," *Proc. Phys. Soc.*, Vol. 59, p. 847, September, 1947.

In many practical cases $h_1, h_2 \ll d$, so that only the first lobe is of consequence; hence the relative field strength can be expressed

$$\left| \frac{E}{E_0} \right| = \frac{2kh_1h_2}{d}.$$

This leads to the familiar rule that the received field strength increases linearly with transmitter height.

Surface Roughness

Criterion of Roughness

The efficiency of specular reflection from a rough surface is a function of the grazing angle and the ratio of the surface roughness dimensions to the wavelength. The most widely accepted criterion involving these parameters is that of Lord Rayleigh. The path difference between two rays such that the phase difference and hence the effective roughness is small and the surface is effectively smooth is arbitrarily taken as $\lambda/4$ (Reference (3)). Thus, quantitatively, if

$$l \sin \gamma < \frac{\lambda}{8}$$

where l is the height of an irregularity of the surface, and γ is the grazing angle, then the surface is considered to be smooth (see Figure 1).

This concept does not lend itself to a quantitative treatment of the effects of roughness. Experiments have shown¹ that if the Rayleigh criterion exceeds $\lambda/2$ the value of the reflection coefficient is of the order of 0.1, and that specular reflection is obtained only for very flat surfaces. (For reflection coefficients of ~ 0.50 , $l \sin \gamma \cong 0.2\lambda$.)

A useful concept of the criterion of roughness is the phase deviation, Φ , in a propagation path as given by Bullington.⁴ Thus

$$\Phi = 2\pi \frac{hH}{H_0^2}$$

³ D. E. Kerr, *The Propagation of Short Radio Waves*, Vol. 13, MIT Radiation Lab Series, McGraw-Hill Book Co., Inc., New York, New York, 1951, p. 411.

⁴ K. Bullington, "Reflection Coefficient of Irregular Terrain," *Proc. I.R.E.*, Vol. 42, p. 1258, August, 1954.

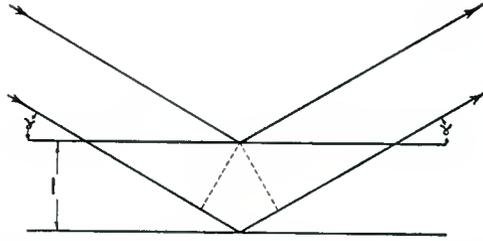


Fig. 1—Pertinent parameters for reflection from a rough surface are the grazing angle, γ , the height of an irregularity of the surface, l , and the wavelength, λ . The phase difference between rays reflected from two levels is shown.

where H is the average clearance,

$\pm h$ is the profile deviation,

$$H_0 = \sqrt{\lambda \frac{x(d-x)}{d}} = \text{the first Fresnel zone clearance} \\ \text{(see Figure 2).}$$

Comparing this with the Rayleigh criterion, specular reflection is expected when $|\Phi| \leq \pi/2$; diffuse reflection occurs for greater phase deviations. On this basis, most overland propagation paths can be classed as rough. It must be kept in mind that the transition from specular to diffuse reflection does not take place at a single point but must occur over a region of increasing roughness. Establishment of this transition range of roughness is still one of the unsolved problems in this field.

Losses Due to Terrain Irregularities

In field-strength measurements over terrain, the amplitude of the local variations depends on the ground contour, the surface, the fre-

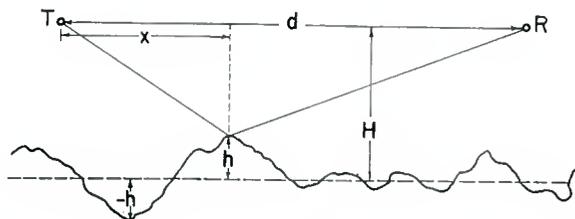


Fig. 2—Parameters defining the phase deviation $\Phi = 2\pi hH/H_0^2$ in propagation over irregular terrain.

quency, and the polarization. Small irregularities in ground contour are sufficient to cause considerable departures from the values of signal computed assuming a flat surface. For short wavelengths the amount of specular reflection is markedly affected by surface roughness. The effects of surface roughness are manifested not only by reduction of the reflection coefficient, but also by its variation over a wide range from location to location. Short wavelengths are sufficiently scattered by most types of earth surfaces so that little or no specular reflection will be observed except at almost zero grazing angle. The critical factor in determining specular reflection is the roughness of the surface, not the conductivity and dielectric constant of the surface.

Measurements by many workers have shown that on most overland paths the reflection coefficient is in the range 0.2-0.4, and seldom does the reflection coefficient exceed 0.5 at frequencies above 1,500 megacycles on near-grazing paths. Small differences in location can produce vast changes in the reflection coefficient so that the magnitude of the reflection coefficient cannot be predicted accurately from the gross features of the path profile. Observations taken along radial lines over both smooth and hilly terrain show close correlation with smooth-earth theory only along smooth radials. The measurements are almost always below the predicted values (sometimes deviations up to 30 decibels are observed). Over hills, the attenuation increases with angle of diffraction and number of diffractions.

It is found that for broadcast coverage the median field-strength deviations from plane-earth field-strength values are approximately independent of distance from the transmitter at all frequencies. However, as the frequency increases so also does the amount by which the median observed field for any distance falls below the smooth-earth theoretical value, and so does the scattering of observed values. In the range 100-600 megacycles, the fluctuations for vertical polarization are about 1 decibel greater than those for horizontal polarization. Fluctuations on both polarizations increase with increasing frequency. Local obstacles at the lower frequency tend to result in smaller variations for horizontal than for vertical polarization. As the frequency increases, this tendency is reduced and both polarizations give approximately the same variation. The median broadcast field strength at frequencies above 40 megacycles, based on data from the U. S. Federal Communications Commission* and RCA,** varies inversely with fre-

* H. Fine, "UHF Propagation Within Line of Sight," FCC, TRR Report 2.4.12, June, 1951.

** G. G. Brown, J. Epstein and D. W. Peterson, "Comparative Propagation Measurements; Television Transmitters at 67.25, 288, 510 and 910 Megacycles," *RCA Review*, Vol. IX, pp. 171-201, June, 1948.

quency with respect to the plane-earth field strength;⁶ this is shown in Figure 3. Measurements in Great Britain⁵ are essentially in agreement. Egli⁶ has published a number of curves on existing data showing the attenuation-frequency dependence expected for various percentages of the locations.

In many shadowed locations, multipath propagation is very much in evidence; strong signals may be observed from many directions. On wide-band systems this can cause effects such as "ghosts" in television reception and cross-talk in multichannel voice systems. The effects of multipath distortion appear to be approximately independent of polarization and frequency in the range 60-3,300 megacycles. As would be expected, more multipath distortion is present in highly built up and mountainous areas where off-path reflections are prevalent. The use of directive antennas appears to be the most effective method of reducing multipath effects.

Effects of Natural Conditions

Space-wave propagation at high frequencies over irregular terrain is subject to many variations and uncertainties. In many instances the experimental data has been used to derive empirical modifications of the idealized theory.⁶⁻⁹ In this section, the effects of various natural conditions on the received power as found experimentally are summarized.

Table I describes propagation over various types of terrain. If the earth is very smooth, the reflection coefficient can be calculated from a knowledge of the electrical constants of the terrain. As the roughness increases, specular reflection and hence the reflection coefficient become steadily less until only diffuse scattering is present. Hills and ridges are opaque obstructions for short radio waves and their losses increase with increasing frequency, terrain roughness, diffraction angle, and number of diffractions.

The effects of trees and buildings on VHF and UHF are difficult to evaluate quantitatively. A sufficiently dense and extensive wood, for example, may approach opacity for VHF, but with less dense woods the transmitted signal through the wood may be greater than that

⁵ J. Saxton and B. N. Harden, "Ground-Wave Field-Strength Surveys at 100 and 600 Mc/s," *Proc. IEE*, Part 3, Vol. 101, p. 215, July, 1954.

⁶ J. J. Egli, "Radio Propagation above 40 Mc/s over Irregular Terrain," *Proc. I.R.E.*, Vol. 45, p. 1383, October, 1957.

⁷ R. S. Kirby, H. T. Dougherty and P. L. McQuate, "VHF Propagation Measurements in the Rocky Mountain Region," *Trans. I.R.E. PGVC*, July, 1956.

⁸ K. Bullington, "Radio Propagation Variations at VHF and UHF," *Proc. I.R.E.*, Vol. 38, p. 27, January, 1950.

⁹ D. W. Peterson and J. Epstein, "A Method of Predicting the Coverage of a Television Station," *RCA Review*, Vol. XVII, p. 571, December, 1956.

diffracted either over or around it when the receiving point is near the wood. The observations shown in Tables II and III are representative of the values to be expected. In general, at the lower frequencies the attenuation due to trees is considerably less, while the relative difference between the values for the two polarizations is more pronounced. The attenuation appears to be approximately proportional to frequency. Effects due to buildings exhibit less frequency dependence than effects due to either trees or hills, but buildings are, nevertheless, rather opaque at frequencies of the order of thousands of megacycles. Also of interest are the natural conditions which change seasonally. For example, the vertical-coverage diagram of a transmitter operating at

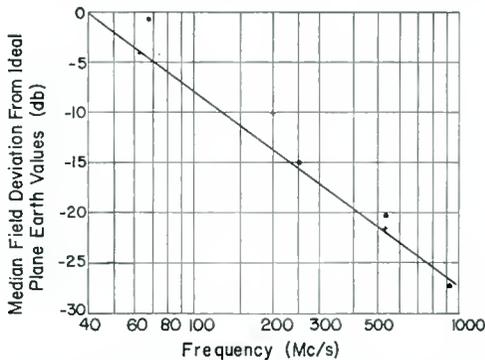


Fig. 3—Median field strength variation with frequency based on data from FCC (•) and RCA (+) compared to field calculated over a plane earth.

high frequencies may be appreciably modified by the presence of a layer of snow and ice. In Table IV the effects of the change of the effective dielectric constant of the earth due to such a layer are summarized. The resulting change in the reflection coefficient can then be estimated from the Fresnel reflection equations. Experiments¹⁰ at VHF over smooth ground have shown variations of 14 decibels with season (dielectric constant values change from 3 to 30). This change is due to the moisture content of the ground which mainly determines the effective dielectric constant of the soil and is relatively independent of the type of ground. The temperature coefficient of the dielectric constant of ground is very small ($\sim -0.005/^{\circ}\text{C}$) and hence no direct influence of temperature on VHF propagation may be expected. An indirect influence may exist insofar as the temperature affects the ground moisture content and hence the dielectric constant.

¹⁰ B. Josephson and A. Bloomquist, "The Influence of Moisture in the Ground, Temperature and Terrain on Ground Wave Propagation in the VHF Band," *Trans. I.R.E. PGAP*, p. 169, April, 1958.

Table I—Reflection Coefficients of Terrain

Experimenters	Frequency (Mc)	Conditions	Reflection Coefficient	Variation	Observations
Ford & Oliver ¹	3,260	Smooth dry ground	R can be accurately calculated from Fresnel's Equations for specular reflection		$\epsilon = 2$; $\eta = 3 \times 10^8$ esu; Attenuation: 36 db/meter
		Smooth wet ground			$\epsilon = 5-25$; $\eta \sim 6 \times 10^9$; Attenuation: 220 db/meter
		Turf-short dry grass 2 cm wet grass			$\epsilon = 3$; $\eta \sim 10^8$; Attenuation: 50 db/meter $\epsilon = 6$; $\eta \sim 5 \times 10^9$; Attenuation: 80 db/meter
Sherwood & Ginzton ¹¹	3,000	Plowed field			Complete scattering
		Grass (4-8 inches)	0.3—0.5 0.05—0.2 0.05—0.2		Horizontal polarization Vertical polarization Both polarizations
Bullington ⁴	4,000	Beet field Brush—trees 3-10 ft., weeds	0.05—0.2		Both polarizations
		Irregular terrain	0.1—0.4		
Epstein & Peterson ¹²	850	Swamp (grass 6 feet high)	0.2	± 1 db	
Saxton & Harden ⁵	102.6 593.6	Irregular terrain Terrain with trees		2-4 db 8-10 db	Shows increased scattering with increasing frequency
McPetrie & Ford ¹³	30-3,000	Hills (bare ridges) (shadow region)	Attenuation in shadow ~ 17.6 db/unit distance D_0 $D_0 = (r^2/\lambda/\pi)^{\frac{1}{2}}$ $r =$ radius of curvature		In some cases vertical polarization shows a lower rate of attenuation. The shadow losses increase with increasing frequency and terrain roughness
Kirke & Rowden ¹⁴	90	Hills (shadow region)			Vertical polarization slightly preferred

- ¹¹ E. M. Sherwood and E. L. Ginzton, "Reflection Coefficient of Irregular Terrain at 10 cm," *Proc. I.R.E.*, Vol. 43, p. 877, July, 1955.
- ¹² J. Epstein and D. W. Peterson, "An Experimental Study of Wave Propagation at 850 Mc," *Proc. I.R.E.*, Vol. 41, p. 595, May, 1953.
- ¹³ J. S. McPetrie and L. H. Ford, "Some Experiments on the Propagation of 9.2 cm Wavelength, Especially on the Effects of Obstacles," *Jour. IEE*, Part IIIA, Vol. 93, No. 3, p. 531, 1946.
- ¹⁴ H. L. Kirke, R. A. Rowden, and G. I. Ross, "A VHF Field Strength Survey on 90 Mc/s," *Proc. IEE*, Part 3, Vol. 98, p. 343, September, 1951.

Table II—Attenuation due to Trees

Experimenters	Frequency (Mc)	Conditions	Polarization	Attenuation	Observations
Trevor ¹⁵	250	500 feet of trees (winter)	Horizontal	.07 db/meter	Summer Winter
			Vertical	.09 db/meter	
	500	500 feet of woods and underbrush	Both	.12 db/meter	
	500	Scrub pine	Horizontal	.09 db/meter	
			Vertical	.05 db/meter	
Saxton & Lane ¹⁶	100	Lime-tree screen	Horizontal	.03 db/meter	Attenuation is less at lower frequencies where polarization effects are also most pronounced. Spatial variation is less for horizontal polarization than for vertical polarization.
		in full leaf	Vertical	.06 db/meter	
	540		Horizontal	.15 db/meter	
	1,200		Vertical	.25 db/meter	
			Both	.35 db/meter	
Josephson & Bloomquist ¹⁰	30		Vertical	.02 db/meter	For longer paths, attenuation (per unit distance) decreases.
	100		Vertical	.04 db/meter	

- ¹⁵ B. Trevor, "Ultra-High-Frequency Propagation Through Woods and Underbrush," *ECA Review*, Vol. 5, p. 97, July, 1940.
- ¹⁶ J. Saxton and J. A. Lane, "VHF and UHF Reception-Effects of Trees and Other Obstacles," *Wireless World*, Vol. 61, p. 229, May, 1955.

Table III—Attenuation due to Buildings

Experimenters	Frequency (Mc)	Type of Walls	Attenuation	Variations	Observations
McPetrie & Ford ¹³	3,260	Brick (dry)	0.5 db/cm		
		(wet)	2.0 db/cm		
		Tiles (dry)	1.1 db/cm		
		(wet)	3.4 db/cm		
		Slate (dry)	2.7 db/cm		
		(wet)	3.3 db/cm		
		Wood (dry)	1.1 db/cm		E-field parallel to grain
(wet)	2.2 db/cm		E-field perpendicular to grain		
Saxton & Lane ¹⁵	30	Brick (dry)	0.2 db/cm		
	3,260	Brick (dry)	0.43 db/cm		
Kirke & Rowden ¹⁴	90	City area, 20 feet high receiving aerial	10-20 db down from open area	± 4 db	Less variation with horizontal polarization
Saxton & Harden ⁵	102.6	Built-up region (many buildings)		6-8 db	
	593.6	Built-up region with trees		15-20 db	

Table IV—Effects of Snow and Ice

Experimenters	Frequency (Mc)	Electrical Properties	Observations
Josephson & Bloomquist ¹⁰	50	Ground with no snow: $\epsilon = 15$ 25 cm snow: $\epsilon = 12$ 50 cm snow: $\epsilon = 9$ 100 cm snow: $\epsilon = 8$	Snow covering causes far greater decrease in dielectric constant (ϵ) than does freezing of the ground.
		Water with no ice on top: $\epsilon = 80$ 25 cm ice: $\epsilon = 50$ 50 cm ice: $\epsilon = 36$ 100 cm ice: $\epsilon = 28$	
Saxton ¹⁷		$\epsilon = 3.05$ (snow)	As ice and snow begin to melt, the presence of even a slight amount of liquid water will greatly increase the absorption. Layers of snow and ice in general: 1. reduce the reflection coefficient for horizontal polarization below the value in the absence of the layer 2. cause considerable variation in the reflection coefficient both above and below the value without the layer.
	10	Specific gravity: 0.01 — 0.3 0.3 spec. grav. snow — 0.007 db/meter attenuation — .7 db/meter attenuation	
	10,000		
Cumming ¹⁸	10,000	$\epsilon = 3.15$ (snow) Loss tangent = 27×10^{-4} at 0°C 6×10^{-4} at -20°C	Loss tangent of snow is a function not only of the density and temperature but also of the crystalline structure.

¹⁷ J. A. Saxton, "Reflection Coefficient of Snow and Ice at VHF," *Wireless Engineer*, Vol. 27, p. 17, January, 1950.
¹⁸ W. A. Cumming, "The Dielectric Properties of Ice and Snow," *Jour. Appl. Phys.*, Vol. 23, p. 768, July, 1952.

PROPAGATION OVER SEA SURFACES (EXPERIMENT)

Forward Propagation

There are often large and rapid fluctuations in the energy reflected from the sea surface; this is particularly true when the sea is rough, so that the position and size of the surface irregularities which cause these variations are continually changing. If, under these circumstances, a reflection coefficient is assigned to this rough reflecting surface, the result is a wide scattering of the measured values (see Kerr³ p. 423, for example). A recent model¹⁹ which appears to fit existing experimental data for forward propagation considers the electromagnetic signals over the ocean to be composed of the vector sum of a constant direct signal, a coherent reflected signal whose amplitude and phase are fixed by the transmission geometry and the sea state, and a fluctuating reflected component of random amplitude and phase. This random incoherent signal is determined by the roughness of the sea and depends on the roughness parameter, $h\gamma/\lambda$. It is found to reach an asymptotic value for an $h\gamma/\lambda$ of approximately one hundred milliradians. The variation of the coherently reflected ray with sea roughness relative to a smooth sea at the same grazing angle is found to fit fairly well the exponential law

$$\frac{C}{D\Gamma} = e^{-2(kh\gamma)^2}$$

where C = coherent reflected-ray amplitude,
 D = direct-ray amplitude,
 Γ = smooth-sea Fresnel reflection coefficient.

In recent years attempts have been made to relate the ocean-wave spectra to the radio-signal spectra.^{20,21} It is found that ocean-wave spectra can be generally classed as singly and doubly peaked. These peaks can occur at frequencies as low as 0.065 cycle per second and as high as 0.12 cycle per second. The half-power "bandwidths" of the wave spectra range between 0.03 and 0.17 cycle per second. The radio-

¹⁹ C. I. Beard, I. Katz and L. M. Spetner, "Phenomenological Vector Model of Microwave Reflections from Ocean," *Trans. I.R.E. PGAP*, p. 162, April, 1956.

²⁰ C. I. Beard and I. Katz, "The Dependence of Microwave Radio Signal Spectra on Ocean Roughness and Wave Spectra," *Trans. I.R.E. PGAP*, April, 1957.

²¹ LaGrove, A. W. Straiton, and H. W. Smith, "Synthesis of Radio Overwater Paths," *Trans. I.R.E. PGAP*, p. 48, April, 1955.

wave spectra on the other hand are broader than the ocean spectra (up to 1 cycle per second or more) with the peaks generally occurring at the same frequency as the peaks in the simultaneous ocean-wave spectra. In many instances, minor maxima occur in the radio spectra which appear to be second and possibly third harmonics of the ocean-wave spectra maxima. (LaGrove²¹ has calculated radio signal strengths which are in substantial agreement with experiment and which contain the frequency of the water-level cycles and also the second and third harmonics of the water-level cycles by assuming reflections from a plane reflecting surface which is rising and falling sinusoidally with time.) The broadening of the radio signal increases significantly with increasing roughness (increasing $h\gamma/\lambda$) and appears to depend more on the swell frequency of the ocean than on the width of the ocean spectra itself.

The doppler shift of radio waves²² reflected from a sea surface measured at a frequency of 13.56 megacycles has shown a small range of doppler frequencies with a principal frequency component which was constant at 0.38 cycle per second irrespective of wind conditions and sea state. Other measurements²³ have shown the percentage deviation from the mean frequency to be dependent on the grazing angle, but the doppler frequency to be relatively independent of sea state. This implies that the instantaneous doppler-frequency variation is a function of the large-scale pattern of the ocean waves, while the instantaneous amplitude variation is dependent on the motion and orientation of small facets. This has led to the suggestion that variable-frequency radio equipment be used as an ocean-wave spectrometer.

Backscattering and Sea Clutter

Sea clutter caused by the reflection and backscattering of the transmitted energy from the many randomly orientated facets of the irregular sea surface may limit the usefulness of radar for targets on or near the surface. If the effective scattering area of the rough sea surface is appreciably greater than the effective cross-sectional area of the target, the target will be obscured.

The radar cross-sectional area (σ°) for such a sea surface is defined as the effective radar cross section per unit area of the mean sea surface. The radar cross section depends upon the grazing angle, the frequency and polarization of the incident radiation, the sea state, and

²² D. D. Crombie, "Doppler Spectrum of Sea Echo at 13.56 Mc/s," *Nature*, Vol. 175, p. 681, April, 1955.

²³ J. C. Wiltse, S. P. Schlesinger, and C. M. Johnson, "Backscattering Characteristics of the Sea in the Region 10-50 kmc," *Proc. I.R.E.*, Vol. 45, p. 220, February, 1957.

the position relative to the sea surface pattern (i.e., upwind, downwind, etc.).

Experiments conducted in recent years^{23,24} have shown that the cross section per unit area (σ°) is an increasing function of γ , the grazing angle. There exists a "critical angle" below which σ° decreases rapidly with decreasing angle and above which it rises much more slowly. The critical angle decreases with increasing frequency. For rough seas, the scattering cross section is increased for grazing angles below about 70° and usually decreased for greater grazing angles. In calm sea, horizontal polarization gives less clutter than vertical polarization, but in rough seas or at large grazing angles clutter is independent of polarization. The cross-polarized values of σ° are much lower (15-20 decibels) than those of the transmitted polarization. For a given sea state, σ° is approximately independent of frequency for large grazing angles ($15^\circ \leq \gamma \leq 90^\circ$) but does depend on the grazing angle. At grazing angles below 15° there is a strong frequency dependence. At 60° grazing angle, σ° is found to be almost independent of wave aspect (upwind, downwind, crosswind).

The percentage fluctuations of the received-signal amplitudes about the mean value are greater for large grazing angles than for small. There is good correlation of instantaneous amplitudes among signals received simultaneously at different microwave frequencies.

At low grazing angles the basic scattering appears to be due to small facets which overlie the main large-scale wave pattern or swell. Facets having a perimeter of $\sim \lambda/2$ backscatter most effectively at small grazing angles, and the backscattering of a facet increases about as the square of its slope. Facets near the wave crests contribute most strongly to the backscatter. The frequency dependence of the backscattering parameter σ° is determined by the size distribution of the facets. At large grazing angles some of the facets are viewed at normal incidence and thus backscatter strongly. The angular variation of σ° is then determined mainly by the slope distribution of the facets. On the basis of optical measurements of slope distribution, σ° at vertical incidence is inversely proportional to wind speed, but at grazing incidence it is approximately proportional to wind speed. The scattering is appreciably greater looking upwind than downwind, and present theories which consider the wind effect predict values far below those which are observed.²⁵

²⁴ H. Davies and G. G. MacFarlane, "Radar Echoes from the Sea Surface of Centimeter Wavelength," *Proc. Phys. Soc.*, Vol. 58, p. 717, November, 1946.

²⁵ M. Katzin, "On the Mechanisms of Radar Sea Clutter," *Proc. I.R.E.*, Vol. 45, p. 44, January, 1957.

At high grazing angles the backscatter appears to be due to large facets, the power being proportional to the fraction of the surface which is "large." If the slope and size distributions are independent, then the angular dependence of σ° is substantially independent of frequency. Some expressions^{25,26} are included to illustrate the limiting values of the radar cross section σ° to be expected.

(a) Grazing incidence and small facets:²⁵

$$\sigma_{0^\circ}^\circ = \frac{N_0}{\lambda} \times 9 \times 10^{-3} W^{3/4} (1 \pm 2 \times 10^{-3} W),$$

where N_0 is a constant (estimated experimentally to be 7.2×10^{-5})

W = wind speed in knots and the + is used for upwind and
- for downwind.

(b) Vertical incidence:²⁵

$$\sigma_{90^\circ}^\circ = \frac{1.53 \times 10^3 N_0}{\lambda W},$$

so that $\frac{\sigma_{90^\circ}^\circ}{\sigma_{0^\circ}^\circ} = 1.7 \times 10^5 W^{-7/4}$.

(c) Smooth surface²⁶ (the antenna beam has been approximated by a Gaussian distribution; this is satisfactory to about 20 decibels down on the main lobe, but neglects all side lobes):

$$\sigma^\circ = \frac{13.1 \times 10^3}{\Psi^2} \exp \left\{ - \left(\frac{10^4}{\Psi^2} \right) \left(\gamma - \frac{\pi}{2} \right)^2 \right\},$$

where Ψ is the half-power antenna beamwidth in degrees,

γ is the radar grazing angle in radians.

(d) Rough surface²⁶ (one which re-radiates power independent of the direction of the incident energy):

$$\sigma^\circ = 2 \sin \gamma.$$

PROPAGATION OVER ROUGH SURFACES (THEORY)

Theoretical Limitations

The general problem of reflection of electromagnetic waves from

²⁶ A. H. Schooley, "Some Limiting Cases of Radar Sea Clutter Noise," *Proc. I.R.E.*, Vol. 44, p. 1043, August, 1956.

rough surfaces is a difficult one. Many rigorous and semi-rigorous attempts have appeared in the literature,²⁷⁻³⁵ but these have yielded complicated expressions which are difficult to apply. Simple expressions are usually the result of many simplifying assumptions which may not adequately describe the true state of affairs. Thus the problem seems to be one of attempting to find a compromise between oversimplification which may lead to erroneous conclusions and excess rigour which may lead to no practical result.

The usual simplifications which are introduced into the theoretical treatments, and hence limit the results, can be listed as follows:

1. The dimensions of the scattering elements are considered either much larger than the wavelength of the incident field or very small compared to the wavelength of the incident field.
2. The far field only, is calculated.
3. Shadowing effects including diffraction are neglected.
4. Interaction including multiple reflections between obstructions is generally neglected.
5. The number of irregularities per unit length is not considered.
6. The treatments are often limited to particular models of roughness (e.g., normal distribution of scattering elements, regular sinusoidal undulations of the reflecting surface, etc.).

Reflection Coefficient of Rough Surfaces

In this section an attempt is made to summarize only those ap-

²⁷ S. O. Rice, "Reflection of Electromagnetic Waves from Slightly Rough Surfaces," *Communications on Pure and Applied Mathematics*, Vol. 4, p. 351, August, 1951.

²⁸ W. C. Hoffmann, "Scattering of Electromagnetic Waves from a Rough Surface," *Quart. Appl. Math.*, Vol. 13, p. 291, October, 1955.

²⁹ C. M. Chu and S. W. Churchill, "Multiple Scattering by Randomly Distributed Obstacle—Method of Solution," *Trans. I.R.E. PGAP*, p. 142, April, 1956.

³⁰ W. S. Ament, "Toward a Theory of Reflection by a Rough Surface," *Proc. I.R.E.*, Vol. 41, p. 142, January, 1953.

³¹ P. Beckmann, "A New Approach to the Problem of Reflection from a Rough Surface," *Acta Technica CSAV*, Vol. 2, p. 311, 1957.

³² H. Davies, "The Reflection of Electromagnetic Waves from a Rough Surface," *Proc. IEE*, Part IV, Vol. 101, p. 209, August, 1954.

³³ M. A. Isakovich, "The Scattering of Waves from a Statistically Rough Surface," *Zh. Eksp. Theor. Fiz. (USSR)*, Vol. 23, p. 305, 1952.

³⁴ Lord Rayleigh, *Theory of Sound*, Dover Publications, New York, New York, Vol. II, Sec. 272a, 1945.

³⁵ M. A. Biot, "Some New Aspects of the Reflection of Electromagnetic Waves on a Rough Surface," *Jour. Appl. Phys.*, Vol. 28, p. 1455, December, 1957; "Some New Aspects of the Reflection of Electromagnetic Waves on A Rough Surface," *Jour. Appl. Phys.*, Vol. 29, p. 998, (Letter to the Editor), June, 1958.

proaches to the problem of scattering by rough surfaces which are readily amenable to practical computation. These are subdivided into the two limiting cases where either the irregularities are large compared to wavelength or the irregularities are small compared to wavelength.

In Table V the reflection coefficients for the various models assumed for the profile of the rough surface are tabulated and, as can be seen, the results are primarily dependent upon the assumed profile. From these, some estimate of the percentage of the incident field which is reflected should be possible, the validity of the values depending on how closely the actual surface approaches any of these theoretical models. The treatment by Beckmann³¹ is attractive, primarily because

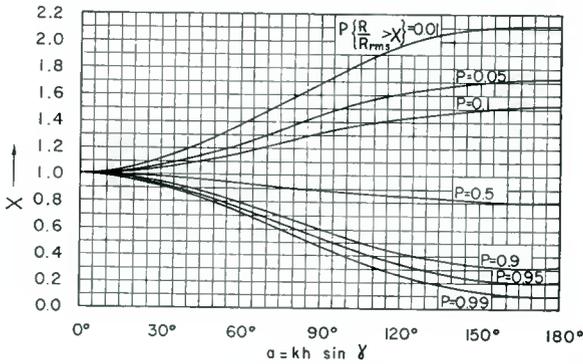


Fig. 4—Plot of the function $X(a,P) = \sqrt{\ln(1/P)} + (1 - \sqrt{\ln(1/P)}) \times (\sin^2 a)/a^2$, where P is the probability that the ratio $R/R_{rms} > X$ (after Beckmann³¹). Thus, as an example, the probability that $R/R_{rms} > 1.3$ when $a = 90^\circ$ is 10%. These curves can be used to estimate the probability distribution of the reflection coefficient.

of its final simplicity and the fact that an estimate of the spread in the measured values can be obtained as shown in Figure 4.

In Table VI some of the effects of small irregularities on the reflected field are summarized. Rayleigh³⁴ was perhaps the first to arrive at quantitative values for reflection from a slightly rough surface (his analysis was limited to normal incidence) and although under these conditions the reflection coefficient remains near unity, the analysis by Biot³⁵ shows that a slightly rough surface affects the phase of the reflected wave, particularly for vertical polarization, and this may be of some consequence for transmission paths where the direct and reflected waves are allowed to interfere. This suggests that the appropriate phase relations between the direct and reflected rays cannot be obtained from geometric optics alone.

Table V—Reflection Coefficient for Irregularities Large Compared to Wavelength

Author	—Ament ³⁰
Model of surface	—very gently rolling surface (horizontal polarization)
Relationship	$R = -\exp(-2k^2h^2 \sin^2\gamma),$ <p>where γ is the grazing angle and h is the average amplitude of height variation.</p>
Author	—Beckman ³¹
Model of surface	—random surface generated by a Markoff chain with a finite number of states and discrete moments of transition.
Assumptions and Definitions:	<ol style="list-style-type: none"> 1. Surface reflects in specular (principal) direction with Fresnel reflection coefficient (F); in other (subsidiary) directions have no reflection, i.e., diffraction and the side-lobes caused by it are neglected. 2. The reflected waves in a given direction are summed with respect to their phases to form a single resultant wave in that direction (far field only). 3. A "horizontal part" of the profile is defined as an interval (not much shorter than a wavelength) in which the random (continuous) curve of the profile stays within a range Δh so that $\Delta a = 2k \Delta h \sin \gamma < \epsilon'$ <p>where ϵ' is the value assigned to the Rayleigh criterion (varies according to author).</p> 4. For a two-dimensional surface, a "horizontal part" of the surface is a region (not smaller than a square of side λ) in which the surface stays within a range Δh as given above; p will then be the limit of the ratio of the sum of the areas of such regions to the total area of the region in which the surface is given.
Relationships:	<ol style="list-style-type: none"> 1. For a uniform distribution of horizontal elements (uniform phase distribution) over the height interval h in which the rough surface exists (i.e., horizontal elements are equally probable between the lowest and highest level of the surface which reflects energy)

$$R = \Gamma(\gamma) \sqrt{p + (1-p) \frac{\sin^2 \alpha}{a^2}},$$

where $\Gamma(\gamma)$ = Fresnel reflection coefficient,

λ = grazing angle,

p = probability of occurrence of horizontal elements (physically this means that for scattering in the principal direction it is the ratio of horizontal elements to the total number of elements when this number tends to infinity),

$\alpha = k h \sin \gamma$.

2. $\lim_{p \rightarrow 0} R = \Gamma(\gamma) \frac{\sin \alpha}{a}$ (case of a sawtooth profile).

3. Scattering in subsidiary directions (other than γ) is always Rayleigh distributed such that:

$$R = \Gamma(\theta) \frac{1}{2} \sqrt{\pi P(\theta)},$$

where $P(\theta)$ = probability of occurrence of elements with favorable slope for the direction (θ) under consideration,

θ = the grazing angle for the slope under consideration.

- Remarks:
1. Finds the amplitude of n unit vectors with random phases distributed from $-a$ to a is Rayleigh distributed whenever $a \gg \pi$ even when a is not an integral multiple of π .
 2. The reflection coefficient is more difficult to calculate if the distribution of horizontal elements is not uniform.
 3. The fact that vertically polarized waves obey the Fresnel reflection laws better than horizontally polarized waves is explained as due to the fact that $\Gamma(\gamma)$ changes slowly for horizontal polarization and rapidly for vertical polarization with the result that $\Gamma(\gamma)$ remains greater for horizontal polarization and when multiplied by the probability factors will result in a greater spread in the observed points.

Table V—(Continued)

4. A method is derived whereby the probability distribution of the reflection coefficient exceeding a specified value a given percentage of the time is approximated. This is shown in Figure 4. Thus for example if $R = 0.50$ under conditions for which $\alpha = 90^\circ$, then using Figure 4 it is found that approx. 90% of the points will lie above $R = 0.50 \times .60 = 0.30$ and 10% of the observed points will lie approx. above $R = 0.50 \times 1.30 = 0.65$.

Author

—Bullington⁴Model of surface
and relationship:

1. Sawtooth profile (uniform height but not necessarily uniform slope, with all values of phase between α and $-\alpha$ equally probable)

$$R = \frac{\sin \alpha}{\alpha}$$

2. Sinusoidal profile:

$$R = J_0(\alpha)$$

where J_0 is the Bessel function of the first kind and zero order.

3. Triangular profile:

$$R = \left(\frac{\sin (\alpha/2)}{(\alpha/2)} \right)^2$$

4. \sin^2 profile:

$$R = \cos \frac{\alpha}{2} J_0 \left(\frac{\alpha}{2} \right)$$

Author	—Davies ³²
Model of surface	—surface represented by a random function which yields a gaussian height distribution and correlation coefficient.
Assumptions:	<ol style="list-style-type: none"> 1. A single arbitrary parameter (m/σ) describes the surface. 2. Assume: $\sigma \gg \lambda$; $\sigma \gg m$; i.e., the surface is rough but not precipitous. 3. The surface currents of the rough surface are the same as for a corresponding smooth surface. 4. All shadowing effects neglected.
Relationship:	$R = \left(\frac{m}{2\sigma} \right) \sqrt{\sec \gamma} \exp \left\{ - \left(\frac{m}{2\sigma} \right)^2 \tan^2 \gamma \right\}.$
Remarks:	<ol style="list-style-type: none"> 1. Choosing (m/σ) values of 50-60 gives reasonable agreement with experiment on sea surfaces up to grazing angles of 40°. 2. Analysis does not hold for large grazing angles.
Author	—Isakovich ³³
Model of surface	—Statistically irregular surface, isotropic in all directions with a normal height distribution.
Assumptions:	<ol style="list-style-type: none"> 1. Surface described by parameter m/σ. 2. Assume no shadowing or multiple reflections, neglect all edge effects, consider an incident plane wave and calculate only far field.
Relationship:	$R = \left\{ \frac{1 - \cos \gamma \cos \psi \cos \phi + \sin \gamma \sin \psi}{(\sin \gamma + \sin \psi)^2} \right\} \frac{m}{\sigma} \exp \left\{ - \frac{m^2 \cos^2 \gamma + \cos^2 \psi - 2 \cos \gamma \cos \psi \cos \phi}{8\sigma^2} \right\}$ <p>where ψ = angle of observation, ϕ = azimuth angle.</p>
Remarks:	<ol style="list-style-type: none"> 1. This is a means of calculating the dispersion of waves due to a random surface. 2. The intensity of the scattering in a given direction depends only on the ratio m/σ in the approximation employed here.

Table VI—Phase Change on Reflection when Irregularities Are Small in Comparison with Wavelength

Author	—Biot ³⁵
Model of surface	—roughness given by a uniform density distribution of hemispherical bosses, the radii of which and distances between bosses are small relative to wavelength.
Assumptions:	<ol style="list-style-type: none"> 1. Incident plane wave. 2. Bosses are of radius r and there are N bosses per unit area define
Relationships:	$\rho = 2\pi N r^3,$ $b = \text{average distance between centers of bosses.}$ <ol style="list-style-type: none"> 1. For a perfectly conducting surface the phase change of the reflected wave is $e^{i2\theta_v, h}$, $\text{where } \tan \theta_v = \frac{1}{2} \frac{\cos 2\gamma}{\sin \gamma} \rho k \left(\frac{\cos 2\gamma}{\sin \gamma} \right) \text{ for vertical polarization,}$ $\tan \theta_h = \frac{1}{2} \rho k \sin \gamma \text{ for horizontal polarization.}$ 2. For an imperfectly conducting surface $\tan \theta_v = \frac{1}{2} \frac{k_1}{\pi^2 r^3} \frac{\sin \gamma}{\pi^2 r^3} \frac{\sin \gamma}{4b^3} \frac{\sin \gamma}{\pi^2 r^3} - (1 + \cos 2\gamma)$ $\text{where } k_1 = 1 + \frac{\rho}{b^3}; \quad k_2 = 1 + \frac{\rho}{4b^3}.$
Remarks:	<ol style="list-style-type: none"> 1. The exact solution including the electromagnetic interaction of the bosses is obtained. 2. For small roughness the reflection coefficient is essentially unity, however important phase changes are predicted. 3. For a perfectly conducting surface at grazing incidence ($\gamma = 0$) a phase reversal is obtained for vertical polarization. For $\gamma = \pi/4$, $\tan \theta_v = 0$, i.e., the effects of the roughness disappear. 4. The effect on horizontal polarization is generally small—it is a maximum at normal incidence and disappears at grazing. 5. The phase angle is determined by the size and distribution of the bosses and the angle of incidence. 6. For an imperfectly conducting surface, the angle at which the effect of the roughness disappears is shifted from $\pi/4$ to smaller grazing angles. As $(r/b)^3 \rightarrow 0$ this angle $\rightarrow \pi/4$.

CONCLUSIONS

An attempt has been made to give a comprehensive picture of the problems of the effect of rough surfaces on the propagation of short radio waves and the "state of the art" in efforts to recognize, understand and predict these effects. Some of the aspects requiring further study can be listed as

1. Establishment of a roughness criterion which relates the reflected field to the reflection coefficient.
2. The study of the region of roughness where specular reflection transcends into diffuse reflection.
3. The experimental determination of whether the arguments applied to the Fresnel zones in optics can validly be extended to the long, narrow, nonuniformly illuminated Fresnel reflecting zones which occur in propagation situations.
4. Further accumulation of experimental data under various conditions which could at least be utilized to establish empirical modifications, on a statistical basis, of idealized theories.
5. There still exists considerable need for further development of theoretical analyses which approach more closely the actual surface properties.

BIBLIOGRAPHY ON MICROWAVE PROPAGATION OVER ROUGH SURFACES

Propagation over Irregular Terrain (Experiment)

G. H. Brown, J. Epstein and D. W. Peterson, "Comparative Propagation Measurements; Television Transmitters at 67.25, 288, 510 and 910 Megacycles," *RCA Review*, Vol. IX, pp. 171-201, June, 1948.

J. Houtsmuller, "Propagation Properties of Meter Waves at Not Very Great Distances from the Transmitter," *Tijdschr. Ned. Radiogest.*, Vol. 21, pp. 103-144, May, 1956.

R. S. Kirby, J. M. Taff and H. S. Moore, "Measurement of the Effect of Irregular Terrain on VHF and UHF Directive Antenna Pattern," *Trans. I.R.E. PGAP*, pp. 167-178, August, 1952.

R. S. Kirby and F. M. Capps, "Correlation in VHF Propagation over Irregular Terrain," *Trans. I.R.E. PGAP*, pp. 77-85, January, 1956.

J. S. McPetrie, "The Reflection Coefficient of the Earth's Surface for Radio Waves," *Jour. I.E.E.*, Vol. 82, p. 214, February, 1938.

J. S. McPetrie and J. A. Saxton, "An Experimental Investigation of the Propagation of Radiation Having Wavelengths of 2 and 3 Meters," *Jour. I.E.E.*, Vol. 87, pp. 146-153, August, 1940.

R. K. Potter and H. T. Friis, "Some Effects of Topography on Short-Wave Reception," *Proc. I.R.E.*, Vol. 20, pp. 699-721, April, 1932.

W. Von Hagen, A. N. MacDiarmid and L. V. Goldenberg, "Path-Loss Testing of the Trans-Canada TD-2 Route," presented at the *Canadian IRE Convention*, Toronto, Canada, October 16, 1957.

Also footnote References (1), (2), (4), (5), (6), (7), (8), (9), (10), (11), (12), (13), (14), (15), (16), (17), and (18).

Propagation over Sea Surfaces (Experiment)

G. W. G. Court, "Determination of the Reflection Coefficient of the Sea," *Proc. I.E.E.*, Vol. 102, Part B, pp. 827-830, November, 1955.

H. Goldstein, "The Frequency Dependence of Radar Echoes from the Surface of the Sea," *Phys. Rev.*, Vol. 69, p. 695, June 15, 1946.

C. R. Grant and B. S. Yaplee, "Back Scattering from Water and Land at Centimeter and Millimeter Wavelength," *Proc. I.R.E.*, Vol. 45, pp. 976-982, July, 1957.

P. Gudmandsen and B. F. Larsen, "Statistical Data for Microwave Propagation Measurements on Two Overseas Paths in Denmark," *Trans. I.R.E. PGAP*, p. 255, July, 1957.

J. C. Johnson, "The Backscatter Coefficient of a Spherical Homogeneous Mixture of Ice and Air at Wavelengths between 1 and 10 cm," *Jour. Meteorology*, Vol. 12, pp. 188-189, April, 1955.

D. G. Kiely, "Measurements of the Reflection Coefficient of Water at $\lambda = 8.7$ mm," *Proc. Phys. Soc.*, Vol. 63, pp. 46-49, January, 1950.

D. G. Kiely, "Some Problems Posed by Wave Propagation at 8 mm. and 3 cm. Wavelengths Over the Sea and Through Rain," *Ann. Telecommun.*, Vol. 11, pp. 233-244, November, 1956 and Vol. 11, pp. 267-279, December, 1956.

F. C. Macdonald, "The Correlation of Radar Sea Clutter on Horizontal and Vertical Polarization with Wave Height and Slope," *IRE Convention Record*, Part 1, pp. 29-32, March, 1956.

R. C. Staley, "Some Problems in Modeling Water Waves Relating to the Back-Scattering of Microwave Radio Energy," Texas University Report #98, February, 1958.

Also, footnote References (3), (19), (20), (21), (22), (23), (24), (25), and (26).

Propagation over Rough Surfaces (Theory)

Ia. L. Al'pert, V. L. Ginzburg and E. L. Feinberg, *Radiowave Propagation*, Gostekhizdat, Moscow, 1953.

W. S. Ament, "Forward and Back-Scattering From Certain Rough Surfaces," *Trans. I.R.E. PGAP*, pp. 369-373, July, 1956.

C. I. Beard, I. Katz and L. M. Spetner, "Phenomenological Vector Model of Microwave Reflections from the Ocean," *Trans. I.R.E. PGAP*, pp. 162-167, April, 1956.

M. A. Biot, "Reflections on a Rough Surface from an Acoustic Point Source," *Jour. Acous. Soc. Amer.*, Vol. 29, pp. 1193-1200, November, 1957.

L. V. Blake, "Reflections of Radio Waves From a Rough Sea," *Proc. I.R.E.*, Vol. 38, pp. 301-304, March, 1950.

H. G. Booker, J. A. Ratcliffe and D. H. Shinn, "Diffraction from an Irregular Screen with Application to Ionospheric Problems," *Philosophical Trans. of Royal Society of London, Series A*, Vol. 292, p. 579, 1950.

M. Bouix, "Study of the e.m. Field in the Vicinity of a Reflecting Surface," *Comptes Rendus Acad. Sci. (Paris)*, Vol. 240, pp. 1763-1765, May 2, 1955.

L. M. Brekhovskikh, "The Diffraction of Waves by a Rough Surface," Parts I and II, *Zhurnal Eksperimental' noi Theorticheskoi Fiziki*, (USSR), Vol. 23, pp. 275-282 and 289-304, 1952.

K. Bullington, "Reflection Coefficients of Irregular Terrain," *Proc. I.R.E.*, Vol. 42, pp. 1258-1262, August, 1954.

K. Bullington, "Radio Propagation Fundamentals," *Bell Sys. Tech. Jour.*, Vol. 36, pp. 593-626, May, 1957.

J. F. Carlson and A. E. Heins, "The Reflection of an Electromagnetic Plane Wave by an Infinite Set of Plates," *Quarterly of Applied Mathematics*, Vol. 4, pp. 313-329, January, 1946 and Vol. 5, pp. 82-88, April, 1947.

E. Feinberg, "On the Propagation of Radio Waves Along an Imperfect Surface," *Journal of Physics (USSR)*, Vol. 8, pp. 317-330, No. 6, 1944 and Vol. 9, pp. 1-6, No. 1, 1945 and Vol. 10, pp. 410-418, 1946.

J. Feinstein, "Some Stochastic Problems in Wave Propagation," *Trans. I.R.E. PGAP*, pp. 23-30, January, 1954 and pp. 63-70, April, 1954.

C. B. Feldman, "Optical Behavior of the Ground for Short Radio Waves," *Proc. I.R.E.*, Vol. 21, pp. 764-801, June, 1933.

D. E. Kerr, *The Propagation of Short Radio Waves*, Vol. 13, MIT Radiation Lab Series, McGraw-Hill Book Co., Inc., New York, New York, 1951.

K. Furutsu, "Field Strength in the Vicinity of the Line of Sight in Diffraction by a Spherical Surface," *Jour. Radio Res. Labs. (Japan)*, Vol. 3, pp. 55-76, January, 1956.

K. Furutsu, "Wave Propagation Over an Irregular Terrain," Part I, *Jour. Radio Res. Labs (Japan)*, Vol. 4, pp. 135-153, April, 1957.

J. Grosskopf, "Propagation of a Surface Wave Over Curved and Stratified Ground," *Hochfrequenztech. Electro.*, Vol. 58, pp. 163-171,

1941; "Propagation of e.m. Waves Over Inhomogeneous Earth," *Hochfrequenztech. Electro.*, Vol. 62, pp. 103-110, October, 1943.

H. S. Heaps, "Reflection of Plane Waves of Sound from a Sinusoidal Surface," *Jour. Appl. Phys.*, Vol. 28, pp. 815-818, July, 1957.

W. C. Hoffman, "Backscattering from Perfectly Conducting Doubly-Trochoidal and Doubly-Sinusoidal Surfaces," *Trans. I.R.E. PGAP*, pp. 96-100, July, 1955.

G. A. Hufford, "An Integral Equation Approach to the Problem of Propagation Over an Irregular Surface," *Quarterly of Applied Mathematics*, Vol. 9, p. 391, January, 1952.

M. A. Isakovich, "Scattering and Radiation of Waves by Statistically Inhomogeneous and Statistically Oscillating Surfaces," *Akusticheskii Zhurnal* (USSR), Vol. 2, pp. 146-149, April-June, 1956.

M. Katzin, "Back-Scattering From the Sea Surface," *IRE Convention Record*, Part I, pp. 72-77, March, 1955.

M. Katzin, "Recent Developments in the Theory of Sea Clutter," *IRE Convention Record*, Part I, p. 19, March, 1956.

M. Katzin, "On the Mechanisms of Radar Sea Clutter," *Proc. I.R.E.*, Vol. 45, pp. 44-54, January, 1957.

A. W. Leporski, "Scattering of Sound Waves by Sinusoidal and Sawtooth Surfaces," *Akusticheskii Zhurnal* (USSR), Vol. 2, pp. 177-181, April-June, 1956.

Y. P. Lysanov, "On the Scattering of Electromagnetic Waves from a Rough Surface," *Doklady Akad. Nauk* (USSR), Vol. 87, pp. 719-722, 1952.

Y. P. Lysanov, "An Approximate Solution of the Problem of Scattering of Sound Waves at An Irregular Surface," *Akusticheskii Zhurnal* (USSR), Vol. 2, pp. 182-187, April-June, 1956.

W. C. Meecham, "Variation Method for the Calculation of the Distribution of Energy Reflected from a Periodic Surface," *Jour. Appl. Phys.*, Vol. 27, pp. 361-367, April, 1956.

J. W. Miles, "On Nonspecular Reflection at a Rough Surface," *Jour. Acous. Soc. Amer.*, Vol. 26, pp. 191-199, March, 1954.

R. K. Moore and C. S. Williams, "Radar Terrain Return at Near Vertical Incidence," *Proc. I.R.E.*, Vol. 45, pp. 228-238, February, 1957.

K. A. Norton and A. C. Omberg, "Maximum Range of a Radar Set," *Proc. I.R.E.*, Vol. 35, pp. 4-24, January, 1947.

K. Norton, "Transmission Loss of Space Waves Propagated Over Irregular Terrain," *Trans. I.R.E. PGAP*, pp. 152-166, August, 1952.

K. Norton, L. E. Vogler, W. V. Mansfield and P. J. Short, "The Probability Distribution of the Amplitude of a Constant Vector Plus a Rayleigh-Distributed Vector," *Proc. I.R.E.*, Vol. 43, pp. 1354-1361, October, 1955.

K. A. Norton, P. L. Rice and L. E. Vogler, "The Use of Angular Distance in Estimating Transmission Loss and Fading Range for Propagation Through a Turbulent Atmosphere Over Irregular Terrain," *Proc. I.R.E.*, Vol. 43, pp. 1488-1526, October, 1955.

G. P. Ohmann, "Universal Curves for the Vertical Polarization Reflection Coefficient," *Trans. I.R.E. PGAP*, pp. 140-142, January, 1957.

L. S. Ornstein and A. van der Berg, "Reflectivity of Corrugated Surfaces," *Physica*, Vol. 4, p. 1181, 1937.

Lord Rayleigh, "Light Dispersed From Fine Lines or Transmitted by Narrow Slits," *Philosophical Magazine*, Vol. 14, p. 350, September, 1907.

L. M. Spetner, "A Statistical Model for Forward Scattering of Waves off a Rough Surface," *Trans. I.R.E. PGAP*, pp. 88-94, January, 1958.

C. T. Tai, "Reflection and Refraction of a Plane Electromagnetic Wave at a Periodical Surface," *Harvard Technical Report No. 28*. Cruft Laboratory, January, 1948.

V. Twersky, "On the Non-Specular Reflection of Plane Waves of Sound," *Jour. Acous. Soc. Amer.*, Vol. 22, pp. 539-546, September, 1950.

V. Twersky, "On the Non-Specular Reflection of Electromagnetic Waves," *Jour. Appl. Phys.*, Vol. 22, pp. 825-835, June, 1951.

V. Twersky, "Multiple Scattering of Radiation by an Arbitrary Planar Configuration of Parallel Cylinders and by Two Parallel Cylinders," *Jour. Appl. Phys.*, Vol. 23, pp. 407-414, April, 1952.

V. Twersky, "Certain Transmission and Reflection Theorems." *Jour. Appl. Phys.*, Vol. 25, pp. 859-862, July, 1954.

V. Twersky, "Scattering Theorems for Bounded Periodic Structures," *Jour. Appl. Phys.*, Vol. 27, pp. 1118-1122, October, 1956.

V. Twersky, "On Scattering and Reflection of Electromagnetic Waves on Rough Surfaces," *Trans. I.R.E. PGAP*, pp. 81-89, January, 1957.

H. C. Van de Hulst, "On the Attenuation of Plane Waves by Obstacles of Arbitrary Size and Form," *Physica*, Vol. 15, pp. 740-746, September, 1949.

J. R. Wait, "Scattering of e.m. Waves from a 'lossy' Strip on a Conducting Plane," *Canadian Jour. Physics*, Vol. 33, pp. 383-390, July, 1955.

J. R. Wait, "Radiation from an Electric Dipole in the Presence of a Corrugated Cylinder," *Applied Scientific Research*, Vol. B6, pp. 117-123, No. 1-2, 1956.

Also, footnote References (27), (28), (29), (30), (31), (32), (33), (34), and (35).

THE EFFECT OF SEVERAL VARIABLES ON PHOSPHOR-DOT SIZE IN COLOR PICTURE TUBES*

BY

N. R. GOLDSTEIN

RCA Electron Tube Division
Lancaster, Pa.

Summary—It was considered desirable to determine the effect on the phosphor-dot size of essentially all of the important variables in the slurry screening process. An empirical equation is established which relates these variables for a particular blue phosphor screen. The equation is most reliable for screens wherein the dot diameter is approximately 17 mils.

INTRODUCTION

PHOSPHOR-DOT screens for color-picture tubes are now produced by the "slurry" process. In this process a solution or slurry containing a phosphor, a light-sensitizing agent, and a binder, is applied to the tube faceplate. The slurried surface is dried, and then exposed photographically through the associated aperture mask in a device known as a "lighthouse" (see Figure 1). It is then developed by means of a spray of water which washes away the unexposed areas. This process is repeated for each phosphor.

Soon after the adoption of the slurry process, it became evident that the sizes of the dots produced were affected by several variables: phosphor composition, amount of sensitizing agent used, and screen weight; mask-aperture diameter, mask-to-screen spacing, collimator-to-mask spacing, and collimator-tip diameter; exposure-light intensity, and length of exposure; pH of the developing water, pressure of the developing spray, and length of development. It was also found that as larger mask apertures, greater mask-to-screen spacings, and smaller gun-to-mask spacings were demanded it became increasingly difficult to meet the criteria for good screens. Additional difficulties were encountered with the introduction of post-acceleration tubes which proved to be even more sensitive than the shadow-mask type to variations in screening parameters. An investigation was undertaken, therefore, to determine the effect on dot size of each of the several variables involved.

The effect of each variable was studied under carefully controlled conditions, with all others held constant. Slurries were very carefully

* Manuscript received February 25, 1959.

compounded, and applied in a manually operated unit having a precisely controlled cycle. Screen transmission was determined, using a standard Weston photocell, and was expressed in percentage of the transmission through a clear faceplate. Masks were inserted as quickly as possible after application of slurry.

Lighthouse parameters were carefully preset to the desired values. Exposure-light intensity was adjusted to the desired value immediately before exposure, and checked after exposure.

Developing water was adjusted to the desired pH by the addition of oxalic acid or ammonium hydroxide, allowed to stand overnight, and the pH checked again immediately before use. The water was then placed in a two-gallon, controllable-pressure tank which had been repeatedly flushed, first with neutral water and then with water having the desired pH .

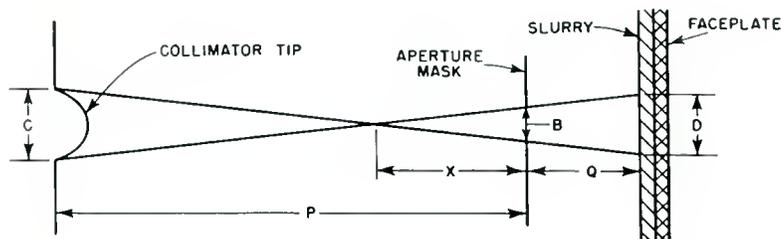


Fig. 1—Geometry of the "lighthouse" exposure system.

A device provided with removable wedges permitted as many as six different exposures to be made on one screen. For development, the position of the spray nozzle, the pressure, and the time of development were carefully controlled. The screens were permitted to soak for exactly three minutes and then very lightly rinsed and dried. The phosphor-dot sizes were then measured.

The first variable studied was the pH of the developing water which an earlier investigation by D. J. Donahue* showed to be an important factor in phosphor-dot size.

Figure 2 shows part of a family of curves of dot diameter D as a function of pH , with exposure time T as a parameter. The curves have the form of a catenary, the equation of which is

$$x = \frac{a}{2} (e^{y/a} + e^{-y/a}),$$

* RCA Electron Tube Division, Lancaster, Pa.

where a is the displacement from the y axis. Since this displacement is 3 scale units,

$$pH = \frac{3}{2} (e^{y/3} + e^{-y/3}).$$

Because $\cosh u = \frac{1}{2} (e^u + e^{-u})$,

$$pH = 3 \cosh \frac{y}{3},$$

and $y = 3 \cosh^{-1} \frac{pH}{3}$.

When the appropriate scale factor is introduced, it is found that $D = B + y/2$. Therefore,

$$D = 15.5 + 1.5 \cosh^{-1} \frac{pH}{3}$$

for the following conditions:

$$T = \text{exposure time in minutes} = 8,$$

$$Q = \text{mask-to-screen spacing in mils} = 560,$$

$$B = \text{mask-aperture diameter in mils} = 15.5,$$

$$C = \text{collimator-tip diameter in mils} = 60,$$

$$I = \text{exposure-light intensity in relative units} = 75,$$

$$T_D = \text{development time in minutes} = 1.$$

EFFECTS OF GEOMETRICAL PARAMETERS

In the investigation of the effects of mask-to-screen spacing, Q , special care was taken to minimize errors which might arise because of faceplate and mask irregularities. Measurements were made on several faceplate and mask assemblies, and duplicate measurements were made using two different masks for each faceplate at each value of Q .

A plot of dot diameter, D , as a function of mask-to-screen spacing, Q , is shown in Figure 3. The plot is straight line having a slope of 4.5, with intercepts at $Q = 500$ and $D = 17.2$. The equation for this line is

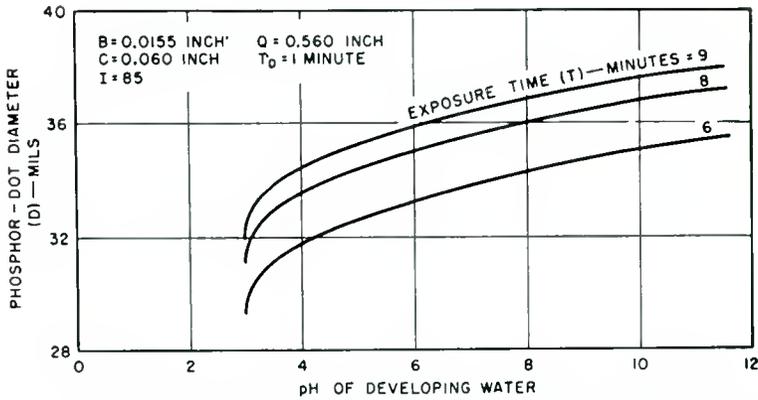


Fig. 2—Phosphor-dot diameter as a function of pH and exposure time, for constant lighthouse geometry, light intensity, and development time.

$$\frac{D - D_1}{Q - Q_1} = m = 4.5$$

$$\frac{D - 17.2}{Q - 0.5} = 4.5$$

$$D = 14.95 + 4.5Q \text{ mils,}$$

with pH = 8.5 and other parameters as previously given.

In the investigation of the effects of collimator-tip diameter, C, it was necessary to take into consideration the variation in light output with C. The effects of this variation were compensated both by adjustment of light intensity and by the use of a series of neutral-

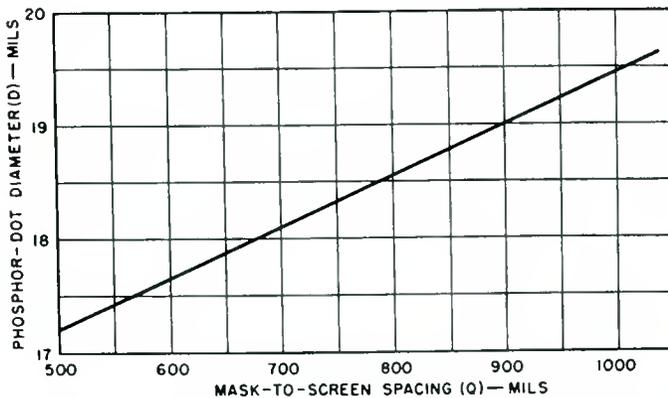


Fig. 3—Phosphor-dot diameter as a function of mask-to-screen spacing, Q.

density filters made of $\frac{1}{4}$ -inch plate glass coated with rhodium.

A plot of dot diameter as a function of collimator-tip diameter is shown in Figure 4. This plot also is a straight line, and has the equation

$$D = \frac{C}{120} + 16.83 \text{ mils,}$$

with parameters as above.

A series of tests were conducted to determine the relation between mask-aperture diameter, B , and dot size. Although the results of these tests are not shown graphically in this paper, it was found that for a given combination of the other parameters $D = B + \Gamma$, Γ being dependent upon the values of the other parameters. It is not surprising that

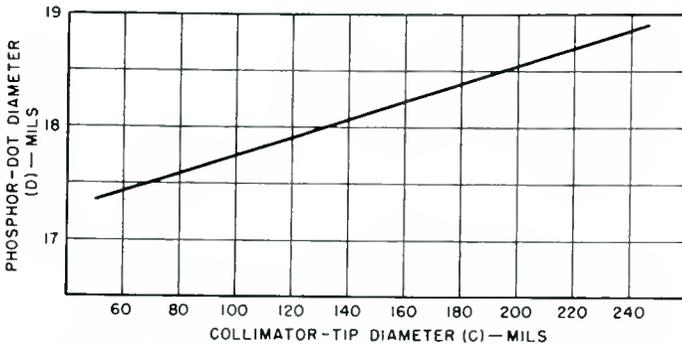


Fig. 4—Phosphor-dot diameter as a function of collimator-tip diameter, C .

the effects of Q , C , and B are linear, if one considers the geometry of the exposure system shown in Figure 1.

By the laws of plane geometry

$$\frac{C}{p - X} = \frac{B}{X} = \frac{D_{\max}}{Q + X},$$

from which

$$D_{\max} = \frac{Q}{p} (B + C) + B.$$

For any tube geometry,

$$\frac{Q}{p} = K_1,$$

where p = collimator aperture to mask aperture spacing.

$$\begin{aligned} D_{\max} &= K_1 (B + C) + B \\ &= B (K_1 + 1) + K_1 C \\ &= K_2 B + K_1 C. \end{aligned}$$

The preceding calculations are a development of the interrelationship of the tube parameters for a theoretical maximum area covered by the light from the collimator tip and, therefore, give a maximum value for D . Because they are based purely on the geometry of the system, they do not take into account any fringe effects. A similar procedure can be used to calculate the minimum area.

$$D_{\min} = K_2 B - K_1 C.$$

Table I shows typical values for D_{\max} and D_{\min} as calculated from these expressions

Table I

Collimator Tip Diameter (mils)	Exp. Phosphor- Dot Diameter (mils)	Light Spot Diameter (mils)		
		Theor. Max.	Theor. Min.	Theor. Ave.
60	160	165	159	162
80	162	166	158	162
100	163	167	157	162
160	168	170	154	162
220	173	172	152	162

$$Q = 535 \quad p = 11.2 \text{ inches} \quad B = 155$$

This derivation neglects not only fringe effects, but also any effects due to scattering of the light by the phosphor itself.¹

EFFECTS OF EXPOSURE PARAMETERS

A slurried screen, like any photographic emulsion requiring development after exposure, can be underexposed or overexposed, underdeveloped or overdeveloped. The size after development of an exposed area is proportional to the product of exposure-light intensity, I , and

¹ S. H. Kaplan, "Control of Fluorescent Screen Dot Size for Color TV," *Jour. S.M.P.T.E.*, Vol. 65, p. 407, August, 1956.

exposure time, T . To obtain dots of a given size for given slurry, lighthouse, and development parameters, the product of I and T must be constant.

A plot of $IT = \text{constant}$ has the form of an hyperbola. In the case of a conventional shadow-mask tube the critical region of this curve can be approximated by a straight line. It was possible, therefore, to predict the exposure times required to produce dots of a given size for various values of I . In the case of post-acceleration tube, the critical region of the IT curve is not a straight line, and must, therefore, be plotted from experimental data.

In order for a plot of $D = f(I, T)$ to have the expected hyperbolic form, D must be a linear function of both I and T . Figure 5 shows

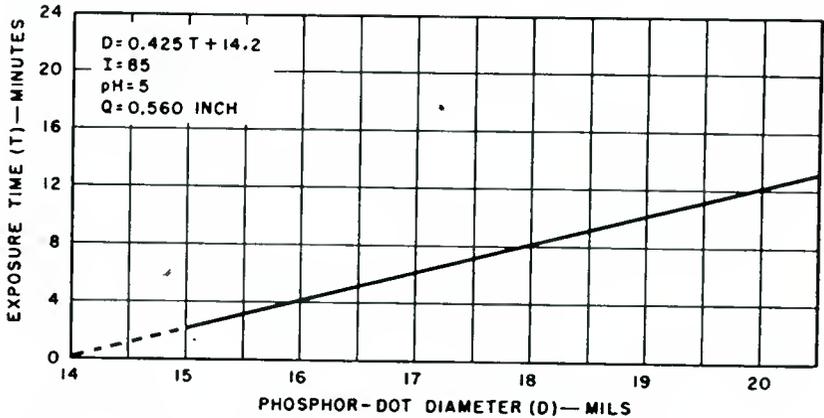


Fig. 5—Phosphor-dot diameter as a function of exposure time for constant lighthouse geometry, light intensity, and developing water pH .

that D is a linear function of T , at least over the range of maximum present interest ($D = 14$ to 20 mils). If we assume that, over this range of values, D is also a linear function of I , we can empirically plot $D = f(I, T)$ for any value of D between 14 and 20 mils. Figure 6 shows the calculated curve for dots 17 mils in diameter, and its relation to experimentally determined values for dots of this size.

The value of the product IT was checked as follows:

Table II, correlating the time and light-intensity units of Figure 6, was prepared. The values in Table II are related to minutes and relative intensity in the following manner:

$$T_{\text{minutes}} = T_{\text{arbitrary units}} + 4,$$

$$I_{\text{relative}} = 10 I_{\text{arbitrary units}} + 35.$$

Table II

Time*	Intensity*
1.03	6
1.40	5
1.90	4
2.54	3
3.28	2
4.10	1
5.30	0

* Units are the scale units of Figure 6.

If it is assumed that $IT = C$, then $K_3 = 4.12$, and $K_4 = 7.23$.

If the factors K_3 and K_4 (which, in effect, relocate the axes of Figure 6) are added to the scale-unit values of Table II, we obtain the values shown in Table III, where T_c represents the corrected value of time.

$$(T + K_3)(I + K_4) = 67.9.$$

Table III

T	I	IT	Dev.	% Dev.	T_c
5.15	13.23	68.2	+ 0.3	0.44	5.13
5.52	12.23	67.6	- 0.3	0.44	5.55
6.02	11.23	67.7	- 0.2	0.29	6.04
6.66	10.23	68.2	+ 0.3	0.44	6.63
7.40	9.23	68.3	+ 0.4	0.59	7.36
8.22	8.23	67.6	- 0.3	0.44	8.24
9.42	7.23	68.1	+ 0.2	0.29	9.40
		Avg. 67.9	Avg. 0.28		

Converting from scale units back to minutes and relative intensity, we find that for a 17-mil dot

$$(T + K_3) \text{ units} = (T + 0.12) \text{ minutes},$$

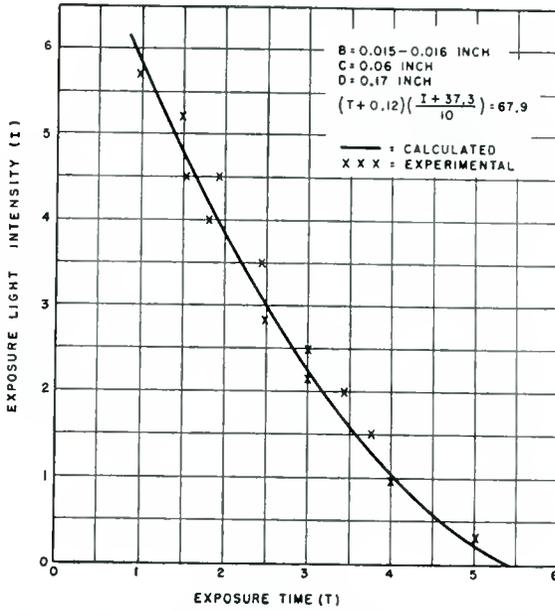


Fig. 6—Exposure-light intensity versus exposure time for phosphor dots 17 mils in diameter.

and $(I + K_4)$ units = $\frac{I - 35}{10} + 7.23$ relative intensity

$$(T + 0.12) \left(\frac{I - 35}{10} + 7.23 \right) = 67.9.$$

EFFECTS OF DEVELOPMENT PARAMETERS

It has long been known that excessive developing-water pressure or unduly long development time will clean all the phosphor off the screen.

The effects of developing-water pressure were studied at pressures ranging from 10 to 80 psi. Although there was no conclusive evidence in favor of either high or low pressure, there were indications that the lower pressures gave smaller dots. At very low pressures the developing water apparently is able to wet and penetrate the phosphor, whereas at higher pressures it tends to bounce off. When the pressure is sufficiently high, the phosphor is removed by the impact of the particles of water. However, there was no sign of dot removal at the highest pressure tested (80 psi).

The results of the development-time tests are somewhat more con-

clusive. Figure 7 is a plot of dot size as a function of development time, and shows a uniform decrease of one mil in dot size per minute of development.

EFFECTS OF SLURRY PARAMETERS

Several theories have been proposed as to the precise effect produced on polyvinyl alcohol by the addition of dichromate. For the purposes of this study it may be assumed that the presence of chromate ions in some manner makes the polyvinyl alcohol sensitive to light in the ultraviolet and blue regions.

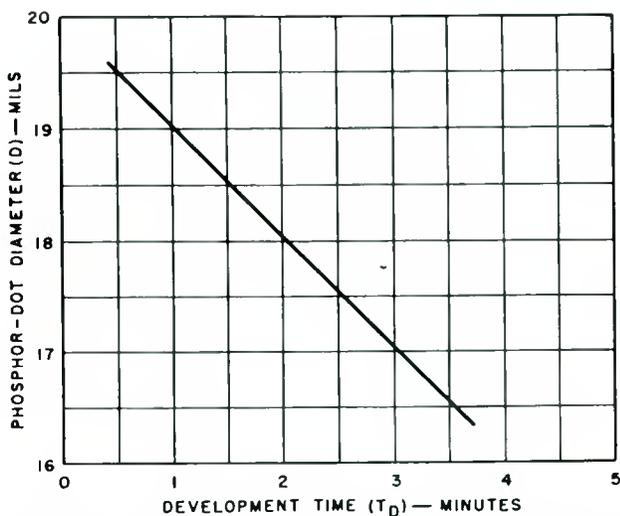


Fig. 7—Phosphor-dot diameter as a function of development time.

Tests show that for a given set of exposure conditions an increase in dichromate concentration results in larger phosphor dots or reduces the exposure time required to produce dots of a given size. This relation holds for dichromate concentrations up to approximately four times the normal value. Any further increase in concentration does not appear to have any effect. If the theory that photosensitization results from penetration of the polyvinyl alcohol lattice by chromate ions is correct, the lattice is saturated at a concentration of four times normal. Table IV shows the effect of dichromate concentration on dot size for constant exposure time.

Table IV

Relative Concentration	Dot Diameter (mils)
1	165
2	180
3	185
4	195
5	195

Because of the diffusing action of the phosphor in the slurry, screen weight greatly affects the dot size. For identical exposure parameters a heavy screen gives a smaller dot than a light screen. However, a light screen can be made to provide smaller dots than the heavy screen, under conditions which would result in no dot on the heavy screen.

This screening contradiction can be explained in terms of the curve for the distribution of light intensity across an aperture. Assuming a given exposure time, in the case of a heavy screen only the center of the light beam penetrates the screen with sufficient intensity to convert the polyvinyl alcohol to the insoluble state; the result, therefore, is either no dot or a small dot. In the case of a thin screen the entire beam penetrates the screen, with the result that the dot has a chance to grow. There is seldom an adherence problem with a thin screen. If the exposure time is shortened, less dot growth takes place, and since there is less diffusion, the phosphor dot is small.

In tests to determine the precise effect of screen weight on phosphor dot size under given conditions, it was found that the very phenomenon which causes growth was being overlooked, namely, the diffusing power of the slurry. On investigation, it was found that two caps which had apparently identical screen weights had totally different diffusing qualities, i.e., a light could be clearly seen through one but appeared very blurred when viewed through the other. Although both screens had the same weight as measured by the light-transmission method, actual weighing of the screen materials showed that the one having the greater diffusion was about 10 per cent heavier than the other. Since analysis showed that the phosphors of both screens had the same particle-size distribution, it was assumed that one was more dense than the other. This phenomenon is now the subject of a separate investigation. Experience has shown, however, that a light-weight screen will give a smaller phosphor dot than a heavy screen.

EMPIRICAL EQUATION FOR A PARTICULAR BLUE PHOSPHOR

It has been shown that phosphor-dot size is a function of several

variables. The geometrical variables such as mask-to-screen spacing, collimator-tip diameter, and mask aperture diameter were shown to have linear effects on dot size, and the theoretical hyperbolic relation of the product of exposure time and intensity was confirmed experimentally.

It has also been shown that the effect of development time is linear over the range studied. The effect of the *pH* of the developing water on phosphor-dot size is an exponential function, although no theory is advanced to support this conclusion.

The effects of the pressure of the developing water, concentration of ammonium dichromate, and screen weight were not determined with sufficient accuracy to permit quantitative conclusions to be drawn.

Consideration of the factors involved in the production of a phosphor-dot screen suggests that the size of an individual dot is dependent on at least four effects;

$$D = f(P), g(G), h(E), j(F),$$

where D is dot size,

P is the phosphor,

G is the geometry,

E is the exposure conditions,

F is the developing conditions.

From the data the following empirical equations can be written:

$$D = (T + 0.12) \left(\frac{I + 37.3}{40} \right) \quad (1)$$

$$D = 14.95 + 4.5Q, \quad (2)$$

$$D = 15.5 + 1.5 \cosh^{-1} \frac{pH}{3}. \quad (3)$$

It was also seen from the data that

$$D = B + \Gamma, \quad (4)$$

where Γ represents increase in dot size.

$$D = \frac{C}{120} + 16.3 \quad (5)$$

$$D = 19.0 - T_D. \quad (6)$$

Re-arranging by functional groups, and combining

$$D = B + 4.5Q + \frac{C}{120} + \frac{(T + 0.12)(I + 37.3)}{271} - \left(T_D - 1.5 \cosh^{-1} \frac{pH}{3} \right) - 4.36. \quad (7)$$

In the following region of values of the variables:

$$B = 15.5 \text{ mils}, \quad Q = 535 \text{ mils}, \quad C = 60 \text{ mils}, \quad I = 75,$$

$$T = 5.75 \text{ minutes}, \quad T_D = 1 \text{ minute}, \quad \text{and } pH = 6$$

$$g(G) = B + 4.5Q + \frac{C}{120},$$

$$h(E) = \frac{(T + 0.12)(I + 32.3)}{271},$$

$$j(F) = T_D - 1.5 \cosh^{-1} \frac{pH}{3}.$$

It should be noted that this equation applies only for a particular blue phosphor. Screen weights used are constant, and the effect of this as a variable is the subject of a separate investigation.

REDUCTION OF CO-CHANNEL INTERFERENCE BY PRECISE FREQUENCY CONTROL OF TELEVISION PICTURE CARRIERS. PART II*

BY

W. L. BEHREND

RCA Laboratories,
Princeton, N. J.

Summary—The visibility of co-channel television interference has maxima and minima at carrier offset frequencies which are multiples of frame frequency. Earlier subjective tests¹ determined the reduction in visibility of co-channel interference which might be achieved by precise carrier frequency control for the condition of one station, at a 10- or 20-kilocycle offset, interfering with the desired station.

Additional subjective tests have been made for the following conditions: two offset stations simultaneously interfering with a desired station; an on-frequency station interfering with a desired on-frequency station; and a series of tests with moving pictures as the desired picture and as the source of visual interference from a co-channel station offset 10 or 20 kilocycles.

INTRODUCTION

REGULATIONS SET FORTH by the Federal Communications Commission (FCC) require that co-channel television stations separated by less than a specified distance must operate with their picture carriers offset so that the frequency difference between carriers is 10 or 20 kilocycles. This serves to minimize the "Venetian blind" effect caused by co-channel interference. Present regulations specify a carrier frequency accuracy of $\pm 1,000$ cycles. If the carrier frequency is much more precisely controlled (± 2 cycles) and the frequency offset is an even multiple of the frame frequency, there is a further reduction in the visibility of co-channel interference. A previously published article¹ describes subjective tests that had been conducted in the laboratory, the oscillators used to control the picture carriers of two television stations, the method by which the offset was set and measured, and the field tests made to confirm that precise control provided a reduction in co-channel interference.

In the present paper, further subjective viewing tests are reported. In one series of tests, moving pictures were used both as the desired

* Manuscript received April 3, 1959.

¹ W. L. Behrend, "Reduction of Co-channel Television Interference by Precise Frequency Control of Television Picture Carriers," *RCA Review*, Vol. XVII, p. 443, December, 1956.

picture and as the source of visual interference from a single co-channel interfering station. In another series of tests, two co-channel interfering stations are simulated.

SUMMARY OF PREVIOUSLY PUBLISHED RESULTS

Two different types of subjective tests are described in Reference (1). The results of these tests are summarized for purposes of comparison and for completeness. One test measured the maximum tolerable interference level as a function of the offset frequency. The other test measured the improvement due to precise frequency control by the "method of equality."

Tolerable-Interference Tests

The observer indicated the maximum amount of interference he

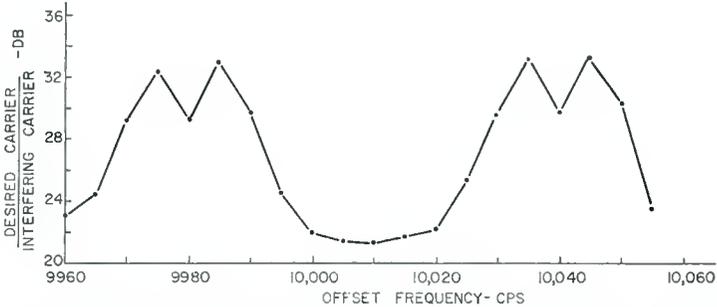


Fig. 1—Carrier ratio for tolerable interference—average of 10 observers viewing a 21-inch color receiver at an 8:1 viewing distance. The desired picture was the sailboat scene; the interfering picture was color bars.

considered tolerable at each offset frequency. Judgments were made for offset frequencies around 10,010 cycles and 20,020 cycles. The observers viewed a still picture on a 21-inch color receiver; the distance from the observer to the screen was eight times the picture height. Ten observers were tested. The average of their tolerable ratios (ratio of peak-of-sync level of desired picture to the peak-of-sync level of undesired picture) as a function of the offset frequency is given in Figures 1 and 2. Modulated r-f carriers were used for the tests. The modulations on the two carriers were completely independent.

Method-of-Equality Tests

In the method-of-equality test the observer had a standard reference picture for comparison with the test picture. He indicated

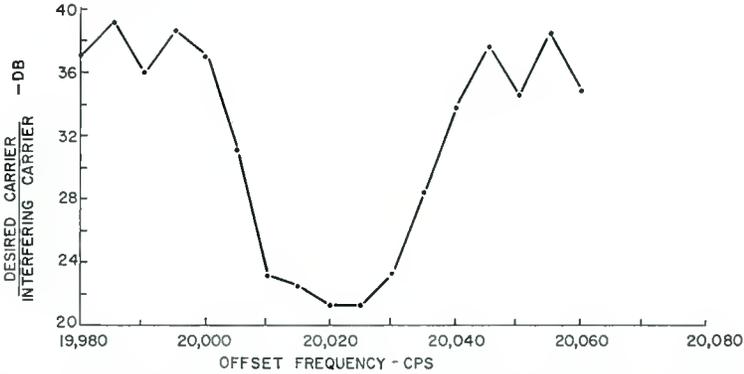


Fig. 2—Carrier ratio for tolerable interference—average of 10 observers viewing a 21-inch color receiver at an 8:1 viewing distance. The desired picture was the sailboat scene; the interfering picture was color bars.

when the pictures were equal in quality, i.e., equally degraded. The viewer was able to switch back-and-forth between the reference picture (with a fixed interference level) and the same scene with the interference at a different offset frequency and an adjustable interference level. Each viewer changed the scene back-and-forth as many times as he desired, as the interference level in the test picture was changed, until he had no preference. The improvement is then expressed as the difference, in decibels, between the signal levels of the two interfering signals.

The tests were made using four different receivers from several

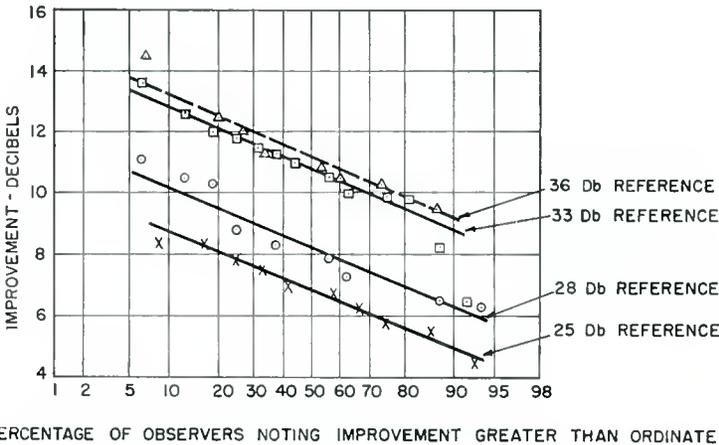
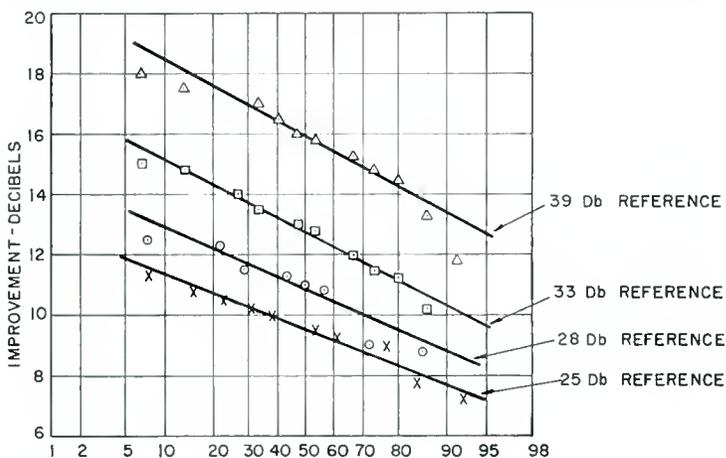


Fig. 3—Reduction in visibility of co-channel interference for a 10,010-cycle offset frequency compared to a 9,985-cycle offset. The desired picture was the sailboat scene; the interfering picture was color bars.

manufacturers—a 21-inch color receiver, a 21-inch black-and-white receiver, and two 17-inch black-and-white receivers. Viewing distance was eight times the picture height. Tests were made with offset frequencies of 10,010 and 20,020 cycles using an interfering picture consisting of color bars.

Figures 3 and 4 show the results obtained using slides. Twelve to fifteen observers were tested for each condition shown. Figures 5 and 6 show the results obtained using off-the-air motion pictures. Eight to ten observers were tested for each condition shown. The off-the-air pictures were detected from two completely independent color broadcasts and remodulated on the r-f carriers used for the subjective tests.



PERCENTAGE OF OBSERVERS NOTING IMPROVEMENT GREATER THAN ORDINATE

Fig. 4—Reduction in visibility of co-channel interference for a 20,020-cycle offset frequency compared to a 19,995-cycle offset. The desired picture was the sailboat scene; the interfering picture was color bars.

NEW TWO-STATION TOLERABLE-INTERFERENCE TESTS

10,010 Cycle and 20,020 Cycle Offsets

The earlier tolerable-interference tests did not include moving pictures. Therefore, such tests were made with moving pictures for offsets of 9,985, 10,010, 19,995 and 20,020 cycles. Two modulated r-f signals, accurately controlled at the various offset frequencies, were used for the tests. The desired picture was generated by the laboratory color-film equipment using the same film for all tests, and the modulation on the interfering carrier was a detected off-the-air color program. In a semi-darkened room, the observers viewed four different receivers in a random order at a viewing distance equal to eight times picture height. A 21-inch color receiver, a 21-inch black-and-white

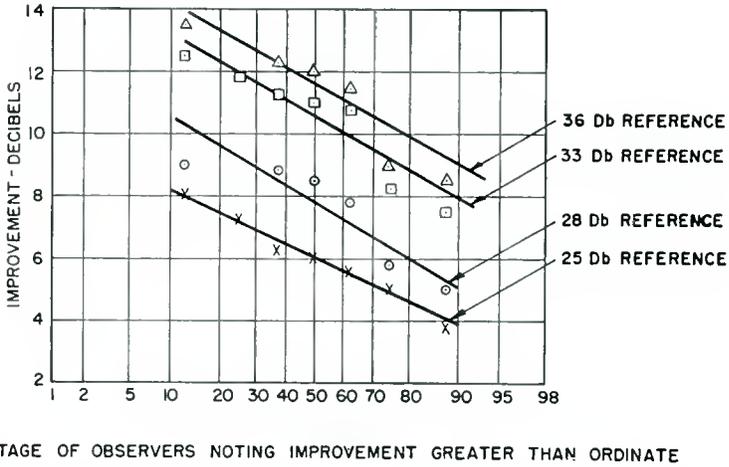


Fig. 5—Reduction in visibility of co-channel interference for a 10,010-cycle offset frequency compared to a 9,985-cycle offset. Desired picture was color program material from a commercial television station. The interfering picture was program material from a commercial television station.

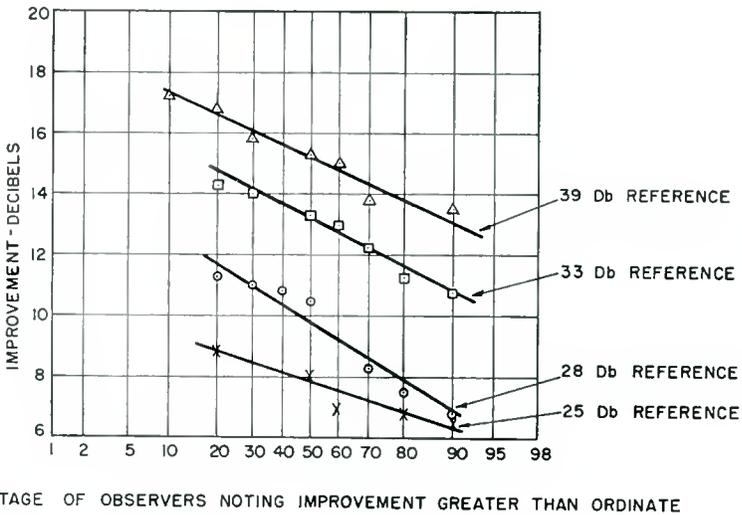


Fig. 6—Reduction in visibility of co-channel interference for a 20,020-cycle offset frequency compared to a 19,995-cycle offset. Desired picture was color program material from a commercial television station. The interfering picture was program material from a commercial television station.

receiver, and two 17-inch black-and-white receivers all of different manufacture were operated at a brightness of approximately 20 foot-lamberts.

Eleven observers were tested for the 9,985 and the 10,010 cycle offsets. The results given in Figure 7 show an improvement of 8.5 decibels for 50 per cent of the observers.

At 10,010 cycles the average tolerable ratio of the eleven observers, averaged for the four receivers, was 22 decibels. The averages for the three black-and-white receivers were within 0.5 of a decibel of the 22 decibels. The value for the color receiver was 20.5 decibels. A 31.4-decibel average was obtained for the 9,985-cycle offset. The averages for the black-and-white receivers were within 1 decibel of 31.4 decibels.

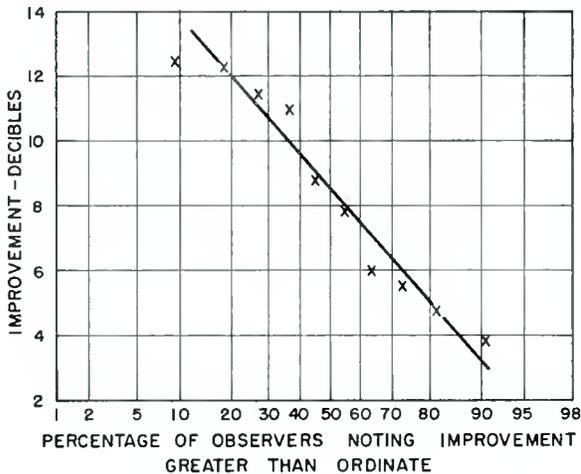


Fig. 7—Improvement in tolerable ratio for eleven observers (average of four receivers) for a 10,010-cycle offset compared with a 9,985-cycle offset. The desired picture was color film from a television studio. The interfering picture was commercial color program.

The value for the color receiver was 29.1 decibels. The results of the earlier¹ "method of equality" tests for moving pictures are given in Figure 5. If one interpolates for a 31.4 decibel reference, an approximate improvement of 9 decibels is obtained.

Fifteen observers were tested at the 19,995 and 20,020 cycle offsets. The results given in Figure 8 show an improvement of 14 decibels for 50 per cent of the observers. The average tolerable ratio of the fifteen observers averaged over the four receivers was 20.5 decibels at 20,020 cycles and 35.3 decibels at 20,045 cycles. At 20,020 cycles the average tolerable ratios for the three black-and-white receivers were within 0.5 decibel of the 20.5 decibels. The average ratio for the color receiver

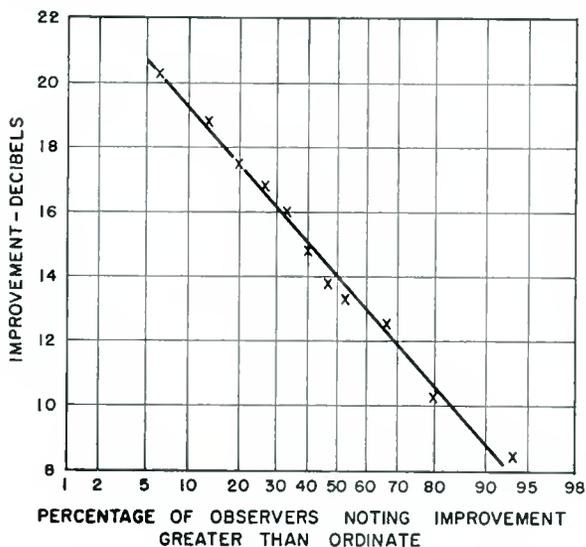


Fig. 8—Improvement in tolerable ratio for fifteen observers (average of four receivers) for a 20,020-cycle offset compared with a 19,995-cycle offset. The desired picture was color film from a television studio. The interfering picture was commercial color program.

was 19.5 decibels. At 20,045 cycles the average tolerable ratios for the three black-and-white receivers were within 1 decibel of 35.3 decibels. The value for the color receiver was 34.2 decibels.

Referring back to the earlier¹ method-of-equality tests for offsets of 19,995 and 20,020 cycles and for moving pictures (Figure 6), and interpolating for a 35.3 decibel ratio, it is seen that an approximate improvement of 14 decibels is obtained.

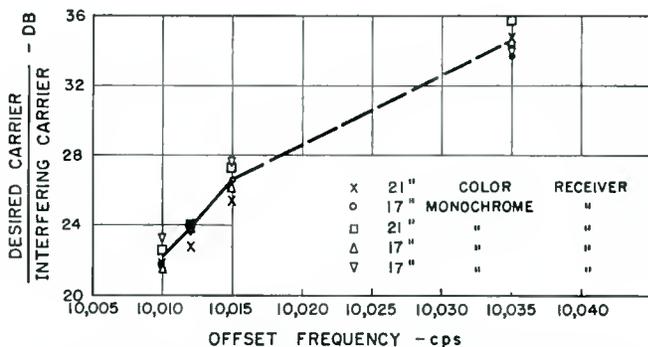


Fig. 9—Carrier ratio for tolerable interference. Average of ten observers. The desired picture was the sailboat scene; the interfering picture was color bars.

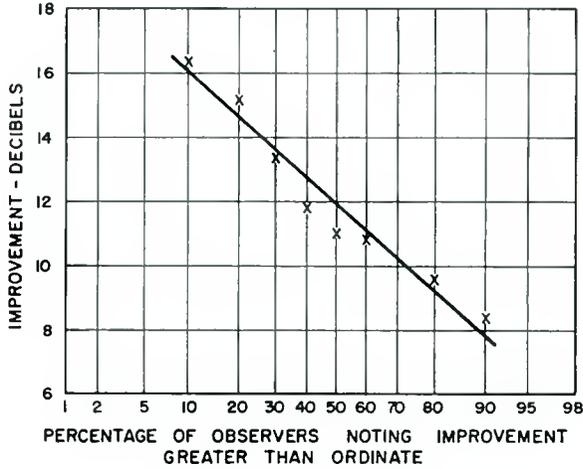


Fig. 10—Improvement in tolerable ratio for ten observers (average of five receivers) for a 10,010-cycle offset compared with a 10,035-cycle offset. The desired picture was the sailboat scene; the interfering picture was color bars.

Under the test conditions listed above, the earlier tolerable-interference tests were repeated—the desired picture was the “sailboat” scene used in the 1956 tests, and the interfering picture was color bars. Ten observers were tested for the 10 and 20 kilocycle offsets. The averages of the ten tolerable ratios, averaged for the five receivers, as a function of the offset frequency are given in Figures 9 and 11. The distributions—improvement versus percentage of observers—aver-

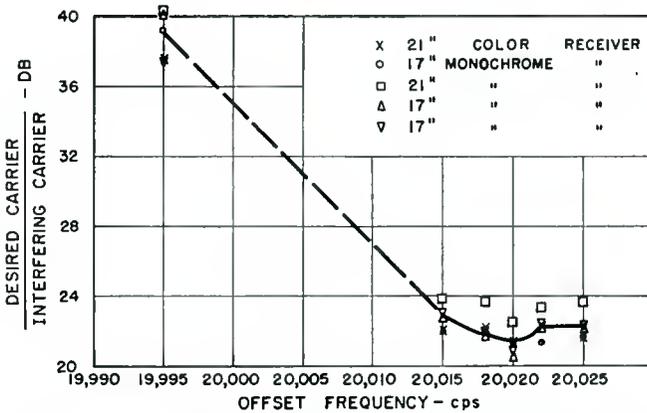


Fig. 11—Carrier ratio for tolerable interference. Average of ten observers. The desired picture was the sailboat scene; the interfering picture was color bars.

aged for the four receivers are given in Figures 10 and 12. According to these results, precise control provides 12 decibels and 17 decibels improvement for 50 per cent of the observers. The improvements obtained in the earlier tests, Figures 1 and 2, were 12 decibels and 18 decibels.

7,867-Cycle Offset

It was desirable to find out if an offset equal to half the horizontal line frequency (7,867 for color) was better than a 10,010-cycle offset. The test conditions were the same as for the previous two-station tests; however, only one person observed the pictures. The offset frequency

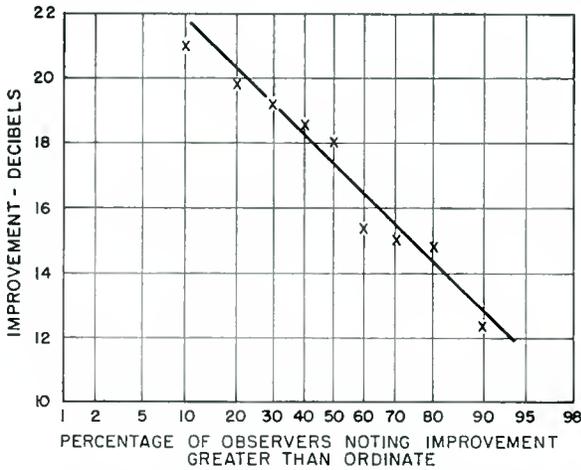


Fig. 12—Improvement in tolerable ratio for ten observers (average of five receivers) for a 20,020-cycle offset compared with a 19,995-cycle offset. The desired picture was the sailboat scene; the interfering picture was color bars.

was held as closely as possible to half the horizontal scanning frequency. The rasters on a few of the receivers had a slight movement for a 20-decibel signal ratio. A one-cycle change in the offset frequency caused a bad flicker that was still noticeable at a ratio of 30 decibels. Therefore, if the frame frequency and the frequency offset are accurately controlled, the 10,010 cycle offset appears to be the better offset.

THREE-STATION TOLERABLE INTERFERENCE TESTS

Tests were made to determine the effect of two co-channel transmitters interfering with the desired transmission. Test conditions were the same as for the previous tests. Three r-f carriers were accu-

rately controlled at the proper offsets, and each carrier was modulated by a completely independent video signal. The desired picture was the "sailboat scene," and the interfering pictures were two different sets of color bars.

Because of equipment limitations, all of the possible interfering conditions were not tested. A test was made for the desired station on-frequency and two interfering stations with offsets of plus 10 kilocycles and minus 10 kilocycles. The other test consisted of the desired station offset plus 10 kilocycles with one interfering station on frequency and the other interfering station offset minus 10 kilocycles. To reduce the number of subjective tests, the levels of the two interfering signals were maintained equal to each other and the levels were varied together.

With the desired station on-frequency, data was taken with the two interfering stations at the following offsets with respect to the desired station: 10,010 cycles and 10,010 cycles; 10,012 cycles and 10,008 cycles; 10,015 cycles and 10,005 cycles. The latter four offsets were determined by three observers as the most undesirable offsets for frequency variations of each of the three picture carriers of plus-or-minus 1.0 cycle and plus-or-minus 2.5 cycles. Tests were also made for the offsets producing the greatest degradation of the picture, 10,035 cycles and 9,985 cycles.

When the two interfering signals are at offsets slightly different from 10,010 cycles, a low-frequency beat is produced in the desired video signal. This beat causes the desired picture to flicker, and, at higher levels of the interfering signals, on some receivers the rasters move at a slow rate.

Ten observers were tested on five receivers for the above conditions. Figure 13 gives the average tolerable ratio of the peak-of-sync level of the desired signal to the peak-of-sync level of one of the undesired signals (the levels of the two interfering signals were equal) versus the offset frequencies. There is an improvement of 10 decibels for the 10,010-cycle offsets compared to the arrangement where one station is offset plus 10,035 cycles and the other minus 9,985 cycles. Improvements of 6 decibels and 5 decibels were measured for frequency stabilities of ± 1 cycle and ± 2.5 cycles on each of the three picture carriers.

The distribution of the observers—improvement versus percentage of observers noting improvement—is given in Figure 14.

An on-frequency station and a station offset minus 10 kilocycles simultaneously interfering with a station offset plus 10 kilocycles cause wide dark and light lines in the picture. The appearance of the lines is the same as that produced at high levels of interference by the

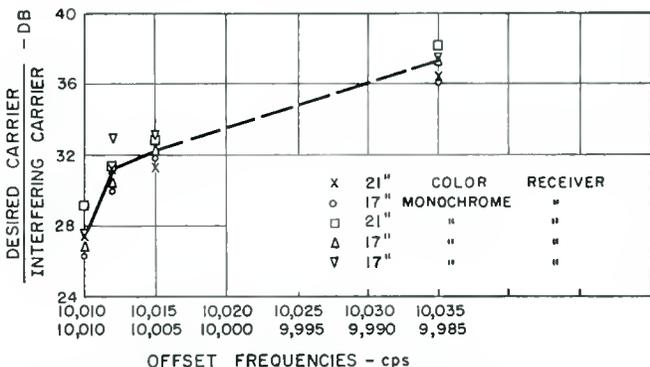


Fig. 13—Ratio of desired carrier to each of two interfering carriers for tolerable interference. Average of ten observers. The desired picture was the sailboat scene; the interfering picture was color bars.

second-order effects when one station offset 10 kilocycles interferes with an on-frequency station. These second-order effects are due mainly to the kinescope nonlinearity. For the three-station case, if the offsets are not exactly 10,010 cycles and 20,020 cycles the wide lines move through the picture.

Three observers were tested to determine the offset frequencies which produced maximum degradation of the picture for frequency changes in the picture carriers of ± 1 cycle and ± 2.5 cycles. Tests were then made at offsets which were 10,010 cycles and 20,020 cycles; 10,012

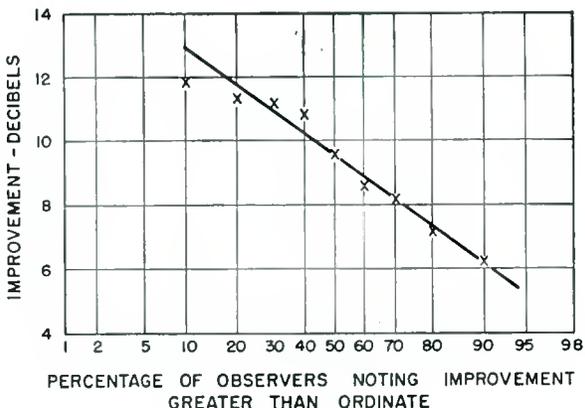


Fig. 14—Improvement in tolerable ratio for ten observers (average of five receivers) for two stations offset 10,010 cycles compared with offsets of 10,035 cycles and 9,985 cycles. The desired picture was the sailboat scene; the two interfering pictures were color bars.

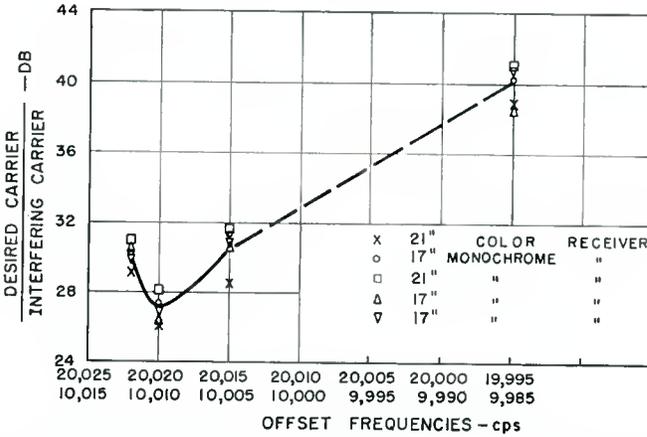


Fig. 15—Ratio of desired carrier to each of two interfering carriers for tolerable interference. Average of ten observers. The desired picture was the sailboat scene; the interfering pictures were color bars.

cycles and 20,022 cycles; 10,005 cycles and 20,015 cycles. The offsets of 9,985 cycles and 19,995 cycles were also tested.

Results for ten observers are given in Figure 15. An improvement of 13 decibels was obtained for offsets of 20,020 cycles and 10,010 cycles compared to offsets of 19,995 cycles and 9,985 cycles. For carrier

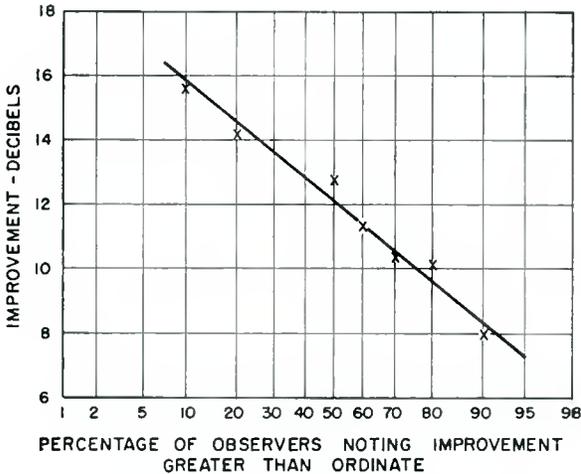


Fig. 16—Improvement in tolerable ratio for ten observers (average of five receivers) for two stations offset 20,020 cycles and 10,010 cycles compared with offsets of 19,995 cycles and 9,985 cycles. The desired picture was the sailboat scene; the two interfering pictures were color bars.

frequency stabilities of ± 2.5 cycles, the improvement is reduced to 10 decibels.

The improvement for offsets of 20,020 and 10,010 cycles compared to the 19,995 and 9,985 cycle offsets versus percentage of observers is given in Figure 16.

Referring to Figures 13 and 15 and to the results for two stations, Figures 1, 2, 9 and 11, one can see that the method of adding two simultaneously interfering signals to determine their subjective effect depends on the offset frequency. If the interfering stations are offset 10,010 cycles and 10,010 cycles or 10,010 cycles and 20,020 cycles, the two signals add approximately in a linear manner as indicated by the fact that the levels of the two interfering stations had to be reduced below the level of one station by 5 or 6 decibels to produce the same subjective effect. However, for the interfering stations offset 10,035 cycles and 9,985 cycles, the levels of the two interfering stations had to be reduced below the level of one interfering station by 3 or 4 decibels to produce the same subjective effect. When the interfering stations were offset 19,995 cycles and 9,985 cycles, it was necessary to reduce their levels by only 1 decibel below that of one interfering station.

An On-Frequency Station Interfering with an On-Frequency Station

The on-frequency tests were made to determine if it would be advantageous, in a network of several stations, to offset the on-frequency stations at some low offset frequency. The tests were not quantitative measurements of tolerable interference. The available laboratory setup could not provide two picture carriers at a small controllable frequency difference, so simulation was used. The simulation was not complete in that it was an analogue only for the case in which the interfering signal is an unmodulated carrier. It consisted of modulating a low-power transmitter with the sum of a video signal and a sine wave. The frequency of the sine wave was equal to the offset frequency. Observations were made by only one person.

At a signal ratio of 26 decibels, observations were made at offset frequencies from 0.5 cycle to 4,000 cycles. A color receiver and a black-and-white receiver of the same manufacture, and three black-and-white receivers all of different manufacture were viewed simultaneously. For a frequency tolerance on the offset frequency of ± 2 cycles, an offset of ± 2 cycles was better than any of the offset frequencies up to 4,000 cycles. The even multiples of frame frequency were better than the odd multiples. Offsets equal to an odd multiple of frame frequency produced a bad flicker on some of the receivers. A spot check of a few offset frequencies approximately half-way

between an odd and an even multiple of frame frequency showed these offsets caused a greater degradation than the odd multiple offsets.

Based on these tests, one might conclude that on-frequency operation is better than operating at a low-frequency offset. However, the extrapolation of these results to the case of color transmissions received on color receivers is of questionable value. It was possible to synchronize the picture carriers, and tests were made using independent modulations on the two picture carriers. Ten observers were tested for tolerable interference. A color receiver and two black-and-white receivers were used. The desired picture was the "sailboat scene" and the interfering picture was color bars. Tests were made for the following three conditions: (1) The synchronizing signals for each of the video signals were locked to independent color subcarriers. The frequency difference between the color subcarriers was 17 cycles. (2) The conditions were the same except that the color subcarriers differed in frequency by one cycle. (3) The sync generator of the interfering signal was locked to the power line and the frequency difference between the color subcarriers was approximately one cycle. The results of these tests are given in Table I from which it is seen that when the frame frequencies are different, the results check approximately with the 32.4 decibels obtained by JTAC, in 1949, for synchronous operation.² The 44.4 decibels obtained for the color receiver, for the condition of two color signals with color subcarriers differing by 1 cycle, and the sync generators locked to the color subcarrier is no worse than the 44.6 decibels found by JTAC for nonsynchronous operation.

For the conditions of test number 2, at high levels of interference, the color picture changed from monochrome to color at a slow rate. At low levels of interference, the red portion (the man's red cap against a blue background) of the picture changed in intensity and saturation at a slow rate. For the monochrome receivers the observers required only a 20-decibel ratio for tolerable interference. If the interfering modulation had been moving pictures the tolerable ratio would probably have been larger.

Tests have been made at the Federal Communications Commission Laboratory³ by means of simulation. The simulation consisted of adding, at the video amplifier of the receiver, a low-frequency sine wave to the video signal. These tests are more complete in that quantitative measurements were made with five observers at various offset

² "Comments on the Proposed Allocations of Television Broadcast Services," *Proceedings of JTAC*, Vol. 4, September 26, 1949.

³ L. C. Middlekamp, "Reduction of Co-channel Television Interference by Very Precise Offset Carrier Frequency," *Trans. I.R.E. PGBTS*, December, 1958.

frequencies from 150 cycles to 450 cycles. The results show an improvement, at an even multiple of frame frequency, of 14 decibels for a stability of ± 1 cycle of the offset frequency, or a stability of ± 0.5 cycle of the picture carriers. This is a stability of 6 parts in 10^9 for the low VHF channels and 2.5 parts in 10^9 for the high VHF channels. These stabilities are easily attained with commercially available oscillators. At 890 megacycles this is a stability of 6 parts in 10^{10} . The units which are installed at WRCA-TV (New York) and WRC-TV (Washington, D. C.) were tested at the laboratory for a three-day period. There were 12-hour periods when the ambient temperature stayed within 2 degrees; during these periods the relative

Table I—Tolerable Ratio (Average of Ten Observers) for Synchronous Carrier Operation

TEST CONDITIONS	Color Receiver 21-inch	Monochrome Receiver "A" 21-inch	Monochrome Receiver "B" 17-inch
Sync generators locked to color subcarriers. Frequency difference between color subcarriers equal to 17 cycles.	33 db	23.6 db	27.3 db
Same as above except for a 1 cycle frequency difference between the color subcarriers.	44.4 db	19.5 db	20.5 db
Sync generator of desired signal locked to the color subcarrier. Sync generator of interfering signal locked to the power line. Frequency difference between color subcarriers equal to 1 cycle.	39.3 db	29.0 db	30.1 db

stability was at least 3 parts in 10^{10} , which was the accuracy of the method of measurement. Better temperature stability could have been achieved if the units had been provided with a fine adjustment to align the operating temperature with the crystal-frequency-versus-temperature characteristic. The oven bridge used for controlling the oven temperature was designed to operate at approximately the nominal operating temperature specified by the crystal manufacturer.

Crystal oscillators have been made⁴ with a relative stability of 1 part in 10^{10} per day. Presently available atomic standards have a stability of 5 parts in 10^{10} .

⁴ Norman Lea, "A Quartz Servo Oscillator," *Proc. I.R.E.*, Vol. 46, p. 1835, November, 1958.

L. C. Middlekamp³ points out the possibility of offsetting the on-frequency stations at a low-frequency difference equal to an even multiple of frame frequency, and if this is done then all stations will be offset at a frequency equal to an even multiple of frame frequency.

As pointed out previously, the on-frequency tests have been too limited to make quantitative conclusions for color transmissions received on color receivers. However, if precise frequency control is applied to the present television system and the on-frequency stations are operated on frequency and on color frame frequency, there would be a large improvement on all receivers for monochrome transmissions. There would be an improvement for color transmissions on the monochrome receivers and more complete tests may show an improvement on the color receivers. However, it appears the on-frequency stations should be offset at a low frequency equal to an even multiple of frame frequency.

RCA TECHNICAL PAPERS†

First Quarter, 1959

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

"Application of RCA Drift Transistors to FM Receivers," J. W. Englund and H. Thanos, <i>Trans. I.R.E. PGBTR</i> (January)...	1959
"The Compatibility of Value Engineering Analysis and Reliability," R. M. Jacobs, <i>The Journal of Industrial Engineering</i> (January-February)	1959
"Detection of Asymmetric Sideband Signals in the Presence of Noise," T. Murakami and R. W. Sonnenfeldt, <i>Trans. I.R.E. PGBTR</i> (January)	1959
"A Feedback Light-Amplifier Panel for Picture Storage," B. Kazan, <i>Proc. I.R.E.</i> (January)	1959
"Focusing Traveling Wave Tubes," D. J. Blattner and F. E. Vaccaro, <i>Electronic Industries</i> (January)	1959
"Four-Terminal Parametric Amplifier," K. K. N. Chang, <i>Proc. I.R.E.</i> (January) (Letter to the Editor)	1959
"Method for Determining Specific Cooling Rates of Plate Materials in Vacuum," C. W. Horsting, I. S. Solet, T. A. Sternberg, and Coauthor, <i>Trans. I.R.E. PGED</i> (January)	1959
"Microphone Technique," Paul Taylor, <i>Trans. I.R.E. PGEWS</i> (January)	1959
"Millimicrosecond Microwave Ferrite Modulator," A. H. Solomon and F. Sterzer, <i>Proc. I.R.E.</i> (January) (Letter to the Editor)	1959
"Reversible Variations of Sensitivity in Certain Cadmium Sulfide and Selenide Photoconductors," R. H. Bube, <i>Jour. Chem. Phys.</i> (January)	1959
"Theoretical Research on Tropospheric Scatter Propagation in the United States, 1954-1957," H. Staras and Coauthor, <i>Trans. I.R.E. PGAP</i> (January)	1959
"Traveling-Wave Cylindrical Antenna Design—A Graphical Synthesis Method," P. Foldes, <i>Trans. I.R.E. PGAP</i> (January)...	1959
"Acoustoelectric Effect," R. H. Parmenter, <i>Phys. Rev.</i> (January 1) ..	1959
"Electrostatically Focused Traveling-Wave Tube," D. J. Blattner and F. E. Vaccaro, <i>Electronics</i> (January 2)	1959
"Magnetic Demodulators for Color TV," M. Cooperman, <i>Electronics</i> (January 2)	1959
"A Design Review Will Catch Those Boners," D. P. Simonton, <i>Product Engineering</i> (January 5)	1959
"Anodic Dissolution of Magnesium Alloys in Aqueous Salt Solutions," R. Glicksman, <i>Jour. Electrochem. Soc.</i> (February)	1959
"Are Plateaus Significant in Scintillation Counting," R. W. Engstrom and J. L. Weaver, <i>Nucleonics</i> (February)	1959
"Automatic Control of Video Tape Equipment at NBC, Burbank," R. W. Byloff, <i>Trans. I.R.E. PGB</i> (February)	1959
"Color TV Service: Fact and Fiction," J. Mohan, <i>Radio & TV News</i> (February)	1959
"Effects of Carrier Injection on the Recombination Velocity in Semiconductor Surfaces," G. C. Dousmanis, <i>Jour. Appl. Phys.</i> (February)	1959

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

- "High-Voltage Photovoltaic Effect," B. Goldstein and L. Pensak, *Jour. Appl. Phys.* (February) 1959
- "Levitation Melting of Ga, In, Au, and Sb," L. R. Weisberg, *Rev. Sci. Instr.* (February) (Notes) 1959
- "Quick-Selection Chart for Rectifier Tubes and Circuits," RCA *Application Note AN-179*, Electron Tube Division, Radio Corporation of America, Harrison, N. J. (February) 1959
- "RCA Integrates Value Engineering, Reliability Programs," R. M. Jacobs, *Production* (February) 1959
- "Use of RCA Power Pentodes 6EH5, 12EH5, 25EH5, and 50EH5 in Low-Cost Audio Amplifiers," RCA *Application Note AN-180*, Electron Tube Division, Radio Corporation of America, Harrison, N. J. (February) 1959
- "The Weekend Special," L. Aurick, RCA *Ham Tips* (February).... 1959
- "Fine Structure, Hyperfine Structure and Relaxation Times of Cr^{3+} in TiO_2 (Rutile)," H. J. Gerritsen, S. E. Harrison, H. R. Lewis, and J. P. Wittke, *Phys. Rev. Letters* (February 15) 1959
- "TV Sound Detector Uses Drift Transistor," M. Meth, *Electronics* (February 20) 1959
- "Effect of Oxide Impurities on the Thermoelectric Powers and Electrical Resistivities of Bismuth, Antimony, Tellurium, and Bismuth-Tellurium Alloys," R. A. Horne, *Jour. Appl. Phys.* (March) 1959
- "The Exponential Gun—A Low-Noise Gun for Traveling-Wave-Tube Amplifiers," A. L. Eichenbaum and R. W. Peter, *RCA Review* (March) 1959
- "Gain-Bandwidth Product for Photoconductors," A. Rose and M. A. Lampert, *RCA Review* (March) 1959
- "Gains, Response Times, and Trap Distributions in Powder Photoconductors," H. B. DeVore, *RCA Review* (March) 1959
- "Improving Stability of AM Directional Antenna Systems," G. H. Brown, *Elec. Eng.* (March) 1959
- "The Influence of Firing Conditions, Composition, and Screening Media on the Zeta Potentials of ZnS Phosphors as Related to Adherence to Glass Surfaces," B. Levy, *Jour. Electrochem. Soc.* (March) 1959
- "Introduction to Junction Transistors, Part II," R. N. Hurst, *Broadcast News* (March) 1959
- "Logical Control of Sampling Saves Computing Time," D. Hammel, *Control Engineering* (March) 1959
- "Magnetics for Computers—A Survey of the State of the Art," J. A. Rajchman, *RCA Review* (March) 1959
- "The Megacoder—A High-Speed, Large-Capacity Microminiature Decoder for Selective Communication," H. Kihn and W. E. Barnette, *RCA Review* (March) 1959
- "Merit Rating and Productivity in an Industrial Research Laboratory: A Case Study," A. G. Grasberg, *Trans. I.R.E. PGEM* (March) 1959
- "New 11-KW VHF Television Transmitter, Type TT-11AH," H. E. Small, *Broadcast News* (March) 1959
- "New 5-KW FM Transmitter Features Unique RF Circuits," D. Bain and J. A. Aurand, *Broadcast News* (March) 1959
- "New 5,000-Watt AM Transmitter," J. Novik and I. E. Skarbek, *Broadcast News* (March) 1959
- "Nonlinear-Capacitance Amplifiers," L. S. Nergaard, *RCA Review* (March) 1959
- "Periodic Electrostatic Focusing of Laminar Parallel-Flow Electron Beams," W. W. Sickanowicz and F. E. Vaccaro, *Proc. I.R.E.* (March) (Letter to the Editor) 1959
- "The Physical Principles of a Negative-Mass Amplifier," H. Krömer, *Proc. I.R.E.* (March) 1959
- "Preparation and Properties of the Systems $\text{LnFe}_x\text{Cr}_{1-x}\text{O}_3$ and $\text{LaFe}_x\text{Co}_{1-x}\text{O}_3$," W. Croft and Coauthor, *Jour. Phys. Chem.* (March) (Notes) 1959

"Properties of Deep Traps Derived from Space-Charge-Current Flow and Photoconductive Decay," R. W. Smith, <i>RCA Review</i> (March)	1959
"RCA Automation Equipment for TV Stations," R. P. Major, <i>Broadcast News</i> (March)	1959
"RCA Production Model TV Tape Recorder Highlights Latest Advancements," H. H. Klerx, <i>Broadcast News</i> (March)	1959
"Recruiting the Technically Creative — A Dual Responsibility," J. Kurshan, <i>Trans. I.R.E. PGEM</i> (March)	1959
"Simple General Analysis of Amplifier Devices with Emitter, Control, and Collector Functions," E. O. Johnson and A. Rose, <i>Proc. I.R.E.</i> (March)	1959
"Some Properties of the Gravitation Field and Their Possible Application to Space Navigation," J. C. Crowley, S. S. Kolodkin and A. M. Schneider, <i>Trans. I.R.E. PGSET</i> (March)	1959
"Transistor Dissipation Ratings for Pulse and Switching Service," RCA <i>Application Note AN-181</i> , Semiconductor and Materials Division, Radio Corporation of America, Somerville, N. J. (March)	1959
"Transistorized Switching Systems Type TS-40, Part I," J. W. Wentworth, C. R. Monro, and A. C. Luther, Jr., <i>Broadcast News</i> (March)	1959
"Traveling Wave Antenna Proves Itself in Service," G. E. Erwin, <i>Broadcast News</i> (March)	1959
"Two New Automatic Turntables," G. C. Weilenmann, <i>Broadcast News</i> (March)	1959
"Unified Representation of Junction Transistor Transient Response," A. Harel and J. F. Cashen, <i>RCA Review</i> (March)	1959
"Use of the 6EM5 in Vertical-Deflection Circuits of 110° Systems," RCA <i>Application Note AN-182</i> , Electron Tube Division, Radio Corporation of America, Harrison, N.J. (March)	1959
"Photoconductor Performance, Space-Charge Currents, and the Steady-State Fermi Level," A. Rose and M. A. Lampert, <i>Phys. Rev.</i> (March 1)	1959
"Transient Behavior of the Ohmic Contact," M. A. Lampert and A. Rose, <i>Phys. Rev.</i> (March 1)	1959
"Magnetic Susceptibility of Photogenerated Current Carriers in Germanium," J. O. Kessler and A. R. Moore, <i>Phys. Rev. Letters</i> (March 15)	1959
"Grinding L-Shaped Bevels," V. E. Califano, <i>American Machinist</i> (March 23) (Reference Book Sheet)	1959
"Elektronik," J. T. Wallmark, <i>Ingenjörshandboken</i> , Stockholm	1959
"High Electric Field Effects in n-Indium Antimonide," M. C. Steele and M. Glicksman, <i>Journal of the Physics and Chemistry of Solids</i> , Vol. 8	1959
"Intermetallic Semiconductors Containing Alkali Metals," W. E. Spicer and A. H. Sommer, <i>Journal of the Physics and Chemistry of Solids</i> , Vol. 8	1959
"Intervalence Band Transitions in Gallium Arsenide," R. Braunstein, <i>Journal of the Physics and Chemistry of Solids</i> , Vol. 8 ..	1959
"Lattice Vibrational Spectrum of Germanium," F. Herman, <i>Journal of the Physics and Chemistry of Solids</i> , Vol. 8	1959
"The Magnetoresistance of Electrons in InP and GaAs," M. Glicksman, <i>Journal of the Physics and Chemistry of Solids</i> , Vol. 8 ..	1959
"Space-Charge-Limited Currents as a Technique for the Study of Imperfections in Pure Crystals," M. A. Lampert, A. Rose, and R. W. Smith, <i>Journal of the Physics and Chemistry of Solids</i> , Vol. 8	1959
"Thermal Conductivity of Germanium in the Temperature Range 300° — 1080°K," B. Abeles, <i>Journal of the Physics and Chemistry of Solids</i> , Vol. 8	1959

AUTHORS



of the Microwave
Institute of Radio
of Physicists.

MORREL P. BACHYNSKI received the B.E. degree in Engineering Physics in 1952, and the M.S. degree in 1953, both from the University of Saskatchewan, and the Ph.D. degree from McGill University in 1955. Until October 1955, as a member of the staff of the Eaton Electronics Research Laboratory, McGill University, he was engaged in investigations on aberrations in microwave lens systems. Since that time he has been with the RCA Victor Research Laboratory, Montreal, Canada, concerned primarily with short-wave propagation and plasma physics problems. He is now Associate Laboratories Director Laboratory. Dr. Bachynski is a Senior Member of the Engineers and a member of the Canadian Association

WALTER R. BEAM received the B.S. degree in Electrical Engineering from the University of Maryland. He pursued graduate work while serving as Instructor in Electrical Engineering at the same university, and received the M.S. degree in 1950. During part of this period he was also engaged in development of radio sounding instruments at Washington Institute of Technology. He joined the Microwave Tube group at RCA Laboratories in 1952. From 1952 to 1956 he was engaged in research on microwave amplifiers, electron guns and electron beam noise. In 1956 he became Manager of the Microwave Advanced Development activity of RCA Tube Division in Princeton. In June 1959, he joined the staff of Rensselaer Polytechnic Institute, Troy, N. Y., as a Professor and Head of the Electrical Engineering Department. Dr. Beam is a member of the Institute of Radio Engineers, Tau Beta Pi, and Sigma Xi, and is a licensed Professional Engineer.



WILLIAM L. BEHREND received the B.S. and M.S. degrees from the University of Wisconsin in 1946 and 1947. During 1953-1955, he did graduate work in Systems Engineering at the University of Pennsylvania. He served as an electronic technician in the U. S. Navy from 1944 to 1946. In 1947 he joined the RCA Laboratories Division, Princeton, N. J., where he is now with the Systems Research Laboratory. He has been associated with the development of antennas, experimental UHF television transmitters, and color television. Mr. Behrend is a member of Sigma Xi and a Senior Member of the Institute of Radio Engineers.



K. K. N. CHANG received the B.S. degree from the National Central University, Nanking, China, in 1940; the M.S. degree in Electrical Engineering from the University of Michigan in 1948, and the D.E.E. degree in 1954 from the Polytechnic Institute of Brooklyn. From 1940 to 1945, he was associated with the Central Radio Manufacturing Works, Kunming, China, working on radio receivers and from 1945 to 1947, he was a radio instructor in the Office of Strategic Service, U. S. Army, China Theatre. Since 1948, he has been at RCA Laboratories, Princeton, N. J., where he has engaged in research on

magnetrons, traveling-wave tubes, beam-focusing devices, and parametric amplifiers. Dr. Chang is a member of Sigma Xi.

JULES M. FORMAN received the M.E. degree from Stevens Institute of Technology in 1940. He was a Teaching-Fellow at Stevens from 1940 to 1941 and attended Newark College of Engineering in 1942. He joined the life Test and Data Laboratory of the RCA Electron Tube Division in Lancaster, Pa. in 1941, and has since worked on the design of complex electro-mechanical electronic test sets, life-test equipments, and special applications of electronic circuitry for electron tubes. He became an engineering leader of Special Equipment Engineering in 1951. Since 1956 his activity has been directly responsible for environmental engineering testing and evaluation of all new and improved tube ruggedization developments in the Lancaster Engineering Laboratory. Mr. Forman is a member of the Institute of Radio Engineers and the Institute of Environmental Engineers, and holds a professional engineering license in the state of Pennsylvania in electrical engineering.



NORMAN R. GOLDSTEIN received the B.S. degree in Chemical Engineering from Northeastern University in 1951 and the M.S. degree in Physics from Franklin and Marshall College in 1956. He joined the RCA Electron Tube Division in Lancaster, Pa. as a development engineer in 1951, and was concerned with the design and development of power and gas tubes until 1954. Since that time, he has worked on phosphor-screen problems in connection with the development of color picture tubes. Mr. Goldstein is a member of Sigma Pi Sigma.

JACK HILIBRAND received the B.S.E.E. degree from the City College of New York in 1951. He was a National Science Foundation Fellow at the Massachusetts Institute of Technology for two years and a Research Assistant at the Research Laboratory of Electronics for two years. In 1956 he received the Sc.D. degree from M.I.T. for work in the analysis of excess physical noises. Since graduation he has been a Member of the Technical Staff of RCA Laboratories in Princeton, N. J., where he is engaged in semiconductor device design and analysis. He also holds the rank of Lecturer at the Graduate Division of the School of Technology, City College of New York. Dr. Hilibrand is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.





GEORGE P. KIRKPATRICK received the B.S. degree in Physics in 1941 from the College of the City of New York and the M.S. degree in Physics in 1945 from Brooklyn Polytechnic Institute. He worked from 1941 to 1944 for the Electrical Testing Laboratories. He was a Research Associate in Physics at Brooklyn Polytechnic Institute during 1944-1945, and was an Instructor in Physics (evenings) from 1945 to 1948. He worked as a physicist for U. S. Radium Corporation from 1945 to 1951. Since 1951 he has been with the Chemical and Physical Laboratory of the RCA Electron Tube Division at Lancaster, Pa. Until January 1, 1959, he specialized in problems in colorimetry, spectroradiometry, and secondary-emission phenomena. Since the beginning of 1959, he has been engaged in a study of the outgassing and gettering properties of electron-tube materials.

HERBERT NELSON received the B.S. degree from Hamline University in 1927 and the M.S. degree in Physics from the University of Minnesota in 1929. From 1929 to 1930 he was an engineer with the Westinghouse Lamp Company in Bloomfield, New Jersey. He transferred to RCA in 1930 and was engaged in electronic research and development at the Tube Plant in Harrison, N. J. until 1953 when he joined the semiconductor activity at the RCA Laboratories in Princeton. Mr. Nelson is a member of Sigma Xi and of the American Physical Society.



LEON S. NERGAARD (see *RCA Review*, Vol. XX, No. 1, March 1959, p. 187.)



FRITZ PASCHKE received the degree of Diplom-Ingenieur in 1953 and the Dr. techn. sc. in 1955 from the Technical University Vienna. Between 1953 and 1955 he was a research assistant at the Institute of High-Frequency Techniques in Vienna. Since April 1956, Dr. Paschke has been a member of the technical staff at RCA Laboratories, Princeton, N. J.

ROBERT A. SCHMELTZER received the B.S. degree from Cooper Union, New York City, in 1955, and the M.S. degree from Stevens Institute of Technology in 1958. From 1955 to 1958 he was engaged in application engineering first with the Electron Tube Division and then the Semiconductor and Materials Division of the Radio Corporation of America. In 1957 he served 6 months in the U. S. Army doing research and development work at the Signal Corps Engineering Laboratories at Fort Monmouth, New Jersey. He is presently on leave of absence from RCA as the recipient of the David Sarnoff Fellowship for the second consecutive year, and is studying full time for the Doctor of Engineering Science degree at New York University.





